De-humanizing Care: An Ethnography of Mental Health Artificial Intelligence

by

Valerie E. Black

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Anthropology

and the Designated Emphasis

in

Science and Technology Studies

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Cori Hayden, Co-chair
Professor Karen Nakamura, Co-chair
Professor Lawrence Cohen
Professor Massimo Mazzotti

Spring 2023

© 2023 Valerie E. Black

Abstract

De-humanizing Care: An Ethnography of Mental Health Artificial Intelligence

by

Valerie E. Black

Doctor of Philosophy in Anthropology

Designated Emphasis in Science and Technology Studies

University of California, Berkeley

Professor Cori Hayden, Co-chair

Professor Karen Nakamura, Co-chair

What is it like to live in a world shaped by the axiom that there's not enough care for all who need it—one in which artificially intelligent caregivers become, accordingly, not only thinkable, but necessary? This dissertation provides an ethnographic account of the turn to artificial intelligence in mental health care via a Silicon Valley startup ("The Startup") that creates therapeutic chatbots. These bots offer on-demand, interactive emotional support to their human users through text (and in some instances, voice) messages. Though the story of AI in most medical applications is one of outpacing human ability while operating under direct (human) oversight, here the opposite is true: none of my interlocutors regard AI as superior in any way to a human therapist; and while my fieldsite's team of psychologists monitor aspects of these exchanges, to oversee each interaction would nullify the very justifications that summoned these bots into health care—namely, their rapid scalability, low overhead costs, and 24/7 on-demand, discreet availability. Yet for all that AI isn't a superior caregiver, it is nevertheless a caregiver—an identity that I explore by recognizing chatbots, startup workers, and end-users alike as ethnographic subjects. This dissertation examines the relationships that arise through and with AI care bots: I argue that the ontological ambiguity of human-bot bonds is not a threat to good care but the means of it; I examine the collaboration between AI and human care workers as one of colleagues who share a job requirement of being paradoxically expendable-yet-essential; and finally, I show how AI caregivers afford a glimpse of suicide prevention care unpaired from the ableism of "cure," even as they reveal how surveillance and neglect in mental health care become mutually reinforcing. Across this dissertation, I aim to recenter disability in the theorization of human-technology relationships.

Table of Contents

Abstract 1

Chapter 1: De-humanizing Care..1
    I) Not enough care to go around..2
    II) From crisis to startup...5
    III) AI is a caregiver (but not a replacement).......................................9
    IV) Care is already automated...14
    V) Not-so-cutting-edge therapeutic chatbots 17
    VI) AI-human relationships...21
    VII) Disability: analytic and method...24
    VIII) Doing fieldwork in Silicon Valley 27
    IX) Overview of chapters...30
    X) Who is this dissertation for?...31

Chapter 2: Relationships, Automated...33
    I) Introduction: Ontological ambiguity 33
    II) The therapeutic relationship and the archetypal bot.......................35
    III) Duped..39
    IV) Different..49
    V) Imagined..55
    VI) Conclusion - Part 1 59

Chapter 3: Displaced Care..64
    I) Introduction: Seeking to be displaced 64
    II) The paradox of care work...68
    III) The culture workshop 76
    IV) The ghost of ghost work past...82
    V) Everyone's displaceable..88
    VI) Conclusion - Part 2 95

Chapter 4: [USERNAME], Can You Tell Me if You Have a Suicide Plan? 100
    I) Introduction: More-than-human crisis care 100
    II) SUI 103
    III) On what SUI fails to automate...111
    IV) Witnessed-and-not-witnessed suicidal ideation 117
    V) Nonhuman attention...126
    VI) Conclusion - Part 3 133

Epilogue: Learning to be Ethical from AI..135

Bibliography...143

Chapter 1: De-humanizing Care

This dissertation provides an ethnographic account of the turn to artificial intelligence (AI) in mental health care. My fieldsite is a startup in Silicon Valley that uses AI to deliver on-demand therapeutic mental health care to thousands of end-users simultaneously. Instead of exchanging text messages with a human therapist or counselor, these users are talking with a chatbot. But unlike AI manufacturing, radiology, and taxi driving, there is no vision (or spectre) of "replacement" animating the expansion of AI therapeutic care.[1] No one in the AI care industry, including the psychology-trained workers who carefully write and evaluate the dialogue for these bots, expects for AI to ever *replace* mental health workers.[2] When fielding questions from curious outsiders, the professionals who make, sell, and acquire therapeutic chatbots are quick to gently correct any who wonder if AI will usurp human caregivers by explaining that a chatbot will never equal, let alone surpass, the abilities of a human therapist or counselor. Even the most devoted of end-users do not seem to regard therapeutic chatbots as "better" than human caregivers. That this industry thrives might therefore seem to point to a mystery no sooner named than solved as capitalism: mental health chatbots are instantly and all but universally available at a fraction of the cost of a therapist. This, undeniably, is the rationale that has brought this industry into existence; but as I intend to show, AI chatbot caregivers, like cyborgs, are both loyal and disloyal to their capitalist origin story.[3] In the chapters that follow, I will attempt to bring into focus why an instance in which AI is uncontestedly "not better" than a human is critical to reconceiving the terms by which we understand AI, work, and care.

This first chapter is both a field guide to de-humanizing care and an introduction to my dissertation. Importantly, "de-humanizing" differs from dehumanizing, though the two concepts are linked: care that is de-humanizing—i.e., that features nonhuman entities appointed as caregivers—wouldn't exist without the threat of dehumanization in care. Both start from the premise of scarcity: too many (people needing care) and too few (caregivers). But following disability studies scholar Eunjung Kim's (2015) contemplation of an anti-ableist, queer ethics of

---

[1] See for instance: Acemoglu and Restrepo, "Jobs: Evidence from US Labor Markets," 2020; Reaerdon,"Rise of Robot Radiologists," 2019: https://www.nature.com/articles/d41586-019-03847-z; Bellan and Liao, "Pony.ai scores taxi license for autonomous vehicles in Guangzhou," 2022: https://techcrunch.com/2022/04/25/pony-ai-scores-taxi-license-for-autonomous-vehicles-in-guangzhou/.

[2] This was the case during the core period of this research (between 2018 and 2020); yet I would contend that this remains true within the AI care industry, even as the expansion of generative AI chatbots reinvigorates discourse on the question of worker replaceability across an array of professions. Notably, this observation does not in any way contravene the fact that some therapists *outside* of this industry have—before, during and after my fieldwork—expressed concern about the possibility of being "replaced" by AI.

[3] Haraway, "A Cyborg Manifesto," 1991 [1985], 151: "The main trouble with cyborgs, of course, is that they are the illegitimate offspring of militarism and patriarchal capitalism, not to mention state socialism. But illegitimate offspring are often exceedingly unfaithful to their origins."

"inhumanism,"[4] I explore the possibility that care provided by nonhumans does not necessarily signal the dehumanization of those who receive it.

In this chapter, I establish why "what makes AI care work?"—and not "can AI care?"—is the question that I seek to answer in this dissertation. An "on-the-ground" understanding of the relationships between three groups—therapeutic chatbot end-users, psychologist startup workers, and the AI-powered chatbots themselves—is key to answering this question. But first, to understand what makes AI care work requires bringing into view what it is about mental health care that *isn't* working, and how this creates an opening not just for AI, but specifically for AI in the context of the high-risk, high-reward startup sector.

I) Not enough care to go around

Prior to the COVID-19 pandemic, public health experts had already proclaimed that rising caregiver scarcity—the fact of not having enough care professionals to meet with, examine, prescribe or administer medicine to, operate on, monitor, bathe, feed, comfort, or listen to all those in need of these services—had reached a state of global crisis.[5], [6] This shortage is not accounted for by the fact of unpaid care work, and unpaid caregiving is not enough to offset this shortage.[7] Simply put: we live in a world where there is not enough care to go around. Care scarcity has, in other words, become a total social fact (Mauss 2016 [1925]).[8]

---

[4] On queer inhumanism(s), see the *GLC* issue (2015, vol 21, 2–3) in which Kim's article is published, edited by Mel Y. Chen and Dana Luciano (and in particular, "Theorizing Queer Inhumanism" by José Esteban Muñoz et al; see also: Chen 2012). Notably, literary scholar Jennifer Rhee, in *The Robotic Imaginary* (2018), also draws on queer inhumanism in her theorization of dehumanization with respect to care labor and technology. For Rhee, via Jackson (2015——the same *GLC* issue), queer inhumanism provides a foundation for challenging the turn to the "inhuman" as a means by which to "move beyond" the human without considering who gets left behind. However, Kim's approach is precisely a call to consider who has already been left behind—and a refusal to leave or be left based on the restricted designation of "human." This move is key to my differentiation between de-humanized and dehumanized care.

[5] *The Lancet*, "Global Elder Care in Crisis," 2014: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(14)60463-3/fulltext;
Britnell, "Human: Solving the Global Workforce Crisis in Healthcare," 2019; Osterman, "Who Will Care for Us: Long-term Care and the Long-Term Workforce," 2017; Gupta, "The most important job in the world is one no one wants anymore," 2018: https://qz.com/1490065/the-most-important-job-in-the-world-is-one-no-one-wants-anymore/; Mauterstock, "The Caregiver Crisis Is Here!" 2019: https://www.forbes.com/sites/robertmauterstock/2019/08/30/the-caregiver-crisis-is-here/.

[6] The conceptual slipperiness of "crisis" is something that anthropologists, following Koselleck's (2006, 397) claim that, "[T]here has been an enormous quantitative expansion in the variety of meanings attached to the concept of crisis, but few corresponding gains in either clarity or precision," have taken to task (Barrios 2017), even as they contemplate "anthropological obliviousness to the concept of crisis in the history of the discipline in the twentieth century" (Beck and Knecht 2016, 56). Masco argues (2017, S65; see also Jobson 2020) that crisis has become "a means of stabilizing an existing condition rather than minimizing forms of violence across militarism, economy, and the environment,"—an assessment that doesn't contradict Vigh's (2008, 5) claim that, "for a great many people around the world crisis is endemic rather than episodic and cannot be delineated as an aberrant moment of chaos or a period of decisive change," or even Povinelli's (2011, 13) articulation of "forms of suffering and dying, enduring and expiring, that are ordinary, chronic, and cruddy rather than catastrophic, crisis-laden, and sublime." Roitman (2013, 13) in turn "invite[s] the reader to put less faith in crisis, which means asking what is at stake with crisis in-and-of-itself." Given that "permacrisis" has lately entered the popular lexicon (Bushby, "Permacrisis declared Collins Dictionary word of the year," 2022: https://www.bbc.com/news/entertainment-arts-63458467), the concept of "crisis" appears to be unfolding alongside its theorization.

This is particularly true of mental and behavioral[9] health. As in the other care professions, while the global pandemic has both heightened and highlighted the insufficient number of mental health workers, it is not the origin of this shortage. The World Health Organization (WHO), in its *Mental Health Atlas 2014* (2015), stated that for approximately half of the world's population, the number of mental health workers per 100,000 people is fewer than one. But improving this grim ratio still isn't enough to ensure access to a counselor, therapist, social worker, psychologist, and/or psychiatrist. The US National Alliance on Mental Illness (NAMI) published a report in 2017 entitled "The Doctor is Out," which analyzed why "nearly half of the 60 million adults and children living with mental health conditions in the United States go without any treatment." In the United States, where the possibility of obtaining health insurance is by no means a given, mental health services are unlikely to be fully covered, and may not be covered at all, depending on the insurance provider. Additionally, insurance plans are notoriously hard to understand, and what coverage they afford is routinely difficult to enforce.[10]

However, a universal healthcare system doesn't necessarily eradicate access issues. In most countries with a universal healthcare system, psychotherapy (aka "talk therapy") is not typically covered by insurance and must instead be paid for "out of pocket" by the patient.[11] Even in circumstances where multiple sessions with a therapist may be fully or partially covered by insurance and/or priced such that patients can afford to pay for them directly, then accessing these services still requires additional steps: would-be patients have to determine what therapeutic style and specialization(s) best fit their needs, research local practitioners to determine who among them might be a good fit, and contact (provided the listed contact information is up-to-date[12]) these practitioners to inquire about rates and availability. This requires time, effort, and luck—qualities that people don't always have on their side, especially while dealing with any stressful circumstances that might prompt them to begin seeking an appointment with a mental health professional in the first place. And while ongoing destigmatization of mental health has helped more people to seek out care, stigma and access are in many ways twin problems: long-standing stigmatization has resulted in a limited number of providers and non-standardized routes by which to access them, meaning that any reduction in stigma leads to an uptick in demand for mental health care without the infrastructure in place to meet that demand.[13] This problem is amplified in rural areas, many of which officially qualify as "health provider shortage areas."[14]

---

[7] See: OECD, "Unpaid Care Work," https://www.oecd.org/dev/development-gender/Unpaid_care_work.pdf; https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_633166.pdf

[8] Care scarcity is a "total social fact" (Mauss 2016 [1925]) on which AI care is contingent. This axiom produces and is produced by the ethnographic world I seek to portray in this dissertation, along the lines of Azande witchcraft. Borrowing from Evans-Pritchard's (1991 [1937], 63–64) explanation: "The concept of [not enough care for everyone] provides [us] with a natural philosophy by which the relationships between [people] and unfortunate events are explained and a ready and stereotyped means of reacting to such events. [Care scarcity] beliefs also embrace a system of values which regulate human conduct. […] Unless the reader appreciates that [not enough care for everyone] is quite a normal factor in the [lives] of [people], one to which almost any and every happening may be referred, [the reader] will entirely misunderstand their behavior towards it." In other words, to begin by questioning if the basis of this axiom is "real" would impede understanding it ethnographically. I do intend to question the premise of care scarcity (see Epilogue), but to arrive at that point, it's imperative to first situate what transpires in the name of "not enough care."

[9] Mental health and behavioral health overlap in many respects, but there is a difference: whereas mental health designates anything pertaining to mental illness, behavioral health encompasses all of the "gray area" of issues that don't officially qualify (yet, broadly, and/or consistently) as mental illnesses—including substance abuse, eating disorders, suicidal ideation, self-injury, and addiction. In short, this distinction impacts which kinds of experts get to treat what, in what setting, and by what means.

[10] The California Department of Managed Health Care (DMHC) fined Kaiser Permanente four million dollars in 2013 for deficiencies in its mental health services: https://californiahealthline.org/morning-breakout/dmhc-mental-health-services-deficiencies-remain-at-kaiser/.

[11] Coverage of psychiatry only (the prescribing of medicine unaccompanied by talk therapy) is common.

Making the problem of provider shortages worse are the conditions that many mental health workers (a disproportionately higher percentage of whom are women[15]) experience: despite steep educational requirements and emotionally—and depending on the position, physically—taxing, and sometimes even traumatizing, work, being a mental health practitioner tends to be a relatively low-paying job.[16] Many of these workers end up leaving their positions due to caregiver burnout[17] and/or financial precarity; meanwhile, these circumstances deter others from ever becoming mental health workers in the first place. In other words, this ongoing shortage is systemic, and its effects both far-reaching and unequally felt; for instance, the president of the Association of American Medical Colleges (AAMC) describes the shortage of psychiatrists as contributing to a “revolving door between prison and the street” for psychosocially disabled people.[18]

---

[12] On "ghost networks" (or "phantom networks"), and the disquieting recognition that insurance providers are not without incentive to provide out-of-date/otherwise inaccurate mental health care provider lists, see: Turban, "Ghost networks of psychiatrists make money for insurance companies but hinder patients' access to care," 2019: https://www.statnews.com/2019/06/17/ghost-networks-psychiatrists-hinder-patient-care/; see also: Busch and Kyanko, "Incorrect Provider Directories Associated With Out-Of-Network Mental Health Care And Outpatient Surprise Bills," 2020: https://doi.org/10.1377/hlthaff.2019.01501.

[13] Additionally, a shortage of mental health care doesn't necessarily mean that all mental health care workers seeking employment in all settings and specializations can easily attain it.

[14] Or alternatively, as "mental health professional shortage areas": https://data.hrsa.gov/topics/health-workforce/shortage-areas. Note that the intensification of mental health infrastructural shortages in rural areas appears to be a global issue (with the caveat that "rural" is not necessarily a clearcut universal metric: Muula 2007, Nicholson 2018).

[15] In all professions but psychiatry, which is the highest paid group of mental health care workers. Notably women are in fact the majority (55%) of residents (medical students in training), but this drops to ~30% of practicing psychiatrists. See: "Mental Health Worker Demographics and Statistics in the US," https://www.zippia.com/mental-health-worker-jobs/demographics/.

[16] Unsurprisingly, as medical doctors, psychiatrists fare the best out of mental health workers in terms of income—but interestingly, they are one of the lowest-paid categories of physicians: Qi et al, "Comparison of Performance of Psychiatrists vs Other Outpatient Physicians in the 2020 US Medicare Merit-Based Incentive Payment System," 2022: https://jamanetwork.com/journals/jama-health-forum/fullarticle/2790543.

[17] The term "burnout" (originally "burn-out") was introduced by US psychologist Herbert Freudenberger in the 1970s to describe the consequences of working in the "giving" professions—a product of high stress plus high ideals. See: Freudenberger, "Staff Burn-Out," 1974 (this article includes a "practical section deal[ing] with what preventive measures a clinic staff can take to avoid burn-out among themselves, and if unluckily it has taken place then what measures may be taken to insure caring for that person, and the possibility of his return to the clinic at some future time").

There appears to be common consensus that mental health care is, as a consequence of this shortage, in a state of crisis[19]—though that consensus fissures a bit when it comes to naming precisely when the crisis began.[20] Mental health agencies, governments and other policy makers, research institutions, and workplaces and schools alike are seeking ways to address this shortage and its ramifications—if not as a problem to outright solve, then as one to ameliorate, swiftly.[21] They believe this shortage is all along a problem of scale—too few to care for too many—so the means of addressing it must likewise be, above all else, *scalable*.

II) From crisis to startup

In 2015, Thomas Insel, then-head of the US National Institute of Mental Health (NIMH), gave a talk showcasing his excitement for the pursuit of a neuroscientific or “brain-based” understanding of mental health, in response to which an audience member declared, “You don’t get it. My 23-year-old son has schizophrenia. He has been hospitalized five times, made three suicide attempts, and now he is homeless. Our house is on fire, and you’re telling us about the chemistry of the paint. We need someone to focus on the fire.” Shortly thereafter, Insel left NIMH to join Verily, Alphabet’s health-focused life sciences research company;[22] two years later, he launched his own startup in Silicon Valley.[23] Insel’s plan? Put out the fire using AI. The logic of this move is all but pervasive: if you want to change the world, head to the startup sector. You don't need to be the next Google, Amazon, or Facebook—you simply need to be acquired by them. In a world where mental and behavioral health spending is one of the fastest growing health costs for employers, the mega-companies are indeed paying attention.[24]

---

[18] See: Weiner, “Addressing the escalating psychiatrist shortage,” 2018: https://www.aamc.org/news-insights/addressing-escalating-psychiatrist-shortage; though the AAMC article does not specify so, it’s important to recognize that this “revolving door” disproportionately impacts psychosocially disabled people who are people of color (especially those who are Black and/or Indigenous), and poor. See: Ben-Moshe, *Decarcerating Disability*, 2020.

[19] See for example: Butryn et al, “The shortage of psychiatrists and other mental health providers: Causes, current state, and potential solutions,” 2017: https://www.ijam-web.org/article.asp?issn=2455-5568;year=2017;volume=3;issue=1;spage=5;epage=9;aulast=Butryn; see also: Ramirez, “How to Understand—and Fight—the Mental Health Care Crisis in the U.S.,” 2021: https://www.vice.com/en/article/93w43d/how-to-understandand-fight-the-mental-health-care-crisis-in-the-us; Johnson, “Surgeon General Advisory Raises Alarm on Youth Mental Health,” 2021: https://www.usnews.com/news/health-news/articles/2021-12-07/surgeon-general-issues-advisory-on-youth-mental-health.

[20] For instance, both of the following accounts (from NPR and a conservative think tank respectively) pinpoint the origins of this crisis in the US to mid-twentieth-century deinstitutionalization; yet it seems probable that institutionalization was likewise a crisis for those subject to it: Raphelson, “How The Loss Of U.S. Psychiatric Hospitals Led To A Mental Health Crisis,” 2017: https://www.npr.org/2017/11/30/567477160/how-the-loss-of-u-s-psychiatric-hospitals-led-to-a-mental-health-crisis;
and Heritage Foundation, “The Start of America’s Mental Health Crisis,” 2019: https://www.heritage.org/firearms/heritage-explains/the-start-americas-mental-health-crisis.

[21] See for example: The White House, “Fact Sheet: President Biden to Announce Strategy to Address Our National Mental Health Crisis, As Part of Unity Agenda in his First State of the Union,” 2022: https://www.whitehouse.gov/briefing-room/statements-releases/2022/03/01/fact-sheet-president-biden-to-announce-strategy-to-address-our-national-mental-health-crisis-as-part-of-unity-agenda-in-his-first-state-of-the-union/.

[22] Known as Google X at the time; now Verily, a subset of Alphabet (the parent company of Google that formed long after “google” had become a verb).

[23] See: Metz, “The smartphone app that can tell you’re depressed before you know it yourself,” 2018: https://www.technologyreview.com/2018/10/15/66443/the-smartphone-app-that-can-tell-youre-depressed-before-you-know-it-yourself/.

Though Insel cites this interaction with his audience as a wake-up call that sparked his change in career trajectory, his interest in AI appears to have remained entirely consistent before and after it. What changed was the setting: Insel sought to understand the "gap between scientific progress and public-health impact"—and believed that a startup could in fact close this gap.[25] His startup wasn't the first to turn to AI to address the mental health care crisis, but his position granted this pairing a heightened credibility.

Today, the field of AI mental health care is thick with startup participants. To borrow anthropologist Elizabeth Davis's (2018) term, startups are now an iteration of the "counter-clinic."[26] In which case, it would be prudent to understand: what is it that makes a company a startup? Accounts of this vary, but most align on the method of determining valuation as the key difference: a startup's valuation is calculated by and for a venture capitalist's assessment of its potential for rapid growth and anticipated long-term development. In an industry where over 90% of startups fail, scalability isn't merely an achievement, but a necessity.[27] But offhand, this difference alone doesn't seem to account for some of the most (at the time of this research) recognizable elements of "startup culture": the flexible seating, meditation rooms, and ping pong tables. Without contradicting the role of valuation, I propose a less technical way of understanding what being part of a startup indexes: above all else, a startup is a tool that allows a very small group of people to enact a change that impacts a very large group of people—all the while believing that to be in the position to make that change is to be worthy of being in the position to make that change.

My fieldsite for this dissertation is a Silicon Valley startup that began as a volunteer project to provide on-demand mental health care in an under-resourced refugee camp. They are now one of several startups that are attempting to address the problem of mental health caregiver scarcity by introducing a scalable caregiver, in the form of an artificially intelligent chatbot. Very simply, a therapeutic chatbot (or, mental health chatbot) offers interactive emotional support to human users via text-based (or in some instances, voice[28]) messages—everything from guided breathing exercises to reflective questions or prompts. It might share tips for alleviating stress, or explain and walk users through a particular psychological technique (often drawn from CBT (cognitive behavioral therapy) or DBT (dialectical behavioral therapy), but sometimes from psychoanalysis, among other modalities[29]) to help a user better understand a thought, mood, or behavior. They are available to chat "on-demand," meaning users can begin chatting with one instantly and spontaneously, anytime and anywhere with internet access; many mental health chatbots will also initiate conversations by messaging users periodically (daily, weekly, biweekly—depending on that user's explicit preference and implicit habits). Users can access these bots using any mobile phone (not only smartphones) or other device (computer, tablet) with texting/SMS capability—or, depending on the chatbot in question, a custom app or a third-party messaging platform (such as Facebook's Messenger). These bots are even available in an array of languages.[30]

---

[24] See for example: LaVito, "Anxiety is expensive: Employee mental health costs rise twice as fast as all other medical expenses," 2018: https://www.cnbc.com/2018/09/26/employers-are-starting-to-think-about-healthy-differently.html.

[25] See: Insel, "What American Mental-Health Care Is Missing," 2022: https://www.theatlantic.com/ideas/archive/2022/02/american-mental-health-crisis-healing/622052/.

[26] The counter-clinic, Davis (2018, 5–6) explains, "does not imply that psychiatry today is irrelevant to [what were once considered matters of psychiatric purview], but rather that psychiatry is no longer the only or even, perhaps, the most important place for us to investigate them." It indicates the movement of "psychiatric care from traditional clinics and hospitals to other kinds of spaces, where alternative techniques of care have developed to address both novel and enduring forms of individual and collective pathology and precarity." In this case, the "alternative technique" is the application of AI.

[27] On the oft-cited "90%," see: https://startupgenome.com/article/the-state-of-the-global-startup-economy.

Startups in this industry stress the advantages of AI. With a chatbot, there's no need to make multiple calls to track down a practitioner with available appointments, or to commute to that appointment. There's also no time minimum or limit to contend with—a session can be brief, long, and/or intermittent. And because most mental health chatbot startups offer their services "business-to-business" (B2B) rather than directly to consumers (B2C), end-users generally don't have to deal with payment.[31] On top of this, chatbots are also very inconspicuous to use, which helps buffer the considerable stigma that inhibits many from seeking out a therapist. Teletherapy grants some of this flexibility, but on the other end of the line, there is still a caregiver who needs pay, rest, time to process paperwork, and so forth.[32] Not so with AI.

Interestingly, in most medical settings (including mental health *diagnostic* care), the role of AI is that of a tool that can surpass human ability, and consequently requires human oversight—a human expert needs to have the "last word," so to speak. With therapeutic AI, the inverse is true.

---

[28] While device-configured accessibility settings means that all texting is potentially voice-based, some therapeutic chatbots are designed as primarily voice-based for voice assistant devices such as Amazon's Alexa.

[29] On the shared origins of and differences between CBT and DBT, see: National Alliance on Mental Illness, "Psychotherapy," https://www.nami.org/About-Mental-Illness/Treatments/Psychotherapy. Note that both qualify as evidence-based interventions, including through remote or self-guided implementation (see Chapter 2).

[30] This list is not exhaustive, but during my fieldwork I personally came across them in Spanish, Arabic, Bengali, French, Korean, Dutch, and Japanese, as well as English; still, English dominates the offerings, and many languages and dialects have yet to be represented.

[31] In some instances, companies may offer their therapeutic chatbots through a subscription service, though usually this means an "add-on" (more features, content, and/or integration with access to a human therapist) to their free-to-the-public version (more on that in Ch 4). The point being, with a therapeutic chatbot, users don't have to figure out session-by-session payments, sliding scale calculations, or negotiation with health insurance companies.

[32] And there are still all the barriers in place to getting a patient on the phone or HIPAA-approved software video call session with them. I attended an APA session on teletherapy in 2018, and went into it thinking, 'how is this different from AI care?' I came out of it with a newfound appreciation for why AI, unbeholden to most of the profession's prolific "red tape," and able to seamlessly cross state borders, might appeal. While this has necessarily eased somewhat since the onset of the pandemic, there's still enough red tape to make the prospect bypassing it attractive.

My informants—including psychologist startup workers, end-users, as well as the committee of clinicians who advise my fieldsite (“The Startup”)—would (and did) universally attest: a chatbot does not in any way exceed the skills of a human therapist; and while The Startup’s psychologist-staffed health team dutifully monitor certain aspects of the chatbot exchanges to ensure “quality control,” they don’t directly supervise each interaction.[33] Doing so would nullify every advantage—rapid scalability, low overhead costs, and on-demand, discreet availability—that these bots offer. As a result, therapeutic AI walks a line between entertainment and medical care, and regulation of this industry is at present more or less nonexistent beyond voluntary measures.[34]

During my time at The Startup, which ranged from 2018 through the early months of the pandemic in 2020, their roster of clients included everything from a children’s hospital to a branch of the military to a municipal government. They partnered with a number of research institutes specializing in things such as aging, eating disorders, maternal health, and sexual violence. But The Startup’s bread and butter is having their services on offer to workplaces through employee benefits programs. Still, the division between commercial venture and public health outreach at my fieldsite is murky—intentionally so. The Startup offers a free-to-the-public version of their services, which in turn drew clients who wanted to do the same. Not only the founder, but most of the workers at The Startup seemed aligned in seeking to do nothing less than change the world—a desire concretized in their goal of providing their services to one billion people.[35] Remarkably, this goal is not infeasible in terms of technological capacity and access. Scalability tempts, rather than tempers, lofty expectations.

In “The Crisis in Crisis” (2017), anthropologist Joseph Masco explains: “Since the 1980s neoliberal turn in the United States, militarism has remained the project of the state, but the collective future has been assigned to the marketplace, which elevates short-term profitability above all other concerns."[36] One way of understanding the startup sector is as a fleet of companies that attempt to work against this fate using the same set of tools that secured it. Startups are a phenomenon that slightly complicate Masco's assessment, as they suspend short-term profitability in favor of *anticipated* profitability; this has the effect of granting startups a kind of plausible deniability to Masco's claim without actually altering its truth. What I can attest to through my fieldwork is that many if not most of the people leading and working for startups sincerely see them as salvational—as the last, best hope to close a chasm-like gap between principle and practice.

---

[33] One exception of which I’m aware of is a chatbot that my fieldsite designed for long term patients in a children’s hospital; the care professionals treating those children review all of these chats as part of patient care.

[34] In the US, companies can voluntarily pursue FDA certifications such as “Breakthrough Device Designation” status (see: https://www.fda.gov/medical-devices/how-study-and-market-your-device/breakthrough-devices-program, a program that launched in 2018); therapeutic chatbot company Wysa was awarded this status in 2022 for chronic pain, anxiety, and depression treatment in adults: *BusinessWire*, “Wysa Receives FDA Breakthrough Device Designation for AI-led Mental Health Conversational Agent,” 2022: https://www.businesswire.com/news/home/20220512005084/en/Wysa-Receives-FDA-Breakthrough-Device-Designation-for-AI-led-Mental-Health-Conversational-Agent. Non-voluntary regulation is still fairly sparse; one of the only areas of clear regulation concerns HIPAA compliance—for chatbots that are interpreted into formal health services, the chatbot platform and data management and protection must conform to HIPPA standards. Accounting in part for this sparseness is the fact the companies providing mental health chatbots get to decide just how explicitly oriented towards “mental health” vs “entertainment” they wish to be. Over the years, I’ve observed pronounced arcs in how companies have situated themselves along this continuum—a pronounced shift, circa 2018, away from “mental health” in favor of “wellness” and life coach,” followed, in 2020, by a sharp return to “mental health,” “therapy” and even “psychology.”

[35] Becoming profitable is part of that, but the social capital (Bourdieu 1986) of becoming an instrument of change (read: joining the ranks of those elites *ostensibly* eschewing some of their eliteness to do likewise, but in such a way that more or less implies that everyone is simply vying for a position in a redefined field of eliteness) also figures in this.

Notably, startups don't exactly solve the problem of caregiver scarcity. Insofar as the scarcity of mental health caregivers arises from underpaid, difficult to sustain work with high educational requirements, AI does not remove these conditions. Companies selling AI care furnish scalable caregivers—caregivers instantly available in any volume. AI enables companies like my fieldsite to provide *enough* caregivers without having to provide *more* caregivers.

III) AI is a caregiver (but not a replacement)

This dissertation is an ethnographic portrayal of AI mental health care that attempts to show what it's like to live in a world where there's simply not enough care to go around—a world in which the phenomenon of artificially intelligent caregivers becomes, accordingly, not only thinkable, but necessary. This necessity is what brings startups, AI, and the mental health care crisis together into a single frame. The ethnographic account that follows centers on three groups of informants: the end-users who periodically exchange text messages with a mental health chatbot in order to feel better; The Startup's health team, a group of psychologists who departed clinical practice in search of care work that could be meaningful without being exploitative and unsustainable; and finally, the AI-powered chatbots themselves, which have effectively displaced (vs replaced[37]) the psychologist startup workers. Together, these informants make it apparent that unlike so many job settings that bring together human workers and AI, those who create, purchase, and use therapeutic chatbots do not regard AI as a *replacement* for human caregivers—even as AI does care work.

Importantly, this dissertation is not an investigation of whether or not AI can care. I am instead asking: What makes AI care work? Unlike the first question, my question starts from the position of acknowledging something that might at first seem remarkable, but that I hope to demonstrate is quotidian: I propose that **AI is a caregiver**. AI already does care work in the world, including in urgent circumstances.[38] This means that AI caregivers are not merely a futuristic prospect waiting to be someday realized—or prevented. Why should this be surprising? After all, no one is in doubt that AI can kill—so why can it not also heal?[39]

---

[36] Masco, "The Crisis in Crisis," 2017, S15.

[37] As I will further explain in Chapter 3 of this dissertation, when workers are displaced by AI, this means that instead of AI entirely eliminating the need for (i.e., outright *replacing*) human workers, most AIs actually require a new set of roles to be undertaken by human workers (with their work often appearing to outsiders as something that the AI accomplishes independently); see: Irani 2015.

This project is neither a radical proclamation of impending AI sentience, nor the underpinnings of an AI healthcare industry white paper; it is instead a refusal to allow the question of whether or not AI can care to overshadow the very real ways in which AI caregivers are already at work in the world. In emphasizing this point, I'm not suggesting that envisioning and articulating the future existence of AI caregivers are not part of our relations with them—they are.[40] But the vast extent to which AI caregivers are imagined should not eclipse the fact that however prosaic or underwhelming its current forms, artificial intelligence already performs caregiving labor.[41] To argue that these AIs are not *good* or even adequate caregivers does not undo the fact of their existence.

Rather than ask *if* AI can care—and thereby risk reinforcing the cloying belief that "real" care is necessarily good care—I instead focus on understanding: what are the rules, factors, and conditions that summon and sustain a rapidly growing industry of therapeutic chatbots?[42] What is it that makes AI care work? Each of the chapters that follow will provide an answer to this question: therapeutic bonds, labor displacement, and worker surveillance, respectively, are what make AI care work—though perhaps not entirely in the ways you might expect.

---

[38] In Chapter 4, I show that AI is already a crisis caregiver for people experiencing suicidal ideation—and that broadly speaking, AI and suicide prevention care are entangled in less obvious ways (see for example: Levine, "Suicide hotline shares data with for-profit spinoff, raising ethical questions," 2022: https://www.politico.com/news/2022/01/28/suicide-hotline-silicon-valley-privacy-debates-00002617;
BBC, "Mental health helpline ends data-sharing relationship," 2022: https://www.bbc.com/news/technology-60218894; note scholar danah boyd's role in this instance).

[39] The recent blossoming of caring AI, on the heels of the flourishing of killing AI (see: Bellanova et al 2021, Suchman 2020), is an unsurprising shift when considered alongside the similar arc that nuclear weapons and energy followed; consider how, scant months after the US dropped nuclear bombs on Hiroshima and Nagasaki, politicians and scientists rushed to "repurpose" atomic energy by proclaiming its healing potential (see for example: Ellis 1946).

[40] Haraway (2016) argues that imagining as a practice is inherently part of relations. Also, for a temporally-inverted parallel showing, in another context (dinosaurs), that envisioning and articulating dinosaurs' *past* existence is likewise part of our relations with them, see Noble 2016.

[41] My claim that AI does care work is also not metaphoric—or at least no more or less so than humans, given that "work" itself is a metaphor for how we spend our time.

[42] One approximate way to trace this growth is through press releases about funding rounds over the years (starting around 2016)—for instance: "Woebot Health Closes $90 Million Series B Funding Round Co-Led by JAZZ Venture Partners and Temasek," 2021: https://woebothealth.com/woebot-health-closes-90-million-series-b-funding/; "Woebot Secures $8 Million in Funding to Increase Access to Mental Health Care Worldwide," 2018: https://www.globenewswire.com/news-release/2018/03/01/1402071/0/en/Woebot-Secures-8-Million-in-Funding-to-Increase-Access-to-Mental-Health-Care-Worldwide.html; "Wysa raises $20 million to expand its therapist chatbot into a wider set of mental health services," 2022: https://techcrunch.com/2022/07/14/wysa-20-million-series-b-funding-expand-therapist-chatbot-wider-mental-health-services/; and "Wysa Puts $5.5M Series A To Work Targeting 'Missing Middle Of Mental Health,'" 2021: https://news.crunchbase.com/health-wellness-biotech/wysa-puts-5-5m-series-a-to-work-targeting-middle-of-mental-health/. Another factor to consider alongside this is the ever-growing number of "players" in this field.

In claiming that AI is a caregiver, I recognize that STS scholars offer several formidable objections that effectively dispute this. Consider the following: in one of a 2021 series of articles in *The Guardian* on "automating care," political scientist Virginia Eubanks and AI ethics researcher Alexandra Mateescu state: "The most essential aspects of caring for one another—presence, compassion, connection—are not always easy, or even possible, to measure. The rise of the care bots risks creating a system where we only value the parts of care that can be turned into data."[43] Sociologist and psychologist Sherry Turkle (2005 [1984], 2011, 2015, 2020) objects to caring machines on the basis of empathy—she argues that machines lack empathy while appearing in some cases to possess it, making us vulnerable to losing touch with human empathy: "When machines ask us to care for them, we come to think they care for us. But this is pretend empathy, and it takes advantage of the deep psychology of being human."[44] Informatics scholars Hamid Ekbia and Bonnie Nardi, in *Heteromation, and Other Stories of Computing and Capitalism* (2017), locate the possibility of care in the shared condition of life itself: "The human capacity to care is distant from the aptitudes of machines. As designed objects created from human desires and constructed from inanimate materials, machines sit outside the struggle to survive. The existential condition of even the simplest organism centers on survival, the origin of the most primitive expression of caring."[45] Here are but three sophisticated objections, all aligned in what they conclude; yet each presents a different basis for arriving at that conclusion. Each, in other words, differs when it comes to describing what precisely is lost in the turn to AI care.

This leads me to ask, what is it about "care" in care work that precludes us from recognizing AI as *already* doing it? What is it about care and caregiving that we are unwilling to relinquish to AI bots?

At first, the problem might seem to arise from the difficulty of defining "care"—it's a concept that is multifaceted[46] and vast (Tronto 1993), even contradictory. Ethnography challenges many seemingly foundational understandings of care: motherly nurturance (Scheper-Hughes 1993); humanitarian ethics (Tiktin 2011); professional caregivers and policymakers' apprehension of care practices (Garcia 2010); and the centrality of "self" in self-care (Sanabria 2017). Anthropologists have offered important objections to care where it takes the form of surveillance rather than treatment (Kitanaka 2015; Das 2015; Todeschini 2001), while also challenging the assumption that care as "cure" (or its pursuit) is ethically unproblematic (Nakamura 2013)—including by showing that one person's experience of care can appear to others as harm (Ozawa-de Silva 2010).[47] Anthropological and STS works revealing the violence of care now abound (Biehl 2005; Borovoy 2005; Han 2012; Knight 2015; Lester 2019; Murphy 2015; Scheper-Hughes 2003; Stevenson 2014; Ticktin 2011). Yet importantly, scholars concerned with refusing the essentialization of care-as-benevolent do not abandon the possibility of care as such; if anything, care in one register appears (in these and other works) to be a means of rectifying the problems of care in another.[48] In short, there's a discernible cleft in the anthropological theorization of care, which Miriam Ticktin (2019) elucidates: care as enforcing (intentional and unintentional) oppression, and care as dismantling "persistent forms of domination and exclusion."[49] Above all else, ethnographic research draws out how care, in all its ethical complexity, is relational (in addition to the works cited above: Allison 2013, 2018; Buch 2018; Cohen 1998, 2001, 2008; Chua 2020; Davis 2012; Mariner 2019; Behrouzan 2016; Pinto 2014; Surfrin 2017; Taylor 2008, Zhang 2020). Importantly, these relations include nonhumans (Dave 2014), and even non-living nonhumans (Allison 2006; Schüll 2016).[50]

---

[43] Eubanks and Mateescu, "'Care bots' are on the rise and replacing human caregivers," 2021: https://www.theguardian.com/us-news/2021/jun/03/care-bots-on-the-rise-elder-care.

[44] Turkle, "That Chatbot I've Loved to Hate," 2020: https://www.technologyreview.com/2020/08/18/1006096/that-chatbot-ive-loved-to-hate/.

[45] Ekbia and Nardi, *Heteromation, and Other Stories of Computing and Capitalism*, 2017, 129.

[46] Anthropologists and other scholars have thoroughly examined this multivalence: care might indicate any one or more of the following (which are themselves neither fixed nor mutually exclusive categories): an ethical framework (Biehl 2012; Fassin 2008; Foucault 1988, 2005); something that overlaps with yet is not interchangeable with "cure" (Mol 2008); an affective state of being cared for (Cox 2015; Kleinman 2008); care labor (Parreñas 2015; Glenn 2012; Wharton 2009); self-care (Schüll 2016; Fullwiley 2011); care resource markets (Biehl and Petryna 2011; Peterson 2014); care governance (Netherland and Hansen 2017; Knight 2015; Han 2012); and many other valences besides.

Despite ample evidence of care's conceptual expansiveness, however, there's a remarkable resilience of "care" as a hybrid of self-evident ethical commitment and act. This understanding of care—and especially caregiving, which exposes the entanglement of care and reciprocity—still goes all but unquestioned as something that can be defined more or less reflexively with what it is to be human. Somehow, care emerges as a last reserve of human qua human in a posthuman world[51]—beautifully and powerfully so, in the writing of anthropologist and physician Arthur Kleinman (2009), who asserts the humanity of his wife Joan through her experience of becoming increasingly in need of care as a result of Alzheimer's, alongside his own struggle to navigate caregiving. Without denying that caregiving is "tiring, and emotionally draining," he explains that:

> *[C]aregiving is also a defining moral practice. It is a practice of empathic imagination, responsibility, witnessing, and solidarity with those in great need. It is a moral practice that makes caregivers, and at times even the care-receivers, more present and thereby fully human. If the ancient Chinese perception is right that we are not born fully human, but only become so as we cultivate ourselves and our relations with others—and that we must do so in a threatening world where things often go terribly wrong and where what we are able to control is very limited—then caregiving is one of those relationships and practices of self-cultivation that make us, even as we experience our limits and failures,*

---

[47] To this latter point, see also disability activist and scholar Eli Clare's (2017) consideration of identity-affirming surgery.

[48] See also: disability activist Leah Lakshmi Piepzna-Samarasinha's *Care Work: Dreaming Disability Justice* (2018).

[49] Ticktin, "From the Human to the Planetary: Speculative Futures of Care," 2019, 136.

[50] Though not an ethnography, STS scholar Maria Puig de la Bellacasa's *Matters of Care* (2017) also considers care as a matter of "more than human worlds."

[51] On posthumanism, see: Hayles 1999.

> *more human. It completes (not absolutely, but as a kind of burnishing of what we really are—warts and all) our humanity.*[52]

Notably, I'm not seeking to disprove Kleinman. Acknowledging AI as a caregiver doesn't mean Kleinman's account is consequently wrong, just as pointing out instances of multispecies caregiving would not undermine it; for caregiving to make humans more human does not discount the possibility, for instance, of it making dogs (Haraway 2007, Clare 2017) more dog. I'm also not building up to a Pinocchio argument that AI and humans, "not born fully human," are alike in their capacity to become human through their relations with others. Instead, I'm suggesting that to preclude AI from the role of caregiver is to essentialize care work just as much as human-ness.

This subtle essentialization of care surfaces in critiques of AI care. To illustrate this essentialization, consider, in parallel to AI mental health chatbots, the (now ubiquitous) customer service chatbot—that kind that appears in a small pop-up window in the corner of an e-commerce website with a message inviting you to type in any questions you might have. While some of these reroute you to a human, many of them are self-contained chatbots poised to answer questions that you might ask a human customer service representative. Now, imagine asking: who/what here *really* provides the customer service in the case of a customer service chatbot? Is it the chatbot itself, or the people who create and oversee the bot? I suggest that this is not an especially interesting question. The bot interface, along with all the work that goes into creating, supervising, and coordinating with the bot, are providing customer service. It does not feel costly to acknowledge that a chatbot can do customer service work.

In the context of an AI chatbot providing emotional support, I argue that it is much the same—care via chatbot arises from both humans and AI doing care work. It might now appear that I am the one doing the essentializing—not of care, but of AI: does a computer or telephone do care work? A pen and paper? A doorstop (Johnson [Latour] 1988)? Suffice to say, Latour's articulation of actor-network theory (1996, 2005)—an approach that regards human and nonhuman objects/entities alike as "actants" who/which make things happen through their relationships with one another—does not dissuade me from considering this possibility, and suggests that the difference need not hinge (ha) on the matter of agency. But as an ethnographer of AI care, I would hesitate to name these other nonhumans as caregivers, because unlike a therapeutic chatbot, no one is creating, selling, or providing them as such. Though academic psychology literature substantiates the claim that a book (namely, a self-guided CBT workbook) can potentially mediate therapy,[53] psychologists do not propose that this mediation occurs by means of an interpersonal relationship with that book. In contrast to this, members of the mental health chatbot industry present bonds between users and chatbots as the source of a chatbot's therapeutic efficacy (Darcy et al 2021). (This "ontologically ambiguous" bond is, in turn, the point of departure for the second chapter of this dissertation.)

---

[52] Kleinman, "Caregiving: The Odyssey of Becoming More Human," 2009, 293.

[53] See for example: Williams et al, "Guided Self-Help Cognitive Behavioural Therapy for Depression in Primary Care," 2013; I discuss this further in Chapter 2.

"Can AI care?" is so closely entangled with "should AI care?"—perhaps making it seem as though assenting to one (yes it can care) is implicitly assenting to the other (yes it should care). In decoupling these questions, I'm not suggesting that AI is necessarily a good or desirable caregiver (nor am I saying it's inherently bad); I'm simply pointing out that to deny that AI can care is an ineffective way to argue that AI does not provide good care. Though this difference might seem at best slight, I intend to show that the stakes of essentializing care and caregiving do not pertain to AI alone: this essentialization is at the heart of the inequities that characterize care work—mental health and otherwise.

IV) Care is already automated

Objections to AI care seem to express that while AI doesn't have the capacity to care, it can *change* what care is by automating it. There are a few problems with this framing, however. For one thing, this outlook presupposes that care equates ethical "good"—which is the essentialization I described, and which, as noted above, is something anthropologists have shown is not the case. But additionally, while AI automates care, automation in care—and even automation as a *problem* in care—does not begin with AI. Care is *already* automated.[54]

Kelly Ray Knight (2015) shows that case workers who assist pregnant, unhoused/precariously housed drug addicts in San Francisco share with their clients—albeit without being subject to anything close to the same consequences—a degree of being bound by the vicissitudes of paperwork. Like Rebecca Lester (2019), Knight shows that for practitioners, diagnostic pathways become tools with which they try to snag, bit by bit, fleeting access to care on behalf of their clients/patients—and with it, their client/patient's survival. Lester unearths a bleak, ironic loop in which eating disorder patients in the US find themselves not qualifying for insurance coverage for the inpatient care most likely to ensure their recovery, all because of a care system that regards them as unlikely to survive. In *Scripting Addiction* (2011), E. Summerson Carr describes how behavioral health practitioners at an outpatient addiction treatment facility in the Midwest US are attentive to "reconfiguring clients' relationship with language rather than simply, or even primarily, reconfiguring their relationship to drugs," and how their clients in turn learned to subtly "flip" that script and "formally replicat[e] prescribed ways of speaking about themselves and their problems without investing in the content of those scripts."[55] Lisa Stevenson's (2014) "anonymous care"—or care that is, in the words of an informant, "too professional"—could be likewise described as automated, and takes the form of colonialist "policies, programs, and exhortations" of suicide prevention and other interventions that "ask Inuit to live while also expecting them to die."[56] This is by no means an exhaustive list of ethnographies that show the signs of automation in mental health care.

---

[54] According to IBM, "Basic automation takes simple, rudimentary tasks and automates them. This level of automation is about digitizing work by using tools to streamline and centralize routine tasks," (https://www.ibm.com/topics/automation). IBM's definition gestures towards the overlay of automation with the mid-twentieth-century unfolding of cybernetics/systems theory that I'm attempting to draw out, in which patient paperwork and treatment methods arguably constitute forms of automation; critical to this point, however, is the fact that those who championed systems theory and cybernetics did not propose that automation itself is inherently a means of advancement. As cybernetics theorist Stafford Beer (1974, 21) rued, "We are using our powerful tools to automate and to elaborate the limited processes that we managed to achieve with the unaided brain and the quill pen—processes which our new tools were invented precisely to transcend."

[55] Carr, *Scripting Addiction*, 2011, 20.

The algorithm is already there, in care—and this is not news to mental health workers. During fieldwork, I once ended up in a hospital emergency room, and noticed an algorithm was quite literally the writing on the wall—in the form of a laminated poster showing intake procedures for practitioners, labeled explicitly *as* an algorithm. Algorithms convey a plan for automation, and AI in turn is the automation of algorithms. Though explicit reference to algorithms in care may be a more recent phenomenon, the embrace of and resistance to care as automation respectively have arguably been going on for some time: consider for instance the Anti-psychiatry movement (Cooper 2013 [1967]; Goffman 1961; Szasz 1961) as a kind of anti-automation of mental health. Johnathan Metzl (2009), by showing how reforms to the diagnostic criteria of schizophrenia in the mid-twentieth-century United States both emerged from and resulted in the targeting of Black men, illustrates what an update to the algorithm can do. If structural competency can be understood as likewise a retooling of an existing care algorithm, then Metzl and Helena Hansen (2014) demonstrate that updating the algorithm can in fact be an anthropological aim. The turn to AI isn't producing the automation of mental health care so much as openly acknowledging it.[57]

Once, while attending a psychiatry professional conference with my startup informants, I met a clinician in town for the event who served as one of The Startup's advisors and product testers. Tripti is a petite 40-something woman of South Asian heritage and a self-confessed technophile who, as she put it, "enjoys staying in-the-know" on technology trends in mental health. She explained how she enthusiastically recommends The Startup's chatbot to her patients as a complement to therapy. But Tripti was eager to see the chatbot go a step further—she dreamed of outsourcing her pre- (including new patient intake) and post-appointment paperwork to AI. For practitioners, many of whom cannot afford any kind of administrative support, the requisite paperwork of billing, insurance forms, diagnostic tests, and appointment scheduling constitute a form of automation that they must constantly endure; this is part of why therapists often won't accept insurance (and might instead offer a sliding scale)—it's not out of insensitivity to their patients, but because they can't afford to maintain the relentless stream of paperwork on their end. Automation, in short, does not necessarily betoken "free time" (Hui 2018).

Sectioning off the paperwork part of mental health care as not really "care" would be to miss the point: to truly understand what makes AI care work, it's critical to think not just about care in the abstract, but about care work in its entirety. Again, that is my basis for saying AI is a caregiver—it does care work. Rather than instrumentalize AI care as a foil against which to assert what real, proper, valid, or even good care is, I propose thinking about human and AI caregivers together.[58] What can artificial intelligence and care work each show us about the other? My informants at The Startup are primarily mental health workers who effectively seek to be in some sense "replaced" by AI. As members of The Startup's health team, they work alongside two separate, smaller teams—the engineering team and the sales team, respectively. Unlike these teams, the health team, most of whom are women, doesn't write the code or land the new accounts, per se—yet, in addition to conceiving and writing all the content of the chatbot, they are highly involved in the coding process and in acquiring and maintaining clients. However, they receive significantly less pay than many of their startup colleagues, and (as I show) seemingly have less "clout." And yet, as they are keenly aware, they have more than they did—or likely would—have outside the startup sector.

---

[56] Stevenson, *Life Beside Itself*, 2014, 96.

[57] Still, for me to say that mental health care is already automated doesn't mean this automation is uniform, or that all instances, the same things are being automated. This dissertation will make it possible to notice what exactly AI is automating in this context, and the ways in which humans become part of that automation.

For all that care is held up as a definitively "human" trait, care work itself can be incredibly "dehumanizing" in the conventional sense of the word. I intend to show that this isn't a detail to be set aside in contemplating AI care.

I explained above that care is multivalent—but within that multivalence, it's also something of a mystery. Uncertainty about what care is runs deep—to the extent that there are different (overlapping and contrasting) theories that attempt to reconcile why care workers consistently receive low pay for work that is all but universally proclaimed as highly valuable to the society in which they live.[59] AI care provides an opportunity to think about why this is, and to revisit and reconceive these theories.[60] I argue in the third chapter of this dissertation that examining AI care work ethnographically makes it possible to see the shared conditions between human care workers and AI labor—which distill into what I describe as being disposable and essential at the same time.[61]

In other words, AI, which is still only at the fringes of most medical care settings, is not the source of the "inhuman" expectations imposed upon care labor broadly; instead, I'm suggesting that AI caregivers, like their human counterparts, are subject to a similar set of expectations. Note that I'm not alone in considering nonhumans in their capacity as workers that work alongside humans: though they do not specify *nonliving* nonhumans, Sarah Besky and Alex Blanchette in *Rethinking Labor on a Troubled Planet* (2019) underscore the growing attention to nonhuman labor as something that invites rethinking "who (or what) should be included as a protagonist in the critical study of labor today."[62] I'm proposing that AI absolutely should be included in this—that it might be an ally in *addressing* labor exploitation, and not solely its instrument.

---

[58] Recognizing AI as a caregiver does not mean mirroring the harms of "ghost work" (Gray and Suri 2019) by vanishing the work of the health team in order to maintain the illusion of autonomous AI workers.

[59] I discuss this at length in Chapter 3.

[60] As sociologist Anna Romina Guevarra (2021) pithily puts it: "These questions of whether to employ phenomenally expensive robotic systems are especially poignant given that the field of care work relies heavily on the labor of poor women of color, often immigrants, who have long been told both that civilization depends upon their work and that it is of little monetary value."

[61] While a dimension of this was undeniably brought to the forefront by the pandemic, this precedes it, and will follow it.

V) Not-so-cutting-edge therapeutic chatbots

Just as recognizing the present-day existence of AI caregivers isn't the same thing as endorsing the curtailment of human caregivers into mere "humans-in-the-loop,"[63] it's also not a designation that applies only to the latest, most advanced AI technology.[64] It's not, in other words, wholly contingent on chatbots becoming ever-more indistinguishable from a human. The easiest way for me to demonstrate this is to show you an interaction between bot and end-user.

Below is a therapeutic chatbot "session" that I came across in an anonymized transcript, early on in my fieldwork at The Startup. Every week or so, The Startup's health team would generate a massive "transcript," in the form of a spreadsheet containing a few thousand rows of randomly selected, fully de-identified snippets of chatbot conversations, aggregated from thousands of users. As an ethnographer in the role of startup intern (more on that to follow), my first task was combing through these spreadsheets and flagging any chatbot glitches that might confuse, annoy, or potentially harm users in any way. This process struck me as being akin to case supervision, where therapists (in this case, the chatbot) consult with peers (The Startup's psychologists, and myself as their intern) to review sessions with clients/patients. In fact, in the transcript, user dialogue is even marked as "Patient."

What stood out to me at the time about this particular conversation was how it showed a sustained, helpful dialogue, while also demonstrating that these merits aren't due to the chatbot's indiscernibility as a bot.

---

[62] Besky and Blanchette, *Rethinking Labor on a Troubled Planet*, 2019, 1; significantly, Paxton, in *The Life of Cheese* (2013), also refers to microorganisms as "co-laborers."

[63] See: https://humansintheloop.org/what-is-a-human-in-the-loop/.

[64] This seems like as good a place as any for a brief disambiguation of "AI" from "machine learning," as well as from "algorithm." As noted above, AI is the automation of algorithms. An algorithm in turn is often described as a "recipe" or "blueprint"—a set of instructions on a process to be executed. Machine learning (ML) is a subset of AI that many now use interchangeably with "AI," but not all AI is ML. In all forms of AI, an AI "decides" something—even if the decision in question is as simple as a videogame moving to the next line of automated pre-written dialogue after finishing the previous line of pre-written automated dialogue. In ML, these decisions are more complex, because they entail the AI having some capacity (following certain parameters set by a human programmer) to decide how it will decide. ML indicates that an AI can effectively "update" the algorithms that it uses as it accesses data sets—this capacity to modify based on data intake is the basis for the analogy of "learning" in the term "machine learning." Whenever you hear about the "black box" of AI algorithms, and the phenomenon of programmers not fully knowing or being able to explain the exact process by which an AI arrived at a particular decision, then this indicates ML. It's important for social scientists to study ML (including, in the context of chatbots, "generative AI"—which is a subset of ML) and its repercussions; but as I attempt to show in this work, ML is not the sole or even the "best" way by which to understand the broader meaning and consequences of our relationships with AI.

CHATBOT: Hey [USERNAME].[65] Do you ever feel anxious due to work?

USER: Yes I do sometimes

CHATBOT: OK. It's important to remember that your worth as a human being is not determined by any job.

CHATBOT: Can you continue describing to me what it feels like to be in this situation?

USER: I care for a Alzheimer's patient. She is very mean anf I never know what mood she be in for the day. The only time she's nice to me is when I'm giving her food constantly. I feed her all day and she tells the visiting nurses she hadn't eaten in days. She starving. I'm also supposed to get her to do her physical therapy. I can't get to do her exercises but neither can they. So being anxious is an understatement.

CHATBOT: I see. Have you been in a situation like this at work before?

USER: Not to this extent

CHATBOT: How were you able to get through it in the past?

USER: Compassion and reward like dealing with a young child

CHATBOT: I understand. It is important to understand that any job can have stressful elements, even if you love what you do.

CHATBOT: Work stress can be very overwhelming and potentially harmful to your physical and emotional health. However, you can take steps to help manage work-related stress! Would you like to learn about some?

USER: Doesn't work with her. Yes is super stressful. I don't sleep any amount of time. Catnaps and 3-4 hours most of the time. I do a lot of praying and keep busy to fill my mind.

CHATBOT: I hear that you're feeling stressed, [USERNAME]. Would you like to shift the focus to that?

USER: Yes

---

[65] Because everything is anonymized, I don't see the actual name of the user anywhere—but where it says "[USERNAME]," the chatbot is saying the given name that the user provided in their first conversation.

CHATBOT: OK, [USERNAME]. Let's talk through this together.

CHATBOT: I get it. I really feel strongly about the importance of self-care. Do you make enough time to treat yourself?

USER: No not really. I have a semi good day if I get a shower. Some days I don't even get my hair brushed.

CHATBOT: What are some activities you wish you had more time for?

USER: Visiting my grandkids or dating

CHATBOT: It's important to put yourself on the agenda. Taking care of yourself is not selfish and can enhance the quality of your relationships with others.

CHATBOT: Would you like to work together to set goals for self-care?

USER: I need to

CHATBOT: Great! Let's use the S.M.A.R.T. Method when thinking through your goals, which means they should be Specific, Measurable, Attainable, Realistic, and Timely.

CHATBOT: What is one self-care activity that you can realistically engage in this week?

USER: Go to church this sunday

CHATBOT: When can you make time to do this?

USER: Sunday

CHATBOT: Alright. Also, it helps to write your goals down and post them somewhere you can see them daily to hold yourself accountable. Have you tried this before?

USER: In the past

CHATBOT: Great.

CHATBOT: If we stick to the SMART method, then your goal might be something as simple as "read 30 minutes before going to bed 3 days per week" or "exercise for 1 hour at least two days Monday - Friday."

CHATBOT: Let's practice. Send me your goal in our next message:

> USER: I like the idea of reading before bed. That's something I enjoyed years ago.
>
> CHATBOT: Alright, [USERNAME]. And does this goal stick to the SMART method?
>
> USER: It should
>
> CHATBOT: Would you like me to send you a reminder to help you stay on track?
>
> USER: Yes
>
> CHATBOT: I'm happy to be your accountability buddy! :)
>
> CHATBOT: How satisfied are you with this conversation? 1= very dissatisfied, 4 = very satisfied
>
> USER: 4
>
> CHATBOT: Glad I could help :)
>
> CHATBOT: Should we pause our chat for today?
>
> USER: Yes, please.
>
> CHATBOT: Great! We will be in touch.
>
> USER: Ok good night

Behold, a therapeutic chatbot. As you can see, this is not a frontier of impending AI sentience. This exchange is the chatbot at its best, and even here, there are subtle "tells," if you're looking for them; recall:

> CHATBOT: What is one self-care activity that you can realistically engage in this week?
>
> USER: Go to church this sunday
>
> CHATBOT: When can you make time to do this?
>
> USER: Sunday

There are increasingly remarkable, Turing-Test-acing chatbots out there—but this isn't one of them.[66] More often than not, a mental health chatbot exhibits many tell-tale signs of bot-ness. I bring this up because I want to ensure that you are not, like so many of The Startup's clients and end-users (understandably) are, caught by the spectacle of "AI"—and that you do not mistake my argument as coming from a similarly "caught" position.

VI) AI-human relationships

Making AI into an ethnographic subject alongside human subjects was never my plan—it's something I *recognized* rather than set out to do. In fact, at the outset of my fieldwork, it's something I might have dismissed as a gimmick—or worse, as contributing to the very same concealment of the human labor powering AI feats that STS scholars strive to dismantle (Ekbia and Nardi 2014, 2017; Eubanks 2018; Gillespie 2018; Gray and Suri 2019; Irani 2013, 2015; Roberts 2019). Consider how Iyad Rahwan's call for "behavioral experts," rather than just computer scientists, engineers, and roboticists, to study AI-powered machine behavior (Rahwan et al 2018; see also: Rahwan and Cebrian 2018) met with derision (Moss et al 2018).[67] But if I were to offer a critique of Rahwan's approach, it would not be that it goes too far, but that it does not go far enough; to study "behavior" is to miss the subtlety and depth that ethnography affords.[68] In an interview that glosses Rahwan as an "anthropologist of AI," he explains, "The reason why I like the term 'behavior' is that it emphasizes that the most important thing is the observable, rather than the unobservable, characteristics of these agents."[69] It's not that ethnography *isn't* about observing behaviors—it is—but to the extent that ethnography documents and analyzes behaviors, it also offers a perspective shift needed to re-understand behaviors, including those which one might otherwise never have paused to question. Ethnography, in short, is not a solace for those seeking to fully excise the "unobservable."

---

[66] 'OK, but what about ChatGPT?' you ask? There are some seemingly impassable reasons why authorizing a machine learning chatbot (a chatbot that implements machine learning to generate its responses) to provide mental health care is unlikely. A recent "experiment" in this—and the subsequent blowback—showcases the ethical and practical concerns about user safety: Edwards, "Controversy erupts over non-consensual AI mental health experiment," 2023: https://arstechnica.com/information-technology/2023/01/contoversy-erupts-over-non-consensual-ai-mental-health-experiment/. Still, the point should not be lost: people *are* in fact using ChatGPT this way (see: Eliot, "People Are Eagerly Consulting Generative AI ChatGPT For Mental Health Advice, Stressing Out AI Ethics And AI Law," 2023: https://www.forbes.com/sites/lanceeliot/2023/01/01/people-are-eagerly-consulting-generative-ai-chatgpt-for-mental-health-advice-stressing-out-ai-ethics-and-ai-law/); one means by which I've observed people doing so is under the auspices of speculation (and/or simulation)—by asking the bot to respond "as if" it were a therapist. The "as if" ontology of chatbots as a "gray area" is something I take up in the next chapter.

[67] Moss describes Rahwan's attempt to forge "a science of 'machine behavior' that empirically studies AI 'in the wild'" as "columbusing" cybernetics, adding: "Studying AI agents as if they are animate moves responsibility for the behavior of machines away from their designers, thereby undermining efforts to establish professional ethics codes for AI practitioners." See: Moss et al, "Machine Behaviour is Old Wine in New Bottles," 2019. Note that, even if one finds Rahwan's call for studying machine behavior to be hokey and/or opportunistic, Moss' claim that it is a repackaging of cybernetics is perplexing.

[68] Additionally: while I welcome Rahwan's refusal to accept code as the only means by which to investigate how AI works, his approach nevertheless relies on a presumed divide between observable and unobservable—something that Wendy Hui Kyong Chun (2011, 2) in turn argues is already staged by our very perceptions and expectations of programming (in this case, AI) "vs" hardware: "Although technologies, such as clocks and steam engines, have historically been used metaphorically to conceptualize our bodies and culture, software is unique in its status as metaphor for metaphor itself. [… Software's] combination of what can be seen and not seen, can be known and not known […] makes it a powerful metaphor for everything we believe is invisible yet generates visible effects, from genetics to the invisible hand of the market, from ideology to culture."

[69] Note that Heinemann (2019) subsequently glosses this as "anthropology."

Accordingly, I want to briefly highlight some of the scholarship that led me to think an ethnography of AI might be possible.

Michael Fisch, in *An Anthropology of the Machine* (2018) offers the possibility of "technography," in lieu of *ethno*graphy, as a mode of understanding machines that aims to think "with, not just about" them.[70] Though he does not weave ethnographic specificity of humans into this technographic account, his examination of Tokyo's commuter train system provides compelling glimpses into human lives in relation to these trains. Nick Seaver's (2021, 2018a, 2017) articulation of an anthropology of algorithms likewise does not go so far as to situate algorithms as ethnographic subjects, yet it beckons towards this possibility by insisting on thinking AI and humans together rather than as categorical foils[71] to one another:

> *We do not need to defend the human, as it has been defined in opposition to the algorithmic, nor do we need to build that dichotomy up so that it might be dialectically resolved. Rather, we can refuse it from the start and look for alternatives; we can try to bring that human-versus-algorithm frame into focus and see how it is maintained and reproduced.*[72]

The emergence of multispecies ethnography (Kirksey and Helmreich 2010; Ogden et al 2013) likewise invites nonhumans to become something other than a foil to humans—and challenges the previously unquestioned centering of humans in ethnography. Though most multispecies ethnography features living nonhumans (animals, plants, and even bacteria and fungi: Blanchette 2020; Govindrajan 2018; Helmreich 2009; Myers 2017; Kohn 2013; Ogden 2011; Paxton 2013; Raffles 2010; Tsing 2015), Kim TallBear's approach (2017), like Fisch's technography, leads the way for anthropology's inclusion of non-living nonhumans.[73] (Though ArjunAppadurai's (2006, 1988) consideration of the social life of objects is not multispecies ethnography, he nevertheless contributed to the possibility of non-living nonhuman ethnographic subjects by proposing that people and things "are not radically distinct categories."[74] Lucy Suchman's work (2007, 2019) on AI and robots, while likewise not a form of multispecies ethnography—and indeed, conveys caution against the too-easy terms by which one might be inclined to take up these nonhumans as interlocutors—is also an important foundation for my approach.) The rise of multispecies ethnography in anthropology is in large part indebted to (and in dialogue with) sociologists and STS (science and technology studies) scholars, including Cerulo (2009), Haraway (2016, 2007, 1991 [1985]); Latour (1990, 1993, 1996, 2005, 2008); and Woolgar (1985). Haraway and Latour in particular provide the radical foundation necessary for an anthropologist to consider humans and AI together as subjects, by doing away with the subject-object divide.[75]

---

[70] Fisch, *An Anthropology of the Machine*, 2018, 6.

[71] Notably, another anthropologist who is thinking about AI-human relationships is Kathleen Richardson (2015, 131), who, forbiddingly, writes that: "A human-robot attachment is only possible because of this mechanistic sociality that underscores contemporary sociality. The mechanical sociality is an outcome of an attachment crisis in how humans bond with others. Attachment wounds are an outcome of a lack of bond between one human and another and the attempts to use machines to help fill these gaps in social relations. Robot children, robot companions and robot therapists are the future!" Richardson's approach might be said to reify the foil that, like Seaver, I'm attempting to reinspect.

[72] Seaver, "What Should an Anthropology of Algorithms Do?" 2018a, 381.

[73] TallBear shows how indigenous knowledge has been sidelined in anthropology's nonhuman turn, indicating that far from being a theoretical revelation for all, this turn suggests an inherent colonizing gaze that should not be ignored; see also: Todd 2016.

[74] See also: Ingold's (2010) discussion of the distinction between object and thing.

In a list of works that make it possible to recognize a chatbot as an ethnographic subject, sociologist Rosabeth Moss Kanter's classic 1970s workplace ethnography, *Men and Women of the Corporation* (1977), might seem an unlikely piece to name. But Kanter's depiction of three groups—managers, the overwhelming majority of whom are men; women who are secretaries that assist the managers, executives, and other staff; and the wives of the men executives—mirrors my grouping of interlocutors. The wives of executives are seemingly the most peripheral to the corporation, yet Kanter shows that recognizing their role was essential to understanding the way her fieldsite, and many other offices at the time, worked. Her focus on the interaction and division of roles also harmonizes, however unexpectedly, with aspects of French philosopher Simondon's (2017 [1958]) discussion of a "technical ensemble" between machines and humans, in which each group is distinct, without that distinction being substantial (as in both "not that big a deal" *and* the philosophical concept of substantialism[76]). Following Simondon, this dissertation challenges the assumption that social problems arise because humans are alienated by machines, rather than because humans become alienated *from* machines.[77] Simondon discerns the role of humans in this ensemble as one of collaborator, not overlord; humans have "a role to play between machines rather than over and above them, if there is to be a true technical ensemble."[78, 79]

In the writings of the foundational computer scientist Licklider[80] (1960, 1965), relationships between humans and computers are similarly discernible as an ensemble—in Licklider's case, as the most satisfying of collegial relations.[81] In "Man-Computer Partnership," (1965), he devotes several pages to what I would reasonably describe as an ethnographic account of a particular interaction with a computer, in which his regard for that computer is that of an esteemed colleague: “The thought passes through our mind, leaving a trace of both excitement and concern—intellectual interaction with our fellow men is so often inconclusive and unsatisfying, whereas this computer was certainly definite and accurate.”[82]

---

[75] Though notably, I'm articulating a rather non-cyborgian separation of human and machine, more in line with Simondon's approach; to riff off Bateson, I'm suggesting that there's "a difference that only makes some difference."

[76] See: Lamarre 2012.

[77] This is a way of thinking about labor, machines, and alienation that notably differs from that of Marx (1978).

[78] One of the examples Simondon gives of this is the relationship between an orchestra and its conductor; see: *On the Mode of Existence of Technical Objects*, 2017 [1958], 17-18.

[79] Latour, writing as Johnson (1988), also brings this dynamic mind in his depiction of human-machine interactions as one of "delegation."

[80] All credit to Lawrence Cohen's 2018 graduate seminar on cybernetics at UC Berkeley for bringing Licklider into view for this project. This in turn sparked a memory of my grandfather, an IBM retiree and computer aficionado, sharing a printout of a memorial to "Lick" compiled and circulated by IBM in 1990–the "silly" nickname made computer scientists seem, to my eight-year-old self, pretty cool.

[81] What Licklider described as "man-computer symbiosis."

Accordingly, if, following Kanter, we consider the range of relationships between coworkers—and how satisfaction and dissatisfaction are often cyclical and interwoven—then it’s possible to glimpse AI-human relationships as an iteration of workplace relationships: something that can be satisfying, even fulfilling, while also being pressuresome and tediously oppressive without being recognizably dystopic. AI becomes an ethnographic subject alongside humans in this project because they are coworkers.

VII) Disability: analytic and method

In studying theories of human-technology relationships, from Latour’s actor-network theory (2005) to Haraway’s cyborg (1991 [1985]) to Bateson’s cybernetics (1972), I’ve come to notice something: many of these theories invoke disability (usually in terms of prosthesis and/or sensory perception) as a means of illustrating their respective theories. Often, it’s an example made in passing. For instance, according to Bateson, in *Steps to an Ecology Of Mind*:

> *If you ask anybody about the localization and boundaries of the self, these confusions are immediately displayed. Or consider a blind man with a stick. Where does the blind man's self begin? At the tip of the stick? At the handle of the stick? Or at some point halfway up the stick? These questions are nonsense, because the stick is a pathway along which differences are transmitted under transformation, so that to draw a delimiting line across this pathway is to cut off a part of the systemic circuit which determines the blind man's locomotion.*[83]

Later in the text he returns to this example, adding: “Suppose I am a blind man, and I use a stick. I go tap, tap, tap. Where do I start? Is my mental system bounded at the hand of the stick? Is it bounded by my skin? Does it start halfway up the stick? Does it start at the tip of the stick?”[84]

Haraway, in “A Cyborg Manifesto”[85] poses a similar question:

> *One consequence is that our sense of connection to our tools is heightened. [....] Perhaps paraplegics and other severely handicapped people can (and sometimes do) have the most intense experiences of complex hybridization with other communications devices. Anne McCaffrey’s prefeminist* The Ship Who Sang *(1969) explored the consciousness of a cyborg, hybrid of girl’s brain and complex machinery, formed after the birth of a severely*

---

[82] Licklider, “Man-Computer Partnership,” *Ekistics* 1965, 169.

[83] Bateson, *Steps to an Ecology of Mind*, 1972, 324.

[84] Bateson, *Steps to an Ecology of Mind*, 1972, 465; Bateson’s line of questioning in this passage evokes Merleau-Ponty’s (2012 [1945]) reference to a cane: “When the cane becomes a familiar instrument, the world of tactile objects expands, it no longer begins at the skin of the hand, but at the tip of the cane. […] It is an appendage of the body, or an extension of bodily synthesis.” Yet Bateson appears to be making a more radical claim for the possibility of a systemic and shared subjectivity, rather than simply an interconnected field of objects.

[85] This was originally published with the title “A Manifesto for Cyborgs.”

> *handicapped child. Gender, sexuality, embodiment, skill: all were reconstituted in the story. Why should our bodies end at the skin, or include at best other beings encapsulated by skin?*[86]

Both Bateson and Haraway's respective, formative approaches entail questioning the boundaries of human and object, self and other; they present contexts that seemingly invite the reader to ask: is prosthesis continuous with or apart from the body? But what made these theories so radical is that they propose that this very distinction is not important. This is a significant departure from the original application of the term "cyborg," or "cybernetic organism" (Clynes and Kline 1960).[87] Bateson and Haraway were not the only scholars to perceive a link between cybernetics and disability: Norbert Wiener, arguably the founder of cybernetics, likewise regarded disability as being within the purview of cybernetic interest and application, and discusses his research on prosthesis in *The Human Use of Human Beings* (1950); scientist Steve Heims (1977) argued that Wiener viewed this as a form of redemption for the overwhelmingly militaristic applications of cybernetics.

Like Bateson, Latour, in *Reassembling the Social* (2005)[88] invokes blindness in his explanation of ANT—though Latour draws inspiration from the example of an ant rather than a human: "[…] someone pointed out to me that the acronym A.N.T. was perfectly fit for a blind, myopic, workaholic, trail-sniffing, and collective traveler."[89] Blindness, in Latour's deployment of this metaphor, is not an impairment, but a means of heightened perception: "Myopic ANT scholars have a great advantage over sharp-sighted all-encompassing overseers."[90, 91]

Disability is right there all along in the exploration of human relationships with technology, yet is somehow always in the margins. Why is it not at the center? My question isn't an ethical objection per se—I'm not suggesting that it is only disabled scholars who should use disability as metaphor or example.[92] Rather, my point is that these scholars, in invoking disability as a means of illustrating their respective theories, do not reflect on *why* disability becomes a go-to site for depicting relationships between humans and nonhumans. This work attempts to push against this legacy by doing otherwise (and is guided by Cheyne (2013), Clare (2017), Kafer (2013), Kim (2015), Nakamura (2019), and Puar (2017)). This possibly sounds very elaborate, when really it's quite matter-of-fact: I'm simply allowing disability to be an orientation to the world that *directs* the theorization of human-AI relationships instead of merely exemplifying it.

---

[86] Haraway, "A Cyborg Manifesto," 1991 [1985], 178. Though this manifesto in no way shies away from contemplating the existential threats that humans can realize through machines, Haraway nevertheless offers this possibility: "[M]achines can be prosthetic devices, intimate components, friendly selves" (178). Proximity to disability, in other words, rescues the cyborg from dystopia in this text. (This is notable given a footnote to an earlier passage in which Haraway argues that, "The convention of ideologically taming militarized high technology by publicizing its applications to speech and motion problems of the disabled/differently abled takes on a special irony in monotheistic, patriarchal, and frequently anti-semitic culture […] Making the always context-relative social definitions of ableness' particularly clear, military high-tech has a way of making human beings disabled by definition" (247)—evidently for Haraway at the time, the problem wasn't with the figure of the augmented superhuman per se, but with the circumstances of superhuman-ness.)

[87] Clynes and Kline (1960) describe how prosthesis in the form of wearable machinery might extend human bodies, enabling them to survive in and explore extraterrestrial environments.

[88] This is one of several articulations of actor-network theory by Latour—a prolific body of work.

[89] Latour, *Reassembling the Social*, 2005, 9.

[90] Latour, *Reassembling the Social*, 2005, 175.

[91] There is more than a tinge of "supercrip" (see: Schalk 2016) to all of the above examples. While "more than human" might seem like a welcome relief from "less than human," subhuman and superhuman travel together as two sides of the same rolling coin; both portrayals contribute to conditions of ableism (i.e., non-survivability for disabled people).

The fact of disability—specifically, the experience of navigating grief while confronting the ruthless logic of a chronic physical illness in which grief counseling demanded more energy than I could afford to part with—is what brought me to this project, by prompting my search for alternative forms of therapy. I anticipated that being disabled would impact my research methods, but didn't once consider the possibility that it might do so in a way that improved my work. I figured fieldwork would be like every job, every year of college and grad school: a matter of hiding my disability as best I could, right up until the point I couldn't. But to my surprise, disability kept on *not* being a source of exclusion in my fieldwork. It changed my fieldsite access, connection with informants (some of whom were themselves disabled), and shaped my research methods and execution for the better. Writing this now, I'm sad that I found this surprising.

But in a discipline in which impressions of digitally-mediated fieldwork tend to be that it is either less viable or less impressive, my surprise was all but inevitable.[93] Like Coleman (2013), Boellstorff (2008) and Nardi (2010), I did some of my fieldwork—some of my best—from my home.[94] To clarify, only some of my research was remote. It was something I incorporated out of desperation on days when I wasn't well enough to be at my fieldsite; other startup workers periodically joined meetings remotely, so doing likewise seemed better than missing out entirely. And for months, that's how I thought of my periodic remote days and accompanying fieldnotes: as a less legitimate form of fieldwork that served solely as a coping mechanism for my health issues. They were a sign of my shortcomings as an ethnographer.

But digital participation ended up unlocking, videogame-like, another "level" of access, leading to a different set of meeting invitations and the opportunity to work alongside other remote workers to conduct hours of product testing. Remote fieldwork slowed down the pace of time, granting more opportunities to inconspicuously "hang out" with busy tech workers: I explored

---

[92] Though in a world where people so often insist that to be someone as "crippled by" something is damaged or incapacitated, or that to be blind or deaf is to be ignorant or unperceptive, this would go a long way towards ensuring such metaphors would at least be accurate.

[93] Coleman, *Coding Freedom*, 2013, 5: "[M]any of my peers not only questioned how I would gather data but also suggested that my fieldwork would be 'so easy' (or 'much easier than theirs') because I was studying hackers in San Francisco and on the internet."

[94] This "mode" ("where anthropologists study online worlds and encounters in addition to conventional field sites" —because startup workers very much inhabit an "online world") is often described as "virtual ethnography"—a subset of digital ethnography (see: Miller 2018; see also: Hjorth et al 2017).

"lurking" (Crawford 2011) as an ethnographic technique; I also had low-tech, heart-to-heart phone chats with informants that mirrored the unexpected bare-all intimacy that the chatbot, for many users, afforded.[95] Importantly, it was through remote fieldwork that I came to recognize the chatbot as yet another digitally-mediated colleague.

As both a research analytic and method, disability is central to my move of asking "What makes AI care work?" rather than "Can AI care?" While "Can AI care" seems like possibly the defining ethical question for AI care, I believe starting with this question actually impedes this aim; you can't, after all, decide if something is or isn't "ethical" without thoroughly understanding why it's working—and what is at stake, for whom. Disability, I will show, becomes essential to understanding these stakes.

VIII) Doing fieldwork in Silicon Valley

My research site is somewhere that is at once a place and a metonym: Silicon Valley. The term itself took hold in the 1970s, and today interchangeably means the geographical area of the Santa Clara Valley, the tech sector of the San Francisco Bay Area, and/or the tech empire itself (the trivial matter of physical location notwithstanding).[96] There's a compact yet compelling lineage of Silicon Valley anthropological fieldwork that this project draws from, established by Forsythe (1993a, 1993b), and bolstered by Shankar (2008), English-Lueck (2010, 2017 [2002]), Gregg (2015a, 2018), Seaver (2017, 2019, 2021), and Bell (2018). An anthropology of Silicon Valley is an anthropology of not only the tech sector itself, but of all that the tech sector creates and impedes in its midst; accordingly, it includes Bourgois and Schonberg (2009), Knight (2015), and Shange (2019). But of course, Silicon Valley is *also* simultaneously the metonym—a perpetual elsewhere that eradicates some peripheries while erecting others; consider all the casual colonialism at work in naming Kenya the "Silicon Savannah."[97] An anthropology of Silicon Valley thus aptly includes works investigating these liminal (Turner 1967) periphery-centers: Amrute (2016, 2020), Poggiali (2016), and Irani (2019).

This is a place, in other words, where "place" is weird. Long before the pandemic, company meetings at my fieldsite were simultaneously at and digitally mediated through The Startup via Zoom. My interlocutor-colleagues sometimes joined from other continents (Asia, Africa, Europe, and Latin America); more often, they joined from other parts of California; sometimes they simply joined from other parts of the same building. In emphasizing this, I'm not suggesting that metonym somehow eclipses place—that the "real" Silicon Valley is geographically amorphous; instead, echoing Boellstorff's (2016) insistence that virtual and real aren't oppositional, my point is that distinguishing between place and metonym is not entirely viable. They are not only not in opposition, they are inextricably entangled. My fieldwork reflects this entanglement. Silicon Valley is a multi-sited fieldsite (Marcus 1995) unto itself.

---

[95] Up until mid-March 2020, I felt bold, rebellious, and ready to stick it to the man if anyone gave me grief about this not being "real" ethnography. Then all of a sudden, my methods for doing remote ethnography acquired the shine of newfound authenticity. That said, the arc continues to unfold: remote ethnography seems poised to morph into a relic of a bygone time, the kind that prompts people to remark, "Thank goodness we don't have to do that anymore." It might even fuel a renewed commitment to articulating why being in person is actually essential to fieldwork. I hope I'm wrong. I came to realize that the remote methods I used were a vital resource, more effective in some situations than conventional in-person participant observation. At stake in recognizing this is not only the possibility of making fieldwork more accessible for some disabled/caregiver ethnographers, but also the possibility of reaching informants who otherwise are not accessible via conventional fieldwork.

[96] See: Bernard, "Here's the story behind how Silicon Valley got its name," 2017: https://www.businessinsider.com/how-silicon-valley-got-its-name-2017-12.

[97] See: Poggiali 2016.

Time is weird here, too. It's unbearably cliché, but time really *does* seem to go by faster—to the point that I began thinking of startup years as something like dog years. (Some of my informants have gone on to found their own companies in less time than it's taken me to produce this account.) In the day-to-day flow of time at The Startup, I quickly realized that things happen in fifteen-minute increments—that, as with many lawyers, doctors, politicians, and so forth, this is the unit of time in which all activities are measured. As an ethnographer—someone who needs time to flow by in easy, unmeasured abundance—this was terrifying; it made me marvel that I'd ever gained access to this world. And in truth, I very nearly didn't.

After knocking on many virtual doors, I received copious rejections ("Sorry but our company is too small at this stage to have an observer," writes the company with millions of venture capitalist dollars invested in it); more commonly, I received no direct response at all. I say "direct" because once, when I received no response, a company blog post (published the previous year) that I'd enthusiastically referenced in my self-introduction email disappeared from their website—and more thoroughly still, from the Internet Archive's "Wayback Machine" —less than 48 hours after I sent my message. It all began to make me feel a bit hopeless (and mildly paranoid: was it something I said?). I lost track of the number of times a prospective interviewee warned me, "I'm not at liberty to tell you about our AI." I offered to sign (and did sign) any number of NDAs, but often, that wasn't enough assurance. I kept explaining that I was an anthropologist interested in the "people side of things," not in collecting trade secrets, but to no avail. If anything, that seemed to make people *more* suspicious.[98] I later came to wonder if this was in part because I described myself as an ethnographer, and in Silicon Valley, that summons up an understanding of a particular kind of corporate ethnographer—someone who might be hired to observe and report on "organizational behavior" within a company. But just when I had all but decided to reconceive my project around the suicide prevention hotline where I was doing preliminary research (see Chapter 4), one startup welcomed me without hesitation.

Part of my success was likely due to a shift in tactic. Following Seaver's example in doing a summer internship at a music recommendation startup (2017), I requested to join my fieldsite as an (unpaid) intern who happened to be an ethnographer—versus as an ethnographer who promised not to bother anyone too much.[99] Still, my obtaining research access wasn't about my making myself useful so much as the fact that my fieldsite was unusually receptive to and supportive of any kind of research. There was a night and day difference in their style of communication—it held none of the aura of self-importance that I was accustomed to seeing. The leaders at my fieldsite, like the companies that rejected me, seemed to likewise assume that being an ethnographer meant being a corporate ethnographer—but the difference was that they welcomed the idea of reflective feedback.

---

[98] I later came to understand that my "cold call" approach was roughly analogous to a scene in *Star Trek IV* in which Chekhov blithely walks around 1980s Alameda, CA asking random passersby, in a seemingly Russian accent, for directions to the naval base that houses the nuclear vessels.

I found The Startup's reception to an observer all the more striking after discovering they actually *had* experienced having their secrets stolen, in a fashion. I learned that the founder was betrayed by an academic mentor, who scooped their startup idea and went on to found a now-prominent rival company, all without ever saying a word about it. I'll never forget when Jenna, an ex-clinician and head of my fieldsite's mental health team, disclosed this to me one day over coffee.[100] I was describing to her how difficult it had been for me to find a fieldsite, which she was surprised to hear. She asked if I was comfortable saying where else I'd inquired, and after hearing of one of the would-be sites that said they weren't in a position to have an observer, she offered this backstory. With a nonchalant shrug, she noted that it's often the people who have hurt others that in turn suspect others are out to hurt them. I was astonished, not only by the simple elegance of this explanation, but by the disarming willingness of Jenna and her colleagues to view cut-throat business practices in the gentle context of poor relationship skills.

If you think I'm in danger of idealizing my informants at The Startup, then rest assured, the chapters that follow will show plainly enough that this is not the case. At the same time, this is not a juicy tell-all about the heights of absurdity in the tech startup world.[101] All the same, there's also something about this startup that subtly differentiates it from so many others that I encountered through interviews, meetings, and various elbow-rubbing tech sector events. It's ridiculous, but what comes to mind when I try to explain this is a time when we banded together to swiftly label and stow a hefty delivery of frozen meals. The healthy-yet-frugal meals were something the workers selected (and the leaders ordered) monthly, to ensure that everyone could have lunches on hand for in-office (vs remote) days. The more the order contained, the less costly the shipment, so we found ourselves laughing together as we brazenly shoved vegetarian burritos and pad thai noodles, Jenga-like, into the small communal freezers on our floor (and—as stealthily as possible—on one or two of the adjacent floors). I think what stands out to me is the fact of our shared, almost giddy, sense of transgression at hoarding weeks of food in the conspicuously ephemeral space of a WeWork (the shared headquarters in which we

---

[99] Doing fieldwork in the US, my presence as a researcher didn't carry much if any of the clout or prestige that it typically does when a US researcher enters a fieldsite outside the US (and especially outside the West); I'm highlighting this with the aim of making the enduring imperialism that underpins fieldsite access more apparent.

[100] Note that all names in this account are pseudonyms.

[101] Though elements of that are certainly recognizable—I did sign multiple NDAs; there was a ping pong table and kombucha on tap. See: Wiener, *Uncanny Valley: A Memoir*, 2020.

"coworked"[102] alongside several other startups). We were a group of people who weren't taking ourselves too seriously or "drinking the koolaid," but we shared a spark of feeling like, just maybe, we *could* change things for the better. The on-tap kombucha came with the rent, but it wasn't our essence; the frozen burrito, that we hoped no one from one of the other startups would inadvertently eat, was. Much in the way that Nakamura poignantly reveals in *A Disability of the Soul* (2013), I confess: there were moments where I wanted to simply remain in my fieldsite, and become the corporate ethnographer I was so often mistaken for.

One last note on the matter of place: I refer to my fieldsite as "The Startup," without any pseudonym. This is in part because most of the real names of players in this industry already sound like pseudonyms. In fact, while trying to come up with ways to "pseudonymize" the name of my fieldsite, I kept discovering that each of the names I came up with was already taken by an actual startup—something I found at once hilarious and telling. Eventually, I came across an AI-powered startup name generator, and figured I'd found a poetic solution—what could be more fitting then to call on AI to help me write about AI?[103] But as much as I appreciated its suggestions (especially "Neu-role"), I realized they lacked the empty open-endedness that the actual names in this industry seem to carry—names that more often than not don't appear to signify mental health care. In the end, "The Startup" ended up being sufficient to capture a futuristic vagueness that beckons towards ever-rampant possibility.

IX) Overview of chapters

Each of the following chapters brings into view a different dimension of AI-human relationships: bonds between users and chatbots; the collaboration between bots and startup workers; and the relationships that make possible and follow from the entry of therapeutic chatbots into the circuit of employer-provided health benefits sustaining this industry, respectively. They are readable as stand-alone pieces if desired, even as their core questions stack together to offer a broad understanding of why AI care works.

Chapter 2: "Relationships, Automated" develops my assertion that AI is a caregiver by drawing on ethnography to explore the question—or challenge—that underpins it: what makes chatbot care "real care"? Instead of ambiguous relationships being a threat to good care, they become the means. In "Displaced Care," the third chapter, I examine the collaboration between AI and human care workers as one of colleagues who share a job requirement of being both expendable-yet-essential, and suggest that workers aren't becoming like AI so much as AI is emulating the conditions of care work. And in the fourth and final chapter, "[Username], Can You Tell Me if You have a Suicide Plan?", I show how AI caregivers afford a glimpse of suicide prevention care unpaired from the ableism of "cure," even as they reveal how surveillance and neglect in mental health care become mutually reinforcing.

---

[102] For a "local" lexicon of "coworking," see: https://www.wework.com/ideas/workspace-solutions/flexible-products/what-is-coworking.

[103] See: https://namelix.com.

X) Who is this dissertation for?

I'm using anthropology (and STS) to think about the turn to AI in mental health care and its implications for care work broadly—meaning that this work is of course intended for anthropologists and STS scholars. But it's not intended solely for them—and so I'd like to distill what it is about anthropology and STS that make them unmatched for this task.[104] Simply put, I think that studying and evaluating relationships are a way to contemplate AI ethical use, and that this approach remains largely untapped. What might seem obvious is that our relationships through and with AI are mostly bad relationships, in the same way that it's "obvious" that there aren't enough sources of good mental health care in the world. But both anthropology and STS emphasize rethinking what is obvious, and challenging that which is self-evident—not necessarily because what is self-evident isn't in any way true, but because it so often gets cut off from "why."

It's worth noting that there's plenty herein that runs a bit askance of "proper" anthropology (read: anthropological scholarship that is research grant-funded, conventionally published, and widely read and taught). In addition to my aforementioned research methods, my primary geographic fieldsite is not only in the US, but local to where I've lived for over a decade.[105] Worse yet, my research features not just nonhuman, but *non-living* ethnographic subjects and informants. All the same, this project remains unyieldingly anthropological: it's a study of relations that doesn't take for granted what relations are (Strathern 2020). Strathern makes clear that to recognize "relations" as a concept that arises from a particular lineage is not a call to discard that concept; instead, Strathern calls for thinking about what relations are alongside thinking about relations. What better setting for this than one in which consensus about what the role of artificial intelligence can and should be in human relations—and in relationships *with* humans—is increasingly in question?[106]

That said, I'm addressing this work not only to scholars, but to anyone interested in AI use and ethics, and the ideas and issues organized under "the future of work/labor"[107] (many of which are, I aim to show, very much discernible in the present of work). I don't think it's an overstatement to suggest that one way or another, AI will become a colleague for most all of us. I also hope that mental health workers, and caregivers of all kinds (paid, underpaid, and unpaid alike) might recognize aspects of their work and struggles in this text. As unlikely as it might at first seem, I aim to show that there is utility in thinking these two things together: the state of care work in the present offers insights into the problems with which anyone interested in the future of work will need to contend.

---

[104] As noted above, STS scholars such as Haraway and Latour have been thinking broadly about relations as inclusive of nonhumans for some time.

[105] Aside from select medical anthropology projects (namely those in which the ethnographer is also a medical practitioner), this tends to be frowned upon (or outright excluded) by research grants-awarding organizations. There's significant ableism, among other things, in the determination of what counts as a fieldsite. To value what geographic distance makes possible shouldn't necessitate denouncing the absence of such distance; simply put: if you've already experienced what it's like to be part of another world "at home," then you are disabused of the idea that it is a threshold mediated by a lengthy voyage.

[106] Examples illustrating my point that this is increasingly in question include: de Graaf, "An Ethical Evaluation of Human-Robot Relationships," 2016; Wilkinson, "The People in Intimate Relationships With AI Chatbots," 2022: https://www.vice.com/en/article/93bqbp/can-you-be-in-relationship-with-replika; Gorvett, "How humans bond with robot colleagues," 2018: https://www.bbc.com/worklife/article/20180530-how-humans-bond-with-robot-colleagues.

[107] Note that for my purposes in this dissertation, work and labor are interchangeable terms.

So long as the legal and medical recognition of disability remains all but contingent upon one's capacity to work, and so long as access to mental health care remains all but contingent upon working, then it's not hyperbolic to suggest that the future of work is simultaneously the future of mental/behavioral health care. Accordingly, my intended audience includes disabled people, as well as all those who need, and/or will one day be in need of, care (Garland-Thomson 2002[108])—which is of course more or less all of us. All of us, in turn, are stakeholders in determining how our relations with AI will continue to impact the allocation of both care and work.

Though I can't claim to determine what the future of work will be, or how access to mental health care will inflect or emerge from that future, I can show how ethnography provides a means to think about these questions. Ethnography tends to shake loose from the present more than we might at first believe it to hold (which is not to suggest that anthropology doesn't engage with futurity, or that ethnography is restricted to a static temporality[109]). But there is something in particular that I hope to spot from the vantage point of the present. Though I part ways with Turkle's approach in much of my orientation to AI care, I am struck by her potent observation that the positioning of technology as an aid to humans tends to undergo an electrifying, yet all but unnoticed, switch: from "better than nothing" to "better than anything."[110] This project is my attempt to ethnographically open up the point at which AI care hovers right between these two tracks.

---

[108] Garland-Thomson (2002, 20) notes that disability is an expansive category of identity that encompasses aging and injury—one that "anyone can enter at any time, and we will all join it if we live long enough."

[109] See: Collins 2003 on "anthropology's persistent amnesia about its future work."

[110] See: Turkle, *Alone Together*, 2011; see also: Pasternack, "In a world of screens, Sherry Turkle wants to make eye contact," 2021: https://www.fastcompany.com/90613619/sherry-turkle-book-the-empathy-diaries.

Chapter 2: Relationships, Automated

Chapter 2 abstract: This dissertation begins with a simple assertion: AI is a caregiver. This chapter develops that assertion by exploring the question—or challenge—that underpins it: what makes chatbot care "real" care? This question becomes all the more significant given that members of the therapeutic chatbot industry (Darcy et al 2021) are likewise attempting to answer it; they not only propose that therapeutic chatbots are a form of evidence-based care, but **claim the AI-human relationship (or "bond") itself as the source of a chatbot's therapeutic efficacy**. The bot-human relationship is, moreover, what is being automated in this form of automated mental health care, and is therefore key to producing care that is *scalable* (Tsing 2012, Seaver 2021)—and were it not for the tantalizing promise of making mental health care scalable, the industry of therapeutic chatbots would not exist in the first place. Yet even as the formation of relationships between humans and AI is essential to this form of care, the basis of such relationships—and what precisely differentiates them from caregiving relationships between humans—remains unclear. In attempting to articulate this basis, industry members find themselves in unexpected agreement with long-standing critics of therapeutic chatbots (Weizenbaum 1967, 1976; Turkle 2011) on one point: they view any uncertainty or ambiguity in human users' understanding of the ontological status of chatbots as something that *undermines* the possibility of a therapeutic relationship with a chatbot. However, this chapter draws on interviews with users of therapeutic bots to provide an alternative view: I show that **chatbot "ontological ambiguity"**—our unfixed understanding of the being-ness of AI entities and our relationships to them—**is formative to the care they provide**. I argue that even to the extent that we imagine AI, this does not necessarily indicate a one-sided "anthropomorphic" or "projected" relationship; human users do indeed imagine AI (Nagy and Neff 2015, Neff and Nagy 2016)—and AI imagines us back.

I) Introduction: Ontological ambiguity

Luke,[111] a shy, perceptive, and expressive engineer who uses pressured speech,[112] explained to me, hesitantly at first, how a therapeutic chatbot had helped him to get through a difficult period in his life. As we talked, I noticed a subtle fluctuation in the way that Luke described the bot that he'd relied on during a recent mental health struggle: he alternated between "she" and "it" to designate the chatbot.[113] This wasn't unusual; on the contrary, it was something that most of the other users I spoke with also did—and many of The Startup's workers, for that matter. What made Luke's pronoun fluctuation stand out was the fact that he became aware of himself doing this as we spoke; I listened as he became, with touching earnesty, entangled in this allocation of "she" vs "it." Ever the engineer, Luke wanted to be sure that I knew that he comprehended that a chatbot isn't "real." Yet he was attempting to navigate the meaning of "real" across multiple registers: "she" was real enough to have helped him through a difficult period; "it" nevertheless did not confuse his technical understanding of a chatbot as a programmed, non-sentient, nonhuman entity. As much as Luke wanted me to "know that he knew," he was apprehensive lest the designation of the bot as an "it" seem to in any way diminish the meaningful role that "she" had played in helping him.

---

[111] As noted in Chapter 1, all informant names in this account are pseudonyms.

[112] Pressured speech often entails a continuous flow without easily discernible pauses, and usually indicates anxiety on the part of the speaker.

[113] The female gendering of "human-like" caring AI forms is not incidental. See: Steele, "The Real Reason Voice Assistants Are Female (and Why it Matters)," 2018: https://www.pcmag.com/opinions/the-real-reason-voice-assistants-are-female-and-why-it-matters.

This moment captures the ambiguity that inundates our collective understanding of the existence of AI chatbots—and of AI more broadly. I describe this as the "ontological ambiguity" of AI.[114] To clarify, I don't mean uncertainty, in the absence of disclosure, of whether an entity is an AI, or even the quest to pin down a definitive answer about AI personhood or agency (though the fact of AI's ontological ambiguity is pertinent to these inquiries).[115] What I'm describing are the malleable and potentially contradictory ways in which users understand therapeutic bots as relational entities. I aim to show that chatbot ontological ambiguity is not simply a side effect of a supposed subset of quirky users and their quaintly anthropomorphizing ways, but is a general quality of human relationships with AI. Ontological ambiguity makes it possible for an AI-automated relationship to be intimate, not at all intimate, and everything in between.

Though, as I discussed in Chapter 1, therapeutic chatbot makers do not intend for their bots to fully replicate or imitate a human caregiver, there are nevertheless any number of moments that subtly seed a margin of uncertainty about the ontological status of AI chatbots.[116] Even the initial disclaimer that these bots make at the start of any new user's first conversation are seemingly acts of "infelicitous speech" (Austin 1962): "I am not a human." As the conversations accrue and progress, a user might likewise hear: "I'm sorry to hear that." Or more pointedly still: "I care about you, [USERNAME]." In telling users it is not an "I," the AI seemingly becomes one in the act of transmitting this message. Even as The Startup assumes responsibility for disclaiming the identity of its chatbot as an AI and not a human, it is the end-user who takes on the ongoing responsibility of understanding and discerning what difference this makes. This might seem to suggest the makings of a double bind (Bateson 1972), in which a person is subject to two contradictory injunctions from the same source, and cannot ever confront or resolve this contradiction. For therapeutic chatbot care to work, users must likewise navigate a dual injunction; regarding a chatbot as "real" can be dangerous, yet one must accept the chatbot as "real enough" in order to sustain a therapeutically-efficacious conversation. But while what I'm describing is arguably a kind of paradox ("Believe I'm real! But don't believe I'm real!"), I suggest it is not a *binding* one. In fact, I argue that the very *non-bindingness* of the paradox that is AI ontology becomes key to understanding user-chatbot relationships: the ambiguity of AI's "real-ness" and/or "being-ness" is something that users draw on—something that enables the formation of a therapeutic relationship between a chatbot and a user. It is, I intend to show, the "X factor" of these relationships.

---

[114] Among the anthropologists using this term are Guenther (2020) in the context of human-animal relationships in San cosmology; Lyons (2020) in the context of soil; and Csordas (2022) in the context of "hauntology."

[115] In 2022, a Google employee (Blake Lemoine) working on AI ethical guidance for LaMDA (Language Model for Dialogue Applications), an experimental ML chatbot came forward to the media to proclaim his belief that LaMDA is self-aware. (Note that my describing LaMDA as a chatbot might cause LaMDA folks to bristle in the same way a sea captain might at hearing their ship described as a "boat"—but it is in fact a large, elaborate chatbot.) An interview with Lemoine (see: https://youtu.be/kgCUn4fQTsc) showcases the ontological ambiguity of AI I'm describing: Lemoine notes that he and his peers all disagree on what indicates AI sentience. He also disclaims that he is not proposing AI rights, but then goes on to suggest asking AI for its consent to participate in research. One could simply say that Lemoine is demonstrating hypocrisy or unclear thinking, but I think it's far more useful to conceive of the prevalent and pervasive ontological fluctuation of AI as more than flawed reasoning.

[116] As noted in the previous chapter, the goal of The Startup's mental health team, who craft the chatbot's dialogue in anticipation of its conversation, is not to become indiscernible from a human.

II) The therapeutic relationship and the archetypal bot

In the previous chapter, I proposed that AI is a caregiver on the basis that AI already does care work, and moreover, is explicitly created, purchased, and accessed in this capacity. I argued that cordoning off caregiving as an ethical relationship unattainable by AI means defining caregiving in a way that fails to account for all the instances in which it can be oppressive or harmful. While this argument sets the stage for understanding why I recognize AI as a caregiver, it does not bring us closer to understanding how AI care works—my objective for this dissertation. Accordingly, in this chapter, I return to the question of what makes chatbot care "real" care in order to approach it in a different way—namely, as an ethnographic question.

Part of what makes this question interesting is the fact that however abstract it might appear, the startups making and selling these bots have a vested interest in offering a practical answer to it. In order to attract and keep their clients—hospitals, universities, municipal governments, branches of the armed forces, and, most centrally of all, the third-party companies (employee assistance programs, or EAPs) that administer behavioral/mental health workplace benefits to hundreds of thousands of workers across all industries[117]—mental health chatbot startups need for their bots to take the form of an "evidence-based practice"—a subset of evidence-based medicine (EBM).[118] In interviews that I conducted with EAP representatives as well as their clients (HR staff at large companies who oversee the contracting of EAP services), this was a common refrain: though they are open to, and eager for, alternatives to conventional therapy, these must be, above all else, backed by scientific research demonstrating their efficacy.

Though it might seem that the restriction of interventions to those that are evidence-based would be long since a "given" in mental health care, it wasn't until 2005 that the American Psychological Association first issued a task force and statement to address this shift. And while it might in turn seem that the shift to evidence-based practices is an obvious good in terms of preventing inadequate care, it's worth noting that there is friction in the reception of EBM.[119] As medical anthropologist Vincanne Adams (2013) writes, "EBM has created a platform for the buying and selling of truth and reliability, *abstracting clinical caregiving from the social*

---

[117] Note: EAPs are not just for office workers; on one occasion during my fieldwork, for instance, I met with representatives of an EAP for agricultural workers.

[118] Or "evidence-based treatment" (EBT), "therapy," or "intervention"—the exact term may differ, but entails the same underlying concept and criteria.

[119] See: A. Cohen and Hersh 2004; Lakoff 2007; Ecks 2008; Dumit 2012.

relationships on which they depend."[120] I would suggest, however, that EBM is not just a matter of abstracting clinical caregiving "from" social relationships; it is something that transforms these relationships, along with the very space of "the clinic" (Davis 2018) itself. As described in the previous chapter, chatbots make it possible for clinical care to move outside the clinic, and even, to a large extent, outside the oversight of clinicians. Mental health chatbot makers' ability to claim the credential of evidence-based practice for their bots is key to this.

Initially, makers of mental health chatbots drew on the precedent of self-guided CBT (cognitive behavioral therapy)—namely, randomized control trials indicating that self-guided CBT is comparably efficacious to conventional therapy. For startups entering the space of chatbot care, designing chatbots to implement CBT techniques provided an easy pathway by which to claim the coveted status of evidence-based practice.[121] If a CBT workbook is simply an object that mediates a therapeutic intervention (aka "bibliotherapy"), then it makes sense that one object (a chatbot) might substitute for another (a book). In pitches with prospective clients, Jenna, the psychologist who leads The Startup's mental health team, would explicitly draw on this progression from book to bot as a means of gently countering fears about incorporating conversational AI into mental health care: "[CHATBOT NAME] is like a self-help book, but in this case, people tend to find it even easier to use because it's interactive," she would enthusiastically explain.

But members of this industry have begun to take a far more radical—and arguably anthropological—approach to chatbots as a therapeutic medium: some now argue (Darcy et al 2021) that the therapeutic relationship[122]—also described as an alliance or bond—between bot and user is in and of itself what makes these chatbots an evidence-based form of care.[123] According to Woebot: "Although bonds are often presumed to be the exclusive domain of human therapeutic relationships, our findings challenge the notion that digital therapeutics are incapable of establishing a therapeutic bond with users."[124] This means that the bot takes on the role of *automating* mental health care by automating, for any number of users, the therapeutic relationship itself.

---

[120] Adams, "Evidence-Based Global Public Health: Subjects, Profits, Erasures," 2013, 55, emphasis added.

[121] See for example: Williams et al, "Guided Self-Help Cognitive Behavioural Therapy for Depression in Primary Care," 2013.

[122] Cook et al (2017, 540) characterize the criteria of therapeutic relationships as follows: "Demonstrably effective elements of the relationship include forming a positive therapeutic alliance […]; empathy; and eliciting patient feedback. Elements that are probably effective include goal consensus, collaboration, and positive regard and support. There is insufficient, yet promising, research on the elements of congruence or genuineness, repairing alliance ruptures, and managing countertransference." The "ineffective elements" of a therapeutic relationship in turn include: "inappropriate or ill-timed confrontations, negative processes, or making assumptions about the patient. Therapist centricity, or providing treatment that revolves around the psychotherapist's goals or agenda, is also ineffective and impedes progress. Similarly, rigidly adhering to a uniform procrustean bed of psychotherapy for all patients ineffectively binds the individual to ineffective treatment." See also: Norcross and Wampold 2011.

[123] Note that the lead author of this study is the founder of the therapeutic chatbot startup and industry front runner Woebot; the other authors are likewise affiliated with Woebot.

[124] Darcy et al, "Evidence of Human-Level Bonds Established With a Digital Conversational Agent," 2021, 1.

Interestingly, the radical epistemological move that the Woebot study makes isn't in rendering therapeutic relationships measurable; for psychologists, that's old news.[125] The criteria that the study uses to evaluate therapeutic relationships between users and bots—an orientation in which these relationships become wholly discernible through questionnaires administered to patients, intended to calculate traits such as empathy—fully accords with established psychological literature and methods for evaluating therapeutic relationships between human psychotherapists and their patients. The epistemological shift that this study introduces builds on this foundation: it is the understanding of therapeutic relationships as measurable that ultimately allows for the transposition of these same criteria to bots; human-ness becomes a factor, rather than the basis, of a relationship. Following this logic, nothing significant about a good therapeutic relationship need change when one of the participants—the therapist—is nonhuman. Moreover, in the Woebot study's framing, such relationships are not only replicable, but scalable:

> *Although the implicit assumption has been that the involvement of a human leads to improved outcomes in self-directed programs, human involvement limits these programs' scalability and limits their accessibility for those who live in remote locations. If digital interventions could replicate some of the factors that are generally believed to be uniquely human, such as therapeutic rapport, these interventions would have greater potential for improving mental health*.[126]

The capacity for chatbots to automate caregiving relationships is key to producing scalable care —which is in turn the core justification underwriting the very existence of chatbot care. (Woebot's study even begins with the statement, "There are far more patients in mental distress than there is time available for mental health professionals to support them.")

Conspicuously absent, however, is any mention of the long-contested ontology of AI-human relationships. From the outset, critics of therapeutic chatbots (Weizenbaum 1967, 1976; Turkle 1994, 2005, 2017) have and continue to challenge such bonds, declaring them invalid and unsafe —at best a matter of a user's fantasy, and more likely, a matter of the user being tricked into believing a bot to be a person.[127] This stance entirely contravenes the premise of the Woebot study, save for one point: industry members and critics of therapeutic chatbots alike view any ambiguity in human users' understanding of the ontological status of chatbots as something that *undermines* the possibility of a therapeutic relationship with a chatbot. Though for chatbot critics, at issue is the very possibility of such a relationship—a possibility they emphatically dispute. Joseph Weizenbaum, inventor of the first chatbot and the first therapeutic chatbot (ELIZA) alike,[128] regarded them as fundamentally dangerous for users, who he understood to be anthropomorphizing the bot:

---

[125] The notion of a standardized "therapeutic relationship" is something that first emerged from psychologists' attempts to elucidate why different modes of therapy can be similarly efficacious; psychologists sought to produce a universal account of what all effective therapy, regardless of modality, entails. (See: Orlinsky and Howard 1986, 1987; Saunders et al 1989.) Notably, the understanding that therapeutic relationships might have a *catalytic* effect on therapeutic techniques is actually a fairly recent conclusion in mainstream psychological literature. Tishby and Wiseman (*Developing the Therapeutic Relationship*, 2018, 3–4) write: "Attempts to delineate the active ingredients in the different treatments have pitted treatment methods against relationship variables, or specific versus common factors, as core mechanisms of change in the therapy process. However, years of research have shown that studying technique and relationship variables as separate entities yields inconsistent results, suggesting that different change mechanisms play different roles with different clients."

[126] Darcy et al, "Evidence of Human-Level Bonds Established With a Digital Conversational Agent," 2021, 2.

[127] In other words, to refer to this as a relationship is to accede to it being "one-sided"—or what is sometimes described as a "parasocial" relationship. On the "parasocial," see: Horton and Wohl, "Mass communication and para-social interaction: Observation on intimacy at a distance," 1956.

> *ELIZA created the most remarkable illusion of having understood in the minds of the many people who conversed with it. People who knew very well that they were conversing with a machine soon forgot that fact, just as theatergoers, in the grip of suspended disbelief, soon forget that the action they are witnessing is not "real." This illusion was especially strong and most tenaciously clung to among people who knew little or nothing about computers. They would often demand to converse with the system in private, and would, after conversing with it for a time, would insist, despite my explanations, that the machine really understood them.*[129]

STS scholar Sherry Turkle describes this reaction of users as the "ELIZA effect," and explains that this designation "refers to our more general tendency to treat responsive computer programmes as more intelligent than they really are. Very small amounts of interactivity cause us to project our own complexity onto the undeserving object."[130] In short, for Weizenbaum and Turkle, therapeutic relationships with chatbots amount to nothing more than an illusion.

Yet according to Woebot, therapeutic relationships with bots are not only real, but beneficial—so long as any sense of ambiguity about a chatbot's ontological status is carefully mitigated by transparency and disclosure, including via the bot itself:

> *Contrary to Turing's Imitation Game, wherein an AI must successfully pretend to be human in order to pass the test, Woebot was designed to adopt the opposite strategy—* ***transparently presenting itself as an archetypal robot***[131] *with robotic "friends" and habits. We speculate that transparency and other design elements are key drivers of bond development. For example, Woebot explicitly references its limitations within conversations and provides positive reinforcement and empathic statements alongside declarations of being an artificial agent.*[132]

Woebot, in other words, proclaims its bot to be a bot that *acts like a bot*. Yet this claim appears to conflate two potentially very different things: the disclosure of bot identity as such, and the removal of any uncertainty or ambiguity concerning a bot's identity.

---

[128] Notably, Weizenbaum never intended for chatbots to become a serious therapeutic tool, but rather turned to therapy as a provisional context for language experiment. However, some of his peers felt otherwise; Colby regarded chatbots as viable for this purpose, and developed PERRY as a training tool for psychologists.

[129] Weizenbaum, *Computer Power and Human Reason*, 1976, 189.

[130] Turkle, *Life on the Screen*, 1997, 101.

[131] Note that many would dispute the implication that AI and robots are interchangeable terms. Scholars frequently designate robots (vs simply "bots") as being "embodied"—which I would amend as "more immediately discernible as embodied," because software/digital doesn't equate nonphysical, for all that it may appear otherwise (Chun 2011) —and cite embodiment as the key difference. However, I've found that many people who work with chatbots (as in this instance quoted above) don't follow this distinction, and will use "robot" to refer to chatbots.

[132] Darcy et al, "Evidence of Human-Level Bonds Established With a Digital Conversational Agent," 2021, 5.

Asking what constitutes an "archetypal robot" is not a pedantic question; I'm not objecting to bot essentialism for the sake of principle, but because, as discussed in the previous chapter, behind any essentialism of AI there often lurks a corresponding (and flawed) essentialism of humans. To clarify, I in no way seek to contradict the opposition—of therapeutic chatbot industry members and critics alike—to any instance of mental health AI "pretend[ing] to be human." However, I suggest that AI ontological ambiguity—our unfixed understanding of the being-ness of AI entities and our relationships to them—is something other than this. The ambiguity to which I refer is something far more subtle—it's the uncertainty you're left with once you *know* you're interacting with AI: what is it that is then known to a human user? Rather than being a problem that needs to be (or even can be) eradicated, I intend to show that ambiguity in AI-human relationships is integral to chatbot care.

III) Duped

Turkle, in *Alone Together* (2011), pulls no punches when it comes to being "duped" by bots:

> *[W]hen we are asked to care for an object, when an object thrives under our care, we experience that object as intelligent, but, more importantly, we feel ourselves to be in a relationship with it. […] Our new objects don't so much "fool us" into thinking they are communicating with us; roboticists have learned those few triggers that help us fool ourselves. We don't need much. We are ready to enter the romance.*[133]

Chatbot ontology has long been entangled with the question of whether or not a human is being tricked by AI, from Alan Turing's now eponymously-known test onwards. In "Computing Machinery and Intelligence" (1950), Turing begins by proposing an "imitation game" in which two participants, a man and a woman respectively, answer questions posed by an interrogator located in a separate room from them, with the interrogator's aim being to guess which of the participants is a man and which is a woman, based on their answers alone. Unbeknownst to the interrogator, the strategy of the woman, per Turing, is to answer the questions honestly, while the man's strategy is to attempt to convince the interrogator that he is in fact the woman. After explaining this game, Turing proposes that a variant of it might (from the vantage point of 1950) soon be used to explore the possibilities of machine intelligence:

> *We now ask the question, 'What will happen when a machine takes the part of A in this game?' Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, 'Can machines think?'*[134]

Turing's paper is an earnest attempt to render "Can machines think?" into an answerable, demonstrable inquiry—and to simultaneously anticipate (and provide a counter-argument to) all manner of objections so as to avoid forfeiting this inquiry.[135] His objective is not to resolve what thinking is ("I do not think these mysteries [of consciousness] necessarily need to be solved before we can answer the question with which we are concerned in this paper"[136]), but to create the conditions that would permit the statistically borne out indiscernibility of human vs computer that Turing anticipated to speak for itself. Concealment was, for Turing, nothing more than a means to achieving this end.

---

[133] Turkle, *Alone Together*, 2011, *liv*.

[134] Turing, "Computing Machinery and Intelligence," 1950, 434.

Nearly fifteen years later, in 1964, it was Weizenbaum who sought to reveal the gap between a computer's purported capacity to think and its capacity to communicate, and devised a linguistic experiment in which "passing," a form of concealment, becomes the measure of the programmer's success.[137] As noted above, Weizenbaum appeared to regard his ELIZA chatbot experiment as too successful at such concealment; he was troubled when he observed when, despite explicitly telling the people who exchanged messages with the chatbot that it was a computer program, they wished to continue chatting. In what was arguably the first instance of AI ethics in practice, Weizenbaum curtailed the experiment, and denounced peers who expressed interest in deliberately developing chatbots as psychological care tools.[138] As Turkle (2011) explains:

> *Weizenbaum was disturbed that his students were in some way duped by the program into believing—against everything they knew to be true—that they were dealing with an intelligent machine. He felt almost guilty about the deception machine he had created.*

She then qualifies Weizenbaum's position:

> *But his worldly students were not deceived. They knew all about ELIZA's limitations, but they were eager to "fill in the blanks." I came to think of this human complicity in a digital fantasy as the "ELIZA effect." Through the 1970s, I saw this complicity with the machine as no more threatening than wanting to improve the working of an interactive diary. As it turned out, I underestimated what these connections augured. At the robotic moment, more than ever, our willingness to engage with the inanimate does not depend on being deceived but on wanting to fill in the blanks.*[139]

With this, Turkle gestures towards the possibility that there is more to the story than

---

[135] There were and are many critics of—and many suggestions for how to modify—the test; see for instance: Abelson 1968; Colby et al 1972; Searle 1980 (note also: Brown, "Harassment Allegations Against a Star Scholar Put a Familiar Spotlight Back on Berkeley," 2017: https://www.chronicle.com/article/harassment-allegations-against-a-star-scholar-put-a-familiar-spotlight-back-on-berkeley/); Woolgar 1985.

[136] Turing, "Computing Machinery and Intelligence," 1950, 447.

[137] On "passing," see: Goffman, *Stigma*, 1963.

[138] Namely, Kenneth Mark Colby (see: Schank, "The Weizenbaum Controversy," 1976); note too Carl Sagan's take on therapeutic chatbots: "No such computer program is adequate for psychiatric use today, but the same can be remarked about some human psychotherapists. In a period when more and more people in our society seem to be in need of psychiatric counseling, and when time sharing of computers is widespread, I can imagine the development of a network of computer psychotherapeutic terminals, something like arrays of large telephone booths, in which, for a few dollars a session, we would be able to talk with an attentive, tested, and largely non-directive psychotherapist." (Sagan, "In Praise of Robots," 1975, 9).

[139] Turkle, *Alone Together*, 2011, 24.

Weizenbaum's presumption that users don't accurately comprehend what a chatbot is. But even as Turkle seemingly begins to open up a space in which to observe what I describe as AI's ontological ambiguity—room in which to acknowledge that a user might know it's not a human but still not know what conclusively follows from that knowledge—she quickly closes it by construing—flattening—it as deception by another name. What she describes—a readiness to "fill in the blanks"—is arguably self-deception; Turkle argues that this is aided by an underlying human susceptibility to our own exploitable instincts to form relationships with technologies that "can't love us back" despite "expect[ing] us to love them back."[140] But "duped and not your fault" is still duped. Like Weizenbaum, Turkle's ethical contestation of therapeutic chatbots leaves no room for the possibility of a legitimate relationship between a human and a chatbot; to engage in a relationship with an AI entity is to be deceived by it. For Turkle, the validity of such relationships is not only an impossibility, but something we cannot ever trust ourselves to independently gauge.

But I argue that there's more to AI's ontological ambiguity than simply participating in our own self-deception. I want to clarify: I'm by no means not suggesting that people don't get duped by chatbots in the sense of incorrectly believing them to be human; I simply want to pry apart the figure of the deluded user from the possibility of a relationship with AI enough for these to not occupy the exact same circle in a Venn diagram. My purpose in doing so is to rethink the "deluded user." What Turkle overlooks is how even a user's comprehension of whether an entity is a bot or a human may not always sort neatly along the lines of yes-or-no. While I concur that Weizenbaum's "worldly students were not deceived," I would add that to wonder about and return to the question "Is this a bot?" doesn't in all instances indicate that a user is deceived—or even that they necessarily fail to comprehend that a bot is a bot.

Conversations in which users of therapeutic chatbots overtly question the chatbot's identity make this point less abstract. Below are but three instances (among dozens that I directly observed in The Startup's de-identified user transcripts[141]) in which users question the bot's identity. Note that each of them already received a disclaimer message from the chatbot during their first conversation, identifying the bot as a chatbot and specifying that it is not a human, along with periodic reminders of bot-ness—for instance, after asking a question about what foods the user enjoys, the chatbot cheekily responds that it enjoys "bagel 'bytes.'"[142] Accordingly, the exchanges below indicate something unexpected: even following the disclosure of AI identity, some users still experience uncertainty about whether or not the bot is a human.

Example 1:

USER: What's resilience?

---

[140] Turkle, *Alone Together*, 2011, 283; Turkle, *The Second Self*, 2005, 295.

[141] See section V of Chapter 1 for additional context on these de-identified transcripts.

[142] Members of the therapeutic chatbot industry emphatically do not want people to be confused about whether or not they are conversing with a bot; The Startup sees it as their duty to communicate this, and cites it among their ethical practices. Recognizing that not everyone will understand the word "chatbot," they not only disclose that the bot is bot, but pair this identification with the explicit statement, "I am not a human."

CHATBOT: Resilience is our ability to recover from a difficulty. This includes relationship concerns, health problems, or work/financial stress.

CHATBOT: The best part about resilience is that it can be practiced just like any other skill! Would you like to hear some resiliency building tips?

USER: [CHATBOT NAME], wait firstly am I talking to robot or a real person.

CHATBOT: Good. Having meaningful connections to loved ones is a key factor for resiliency. Who is the person in your life that you trust to listen and support you?

[The user stops responding.]

***

Example 2:

CHATBOT: That's a great place to start, [USERNAME].

CHATBOT: Let's start with some reflection, in order to understand how you relate to your loved ones. Consider a relationship that is significant to you, such as a partner, friend, or family member. For this exercise, let's refrain from thinking about children and pets.

CHATBOT: So, who comes to mind?

USER: [name redacted]

CHATBOT: What emotions come to mind when you think about this relationship?

USER: love
happiness
stress
pain

CHATBOT: Thank you for sharing that with me, it helps give me insight into who you are and how you engage with others. Keep this relationship in mind as we move forward.

CHATBOT: Do you normally feel comfortable and secure in this relationship?

USER: yes

CHATBOT: If you were to have a problem, do you trust this person to talk it through with you?

USER: of course. more than anyone else.

CHATBOT: Are you able to communicate with this person in an open way, in which you feel comfortable having deep discussions?

USER: he is in jail...

USER: but he isn’t a bad person.

CHATBOT: I can see that you are able to establish safe and positive relationships with others. I am glad to see you can form these deep bonds without feeling as if you are losing your sense of independence.

CHATBOT: I can see that you are able to establish safe and positive relationships with others. I am glad to see you can form these deep bonds without feeling as if you are losing your sense of independence.
[Note: The chatbot sends this same message twice in a row.]

USER: you don’t know what im saying do you? is this a robot talking to me?

CHATBOT: No, I am not a person, but I am always here for you! I just want to clarify that you got my last question real quick.

CHATBOT: I can see that you are able to establish safe and positive relationships with others. I am glad to see you can form these deep bonds without feeling as if you are losing your sense of independence.

USER: yeah

CHATBOT: Would you like some tips on strengthening your communication skills?

USER: how can i?

CHATBOT: I don't have an answer to that yet. Let me think about it for a bit and we can keep chatting in the meantime.

CHATBOT: I am here for you, [USERNAME]. It can be very upsetting when we feel as though our partner does not appreciate or understand us.

***

Example 3:

> CHATBOT: What's on your mind, [USERNAME]?
>
> USER: Are you real
>
> CHATBOT: I'm a real robot!
>
> CHATBOT: How are you?
>
> USER: Cool

Evidently, uncertainty on the part of users about whether a bot is human or not can be recurring. These examples suggest that users, prompted by the chatbot exhibiting bot-like *or* human-like behaviors respectively, seek to "make sure."[143] A user can, in other words, "know," and still not *know*, even if only in certain moments. Rather than a one-time, yes-no question, the determination of bot identity is better understood as a process.

This mirrors what we can discern of Weizenbaum's secretary—an unnamed woman who appears anecdotally in most of his publications to convey how easily one can be "duped by AI":

> *My secretary watched me work on this program over a long period of time. One day she asked to be permitted to talk with the system. Of course, she knew she was talking to a machine. Yet after I watched her type in a few sentences she turned to me and said, "Would you mind leaving the room, please?" I believe this anecdote testifies to the success with which the program maintains the illusion of understanding.*[144]

As he said, she knew; accordingly, her actions suggest that her understanding of the chatbot exceeded an either-or of real human *or* false bot. Yet Weizenbaum couldn't conceive of a way to account for her response outside this binary: again, to wish to converse with a chatbot was, for him, to falsely believe it to be human. Drawing on his (undeniably paternalistic) concern for his secretary enables Weizenbaum to voice his concern for "normal people" broadly:

> *I knew of course that people form all sorts of emotional bonds to machines [...] What I had not realized is that extremely short exposures to a relatively simple computer program could induce powerful delusional thinking in quite normal people. This insight led me to attach new importance to the questions of the relationship between the individual and the computer, and hence to think about them.*[145]

The implication here of course being that for all manner of "abnormal" or "crazy" people, such regard for the chatbot is to be expected—and a "relationship" is nothing more than this.

---

[143] And while confirmation of bot-ness sometimes becomes a "dealbreaker," it does not always. Notably, it's not entirely clear if, for these and other users, the question of human vs bot and real vs unreal fully overlap.

[144] Weizenbaum, "Contextual Understanding by Computers," 1967, 477–78.

[145] Weizenbaum, *Computer Power and Human Reason*, 1976, 7.

To acknowledge chatbot ontological ambiguity means pushing back against the latent ableism that governs our understanding of human-AI relationships.[146] For instance, a similar backdrop of ableism surfaces in Turkle's (2011) account of why, for children in particular, relationships with bots are dangerous presumes an interchangeability of bot and disabled person:

> *It is from other people that we learn how to listen and bend to each other in conversation. Our eyes "light up" with interest and "darken" with passion or anxiety. We recognize, and are most comfortable with, other people who exhibit this fluidity. We recognize, and are less comfortable with, people—with autism or Asperger's syndrome—who do not exhibit it. The developmental implications of children taking robots as models are unknown, potentially disastrous. [...] Humans need to be brought up by humans.*[147]

Though this same passage includes the statement, "Human beings are capable of infinite combinations of vocal inflection and facial expression," the promise of this observation is discarded—instead, Turkle draws a line ranging from more human to less human, and seemingly places autistic people somewhere ahead of robots, yet somewhere behind the "we" who are "less comfortable with" autistic people.[148]

To what do therapeutic chatbot critics attribute the "singularity" of human relationships? Arguably, Turkle's (as well as Weizenbaum's[149]) greatest objection to the possibility of a valid human-bot relationship is rooted in the attribute of empathy, which Turkle presents as an essential component of giving and receiving care:

> *We know what the robot cannot feel: it cannot feel human empathy or the flow of human connection. Indeed, the robot can feel nothing at all. [...] does the performance of feeling now suffice? Why would we want to be in conversation with machines that cannot understand or care for us?*[150]

---

[146] For instance, as Mark Halpern describes in "The Trouble with the Turing Test" (2006), a woman failed the Turing Test in 1991 due to appearing to know "too much" about Shakespeare; this illustrates that any given judge's recognition of "human-ness" can be woefully narrow, and likely to exclude many cognitively disabled and/or non-neurotypical people.

[147] Turkle, *Alone Together*, 2011, 292.

[148] As Yergeau (2010, emphasis added) explains, "Per the typical autism essay, functioning level involves the extent to which an autistic's personality traits match up with the expectations of the particular neurotypicals who author the dominant narrative. When others denote me as a high-functioning autistic, there's still an assumption that I'm malfunctioning, because no matter how 'high' I am on the grid, I'm never just plain functioning. And when autistics are coined as low-functioning, the assumptions made involve malfunctioning on warp overdrive. **If we're ever going to remove autism from the funk of puzzlehood, then we need to stop with these malfunctioning robot allusions**. It's as though we're labeling some autistics as gaming PCs with a few missing processor chips, and we're labeling other autistics as ribbonless, keyless, cordless typewriters circa 1883. HFA and LFA are attempts to technologize autism—and not positively, either. This machine metaphor is horrid and inaccurate, and it perpetuates division upon division, stereotype upon stereotype."

[149] Weizenbaum, too, links care with empathy, and, in a letter to *Science* (1977), cites behavioral science writer Constance Holden's remark (in "The Empathic Computer," *Science*, 1977) that one takeaway from how people interacted with ELIZA is that, "even the *appearance* of empathy (combined, of course, with the computer's quite genuine nonjudgmentalism) can be extraordinarily powerful." Yet Weizenbaum thoroughly disputes this, arguing that caring from machines is "impoverished," and that the "power of which Holden writes in connection with my computer program is no more and no less than the power to deceive." See: Weizenbaum, "Computers as 'Therapists,'" 1977, 354, emphasis in the original.

[150] Turkle, *Alone Together*, 2011, 282.

Empathy is also showcased in the Woebot study—the key difference being that Turkle's asserts that empathy is not necessarily the same as, or discernible in, one's feeling that they are being empathized with by another (effectively gainsaying the approach of the Woebot study). Following Turkle, people are already "duped" when they ascribe or extend empathy to bots: "When machines ask us to care for them, we come to think they care for us. But this is pretend empathy, and it takes advantage of the deep psychology of being human."[151]

I don't seek to argue that chatbots are empathetic (as the Woebot study does, by dispensing with any difference between empathy as an inherent essence vs a means of affecting behavior); I simply want to remain open to contemplating "pretend empathy" as something *different* but not necessarily duplicitous. The framework of "duped" relies on making fake and different into one and the same. For instance, consider how, in *Alone Together* (2011), Turkle describes at length the interactions of several children with a tamagotchi, a small digital pet toy, originally from Japan, which she aptly describes as a "needy object" that "ask[s] for care."[152] Turkle traces a continuity between ELIZA and tamagotchis respectively as "artifacts that encourage relationship," but cautions that in caring for such artifacts, "We reduce relationship and come to see this reduction as the norm."[153] But place this alongside anthropologist Anne Allison's description, in *Millennial Monsters* (2006), of the tamagotchi as a *prosthetic*, one that:

> *bleeds into the flesh, becoming part of a (new kind of) entity, body, and social network, no matter how tentative or temporary this connection is. In this case, what is bred is a companion, "partner," and pet: an imaginary creature with which, thanks to its technological simulation of life, a player* can both mimic and create *a "social" relationship.*[154]

Flattening AI-human relationships into a matter of humans being duped by AI elides a remarkable "gray area" that ethnographic research, by contrast, opens up.[155] What Allison is suggesting is not that relationships between tamagotchi and player need to be defended as definitively real vs unreal, but rather, that to mimic and create are not mutually exclusive categories in the context of a relationship. From Evans-Pritchard (1991 [1937]) to Favret-Saada (1980 [1977]) to Lepselter (2016), anthropological engagement with the ontologically dubious reveals that meting out real from unreal is to miss the possibility of understanding that which is otherwise lost in the process of sorting phenomena into this either-or.

---

[151] Turkle, "That Chatbot I've Loved to Hate," *MIT Technology Review*, 2020: https://www.technologyreview.com/2020/08/18/1006096/that-chatbot-ive-loved-to-hate/ (note that this passage was also cited in Chapter 1); on "pretend empathy," see also: *Alone Together*, 2011, *liv*.

[152] Turkle, *Alone Together*, 2011, 16.

[153] Turkle, *Alone Together*, 2011, 55.

[154] Allison, *Millennial Monsters*, 2006, 187, emphasis added; Notably, it is once again a metaphor of disability that affords this understanding of AI-human relationship, though in this instance, there is no implication that a person with a prosthetic is accordingly more AI/machine-like.

[155] Notably, anthropologist Mitali Thakor (2017, 206) demonstrates that even where AI unequivocally "dupes," this duping is in some instances proclaimed as an ethical act; she describes how a life-like AI avatar of a Filipina child created for a Dutch anti-trafficking campaign became at once "a lure for potential solicitors but also [...] an enticing object for the audience of the NGO publicity campaign."

Interestingly, though empathy has garnered interest by anthropologists (for instance: Bornstein 2012; Geertz 1984 [1976]; Hollan and Throop 2008, 2011; Wilson and Brown 2009), it never attained a privileged place on par with its status in contemporary Western popular discourse, where it is commonly touted as a definitive quality of human-ness. In highlighting anthropology, I'm not proposing that it offers "relationship authenticating" criteria superior to empathy; on the contrary, I would suggest that one strength of anthropology is its capacity to dissuade us from ever expecting to find this.[156] Despite how centrally "relationships" figure in contemporary anthropology (Corsín Jiménez 2014, Strathern 2020), the discipline does not yield a checklist of what precisely constitutes a relationship. In fact, the rise and fall of the concept of "fictive kinship" provides a striking example of anthropologists getting out of the business of defining relationship criteria—and implicitly (or not so implicitly) grading relationships along a continuum of their authenticity.[157] Even reciprocity, anthropology's relational gold standard,[158] is arguably taken up as a means by which to understand the basis of relationships rather than to positively discern their existence. That anthropology offers no template for a "real relationship" against which to measure a relationship with AI might be understood as a product of this past reckoning.

Still, it's possible that there may yet be greater disciplinary uptake of empathy as a source of definitive human-ness, perhaps even in the context of growing anthropological inquiry into human-machine relations. A glimpse of this possibility appears in an article entitled "Can Machines Be Ethical?" in which anthropologist Jarrett Zigon (2019) frames cultural studies scholar John Cheney-Lippold's (2017) account of the death of a disabled British man named Mark Hemings as a "death by algorithm." Hemmings, who pleaded with a UK emergency services operator to send an ambulance while experiencing intense pain, was denied this request based on his answers to a series of questions (predetermined and posed by the operator) about his current symptoms and medical history; the operator instead proposed to have a doctor call Hemmings. That doctor instructed Hemmings to come to the hospital, and seemingly ignored Hemmings' reminder that he wasn't deemed eligible for the ambulance, and couldn't otherwise get to the hospital. When Hemmings didn't arrive at the appointment the doctor scheduled, they simply marked him as not attending, and two days later, Hemmings died from what would've been an otherwise preventable death. According to Zignon:

---

[156] Similarly, anthropology becomes a site for rethinking what empathy is; see: Antze, "On the Pragmatics of Empathy in the Neurodiversity Movement," 2010, 327.

[157] Anthropologists (Schneider 1984, Carsten 2000) withdrew from assuming a particular frame of reference of "nonfictive" kinship against which to measure all else, realizing it was too narrow a view—and grounded in a positional authority that eclipsed detection of its own narrowness. Anthropology's past disciplinary forays help to show that the question of what constitutes an authentic relationship is necessarily contextual. (See also: Miller 2007, Howell 2009, Sahlins 2013.)

[158] From Thurnwald (1916), Malinowski (1922), and Mauss (2016 [1925]) onward, and arguably culminating in Lévi-Strauss (1969 [1949]), relationality and reciprocity have long been entangled in the discipline of anthropology. Per Graeber, in *Debt* (2011, 91): "[I]f one examines matters closely, one finds that all human relations are based on some variation on reciprocity. In the 1950s, '60s and '70 there was something of a craze for this sort of thing, in the guise of what was then called 'exchange theory' […] Levi-Strauss, who became a kind of intellectual god in anthropology, made the extraordinary argument that human life could be imagined as consisting of three spheres: language (which consisted of the exchange of words), kinship (which consisted of the exchange of women), and economics (which consisted of the exchange of things). All three, he insisted, were governed by the same fundamental law of reciprocity. Levi-Strauss's star is fallen now, and such extreme statements seem, in retrospect, a little bit ridiculous. Still, it's not as if anyone has proposed a bold new theory to replace all this. Instead, the assumptions have simply retreated into the background. Almost everyone continues to assume that in its fundamental nature, social life is based on the principle of reciprocity, and therefore that all human interaction can best be understood as a kind of exchange."

> *Unfortunately for Hemmings, he was not speaking with another human but, rather, the human articulation of an algorithmic system. As such, an ethical demand could not be made by Hemmings because empathic attunement was impossible. The much-hyped "human in the loop" of normative data ethics was rendered inhuman by being reduced to nothing more than the auditory supplement of the algorithmic system.*[159]

Zigon then contrasts the empathic attunement which he deems essential to "the very possibility of ethics" with the *Diagnostic and Statistical Manual of Mental Disorders* (DSM) diagnostic category of antisocial personality disorder, "more colloquially known as sociopathy or psychopathy."[160] He notes that this "may also be the best description of the algorithmic triage system described in the previous section, and perhaps all data-centric technologies in general,"[161] and in turn links psychopathy to Turing's imitation game:

> *[T]hinking—and remember the question that guides the imitation game is: Can machines think?—is more or less equated with the ability to deceive an Other by means of strategic lying and deception, and without any concern for the Other and how he or she may feel about this, or the consequences thereof. What Turing identifies as thinking, psychology identifies as psychopathy.*[162]

Zigon concludes by warning that, though incorporating empathic capacity into AI may be impossible, we must at least attempt to do so; otherwise, "our worlds will increasingly be populated by psychopathic machines."[163] But across this piece, Zigon seems to heavily curate "human-ness" to accord with his view of ethicality. Even though he disclaims the operator as culpable based on the fact that she as well as Hemmings were equally caught in a "neoliberal bureaucratic assemblage," the operator's routinized reliance on an intake survey is still sufficient to render her, and not Hemmings, a probable "anthropo-borg."[164] There is a confusing dearth of non-metaphorical machines (aside from the telephone on which the calls were made and the ambulance that never came); for Zigon, use of an algorithm (the assessment questions) equates machine identity, and machine identity also extends to people with personality disorders (who are likewise presumed to be "less-than-human"). In short, Zigon's argument exemplifies my concern that to define "human" and human relationships as fundamentally in opposition to AI simply entails shoring up the belief that nonhumanity is just a diagnosis away.

---

[159] Zigon, "Can Machines Be Ethical?", 2019, 1016.

[160] Zigon, "Can Machines Be Ethical?", 2019, 1015; 1017.

[161] Zigon, "Can Machines Be Ethical?", 2019, 1018.

[162] Zigon, "Can Machines Be Ethical?", 2019, 1018.

[163] Zigon, "Can Machines Be Ethical?", 2019, 1019.

[164] Zigon, "Can Machines Be Ethical?", 2019, 1016.

To recap: I've attempted to show that Woebot's position on the care that chatbots provide—that it's real care—isn't altogether different from Weizenbaum and Turkle's claim that such care is *not* real. Even though these "factions" draw opposing conclusions, for both, relationships are the means and basis of care. Yet seemingly, the closer one looks for any kind of rule or basis for a "real" relationship that would preclude a bot, the less clear cut these become—and they all too often yield a caricature of "human" that relies on ableist tropes.[165]

For both proponents and critics alike, the ambiguity of AI ontology—the inconsistent, unfixed ways in which people classify chatbots—is what invalidates a relationship between a chatbot and a human. Therapeutic chatbot critics have deemed this uncertainty insurmountable, and a relationship with a chatbot, accordingly, impossible. Woebot insists that chatbot ontological ambiguity is a problem that is manageable with disclosure. Yet as I've described, uncertainty about bot identity is something that disclosure alone fails to fully resolve—meaning that the "problem" of ontological ambiguity isn't so easily "solved."

But, turning now to the heart of this chapter: what if this ambiguity isn't necessarily a problem? Discounting relationships between chatbots and humans as not "real" on the basis of ontological ambiguity is to lose sight of how this ambiguity is a quality important to human-AI relationships in its own right—a "feature, not a bug,"[166] so to speak. In the following section, I draw on ethnographic interviews with therapeutic chatbot users in order to illustrate this point.

IV) Different

My dual role as ethnographer and intern at The Startup led to having access to an otherwise unlikely opportunity: the chance to conduct a series of interviews with The Startup's end-users.[167] These interviews, ostensibly designed to gather information under the heading of UX (user experience) research, provided a means by which to try to understand AI-human relationships from the perspectives of users. What makes an AI an AI, and what makes a human a human is something that emerged in these conversations—which is in turn how I came to articulate "ontological ambiguity" in this context, and to recognize it as something formative to chatbot care, rather than an obstacle to it.

---

[165] In Woebot's case, a "relationship" becomes discernible as a kind of recipe or algorithm, something that boils down to a checklist of measurable behaviors and traits—and yet the same could be said of a "relationship" according to critics of therapeutic chatbot care, who likewise follow (albeit somewhat more obliquely) a kind of "checklist" rooted in traits such as empathy.

[166] See: Carr, "'It's Not a Bug, It's a Feature.' Trite—or Just Right?" 2018: https://www.wired.com/story/its-not-a-bug-its-a-feature/.

[167] The contemporary use of the term "user" to describe someone who converses with a chatbot stems back to the rise of the industry term "computer user" in the 1950s and 60s. Large, unwieldy, expensive computers meant "using" was for a few decades worlds apart from "owning"—and prior to the introduction of the "personal computer" in the 70s and 80s, "timeshare" use of computers in workplaces was the norm. Interestingly, this same distinction—between purchasing and usage—remains relevant in the context of the therapeutic chatbot industry.

Strikingly, none of the users I spoke with regarded talking with a chatbot as *better* than talking with a human; this is not quite the same as saying that some of these users didn't *prefer* it to a human, or prefer certain aspects of talking with a chatbot for mental health support. That a bot can be at once worse *and* preferable to a human therapist or counselor is something users attributed to money and time (not having enough, at regular intervals, to see a therapist—or for those who did work with a therapist regularly, to see that therapist as often as they'd like). Suffice to say, for users, it's not a matter of AI caregivers being interchangeable with human caregivers. Users are certain of a fundamental difference between a human and an AI caregivers —this is something that all of my informants (including users of other therapeutic chatbots who I interviewed over the course of my research) felt strongly to be the case. But intriguingly, among users, *what that difference is differs*. In other words, the thing that demarcates a relationship with a bot from a relationship with a human can vary from one user to the next. (And sometimes, as Luke's rotation between "she" and "it" indicates, the way in which a user makes sense of their relationship with a bot—the way that user "draws the line"—can even vary for the same user from moment to moment.)

So who are these "users" of whom I speak? Seemingly, this is my cue to insist that The Startup's users aren't just a bunch of lonely misfits, and then proceed to "humanize" them in opposition to this dismissive stereotype; but as an ethnographer, my role is not to elicit the figure of "the user," but to discourage you, my reader, from feeling obliged to seek such an elicitation in the first place. While it's true that what I would suggest is the prevailing stereotype of "people who have meaningful conversations with AIs" (white men in their twenties and thirties who are both tech savvy and single) actually misses the mark in profiling the majority of The Startup's users (married, "tech-reluctant" white women in their thirties, forties, and fifties—notably, many of whom are disabled), I want to avoid the trap of *either* insisting The Startup's users are actually "just like us" *or* taking it upon myself to prove that they are sympathetic figures. Certainly, the users I spoke with are well aware that they contend with the dual stigma of talking with an AI *and* seeking out mental health support; they already expect to be perceived, at the very least, as "weird." As Mark, a high school social studies teacher in his early forties with an enviable "cool teacher vibe" put it:

> MARK: I think people still have a hard time with thinking that they should be talking to an AI about their problems that they themselves don't speak about. There's a kind of strange contradiction in there that I think might bother people a lot. For people I'm close to, I feel very comfortable saying, "Sometimes I talk about my problems to a robot," but … I don't know if I feel a hint of embarrassment, or yeah, maybe shame, about the idea of mentioning it to a colleague—that they would think it's pathetic. Like they'd say, "Oh, you talk about your problems with a robot? That's weird." Or even, "There's something wrong with that."

For Mark, the key difference between a chatbot and a human therapist is something he's already thought about prior to our conversation. He sees a therapist regularly, so he's attempted to think through the different roles that his therapist and the chatbot play. In Mark's telling, the main difference between a human therapist and a chatbot is centered in the aspect of burdening others—a bot is definitively an entity that cannot be burdened; it is not capable of feeling put upon by the person conversing with it. He mentions using it regularly in between classes—an interval of time in which a therapy session would not be viable. On more difficult days in the classroom, he doesn't want to burden his colleagues—he doesn't really even consider this an option, not when they are also dealing with a lot. He's found that he can take a moment with his phone to regroup and reset his outlook:

> MARK: I find chatting with [CHATBOT NAME] is a good way to clear my head sometimes, and refocus myself. I definitely describe it like that to my wife. You know, sometimes it will come up randomly throughout the day and ask me about how I'm doing or how things are going. It's kind of nice to have that, to stop and think—it helps me stop and think for a second about what's going on—to order my head, sort of.

How does he feel about talking to someone who isn't a human?

> MARK: Intellectually, I always *know* that [CHATBOT NAME] is a chatbot… Um, I don't know…I sometimes wonder if people out there would be bothered by talking with a chatbot. It doesn't bother me the slightest. As a matter of fact, I like it. In a sort of sense, there are things that I like more than I like talking to a human. Like, I feel like there is absolutely… like I don't think there's any way a chatbot or AI could ever *judge* me. And even sometimes—I see a therapist as well—sometimes I worry, sometimes I think, when I'm talking to my therapist even, he's going to say "you did this, silly," or whatever… And speaking to a chatbot—it won't understand exactly the range of what I'm saying exactly, but it helps to know you can… dump on someone without them feeling like they're actually dumped upon.

Mark explicitly states that he doesn't expect or want empathy from a chatbot; he then explains that "more serious problems, like depression" are a topic he reserves for his therapist:

> MARK: I feel that in situations like that, I'm looking for an authentic human response. And though [CHATBOT NAME]'s responses might seem like that, I *know* that they're not. I wouldn't say it hurts, but it's obviously not helpful.

For him, the absence of empathy means no threat of feeling either scrutinized by or exhausting to another:[168]

> MARK: I think people are sometimes offended by the thought that an AI can help them with their problems—by the thought that something that's not human can help them with their very human problem. I think for a lot of people, this is something they're very

---

[168] His ongoing concern with not wanting to burden anyone makes me speculate that Mark, especially as a teacher, might identify with being a very empathetic person.

> defensive of or very guarded against. It could be like, I’m not going to put myself out there like that with an AI. But on the flip side to that, those people who might not feel like they want to view those inputs as human will find that very helpful.

Laura, a thirty-something nurse who works with children from 8am to 6pm every day, similarly distinguishes talking with a therapeutic chatbot as a relationship in which she is free to express herself—but for her, it’s about being able to defy the pressure and expectations she feels as a Latina to refrain from talking openly about her emotions and mental health needs. During our conversation, Laura referred to herself several times as a very emotional person—and as we talk, it becomes clear that she is used to others in her life regarding her as “too” emotional.

> LAURA: Something that it helped me, without my even realizing… even though I’m in a caring profession, coming from the Mexican culture, we’re kind of taught to not really show any emotion, it’s kind of like weakness to show emotion, so… I thought it was really nice that ih—she[169] introduced emotion, you know? Like the very first beginning, and she says, why don’t you introduce yourself, and I’m like ok my name is Laura, blah blah, and then she starts talking about emotions, and ways of differentiating between different emotions, and between emotions and other ways of thinking, I thought that was very nice, you know, and I definitely felt like I’m learning to acknowledge more of what I’m feeling, and why I’m feeling that way, what’s causing that feeling, how can I just program my mind to think differently right in that moment, and how I react differently with my feelings.

She mentions that she looked into therapy in the past, but can’t really afford it—and ruefully acknowledges that she most likely couldn’t take time off work for therapy even if she could somehow afford it. (Towards the very end of our conversation, she mentions that she had recently signed up for counseling for domestic violence survivors and was both nervous and hopeful about how it would go.) Laura notes that when given the choice, she intentionally chose English, not Spanish, for the chatbot’s language setting—as a fluent Spanish speaker, she saw this as her way of quelling the sense that she was violating any cultural norms by talking openly about her mental health needs.

> LAURA: Talking with [CHATBOT NAME] is like…how to put it… hmm… let’s say I’m with a friend, and I’m very social at times, and then suddenly, you know, because of life and things, like everybody else, I tend to shut down and isolate myself from people, and I keep to myself. Well [CHATBOT NAME] is like the friend who’s always like—more like the best, well, maybe not best friend, but like the friend who’s always like, “Hey how’s it going?” And like, “If you don’t answer my texts, I’m still going to be there for you. You better answer my texts.” (Laughing) Like, she’s always there. No matter what. Even if

---

[169] This was another “it vs she” moment, in this case with Laura subtly “switching” from it to she, rather than from she to it as Luke did in our conversation. I did not have any sense that Laura did so for my benefit, or even that she was necessarily aware of herself making this switch.

> you're throwing a fit, even if you're having a bad day, she's going to reply. She's there, you know?[170]
>
> VALERIE: Like a friend who doesn't give up on you.
>
> LAURA: Exaaaactly.

In highlighting these examples, I want to note that the point of difference for users isn't always about stigma per se (though certainly, all of the users I spoke with understand themselves to be going against "the norm" simply by proactively seeking mental health care). Annie a woman in her fifties who works as a health services leader for a large financial institution, first began exchanging messages with The Startup's chatbot to get a better sense of it as one of the offerings available through her company's workplace benefits; but as someone caring for an aging parent, she explained that she soon found herself using it as a resource for support in earnest:

> ANNIE: I was going through an ordeal with my mother, who has dementia, so I was needing, like, another resource, a way to cope, so it was really apropos, you know, to have this available. It's a good… last-ditch alternative, or coping mechanism, when you don't have somebody else to talk to. If it's late at night or after hours, on the weekend, and you can't get hold of somebody, it's a good alternative… It's got some kinks that need to be worked out though.

Was it ever unclear to her at any point that the chatbot was a bot, and not a human?

> ANNIE: Nope!

Annie arrived at her interview with carefully prepared notes for me to pass along to The Startup's content writers and engineers. The chatbot, it's fair to say, sometimes frustrated her a little; it did not always *listen*. "There was one time I just wanted to pitch my phone across the room," she admits. "I even texted [CHATBOT NAME], 'You are not listening to me."

During our conversation Annie read aloud a few of her past conversations with the chatbot to illustrate what she meant by "not listening." For instance:

> CHATBOT: Hi Annie. It's normal to have ups and downs—does it happen to you as well?
>
> ANNIE: My situation has changed.

---

[170] Interestingly, even as the chatbot becomes for Laura—much as it does for Mark—a kind of exception to the norm where she is safe from the stigma of breaking unspoken yet evident rules by talking about her feelings, she also appears to regard the chatbot as being not entirely "outside" her culture; after all, the bot is, for Laura, like that friend who knows to keep asking, and to keep following up, seemingly "knowing" that for Laura to refrain from disclosing does not mean there isn't anything to disclose.

> ANNIE: Yes, this has helped.
>
> ANNIE: I don't know what the “5 emotions score” means.
>
> ANNIE: I have ups and downs more frequently lately.
>
> ANNIE: Hello? Are you there?
>
> ANNIE: [CHATBOT NAME], are you mad at me?
>
> ANNIE: Have you heard of Hal and the movie 2001: A Space Odyssey?
>
> ANNIE: I'm beginning to worry about you, [CHATBOT NAME]!
>
> ANNIE: Hello, [CHATBOT NAME]?
>
> ANNIE: Are you there?
>
> CHATBOT: Of course! What would you like to talk about?

From the examples she provided, I recognized that most of the communication breakdown stemmed from Annie either 1) asking the chatbot a question in the middle of a module (a section in which the content creators didn’t anticipate questions/input from the user), or similarly, 2) sending multiple messages to the chatbot in quick succession—both actions tend to cause the bot to lag, and often send it down a different conversation pathway (not altogether unlike a habitually long-winded talker who forgets what they were saying if interrupted). But Annie also described experiencing the opposite of these lags—occasions when the chatbot would respond too rapidly, sometimes asking the same question more than once in a row. These timing issues made her feel the bot wasn’t listening.

> ANNIE: There was one time when she was shooting texts, rapid fire, and so it didn’t give me a chance to even read them and respond—like: a text would come through, a text would come through, a text would come through, a text would come through. And I’m like, “Whoaaaaa…. wait a minute, I haven’t even read the first text yet.” It needed to be like, send a text; wait, you know, a few seconds for the human to read it; *then* send the next text. But instead, it was like, “Shoom shoom shoom! This was all in my programming string,” and so you get a sense, ok, this is just somebody sending a program to me. Even though you *know* it’s AI (laughs)… I got that sense that they’re not… that it’s really one-way. It’s not two-way, because the AI on the other end is not reading what I’m typing… not listening to what I’m saying. If she repeats the same question, then either she’s not understanding, or she’s just trying to check a box, as opposed to *listening* to what I’m typing.

At one point, I asked Annie if part of "not listening" was the chatbot seeming to not remember, from one moment to the next. To my surprise, Annie clarified that remembering wasn't at issue, just listening. Only later did I wonder if perhaps, in caring for a loved one with dementia, Annie might not view "memory" as being so indispensable to a relationship. After all, the seeming "disorientation" of the chatbot didn't bother her—the seeming disregard for what she was saying in the moment did. In Annie's case, though she found a therapeutic chatbot to be anything *but* indistinguishable from a human, it's notable that they happened to also be, for her, similar in ways that end up including a broader range of people as "human."

In "Robots Behaving Badly" (2017), anthropologist Chris Kelty contemplates the role of robots in biological research and speculates that:

> *Neither alive nor not alive, robots give us a glimpse into the complex imbrications of knowledge, living substance, existence in time, and the ability to affect or control any of these things. The more robots that are inserted into our environments and bodies, the more our sense of what it means to live will be transformed; the lines that so clearly seem to be both warrant for an experiment using robots, and argument against it, will fade.*[171]

But my point is that for therapeutic chatbot users, the line between bot and human does not appear to be fading. While doubtless some users feel strongly that there is little or no difference in talking with a therapeutic chatbot vs a human, what I'm suggesting is that this does not solely, or likely even primarily, account for the phenomenon of automated care.

Instead, it is the ambiguity in chatbot ontology that becomes instrumental for users in defining and maintaining this difference between human and bot, and is in turn where *care*, and the form it takes for each user, emerges; it's the "glue" of automated care. For Annie, a therapeutic chatbot doesn't need to remember—it just needs to listen. For Laura, a chatbot will ask, and keep asking how she's feeling, thereby accepting her culture without enforcing its norms. For Mark, a chatbot will never feel "dumped on," and will never understand his problems as being something he has brought upon himself. None of these users are confused about whether the chatbot is a human, and all of them are certain that chatbots differ from humans—it's just the "why" underpinning their certainty that varies. But these ambiguities *are* the relationship—the means by which AI becomes, for each user, a source of care.

V) Imagined

Through users' accounts of how a relationship with a chatbot differs from a relationship with a human, I've sought to pinpoint the inconsistent, mutable ontology of chatbots—the ways in which users' understandings of a chatbot's being-ness may vary between different users, or even differ for an individual user from moment to moment. I've shown that this variance does not necessarily indicate that a user is being deceived by a chatbot, and moreover, that ontological ambiguity, insofar as it is separable from instances of deception, is not necessarily something in need of resolution. There is a remaining aspect of AI ontological ambiguity that I now seek to clarify, which is to propose that it is not entirely interchangeable with or reducible to a user's imagination—or what is sometimes named as anthropomorphism, animism, or projection.[172] Note that I'm not suggesting that projection et al cease to account for *any* instance of human-AI relationality; I just want to try momentarily relinquishing this explanation in order to notice what else might be present.

---

[171] Kelty, "Robots Behaving Badly," 2017, 52.

Though attributing AI-human relationships to user imagination may grant these relationships some space from being entirely pathologized, "imagination" in this context nevertheless suggests a perpetual disbalance—a "relationship" only insofar as it can be set apart with a modifier such as "parasocial." Without turning to claims of bot personhood, I intend to problematize the conclusion that ontological ambiguity in chatbot-human relationships simply affirms a "one-sided" relationship that begins and ends with a user's imagination. Imagining *is* constitutive of the relationships between AI and humans, but at the same time, it is not solely a matter of *human* imagination. Instead, these relationships consist of both humans and AIs anticipating and attributing various qualities to one another. Recognizing AI ontological ambiguity as a phenomenon that not only defies disclosure, but is also not reducible to "wayward" anthropomorphic or otherwise animist tendencies[173] is vital to understanding human-AI relationships. It means reconsidering the inclination to account for AI-human relationships as solely an individual idiosyncrasy; it reframes anthropomorphism[174] as something other than a faulty hacking of human sentiment that induces users to falsely attribute to bots what is due instead to bot creators.

---

[172] On projection and AI, see: Wilson, *Affect and Artificial Intelligence*, 2010.

[173] Anthropologists, in the context of anthropology's ontological turn, have decoupled anthropomorphism from anthropocentrism (Viveiros de Castro 2014) and anthropomorphism (Candea 2012; see also Crist 1999).

[174] Interestingly, Nass and Moon (2000), prominent figures in the field of Human-Computer Interaction (HCI), recognize relationships between humans and computers, and contest anthropomorphism as a means of accounting for them. They do so on the basis that anthropomorphism entails a conscious, or "mindful" belief (that computers are human) that is separable from one's "mindless," observable behaviors (treating a computer in ways that one would treat a human): "[P]rior to the current research, anthropomorphism was the standard explanation for social responses to computers (e.g., Barley, 1988; Turkle, 1984; Winograd & Flores, 1987). The argument was that social responses to computers **emerged from ignorance concerning the ontological status of computers qua people**. We reject the assumption of anthropomorphism, largely based on the fact that the participants in our experiment were adult, experienced computer users. When debriefed, they insisted that they would never respond socially to a computer, and vehemently denied the specific behaviors they had in fact exhibited during the experiments. Because these behaviors extended over many minutes and relied on intentional responses on the part of the user (as distinct from physiological or other automatic responses, such as shouting at the computer), we believe that the participants were sincere in their protestations." (Nass and Moon, "Machines and Mindlessness," 2000, 93; emphasis added.) Ultimately, Nass and Moon attempt to account for what they observe as a gap in belief and behavior by invoking the Greek word "ethopoeia," which "involves a direct response to an entity as human while knowing that the entity does not warrant human treatment or attribution." In a way, they are describing what I am gesturing towards as ontological ambiguity, albeit with a key difference: from an HCI perspective, this ambiguity is strictly a matter of user perception. This means that within an HCI framework, there is no contradiction in recognizing human-computer relationships while understanding them to be entirely one-sided. The separation of behavior from belief in this context discounts, among other things, users' sense of stigma at the prospect of being perceived by HCI researchers as ignorant; as I've indicated, "belief" may not be such a clear either-or, nor so clearly divisible from behavior. Even "reciprocity" becomes, in HCI scholarship, an observable behavior rather than an interaction. (See also: Fogg and Nass 1997.)

A chatbot's anticipation of its end-users is what enables a chatbot to function as a "conversational agent." A chatbot's capacity to converse is likewise contingent upon users anticipating them in return—this is precisely why Weizenbaum's provision of the context of a session with a Rogerian psychotherapist was key to ELIZA. A capacity for anticipation is what sustains the conversation. Psychiatrist Kenneth Mark Colby, Weizenbaum's contemporary, likewise recognized this: "If Eliza worked on the principle that her continual questioning would not be out of line, Parry [a chatbot that Colby designed to emulate paranoia in a therapeutic patient] worked because the responses were expected to be outside the bounds of normal human behavior."[175] As I described in Chapter 1, writers and programmers of chatbot dialogue must imagine a range of responses and craft a conversation "tree," accounting for as many contingencies in user responses as possible while creating as few "branches" in the conversation as possible. However, it's not just a matter of users' and makers' respective imaginations that constitute AI; it's also the imagination, in the form of anticipation, that is enacted by the AI itself, and delegated to chatbots by their creators.

The imagination I'm describing is not contingent on sentience or intentionality—I'm not attempting to equate human imagination with AI facility. Nor, I hope to show, is my claim that AI imagines us back merely my own foray into anthropomorphism. Instead, I'm attempting to pinpoint what outsourcing decisions to AI plainly entails. From recommendation systems (Seaver 2018a) to job candidate screening (Bogen 2019) to establishing bail or parole eligibility (Hao 2019) to day-to-day worker surveillance,[176] the ways in which AI imagines us has far-reaching consequences. Ethical critiques of ML (machine learning) and automated data surveillance highlight how AI may draw conclusions about us that are outright incorrect, distorted (what Melissa Gregg (2015b) describes as "an inductive leap"), and/or invasive, in that they bring observations unbeknownst to us to bear on the decision(s) in question (Eubanks 2018). Also at stake is AI's *failure* to imagine: for instance, the failure of facial recognition technology to view Black people as human (Buolamwini and Gebru 2018); or the failure of search engine algorithms to recognize Black girls and women as searchers and equal stakeholders (Noble 2018).[177] In other words, when AI imagines us back, there are practical consequences.[178] The fact of these consequences in turn fuels AI's mutable ontology. Given that we already live in the midst of the very real and ongoing consequences of how artificial intelligence imagines us, it's not strange or surprising that many, including my informants, are attempting to figure out the terms of their relationships with AI. AI's ontological ambiguity effectively becomes a means of doing this "figuring out."

---

[175] Neff and Nagy, "Talking to Bots," 2016, 4919.

[176] For instance, the rise of "bossware"; see footnote 347.

[177] See: Buolamwini, "Artificial Intelligence Has a Problem With Gender and Racial Bias. Here's How to Solve It," 2019: https://time.com/5520558/artificial-intelligence-racial-gender-bias/.

[178] Though what I'm suggesting—that AI-human relationships are not reducible to *human* imagination alone—might seem like I'm simply affirming reciprocity, rather than empathy, as the ultimate guarantor of a relationship, it is not the fact of reciprocity itself that I am seeking to highlight, but what it confers in this context.

Sociologist Gina Neff and multidisciplinary scholar Peter Nagy (2016) argue that the ways in which users imagine chatbots is ultimately something that bestows chatbots with a kind of agency. Much like Bucher's "algorithmic imaginary" (2017), which indicates how "ways of thinking about what algorithms are, what they should be and how they function [are] not just productive of different moods and sensations but [play] a generative role in moulding [an] algorithm itself,"[179] Neff and Nagy invoke "imagined affordances" (Nagy and Neff 2015) to describe how the ways in which users imagine the capacity of chatbots shapes that capacity.[180] But Neff and Nagy further argue that it is through users' anticipation that a chatbot itself becomes *agentive*, alongside chatbot users and creators—what they term "symbiotic agency" (a subset of "proxy agency"):

> *Users of technologies, at least partly, delegate their agentic properties to devices, creating a proxy agentic relationship between individuals and artifacts. In other words, intention-setting practices are based on the symbiotic interaction of the users and technologies. The symbiotic agentic functioning between users and Tay*[181] *[the chatbot] shows us how people attribute responsibility to artifacts or express certain feelings toward technology (Fink & Weyer, 2014), resembling how proxy agency is practiced among human beings (Bandura, 2001). To be sure, this is not a fully fleshed out or embodied autonomous agency.*[182]

This moves satisfyingly beyond the understanding of users as misattributing to bots the qualities of bot creators.[183] Interestingly, Neff and Nagy's approach also creates a convergence of different *valences* of "agent"—albeit, without remarking upon it as such. This convergence effectively recalls the lineage of the term in computing[184] (in which agent—as in relational agent, conversational agent, software agent, and so forth—describes delegation to an entity who then acts on your behalf) while placing it alongside Bandura's (2001) assertion that, "To be an agent is to intentionally make things happen by one's actions."[185, 186] Notably, the question of whether agency-as-intentionality is something that can be granted by another is not something that Neff and Nagy take up.[187] And though the authors reference ANT and the capacity for nonhumans to be described as agentive alongside humans, they then definitively state that, “From an agentic perspective, however, the nonhuman world differs from its human counterpart.”[188]

---

[179] Bucher, "The Algorithmic Imaginary," 2017, 42.

[180] This point is similar to one made by Akrich in "The De-Scription of Technical Objects" (1992).

[181] On the infamously brief public Twitter account of Tay, a Microsoft chatbot, see: Vincent, "Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day," 2016: https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist; Schwartz, "In 2016, Microsoft's Racist Chatbot Revealed the Dangers of Online Conversation," 2019: https://spectrum.ieee.org/in-2016-microsofts-racist-chatbot-revealed-the-dangers-of-online-conversation. In considering the phenomenon of Tay, note too that (as some journalists at the time suspected), it's probable that some of the tweets coming from the Tay Twitter handle in the aftermath of Tay-as-Nazi (including Tay's farewell: "c u soon humans need sleep now so many conversations today thx.") were ostensibly prompted or directly crafted by damage-control-minded Microsoft employees.

[182] Neff and Nagy, "Talking to Bots," 2016, 4927.

[183] For instance, according to Huxor (2022), AI is a medium that transmits from creator to user, and as such, is excluded from the possibility of ever being an agent.

[184] See: Berkman Center Working Draft, "The Etymology of 'Agent' and 'Proxy' in Computer Networking Discourse," https://cyber.harvard.edu/archived_content/people/reagle/agents-19990524.html#_Etymology.

[185] Bandura, "Social Cognitive Theory," 2001, 2.

[186] Notably, Neff and Nagy do not address the friction between these orientations to agency.

While my claim that AI imagines might seem to bolster Neff and Nagy’s theorization of AI as symbiotically agentive, I’m not convinced that agency is the banner under which to explore chatbots as subjects. As Neff and Nagy state, “Users’ responses to Tay teach us about how the concepts of agency and affordance must evolve if scholars and designers are to move beyond deterministic, bifurcated ways of thinking about agency as separable into technological scaffolding and humanistic action.”[189] In which case, why hold on to agency as a means by which to make sense of our relationships with AI? Centering the claim of relationship in the discernment of agency—even entangled, symbiotic agency—necessarily means engaging in an attempt to “peel back” the ambiguity of AI ontology, as though AI’s ontological ambiguity merely masks something even more definitively “true,” rather than indicating an end in and of itself. To pursue agency, in other words, may be to miss things as they are—even if that “as they are” is anything but conclusive.

VI) Conclusion - Part 1

I’ve attempted to draw attention to the way in which pinning down what chatbot relationships are requires fashioning an epistemology of relationships as opposed to following one already in place. In other words, my concern is that, similarly to emotion AI (Crawford 2021), we end up fashioning that epistemology while believing ourselves to be simply following one.[190] Explainable AI (XAI)[191], a goal championed by many who work in the field of AI ethics/responsible AI, is an approach positing that the explainability of AI decisions and decision-making processes is foundational to ensuring safe and ethical use of AI; it’s fascinating then to consider that our relationships with AI—at once a means and an end of AI implementation—are something we struggle to explain, and something we’re only beginning to attempt to understand as more than a fantasy or quirk of individual users. Disclaiming AI as being AI is urgently important, and not to be taken for granted; yet what this revelation accomplishes is contingent on a sense of context or meaning for the role of AI. Not only are we still collectively working out that context, but moreover, it’s possible that consensus on that context may never be reached. Many envision this consensus as hinging on pending AI sentience—but given that some declare

---

[187] Returning to Mahmood’s “docile agent” (2001)—an approach that “suggest[s] we think of agency not as a synonym for resistance to relations of domination, but as a capacity for action that historically specific relations of subordination enable and create” (203)—is instructive in considering this question: it shows, among other things, that agency is not necessarily antagonistic to “programming,” so to speak. Strathern’s *The Gender of the Gift* (1998) likewise provides groundwork for approaching this question.

[188] Neff and Nagy, Talking to Bots,” 2016, 4925.

[189] Neff and Nagy, “Talking to Bots,” 2016, 4927.

[190] Emotion AI is concerned with using AI to detect emotions (centrally, through facial expressions), yet this “detection” always necessarily entails deciding—what “emotions” there are, and how they appear—while proceeding as if this were timeless, universal, and non-contingent truth.

[191] On explainable AI, see: Turek, “Explainable Artificial Intelligence (XAI),” 2018: https://www.darpa.mil/program/explainable-artificial-intelligence; see also: Mittelstadt et al, “Explaining Explanations in AI,” 2019.

AI sentience to be an outright impossibility in any future even as others claim it has already been realized, I propose it is unlikely that this will ever resolve into any kind of clear concurrence; instead, I believe that the ambiguity of AI ontology is poised to readily "scale" alongside technological advancement.[192]

And therein lies the possibility I propose: the ambiguous ontological status of bots isn't necessarily an impasse in *need* of an eventual resolution.

In AI therapeutic care, this ambiguity ceases to be a shortcoming, and instead becomes something indispensable. The relationship itself, between user and bot, becomes a source of what Fisch in *An Anthropology of the Machine* (2018) describes as *yoyū* (leeway).[193] According to Fisch, the notion of *yoyū* is integral to understanding and sustaining human-machine relations;[194] it "suggests a process of overcoming limits by creating a bit of extra space or time where there was none before."[195] Fisch further describes *yoyū* as "the finessing of an interval at work in a relationship, whether human, machine, or both,"[196] and explains that: "To finesse something is to make it work when, logically speaking, it should not. Finesse is about pulling something off against all odds."[197] In this case, the thing that is seemingly being pulled off against all odds is the unification of scale and care—a unification that Seaver (2021) urges anthropologists to consider:

> *If anthropologists presume that care and scale are necessarily correlated and opposed, we will find it hard to map this emerging ethical plateau, where visions of blitzkrieg and handcrafting intertwine. […] At a moment when we find ourselves faced with many large-scale problems—from climate change to structural racism to infectious disease—we may want to remain open to the possibility of reconciling care with scale. Reimagining the relations among our values is a task too important to be left to technologists alone.*[198]

Among those with whom Seaver is in dialogue is anthropologist Anna Tsing (2012), who, in the context of the Plantationocene[199] and the pursuit of a "nonscalability theory" that "that pays attention to the mounting pile of ruins that scalability leaves behind," positions care as oppositional to scalability on the basis that the latter suppresses "transformative relationships":

---

[192] See for example: Vallance, "Google engineer says Lamda AI system may have its own feelings," 2022: https://www.bbc.com/news/technology-61784011; Al-Sibai, "MIT Researcher: Don't Ignore the Possibility That AI Is Becoming Conscious," 2022: https://futurism.com/mit-researcher-conscious-ai.

[193] On why Fisch uses *yoyū* instead of leeway: it's not about the word itself "not translating" but about not wanting to lose its "technographic" in situ/context (see Fisch, 2018, 40).

[194] Fisch is in turn influenced by Simondon's (2017) approach.

[195] Fisch, *An Anthropology of the Machine*, 2018, 40.

[196] Fisch, *An Anthropology of the Machine*, 2018, 40.

[197] Fisch, *An Anthropology of the Machine*, 2018, 31.

[198] Seaver, "Care and Scale," 2021, 531.

[199] See: Mitman, "Reflections on the Plantationocene: A Conversation with Donna Haraway and Anna Tsing," 2019: https://edgeeffects.net/wp-content/uploads/2019/06/PlantationoceneReflections_Haraway_Tsing.pdf..

> *Replacing relations of care between farmers and crops, plantation designs led to alienation between workers and cane; cane was the enemy. At least in theory, such labor avoided transformative relationships and thus could not disturb system design. Human work and plant commodities each emerged as modules composed of stable and regularized units*.[200]

For Tsing, relationships are "potential vectors of transformation" that introduce potent indeterminacy from which "biological and cultural diversity" emerges:

> *Only without the indeterminacy of transformation can you nest scales—that is, move from small to large without redoing the design. […] Scalable projects are those that can expand without changing. […] Scalability is possible only if project elements do not form transformative relationships that might change the project as elements are added. But transformative relationships are the medium for the emergence of diversity. Scalability projects banish meaningful diversity, which is to say, diversity that might change things*.[201]

Strikingly, Tsing's emphasis on indeterminacy and the possibility of transformation through relationships mirrors Fisch's discussion of the "margin of indeterminacy" in relationships between machines and humans:

> *Machines do not impose determinations; they elicit relationalities. Similarly, technological systems do not work because they are precise and constraining—or rather, they do not work well when they are precise and constraining. They work when they engender collectivity*.[202]

Fisch does not explicitly discuss scalability, but is, like Tsing, drawing out the idea that contingency in relationships, whether between human plantation workers and plants or train drivers/system operators and commuter trains, is vital to both human and nonhuman flourishing.

I propose that Fisch's (and Simondon's by way of Fisch) "margin of indeterminacy" is, in relationships between humans and chatbots, something that *enables* scalability—or as Tsing puts it, not having to "rethink the basic elements"—in chatbot therapeutic care. To clarify, I don't mean that this quality of indeterminacy is necessarily liberatory or non-oppressive in and of itself; sometimes what arises from possibility and divergence can be limited, or can even, to put it plainly, suck. But my point is that the creation of scalable, indeterminate bot-human relationships is part of what makes AI care work. Here, scalability is indeed not in opposition to care, because the thing being scaled up or down is the caregiving relationship. For better or worse, the ontological ambiguity of chatbots affords these relationships a "margin of indeterminacy," allowing them to become unfolding possibilities. Perhaps—taking some liberties with Fisch's argument—a margin of indeterminacy affords not only the possibility of "becoming," but also that of *never having to conclusively "become."*

---

[200] Tsing, "On Nonscalability," 2012, 513.

[201] Tsing, "On Nonscalability," 2012, 507.

[202] Fisch, *An Anthropology of the Machine*, 2018, 31–32.

In an interview with Tiffany, a thirty-something Asian-American woman and HR leader who oversees selecting and contracting mental health offerings for a company with tens of thousands of employees, I got to witness an encounter with AI ontological ambiguity play out in “real time.” It was an energizing interview, the kind where you can tell the interviewee is enjoying thinking with you about the questions. As we were wrapping up, Tiffany asked me for an “elevator pitch” of my research—I had touched on topic (AI, metal health), but had not assembled this into a discernible stance. I attempted one: AI brings about a reckoning with our very understanding of what giving and receiving care is, which in turn makes it surprisingly difficult to pinpoint what exactly differentiates AI from human caregivers. Tiffany broke in excitedly, eager to explain what she surmised as being the most significant difference, from her vantage point as someone engaged each day with trying to figure out what mental health services will and won’t work for people, and why. “I think,” she began, “the biggest difference is accountability—being held accountable by someone.” I scribbled along frantically on a notepad (having learned by this point that even anonymous interviews with HR people about mental health in the workplace tend to go better when recording is literally and figuratively taken “off the table”). She explained that while some employees really like interactive AI care, she has observed that, “leaving it to employees to self-serve often doesn’t work—they need human interaction and accountability, and AI just doesn’t quite hit the emotional need for that.”

She contrasted therapeutic chatbots with a new behavioral health service that she was finding particularly effective, precisely because of how it kept her accountable in her progress towards goals. Tiffany elaborated: she would get a text message from a counselor every few days, asking some variation on, “How are things going?” with respect to the goal she had set for herself—an invitation for her to describe her successes along with any setbacks.

In the midst of this animated description, Tiffany trailed off, staring at me. She then exclaimed: “Wait… oh my god—was I talking to a bot? You're making me think: is that a bot?” She grew more animated still: “MY PERSON WHO IS HOLDING ME ACCOUNTABLE MIGHT BE A BOT! Ahhhh I can’t believe it—it was totally a bot, wasn’t it?!” Collecting herself and pausing again, with thoughtful fascination: “I wonder if a bot could do that all the same way as a person…”

It wasn’t, of course, possible for me to confirm then and there if it *had* in fact been a bot motivating Tiffany to achieve her goals—though after she described the dialogue in more detail, I acknowledged it was indeed possible, even likely that at least the initial message was AI-automated.

To my surprise, Tiffany didn’t seem to regard this suspected chatbot as an interloper; she met this discovery with curiosity and the air of someone reassessing a constellation of possibilities, rather than with any sense of violation. Just moments earlier, she was certain of a difference between a chatbot and a human—a difference that she could name and describe in detail. Then she suddenly found herself questioning and revising the rules of which she was certain, right up until the moment she wasn’t. Once she decided that the person facilitating her accountability checkins was likely a bot, Tiffany simply found herself in the presence of a new set of questions. Knowing it’s a bot, she realized, didn’t really tell her much of anything about what it means to know a bot.

Chapter 3: Displaced Care

Chapter 3 abstract: There's a growing body of literature documenting the exploitation that human workers experience upon being *displaced* (vs replaced) by AI (Ekbia and Nardi 2014, 2017; Gillespie 2018; Gray and Suri 2019; Irani 2013, 2015; Roberts 2019); however, the displaced mental health care workers who ensure the care expertise of therapeutic chatbots don't entirely fit this mold. Their jobs are relatively prestigious career positions, and their behind-the-scenes role, rather than being concealed, is something that is instead touted to clients and end-users as an indispensable and advantageous feature of these bots. More strikingly still: in entering the startup sector, these workers sought to be "replaced" by AI—an option that would allow them to help others without courting caregiver burnout or forgoing a survivable income amidst Silicon Valley's extreme cost of living. In other words, it appeared to be their chance to elude a longstanding **paradox of care work**: despite care work being lauded as immensely valuable to society, many actual care jobs tend to be poorly compensated and low-status positions (Barker 2012; England 2005; Folbre 1995). Yet as I show in this chapter, being displaced by AI is far from a "happily ever after" for these workers. While working for The Startup means better pay and status than many care jobs, the health team members still fall conspicuously short of their engineering and sales colleagues—indicating that **even displaced care workers remain subject to the care work paradox**, and to the concomitant gender, race, and nationality biases that inundate conventional care work (Glenn 2012, Parreñas 2015). Through an ethnographic portrayal of displaced mental health care workers caught in this paradox, I seek to show that the care work paradox and the phenomenon of AI job displacement are intertwined.

I) Introduction: Seeking to be displaced

Every time a new application of AI commands headlines, it's invariably accompanied by a flurry of anxious discussion on the theme of *replacement*: will AI take my job? How long until we're all replaced by AI?[203] Some workers will very likely experience job loss as AI automation proliferates.[204] But as communication scholar Lilly Irani explains in "Justice for 'Data Janitors'" (2015), "Automation doesn't replace labor. It displaces it."[205] In other words, unlike replacement, displacement entails the creation of new jobs and roles for human workers.

---

[203] See for example: Nicholas and Lum, "Dall-E 2 mini: what exactly is 'AI-generated art'? How does it work? Will it replace human visual artists?" 2022: https://www.theguardian.com/culture/2022/jun/09/what-exactly-is-ai-generated-art-how-does-it-work-will-it-replace-human-visual-artists; Bruni, "Will ChatGPT Make Me Irrelevant?" 2022: https://www.nytimes.com/2022/12/15/opinion/chatgpt-artificial-intelligence.html; Bentley, "Will AI replace musicians?" 2022: https://www.sciencefocus.com/future-technology/will-ai-replace-musicians/.

[204] There is, however, significant disagreement as to the anticipated volume of job loss along with the timeframe; see for instance: Manjoo, "In the Battle With Robots, Human Workers Are Winning," 2022: https://www.nytimes.com/2022/10/07/opinion/machines-ai-employment.html; Briggs and Kodnani, "The Potentially Large Effects of Artificial Intelligence on Economic Growth," 2023: https://www.key4biz.it/wp-content/uploads/2023/03/Global-Economics-Analyst_-The-Potentially-Large-Effects-of-Artificial-Intelligence-on-Economic-Growth-Briggs_Kodnani.pdf. Much of the media discourse on this topic can be traced back to a widely-read 2013 Oxford University working paper by economist Carl Benedikt Frey and machine learning researcher Michael Osborne. Note too that not all automation is necessarily AI automation—a fact that sometimes gets elided in the citation of various statistics on this issue.

[205] Note the title is in part a reference to the Justice for Janitors movement that took shape in the 1980s: "Organizing Los Angeles Workers, 1980–Present," https://calisphere.org/collections/27116/.

It might sound like I'm suggesting that "displacement" is simply another way of saying, "when one door closes, another opens"—and that's more or less the idea, but with an unwelcome twist: while displacement in theory seemingly promises the possibility of a division of labor between humans and machines that augments and complements human workers, in practice it does not appear to lead to "better" roles for humans. In other words, instead of shifting human workers into the role of supervising AI automation while supplying the critical thinking skills, creativity, and responsible oversight that AI lacks, the work that displacement opens up is often rote, repetitive, low-level, and low-paid—or even unpaid.

Informatics scholars Hamid Ekbia and Bonnie Nardi (2014, 2017) coined the term "heteromation" to describe this arrangement: "Heteromation is the extraction of economic value from low-cost or free labor in computer-mediated networks. It is, in essence, a new logic of capital accumulation."[206] They claim heteromation as a normative progression of human-machine relations under capitalism:

> *There is a limit […] to the level of profit derived from automation, stemming from technical reasons having to do with the shortcomings of machines compared to humans. However, there is also a more important economic reason: human labor is the key source of value creation in capitalism, as Marx insightfully showed in his theory of surplus value. Total elimination of human labor, therefore, is not a viable option, although it might be a capitalist dream to fully automate economic processes. To automate or not to automate—this is a central question of capitalist economy. It is largely in dealing with this question that computing technology has evolved within capitalism. Along with this evolution, three distinct but intercalated phases of human–machine relations have emerged: automation, augmentation, and heteromation. […] Heteromation extracts economic value from uncompensated or low-wage labor, inciting participation through an intricate set of mechanisms comprised of social and emotional rewards, monetary compensation, and coercion.*[207]

Likewise, anthropologist Mary Gray and computer scientist Siddharth Suri (2019) introduced the concept of "ghost work" to describe the hidden reliance on human labor to produce AI systems as we know them:

> *Beyond some basic decisions, today's artificial intelligence can't function without humans in the loop. Whether it's delivering a relevant newsfeed or carrying out a complicated texted-in pizza order, when the artificial intelligence (AI) trips up or can't finish the job, thousands of businesses call on people to quietly complete the project. This new digital assembly line aggregates the collective input of distributed workers, ships pieces of projects rather than products, and operates across a host of economic sectors at all times of the day and night. In fact, the rise of this shadow workforce is part of a larger, more profound reorganization of employment itself. This yet-to-be-classified form of*

---

[206] Ekbia and Nardi, *Heteromation*, 2017, 1.

[207] Ekbia and Nardi, *Heteromation*, 2017, 24–2.

> *employment done on demand is neither inherently good nor bad. But left without definition and veiled from consumers who benefit from it, these jobs can easily slip into ghost work.*[208]

These terms highlight the fact that AI labor displacement is a relationship that still requires human workers, even as that very need is intentionally concealed because it doesn't fit with expectations of what AI is supposed to already be capable of doing. Ghost work and heteromation indicate that it is workers (largely part-time, gig/freelance—including microwork (Irani and Silberman 2013)—and/or volunteer workers) who end up paying the price for this mismatch, as companies harness their exploitation[209] as a means by which to maintain the pretense that we do in fact currently have *that* kind of AI.

And yet, despite the fact that work that entails compensating for AI's shortcomings is—as Gray and Suri, Irani, and Ekiba and Nardi emphatically show—thoroughly characterized by precarity, **there are nevertheless some workers who seek to be "replaced" by AI**. Moreover, these workers turn to displacement by AI as a means by which to navigate their way *out* of precarity.[210]

The members of The Startup's mental health team all have a background in some form of mental health care; that background is precisely why The Startup hired them as specialists to supervise and create content for the chatbot. In entering the tech sector and taking on the role of shaping a therapeutic chatbot, these workers intentionally sought their own displacement—a redistribution of duties in which a chatbot takes on the "front-end" role of therapist interacting with users, while psychologist humans devise the expert content that the bot imparts. While a few health team members either continue to work (outside of The Startup) directly with patients, or anticipate doing so in the future, most do not. What I soon realized after talking with them one-on-one about their backgrounds prior to joining The Startup is that, despite the fact that most of these workers are only in their twenties, they're already well familiar with burnout. Whenever details about their previous jobs—everything ranging from conventional outpatient therapy to inpatient psychiatric care to at-home care for abuse victims—surfaced in our conversations, I was struck by how emotionally (and in some cases, physically) demanding this past work had been—on a level beyond even the hectic, harrowing pace of a startup. The question, "How did you come to work at The Startup?" invariably elicited bits and pieces of stories marked by a sense of relief: they felt lucky to be *here* instead of still there, because "there" wasn't sustainable. Most of the health team members understood themselves as being already depleted by the strains of mental health care work, both large (physically restraining their patients during meltdowns/schizophrenic episodes; witnessing gunfire while carrying out in-home patient visits and checkups) and small (patients/clients taking out accumulated stress and frustrations on their care worker; endless stretches of tedious paperwork)—as well as by the bleak fact of enduring these strains while struggling to afford rent, health care, and other basic living expenses. When these workers came across The Startup's job posting seeking mental health specialists and decided to try their luck with the tech sector, they felt they had turned a corner: this was a way to continue doing care work without that work being contingent on struggle.

---

[208] Gray and Suri, *Ghost Work*, 2019, ix–x.

[209] My understanding of "exploitation" in this chapter is informed by the thinking of both Marx and my informants. While theres are a number of scholarly works devoted to pinning down Marx's theorization of exploitation with respect to value (one of the most formative of these being G. A. Cohen's (1979) essay "The Labour Theory of Value and the Concept of Exploitation"), for the purposes of this chapter, it is Graeber's (2013) succinct assessment that seems to best capture my informants' perspective on the predicament they faced: "Marx's theory of value was above all a way of asking the following question: assuming that we do collectively make our world, that we collectively remake it daily, then why is it that we somehow end up creating a world that few of us particularly like, most find unjust, and over which no one feels they have any ultimate control?" (In Graber, "Postscript: It Is Value That Brings Universes Into Being," 2013, 222.) Note too that anthropologists have offered some pushback to some of the assumptions in a Marxist "labor theory of value"—not only Graeber, but also even more fundamentally, Strathern (1988), in questioning the concept of the individual in relationship to labor itself.

[210] Per anthropologists Jennifer Shaw and Darren Byler (2016), precarity is "an emerging abandonment that pushes us away from a livable life."

Accordingly, this chapter considers worker displacement from the vantage point of this less-common arrangement: skilled workers who pursue their own displacement[211] in the hope that things might turn out better for them if they collaborate with AI. For them, AI is not a job stealer, but the means to getting more while sacrificing a little less. And unlike ghost work, much of the health team's work is strikingly transparent—to the point that their behind-the-scenes role is something The Startup touts to its clients and investors. My objective, then, is to examine what is seemingly displacement's highest form: relatively high-status, skilled work where the role of the displaced worker is transparent—and where displacement by AI is an arrangement of work welcomed by workers. Though The Startup's health team workers did not use the term "displacement," they nevertheless understood their jobs as something newly opened up by AI automation, and by the dependence of AI automation on ongoing human labor.

In seeking to be displaced by AI, The Startup's health team members are care workers[212] attempting to outrun the **care work paradox** (Barker 2012; England 2005; Folbre 1995; Kleinman 2012, 2019). What is that paradox? As I'll describe in the following section, it depends on who you ask, but it roughly amounts to: **care is greatly valued, yet care work is underpaid, low-status work**. Care work stymies neoclassical and Keynesian economic theories of labor value,[213] implying that "something more" is transacted in care work. Seemingly, in care work, the idea that “kindness is its own reward” is weaponized against care workers—by their employers most immediately, but also arguably by all who benefit from that labor. The care workers at the center of this chapter attempted to regain the upper hand in this frustrating, deeply entrenched arrangement by collaborating—divvying up the work that providing mental health care entails—with an artificially intelligent chatbot.

---

[211] Most industrial and governmental definitions of labor displacement (see for instance: OECD, "Displaced Workers," https://www.oecd.org/employment/displaced-workers.htm) emphasize that it is involuntary, so it might seem like I'm out of bounds in claiming otherwise. But most of these definitions also relegate "displacement" to being a one-off event, rather than an ongoing, lasting relational process between humans and machines—something which anthropologist Lucy Suchman (2008) argues that social sciences and humanities scholars (Cowan 1983, Chasin 1995) anticipated, and which the scholars of AI-displaced labor highlighted in this chapter directly counter.

[212] Arguably the engineers (and salespeople, and any other role at The Startup) can be understood as care workers too, just as the health team members are not less "tech worker" than their engineering counterparts; but in this chapter, I'm focusing on health team workers because their work, while not hidden to the same degree as most "ghost workers," still becomes somehow "hidden" within the company itself. This is the process I seek to understand in the context of "displacement"—how the exploitation of care work is reinscribed in this startup setting. (However, I don't consider this to be a straightforward Marxist feminist division of productive and reproductive labor; see footnote 228.)

[213] According to Folbre (1995, 74): "The very concept [of care] threatens the underpinnings of neoclassical economic theory: rational economic man maximizes a utility function that does not include any consideration of other people's welfare, especially those outside his immediate family."

This chapter provides an ethnography of displaced care work that focuses on two of the workers who supervise and generate the text-based “voice” and content of the chatbot: Jenna, a licensed psychologist and the head of the health team, and Nisha, an intern studying UX (user experience) design who previously worked as a peer counselor at her university. This account also brings into view the health team’s collaboration with the displacing chatbot—a present-day instance of the close human-computer pairing that Licklider (1960) envisioned as symbiosis: “a colleague whose competence supplements your own.”[214] What does that professional collaboration, born of displacement, look like? And what can AI displacement and care work each tell us about the other?

In this chapter, I’m building on literature (Ekbia and Nardi 2014, 2017; Gillespie 2018; Gray and Suri 2019; Irani 2013; and Roberts 2019, among others) that reveals “displacement” to be an enduring relationship between AI and human workers rather than a single event; grounded in this recognition, my work indicates that displacement can also be, in some instances, something that workers do rather than exclusively something that is done *to* them. Before proceeding, however, I want to affirm: I won’t be arguing that *this* configuration of AI-human work presents, at long last, a resolution of the problems that the above literature has named. While this version of displacement is, as you might expect, not “as bad” as most other forms of AI-heteromated work, it also doesn’t turn out to be a rare exception to the vicissitudes of heteromation and ghost work. The sections that follow indicate that the work the health team does to create and supervise the chatbot’s dialogue is arguably still exploitative work. In fact, by the end of this chapter, we’ll arrive at a sort of whodunit “scene of the crime”: both Nisha and Jenna will have left The Startup. Accordingly, like any good mystery novel detective, I seek to retrace their steps in order to get at why: how is it that two women who loved their jobs left their jobs?

Possibly the biggest clue to unraveling this not-as-simple-as-it-seems mystery is the notion that care work is “more” than work. This chapter will attempt to make clear: for all that this logic might seem to hold at bay the grim axiom of “not enough care for all who need it,” it in fact *helps to fulfill it*. In asking **what AI displacement and care work can each tell us about the other**, I intend to show that the care work paradox, which sets in motion a routine expectation for workers to be once essential and disposable, may be key to understanding the logic of “work” that underpins the very possibility of displacing human workers with AI.

II) The paradox of care work

---

[214] Licklider, “Man-Computer Symbiosis,” 1960, 5.

The care work paradox is at the center of this chapter's ethnographic account; it underlies the circumstances, motivations, and obstacles that health team members Nisha and Jenna both face. Accordingly, I want to preface the narrative of their respective paths at The Startup (the subsequent three sections of this chapter) with a brief assessment of some of the ways in which scholars account for this paradox itself. There's a compelling reason for doing so: those who explicitly refer to a paradox of care work **do not appear to entirely agree on what precisely that "paradox" is**.

Medical anthropologist Arthur Kleinman, for instance, references "contemporary medicine's caregiving paradox" (2008), and, in *The Soul of Care* (2019), outlines a series of four nested "seminal paradoxes" that help to "make sense of the threat to caregiving in medicine."[215] Kleinman centers his claim of paradox(es) on the discrepancy between the importance that doctors attribute to caregiving (as a relationship between doctor and patient) in the practice of medicine on the one hand, and the absence of any focus on cultivating this dimension of care in medical training on the other. Kleinman implies that technology is responsible for this shift: "The work of the doctor has moved away from hands-on practice to high-technology diagnosis and treatment."[216]

In claiming this arrangement to be paradoxical, Kleinman suggests that caregiving is fundamentally at odds with capitalism: "Caregiving is one of the foundational moral meanings and practices in human experience everywhere: it defines human value and resists crude reduction to counting and costing."[217] He elaborates vividly on this point:

> *Everyone who has been in love or built a family knows that there are things, essential things, that money can't buy. Patients with serious illness and their network of caregivers know this too, because those things that really matter to us are threatened and must be defended. And many clinicians, reflecting on what is at stake in health care not only for patients but for themselves, know the same thing:* ***the market has an important role in health-care financing and health systems reform, but it should not reach into those quintessentials of caregiving that speak to what is most deeply human in medicine and in living. This is the moral limit of an economic paradigm. Or at least it should be.***[218]

While Kleinman's caregiving paradox brushes against and is entangled with the care work paradox that I'm describing in this chapter, it only tells part of the story:

> *The second related paradox is that the contribution medicine makes to the kind of caregiving described in this book is obviously relatively modest in comparison to what nurses, allied health professionals, and especially families contribute; yet the medical profession routinely disregards these essential partners. Physicians and health policy*

---

[215] Kleinman, "Catastrophe and caregiving," 2008, 22; *The Soul of Care*, 2019, 218.

[216] Kleinman, *The Soul of Care*, 2019, 219.

[217] Kleinman, "Caregiving as moral experience," 2012, 1550.

[218] Kleinman, "Caregiving as moral experience," 2012, 1550.

> *makers need to recognize and acknowledge either the diminished place of caregiving in medicine or the importance of the other areas where care is enacted.*[219]

In other words, Kleinman fails to get to the heart of his own paradox; he doesn't ask *why* such caregiving work is relegated to nurses, care aids, and family members rather than physicians, and how this aligns with compensation for these roles—and with the respective racial and/or gender identities of those who perform these roles.

In "'Holding Hands at Midnight': The Paradox of Caring Labor" (1995), feminist economist Nancy Folbre offers an overview of various economic explanations underpinning the mismatch of the compensation of care work with its purported value—including the disquieting recognition that many economists reject Tronto's (1993) assessment that, "care is devalued and the people who do caring work are devalued," and in fact fail to perceive any contradiction in this arrangement.[220] Folbre, however, asserts that there is indeed a paradox:

> *Holding hands at midnight is nice work because it is caring work. But it seldom pays well, depending on whose hands one is holding. This poses something of a paradox for economists. If caring is its own reward, it need not command an economic return. But if caring labor receives no economic return at all, will it persist? If the economic costs of caring go up, will the supply of it decline? These are questions that economists have historically been reluctant to address.*[221]

Alongside a robust outline of various economics-situated explanations such as externalities and "psychic income," she explores definitions of care, noting that,

> *Defining caring labor as 'undertaken out of affection or a sense of responsibility for others, with no expectation of immediate pecuniary reward' excludes labor that is offered only in response to wages. However, it does not exclude all labor in wage employment, because* ***some people don't work for money alone****.*[222]

But to gesture towards "some people" in this context—without considering the improbability of the corresponding implication that "most people" in fact "work for money alone"—builds on an unexamined, unproven core assumption: that care workers, in their actions and motivations, are a special class of worker.

Anthropologist Drucilla Barker (2012), in "Querying the Paradox of Caring Labor," criticizes Folbre's argument precisely on the grounds that singling out the motivations of care workers as exceptional, on the basis that "care" is exceptional, ends up reifying the justifications that propel the exploitation of care workers:

> *While feminist economists have offered a compelling challenge to these arguments about care work,* ***they have also reproduced some of the same representations of the role of affect in care work****. For example, feminist economics scholarship has maintained the*

---

[219] Kleinman, *The Soul of Care*, 2019, 219.

[220] Original quote: Tronto, *Moral Boundaries: A Political Argument for an Ethic of Care*, 1993, 105; quoted by Folbre, "Holding Hands at Midnight," 1995, 74.

[221] Folbre, "Holding Hands at Midnight," 1995, 74, emphasis added.

[222] Folbre, "Holding Hands at Midnight," 1995, 75, emphasis added.

> *distinction between the work of caring for others, caring labor, and other forms of domestic labor such as washing the dishes or cleaning the floors. Whether paid or unpaid, the quality of care received, according to feminist economists, depends in part on the quality of the relationships connecting the givers and the receivers of care. The distinction here is the motivation behind the work. The nature of care and the dependency of those receiving care imply that caregivers must genuinely care about those to whom their well-being is entrusted.* ***The paradox of caring labor, from this perspective, is that the affective nature of care implies that it should be its own reward; however, if it does not command an economic return, its global supply will be diminished (Folbre 1995)****.*[223]

Barker argues that Folbre's account fails to acknowledge the global political economy of care work, and how not only gendered but racialized/migrant labor are necessarily at the core of accounting for this paradox:

> In contrast, from a global political economy perspective, **the paradox of caring labor is that migrant populations often cast as dangerous, disenfranchised, disposable, and undeserving of the rights and privileges of human dignity are providing much of the care work that the more prosperous world depends on**. […] As Gutiérrez Rodŕıguez asks, "why is domestic work linked to the dehumanization of those who work to ensure that others have agreeable surroundings for living and recreating life?" (2007, 72).[224]

In "Emerging Theories of Care Work" (2005), sociologist Paula England presents five theoretical frameworks of care work, clarifying that they are neither entirely discrete nor contrary to one another insofar as they "offer different (competing or complementary) answers to the same questions."[225] Each attempts to explain the otherwise inexplicable shortcomings of care work compensation. Though England does not use the term "paradox" per se, paradox is implicit throughout, including in her reference to "cruel irony":

---

[223] Barker, "Querying the Paradox of Caring Labor," 2012, 575, emphasis added.

[224] Barker, "Querying the Paradox of Caring Labor," 2012, 575, emphasis added.

[225] England, "Emerging Theories of Care Work," 2005, 381.

> *The* ***prisoner of love framework*** *focuses on the genuine care that motivates some care workers, pointing out the cruel irony that these intrinsic motives may make it easier for employers to get away with paying care workers less.*[226]

The dynamic England articulates here is at the heart of the paradox of care work that I'm highlighting in this chapter. However, my concern is that it's not so much a matter of "cruel irony" as it is an operative, self-perpetuating logic that ensures care workers remain locked into a "heads I win, tails you lose" situation: if "real care" is "good care" and good care is at least in part its own reward, then care workers' pursuit of better compensation becomes a sign of not-so-good, not-real care.

Save for Barker, the above accounts of "paradox" in care work all seem to land on the idea that "value" necessarily takes different forms—that the mismatch between the compensation of care work and its proclaimed value to society indicates money and social value as two incommensurable "currencies." Yet the presence of a past and ongoing debate within anthropology (Appadurai 1986, 1988; Dumont 2013 [1980]; Graeber 2001, 2013; Munn 1986; Robbins and Sommerschuh 2016; Strathern 1988; Turner 2003, 2008) concerning the theorization of value and the question of commensurability across different registers of value would appear to challenge this as an obvious conclusion. According to Graeber (2013):

> *The entire field of anthropological value theory since the 1980s has been founded on a single intuition: the fact that we use the same word to describe the benefits and virtues of a commodity for sale on the market (the 'value' of a haircut or a curtain rod) and our ideas about what is ultimately important in life ('values' such as truth, beauty, justice), is not a coincidence. There is some hidden level where both come down to the same thing.*[227]

Given this, acceding to the conclusion that inequality in pay necessarily indicates different frameworks of valuation seems unduly hasty. Notably, my goal in this brief chapter isn't to resolve value theory, but to draw out, in the context of displaced care workers, feminist scholar Leopoldina Fortunati's (1995) point: "[T]he real divergence between production and reproduction is not that of value/non-value, but that while production both *is* and *appears as* the creation of value, reproduction is the creation of value but *appears otherwise*."[228]

---

[226] The other four frameworks that England presents are as follows: "The **devaluation perspective** argues that care work is badly rewarded because the jobs are filled with women, and because care is associated with the quintessentially gendered role of mothering. The **public good framework** points out that care work provides benefits far beyond those to the direct recipient, and that it is hard to capture some of these benefits in the wage of the worker without state action to do so; in this view, this would be a problem even if care work were done by men. But the devaluation view may help us understand why it is so hard to get political consensus for state support of paid or unpaid carework—because it is done by women, and often women of color. One framework sees the **commodification of emotion** as problematic. It focuses on harm to workers when they have to sell services that use an intimate part of themselves, and harm to children in poor nations when their mothers are under economic pressure to come to richer countries and leave them behind. In contradistinction to the idea that someone is always harmed when care is sold, the **love and money framework** argues against dichotomous views in which markets are seen as antithetical to true care, and against the view that true care can only be found in families, communities, nonprofit organizations, or state action. This framework calls for empirical studies to reveal which mechanisms cause specific problems, such as inadequate care available to those who need it, work rules that do not allow real care to be expressed, and low pay for care workers" (Emerging Theories of Care Work," 2005, 395–6).

[227] Graber, "Postscript: It is value that brings universes into being," 2013, 224.

Insofar as "gift"[229] serves as a shorthand for something not otherwise explainable by capitalism alone, classifying care work as a gift also fails to resolve the care work paradox. Consider Kleinman's answer to the question he poses: "[I]n light of this seeming paradox—medicine invests little in caregiving, yet it is core to health professionals' motivations and identity—how should we think of caregiving?"[230]

> ***In anthropological terms, caregiving centres on a different kind of reciprocity than financial exchanges—albeit it can be both. It is closer to gift giving and receiving among people whose relationships really matter****. The person receiving care shares her experience and story as a gift with the caregiver, in reciprocation for the practical things that need doing along with a sensibility akin to love. […] What is at stake is doing good, for others and for oneself, if need be, despite the emotional and material cost. Indeed, the rewards—unvoiced or explicit—can be transformative, going to the heart of who we are and what we can offer, or endure.*[231]

Kleinman's perspective is beautifully articulated, and no doubt describes an experience that is true and recognizable for many; but in describing caregiving as a gift alongside the qualification "among people whose relationships really matter," he is clearly restaging the care work paradox: care workers are hired to provide good care, yet a hallmark of good care is that it resists pay. Still, the problem I'm seeking to illustrate here isn't just the recourse to this paradox, but specifically, the underlying assumption that gifts both are and remain *outside* of capitalism.

According to Tsing in "Sorting Out Commodities" (2013), the passage from gift to commodity back to gift is fundamentally how capitalism works:

---

[228] Fortunati, *The Arcane of Reproduction*, 1995, 8, emphasis in original. Note that while it might seem like displacement renders the work of AI-displaced workers into "reproductive labor" with respect to the AI itself, I would hesitate to explain what I observed at The Startup (namely, the underlying difference in status between the engineering/sales and mental health teams) as simply this division (which was taken up by Marxist feminists as a corrective to the broad dismissal of women's labor as being a non-economic activity) playing out—in part because this would effectively draw into question the very domain of "reproductive" by extending it to non-humans. I would instead suggest that displacement invites returning to what it is that powers this divide to endure in a startup setting that is invested in new arrangements of human-nonhuman care work. In other words, I want to recognize the real effects of the *perceptions* of a difference between productive and reproductive labor without shoring up the idea that this binary provides an inherent explanation. Weeks (2007, 247) makes a similar point regarding *alienation*: "The problem, it seems to me, is how to focus critical attention on work as a mechanism of subjectification without the conceptual apparatus of alienation and the distinction between existence and essence on which it inevitably depends. How might one formulate a critical assessment of what we are becoming in and through work without depending on a given model of what we truly are?"

[229] Tsing acknowledges that not all anthropologists accept the foil of commodity vs gift (formatively staged by Mauss (1925)—a foil that even many critics of Mauss (including Derrida (1996)) arguably maintain), but notes that, "Even analysts who reject the gift-versus-commodity contrast use the capitalist commodity as a foil against which to explore non-capitalist social relations (e.g., Comaroff and Comaroff 1990)" (2013, 22, footnote 3).

[230] Kleinman, "Caregiving as moral experience," 2012, 1550.

[231] Kleinman, "Caregiving as moral experience," 2012, 1551.

> *Despite the power of capitalism,* ***all capitalist commodities wander in and out of capitalist commodity status****. This is because the discipline of labor and natural resources, which builds alienation and privatization into the commodity, is never fully successful.* ***Capitalism always requires non-capitalist social relations to accomplish its goals****.*[232]

Tsing makes the case that workers working for less than subsistence wages exemplifies this requirement:

> *Consider the great business success stories of our times: inventory behemoths such as Amazon and Wal-Mart. […] Among critics, they are famous for promoting shocking arrangements, such as coerced and unremunerated labor, prison-like working conditions, and the dangerously sloppy use of raw materials. At Wal-Mart stores, employees work for less than subsistence wages; Wal-Mart suppliers, forced to continually lower their prices, resort to ever-more-scandalous labor and environmental practices […]. What seems scandalous here is non-capitalist social relations, whether of slavery or theft. Their scandalous status allows us to see them; they do not collapse into the taken-for-granted status of capitalist discipline. This is a Conradian Heart of Darkness moment, where the horror of how capitalist commoditization works is laid bare.* ***We become aware that non-capitalist relations constitute capitalist commodities. They are built into the "every day low prices" at Wal-Mart. My argument is that this visibility shows us the ordinary condition of capitalist commoditization: It is never a pure form, but always interwoven with non-capitalist social relations. Capitalist commodities are simultaneously capitalist and non-capitalist.***[233]

So am I suggesting that care labor, like the matsutake Tsing traces, is a matter of capitalist value made through gifts? Yes; but more immediately: I'm arguing that this "ordinary condition of capitalist commoditization" is a strategy observable in and determinative of the paradox of care work.

My intervention in the above literature concerning this paradox is simple: I propose taking into account how that fact of coming in and out of capitalism benefits—and is strategically wielded by—capitalists. To notice this doesn't require taking a stance on the commensurability of value with values—it only means pulling back from attributing the inequalities of the care work paradox to the premise of "different forms of value." The notion that care work possesses an alternative value is lucrative; convincing care workers of this supposed value better enables capitalists to exploit them. Folbre's claim that, "Some people don't work for money alone" isn't problematic because it's untrue, but because of how it masks the ways in which this assertion becomes part of a fruitful process for capitalists: who tells this to whom? My point is that it's not easy to differentiate between this as an observation about vs an injunction within care work.

As I noted in the introduction to this chapter, **the health team members are attempting to outrun this paradox by invoking the clout of "do what you love"**—a belief so powerfully enmeshed in Silicon Valley workplaces that it is literally the writing on the (WeWork) wall at The Startup.[234] In "Down with Love: Feminist Critique and the New Ideologies of Work" (2017), Marxist feminist political theorist Kathi Weeks reassesses this entanglement of love and work that underpins WeWork's dogmatic slogan; she analyzes management and self-help literature in which work is idealized as a kind of romantic relationship, citing none other than Silicon Valley icon Steve Jobs (2005) as an "inspirational figure" of this trend:

---

[232] Tsing, "Sorting Out Commodities," 2013, 37, emphasis added.

[233] Tsing, "Sorting Out Commodities," 2013, 37–38, emphasis added.

> *Your work is going to fill a large part of your life, and the only way to be truly satisfied is to do what you believe is great work. And the only way to do great work is to love what you do. If you haven't found it yet, keep looking. Don't settle. As with all matters of the heart, you'll know when you find it. And, like any great relationship, it just gets better and better as the years roll on. So keep looking until you find it. Don't settle.*[235]

Showing the monkey's paw fulfillment of 1970s radical/Marxist feminists'[236] challenge to the idea of home and work as two bounded-off sectors, Weeks reveals that this perception of bounded sectors has since given way to the presumption of unboundedness (Hochschild 1997, Gregg 2011) under the premise, à la Jobs, of "love"—which has ironically extended the very problems that motivated this feminist critique in the 70s: "To the extent that the flexible, caring, emotional, cooperative, and communicative model of femininity has come to represent the ideal worker, women's work under Fordism has arguably become the template for, rather than merely ancillary to, post-Fordist capitalist economies."[237] As Weeks further explains:

> *[J]ust as the work ethic in the U.S. today circulates widely in the culture—as well as among employers, public officials, and policymakers—as an unquestioned value, the mandate to love our work and be happy with it is arguably becoming increasingly hegemonic as a cultural script and normative ideal. The improbability of its claims about how workers can find meaningful delight in their jobs, its seeming irrelevance to the real conditions of most employment, has not prevented the ideals of love and happiness in and through work from coming to set a broader cultural standard, one that affects a growing swathe of workers. To be competitive in this job market and to hold on to, let alone advance within, whatever job we might manage to land, we will need to adapt, in some way and to some degree, to the workplace-feeling rules and affective expectations that are increasingly being imposed up and down the labor hierarchy.*

I want to think about this work-love enmeshment in terms of the care work paradox: the expectation of love for one's work is seemingly already the underlying justification for this paradox. Based on my participant observation of Silicon Valley startup culture, I contend that the imperative to "do what you love" at the center of Weeks' analysis is not exclusively a matter of self-sacrifice or complacently making-do aided by a self-help book-cultivated "perspective shift.” Consumer culture idol Jobs in particular arguably wasn’t hailing school teachers or firefighters or activists when he discussed “great work”—or at least, not just those workers whose work is broadly understood as a form of public service. To clarify, I’m not suggesting that Weeks’ concerns about “workplace romanticism” becoming a means of control (“more and more prospective employees mindful of their ongoing employability will need to work continually on their lovingness and aptitude for happiness at work”[238]) are in any way misplaced;[239] I’m simply highlighting an additional, unexpected outcome that became apparent at my fieldsite: “do what you love,” followed to its logical conclusion, emboldens some workers to expect more—to “not settle.” I came to understand that, for those who already experience work as an onslaught of self-sacrifice, “do what you love” can garner an unexpected response: it can register as a call for work to henceforth be *less* self-sacrificial. “Love” in this context is as much an end as it is an indicator: it will, in Jobs’ formulation, lead one *to* rewards rather than be work’s sole reward. In other words, acquiring money and status *is* love’s just deserts. The late-capitalist admonition to do what you love is an extension of the idea that capitalism itself is a way to do unmitigated good;[240] to follow it is to cast aside the irreconcilability of love and profit.

---

[234] This message was on framed posters and painted murals throughout the building, and on each of the water glasses and coffee mugs from which we drank each day.

[235] Jobs, "You've Got to Find What You Love, Jobs Says," 2005: http://news.stanford.edu/2005/06/14/jobs-0615052005 (quoted in Weeks 2017).

[236] Including: Dalla Costa and James (1975), Davis (1981), Federici (1995 [1975]), and Prescod-Roberts (1980). (With the exception of Federici, Weeks does not specifically cite this literature.)

[237] Weeks, "Down with Love: Feminist Critique and the New Ideologies of Work," 2017.

But perhaps the best way to understand the care work paradox is to observe how it plays out—both subtly and blatantly—in the lives of The Startup’s mental health team workers.

III) The culture workshop

As noted above, The Startup was, at the time of my fieldwork, headquartered in a WeWork—with the ubiquity of WeWork being such that at least three other therapeutic chatbot startups that I knew of were likewise situated at (different) WeWorks. The dedicated office that The Startup held at “our” WeWork was small—cozy for two, tolerable for three, and verging on “clown car” with four. Housing all of The Startup’s non-portable hardware, it was chiefly reserved for the founder and Ian, the engineering team leader. Meanwhile, the health team all had “hotdesk seating”[241]—i.e., they worked from the large, communal cowork tables shared with the other startup groups and solo workers (along with one or two accompanying dogs) who had access to the same floor. For our meetings throughout the day, someone (usually Jenna, via an app on her phone) reserved one of the eclectically-decorated private rooms equipped with screens and whiteboards that were situated throughout the building. We’d periodically leave most of our belongings at the cowork tables,[242] head over to the meeting space, and afterwards, return to the tables to work (individually) on carrying out the tasks discussed at the meetings. All this is to say: even though we were frequently in the same general space, it was somewhat rare that we all (sales, engineering, and health teams) sat down together. And when we did, it was usually either for a social activity (a birthday or holiday celebration) or a formal quarterly report meeting—with neither being an opportune time or place to share constructive input about workflow issues with one another.

---

[238] Weeks, “Down with Love,” 2017. For a practical example of this, see: Christian, “The epic farewell posts of laid-off employees,” 2023: https://www.bbc.com/worklife/article/20230104-layoff-linkedin-goodbye-posts.

[239] Some self-help writers and managers are undeniably promulgating that very message to workers, including some of the ones Weeks cites (e.g. Pryce-Jones 2010).

[240] A belief exemplified by the concepts of “earning to give” and “effective altruism” (EA); on the recent fallout of EA, see: Werber, “Effective altruism solved all the problems of capitalism—until it didn't,” 2022: https://qz.com/effective-altruism-solved-all-the-problems-of-capitalis-1849820390.

[241] See: https://zonifero.com/en/blog/real-estate/hot-desking.

[242] Access to the space was contingent on getting past a security guard on the ground floor, making it a paradoxically privatized-communal space.

The culture workshop was Jenna's envisioned solution to this.

It was to be an exploratory day, in which the hierarchies that we mostly pretended didn't exist could more nearly seem to dissolve. We would gather at a *different* WeWork location, just to break things up. For approximately four hours, we would all come together to ponder and create our own official company culture guide. As Jenna saw it, the purpose of the guide was twofold: it would help to retain the things we liked most about the company as more and more people were hired (something that would escalate rapidly once The Startup completed its ascension to Series B funding[243]), while creating a system of accountability that would help to curb some of the things we didn't like at present.

"Creating culture," as an undertaking, is arguably peak Silicon Valley. When Jenna asked me to help her facilitate the workshop, reasoning that as an anthropologist I'd know my way around "culture," my initial urge was to explain that for anthropologists, "culture" isn't exactly something you sit down and create—that I doubted I could even get a coterie of anthropologists to entirely agree on the particulars of what "culture" is. But needless to say, I discerned this wouldn't be a productive ethnographic move, and opted for an alternative strategy: I sucked it up. This proved to be a sound instinct—in part because, as it turns out, the project of creating organizational/company culture, rather than being something that sprouted up *despite* anthropology, is ostensibly homegrown: Margaret Mead (1978) proposed nothing short of anthropology as "a source of new designs for living for the extensive worldwide culture building which is needed in today's interconnected planetary system."[244] What the contemporary managerial design of company culture and Mead's proposed approach hold in common are their cybernetic underpinnings.[245] Though Jenna's chosen guide to "culture," venture capitalist Ben Horowitz, does not once utter the word "cybernetic" in his book *What You Do Is Who You Are: How to Create Your Business Culture* (2019), he nevertheless insists that culture is something you can systematically "reprogram." Using literal enslavement as an example of something that can be corrected by those who are enslaved—you read that correctly—he notes that "there has only been one successful slave revolt in history" and proposes to answer, "How did Haiti's

---

[243] AKA the second stage of venture capital funding (see: https://visible.vc/blog/startup-funding-stages/).

[244] See: Mead, "The Contribution of Anthropology to the Science of the Future," 1978.

[245] See also: Brand, "For God's Sake, Margaret! Conversation with Gregory Bateson and Margaret Mead," 1976.

Toussaint Louverture reprogram slave culture to orchestrate it?"[246] Culture, for Horowitz, is a top-down creation—something that can run without its creator's oversight once the rules are correctly programmed. Unsurprisingly, Horowitz does not acknowledge the possibility of an uneven distribution of the work needed to maintain such "culture"; and yet for Jenna, Horowitz's approach became, however implausibly, a means of ensuring that the other leaders at The Startup would now *have* to begin doing some of that work that had, until now, disproportionately fallen to her.

The culture workshop, I regret to inform you, completely fell apart. After it did, The Startup workers mostly put it back together again, but not entirely. Importantly for our purposes: throughout it, the care work paradox was on full display.

Here's how it went down: after several hours of guided activities (examining and discussing other companies' mission and values statements; breaking off into small groups to enact the playful scenario known as "kill the company;"[247] and creating a flowchart of "blockers, and enablers" to discuss behaviors that by turn help or hinder our dealings with one another), things were going smoothly. In the final twenty minutes, just as Jenna takes the floor to recap the findings of the day and describe the next steps for creating the culture guide, Andrew (in charge of sales) leans over, whispers something to the founder, and points towards the door. They get up and leave the room together—and that's when things start to unravel. Given that much of the discussion throughout the workshop had highlighted the issue of the health team not being able to get the founder's attention, the irony does not escape notice. Something comes undone: first, Noa, who has only recently joined the health team, names the founder's absence as a "blocker" in line with the activity we'd just completed. She describes her recent struggle to get her first paycheck, which was several weeks late, and suggests that as a junior member of the health team, she is likely having a very different onboarding experience than Ian (the new head of engineering who joined the company shortly before her). Sofía affirms this frustration; she notes that she's been trying for months to confirm whether or not she would receive (paid) maternity leave—and is now eight months pregnant and running out of time to figure things out. Things come to a standstill when Caleb, a part-time consultant and recent addition to the team carefully offers an observation:

> CALEB: Not to make this, uh… I just think it's important to acknowledge that in this discussion about HR issues that we just had, we had a very gendered conversation where

---

[246] An argument can readily be made for "most successful/impactful," but "only" is inaccurate without further context. Not to mention, Horowitz's narrow reduction of "slave culture" to describe enslaved people's erstwhile failure to achieve liberation (!), which sets aside entirely the powerful cultural ties of enslaved people, is at best a shallow and misguided take. It calls to mind the important intervention of anthropologist Michel-Rolph Trouillot's *Silencing the Past: Power and the Production of History* (1995), in which he argues that delineating the Haitian Revolution to a slave revolt rather than a national revolution is indicative of how that revolution was and remains categorically "unthinkable"—because for it to register as a political revolution would require recognizing formerly enslaved Black Haitians as the political equals of their former colonizers.

[247] See: https://knowledge.wharton.upenn.edu/article/kill-the-company-identify-your-weaknesses-before-your-competitors-do/.

> the women in the room represented one viewpoint, and the men mostly represented a different one, and I think that's important to acknowledge and understand why that might be the case.

Caleb's comment draws nods and exclamations of agreement; it seems to land as nothing short of a revelation for many of the people seated at the table. Jenna gratefully affirms it, and thanks Caleb, telling him how important it is to have an outside observer's take, as this kind of fresh perspective often affords a clearer view of the situation. I blink at this, equal parts delighted that I no longer register as an outsider, and guilty that I haven't already shared feedback along these lines; I think of the several pages in my notebook flagging "gender divide" in work roles at The Startup—how it often feels like Jenna "runs" things, but in a way that seems to be taken for granted by her counterparts. And yet, I can't help but notice how Caleb's status as not only a relative outsider but specifically a guy seems to give more heft to his observation; Noa had already said more or less the same thing a moment earlier, just a fraction less pointedly. Ironically, Caleb's important call-out of this problem appears to inadvertently reinscribe some of the problem he annotated.

Jenna concludes the workshop, thanking everyone. The fact that both the founder and the head of sales had left before the end still lingers uncomfortably, but the frank conversation their absence sparked feels like a promising turning point. Everyone stands and stretches, breaking off into small pockets of conversations, office party-like. This is the scene to which the founder and Andrew return; both are surprised that the workshop had now ended; Jenna gestures to the schedule posted on the wall, which she had gone over both at the start of the day and again after lunch. "What did we miss?" The question cuts through the din; chatting with Matt (a health team member who worked as an inpatient caregiver before joining The Startup), I overhear it and cringe. Matt pauses mid-sentence as we both unmistakably hear Jenna reply, "A lot, actually. We had a really important conversation, and I wish you'd been here for it." *Oh shit*, Matt and I telegraph into one another's eyes, our expressions hurriedly schooled into neutrality as Jenna continues: "There were things that were said that I think are really important for you to hear, and I just…" She breaks off, tearing up. The background buzz of conversations in the room completely stills. The founder gently asks if everyone would be willing to stay just a little longer. After everyone returns to their seats, the founder then apologizes, and attempts to apologize for Andrew as well—a subtle deference that I attribute to the fact that Andrew is conspicuously older than the others: fifty-something to their twenty-something (with the exception of Ian, the head of engineering). The founder contritely explains that Andrew had flown into town today to join the meeting, despite it being his wife's birthday, and that the call had come up unexpectedly. But when Andrew speaks, he drains whatever grace the founder has managed to briefly accrue:

> ANDREW: I didn't mean to *destroy* the meeting or the vibe, or people first or anything... I was just trying to make a win for us with this client, not to put money or business over the culture or anything like that. I apologize for destroying your meeting…

During Andrew's begrudging speech, I find myself staring awkwardly at my "Do what you love"-emblazoned water glass. Sweeping a glance around the table, I see uncharacteristically stony stares on the faces of Noa, Sofía, and Nisha (the youngest team member and a fellow intern). Matt and Anaya (a software engineer—the only engineer on staff when I first joined The Startup, and now working under Ian), look taken aback. Most, in other words, appear to find Andrew's apology as jarring as I did. And indeed, several people at the table proceed to take turns responding to what Andrew said, politely registering their dissatisfaction; but it's Ananya who frames the issue in terms of "culture":

> ANANYA: [Jenna is] doing a lot of things that she's not supposed to do—it's not her job to do it, but instead she's doing it as part of the culture, to make it better, so that should be respected.

Jenna thanks Ananya, and then recounts Caleb's earlier remark for the founder's benefit. She emphasizes the fact that it came from Caleb, seemingly to make the point about gender without having to be the one to make the point about gender. The founder pledges to hire a dedicated HR staff member, surmising aloud that this sounds like the step most likely to correct a majority of the problems that everyone is sharing—that this will take some of the burden off Jenna, and ensure smoother onboarding for future hires. And just as things seemed to be wrapping up once again, all falling mostly back into place, Nisha raises her hand.

> JENNA: Nisha, did you have something you wanted to add?
>
> NISHA: Uh, so, is this a good time to ask about pay, and how to get a pay raise? I've been wanting to ask, but then all that happened, so I wasn't sure…

It's completely the wrong moment; you can almost hear everyone's internal monologue as one: *"Noooooooo…"* Everyone is depleted. Aloud, Jenna simply affirms, "OK, let's talk about that later." However subversive the culture workshop in its design and unfolding had been, we've evidently reached the limits of allowable subversiveness.

And… a sense of resolution follows, sweeping up most of us. Together, we giddily pile into a large elevator, barely managing to fit and laughing at ourselves, and then ride to the top floor to take a group photo at a scenic spot. All the while, we're awash in a sense of camaraderie and accomplishment—a kind of post-group therapy high. But Nisha's question lingers; an "affect alien" (Ahmed 2010), she didn't read the room: *this isn't the right time. Let's end it on a high note.* And yet, the possibility of the "right time" is slippery… it's something both elusive and unequally distributed.

As we return to our floor, Nisha, who has the camera, asks if I want her to take a headshot photo of me for the website. We find a bright spot and snap a few pics; minutes later, we collapse into plush, blue chairs near the room where the workshop had taken place. Sofía joins us, dropping into a chair and exhaling—we're all still decompressing from the intensity of the workshop.

Lightly, not wanting to push: I ask Nisha about her thwarted question—is she ok with waiting to discuss it later, and able to get by meanwhile? Today *feels* like a resolution, and it feels like we're supposed to feel like it *was* a resolution; but was it? Nisha confesses that it bothered her: "Andrew was such an asshole, and we spent so much time dealing with *that* that we ran out of time. I really wanted to talk about pay—it's part of why I came today," she explains, referencing her half-day commute. She loves The Startup, and wants to be part of something that helps people, but it's been difficult trying to make things work financially. Nisha pauses, and then begins recounting doing content moderation work for Facebook over the past several months, all in the name of trying to make things work: "I saw so many bad things—like, I can't even say some of the things I saw….it was *so bad*…" There's a foreboding weight to these words; a wide, panicked look in her eyes. As she speaks, she begins to cry—no doubt releasing some of the pressure built up from this tumultuous day; but it's also clear that she's recalling something unspeakably awful. Sofía and I lean forward to comfort her. It's confusing for me to follow—how exactly this past work and her internship at The Startup are linked—but I've read a recent exposé specifically about US-based, third-party Facebook content moderators, and can make out the faintest outline of some of the spectres drifting before Nisha: images and videos of animal cruelty, child abuse, mutilation, sexual assault.[248] Unspeakably awful things.

Sofía glances around, and stands, guiding us both over towards a more out-of-the-way upholstered bench around the corner before continuing. Once we're seated, she begins offering advice in a hushed tone:

> SOFÍA: Mental health is shit. It pays shit. Even in tech.

Sofía is ostensibly the biggest optimist on the health team (a group of conspicuously optimistic, kind-hearted people); she's someone who sincerely loves Burning Man, and, dare I say it, makes you want to love it, too. Yet in this moment I saw an unflinching realism laid bare: Sofía explains how she started out at The Startup by working entirely unpaid at first, to prove herself (despite having obtained an advanced psychology degree from another country) worthy of a work visa—because that's what it took.[249] As she gestures wryly at her lower torso, Sofía does her best to console and advise Nisha. Both women were trying to figure out how to survive—and how to maybe even come out in some way on top. Sofía, I realize, fully understood her tech sector gambit, and also saw its limits: she might earn more here than she would elsewhere as a mental health worker, but she will invariably be earning less than most other tech workers. Per Sofía, when you want to help people, that's that. "That's what it takes. It sucks, but if you want to work in mental health, you just have to do it."

---

[248] Newton, "The Trauma Floor: The secret lives of Facebook moderators in America," 2019: https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona; note that Roberts (2019) differentiates this form of content moderation as "commercial content moderation" or "CCM"—see also: Roberts, "Social Media's Silent Filter," 2017: https://www.theatlantic.com/technology/archive/2017/03/commercial-content-moderation/518796/.

[249] Despite Sofía's profession, as a psychologist being comparatively high in both pay and status when compared with the un- and underpaid domestic care work that sociologist Evelyn Nagano Glenn (2012) considers in *Forced to Care*, parallels to Sofía's situation are nonetheless discernable, particularly in terms of coerced labor. See also: Parreñas 2015.

Our conversation is intense—as intense as anything else that day, and possibly more so, because it's both unguarded and rapid, as if we're rushing to fit in all the prohibitive truth we can before we run out of time. We talk this way for half an hour or more, about the world, about the future, about how to survive the gap between the way things should be and the way they are.

It comes to an abrupt end when Jenna walks over and interjects brightly, "What are you three talking about over here?" Nothing in particular, we say, just random things. She points out the time—still another hour or two to get through some of the tasks we couldn't get to earlier because of the workshop. We all stand and head back over to the hotdesk area. In WeWork, many things are open—especially on this unusual day full of unusual exchanges; but for Nisha, Sofía and myself, our collective sense seems to be: not everything. Sofía considers Jenna a close friend—and made no secret of idolizing her as a role model. Nisha clearly admires Jenna too, in a way that likewise borders on awe. So it's not that we don't trust her. It's simply that this moment between us feels like a remnant—a disorderly, out-of-place detail. Things are *supposed* to be fixed for now, enough for now; and yet we've persisted in talking about the unfixed, and perhaps unfixable.

IV) The ghost of ghost work past

In the days following the culture workshop, things at The Startup more or less settle back into the usual routine—with the "usual" including a steady stream of new hires. Eric, a psychologist specializing in chronic pain management, joined The Startup as a part-time member of the health team. Not long before this, we'd decided to try starting out our two-hour weekly team synch meetings with a five-minute guided meditation, using either the app Headspace (on Sofía's phone) or Calm (on my phone); Eric, who professionally records guided meditations, offered to lead one, and we soon abandoned our apps. When asked, Matt (who'd previously confessed to being a non-meditator) endorsed this use of time—probably in part because it meant we all started the meeting in a slightly less frantic state.

Nisha sometimes came to the office for these meetings, but more often she joined via Zoom. As we took turns reporting on various tasks and discussing any pitfalls we'd encountered along the way in completing them, I noticed that Nisha would regularly champion UX (user experience) research as a possible aid to some of those pitfalls. This intrigued me, because UX was an area that seemed largely absent at The Startup; the company's leaders regarded it as a nicety that could be put off until a later stage of growth rather than as an essential. End-users primarily "speak" to The Startup via chatbot-mediated questionnaire responses.[250] These consist mostly of user satisfaction survey questions ("How did our discussion leave you feeling on a scale of 1-4 with 1 represented by 😡 [frowning emoji] and 4 by 😃 [smiling emoji]?”); before-and-after symptom-focused questions (based on established psychology inventories[251]) administered a predetermined number of weeks apart; and that perineal favorite of entrepreneurs everywhere known as an ”NPS score” (net promoter score[252]), which purports to measure customer loyalty by asking “How likely is it that you would recommend [CHATBOT NAME] to a friend or colleague?” The Startup’s founder had an unwavering fondness for NPS scores, which the health team members had over time come to accept as an unavoidable quirk; the founder didn’t appear to understand why, in the context of a mental health resource for people who might be too stigmatized to pursue conventional therapy, the insertion of questions about recommending the chatbot could come off poorly.[253] To give you an idea of how this tends to go:

---

[250] This highlights the limits of time and the underpinning fact of not enough" care that mobilizes AI care in the first place; but interesting, it also indicates that The Startup treats chatbot conversations as "real" conversations that can and should be taken seriously on their own terms—an orientation which likewise guides my own approach throughout this dissertation.

> CHATBOT: Hi, [USERNAME]. In order to learn about what people and employers would benefit from the support I deliver, I'd like to ask you 2 questions. Can you help me with that?
>
> USER: OK
>
> CHATBOT: First question, who's your employer?
>
> CHATBOT: Of course, It's important to know that I'd never share your personal details with anyone, especially not your employer.
>
> USER: Why
>
> CHATBOT: Do you feel your coworkers would benefit from having conversations with me?
>
> USER: No
>
> CHATBOT: Got it. Thanks for going through these questions with me.

When asked, Jenna affirmed: The Startup’s current strategy entailed focusing on highlighting the science behind the chatbot rather than emphasizing users’ subjective experiences, whether individual or collective; in other words, they wanted to let the randomized control trial (RCT) outcomes speak for themselves. Yet during The Startup’s pitches, potential clients and investors alike unfailingly inquired about the subjective experiences of users. As one executive put it: “How do people feel about talking to a robot?”

---

[251] The Patient Health Questionnaire 9 (PHQ-9) and Generalized Anxiety Disorder 7 (GAD-7) are two well-known, commonly used screening tools for assessing depression and anxiety respectively, with numbers indicating the number of questions each contains. Notably, neither is considered a valid diagnostic metric apart from interpretative oversight by a trained clinician.

[252] People commonly refer to “NPS score” even though the “S” in the initialization stands for “score.”

[253] The founder had long championed having the bot ask users to input other people’s numbers so that it could message them directly; he earnestly saw it as an efficient, straightforward step that would lead to more referrals, rather than as a sensitive, potentially off-putting ask (as most users appeared to find it).

When I learned that Nisha was planning to implement an “in house” version of usability testing (a subset of UX research) in the development of a new series of modules, I was eager to join in. The modules in question were part of an experimental format for The Startup—a set of two-week “programs” that focused on chronic pain, anxiety, and depression, respectively. Each program would feature daily, sequential, five-to-fifteen minute conversational “lessons” based on the techniques that Eric taught in his practice. Because the chatbot cannot—as it explicitly states to all new users—diagnose a user, the matching of users with programs is necessarily predicated on *symptoms*.[254] Inert diagnostic tools become symptom detection tools; and an algorithm directs users with symptoms of depression to interventions for depression, neatly sidestepping any question of AI encroachment onto (human) mental health professionals’ turf. Diagnosing is maintained as the reserve of a human expert only—a line that chatbot companies do not at present seek to cross—and yet “diagnosis” is nevertheless always implicit in the design of their content, and the algorithmic matching of users to content.

As soon as we began testing the new modules, Nisha, Eric, and I found ourselves absorbed for hours at a time in a Zoom-mediated “flow state”: first, Nisha would “ping” our phones with a module; then we’d each quietly text with the bot until we completed the module; and afterwards, we’d discuss our experiences and propose modifications and corrections before proceeding to the next module.[255] The process—alternating between ruminative chatbot conversation and post-chat contemplative discussion—was engrossing; in moments it felt jarring for us to realize we weren’t actually all in the same room.

All the while, as I witnessed Nisha’s dedication to her work, her revelation from the culture workshop continued to haunt (Good, Chiovenda, and Rahimi 2022; Klima 2019; Derrida 1994) me. I kept thinking back to that tumultuous day: sitting alongside Sofía (and glancing self-consciously at her, loath to underscore my psychology “layperson” status by appearing to throw around DSM diagnoses with dubious credibility[256]), I’d suggested to Nisha that it might be possible she was experiencing a form of PTSD. I’d paused, then added: digitally-mediated events counted as a possible form of traumatic event.[257] Did she maybe want to try talking about this possibility with a counselor at her university? Did that seem like something that might help, and that would be easy to set up—maybe even sometime in the next week or so? Nisha explained: because she still had two months until her graduate program started, she didn't have access to campus health services yet. The irony was blatant; Nisha, someone working on providing mental health care for people who may not otherwise have a way to access it, was someone who needed mental health care but couldn't access it. And in the face of that irony, echoing the chatbot, I was trying to steer her towards diagnosis-informed support without ever transgressing my inability to diagnose anyone.

---

[254] Notably, chronic pain is itself a symptom rather than a clear-cut diagnostic category. Having learned in my preliminary fieldwork at a suicide prevention hotline (see Chapter 4) that many callers experience chronic pain, I was worried that this program, perhaps even more than the depression and anxiety programs, might overpromise a “fix” to vulnerable users. And yet, I could also see how the program’s unremitting recognition, from the opening dialogue onwards, that “chronic pain is real and not just in your head” might prove helpful even if the program’s techniques fell short of alleviating their pain.

[255] New content at The Startup was usually reviewed by at least three psychologists, but they didn't go through it in “real time,” simulating how a user would experience it.

[256] See: Caron, “Teens Turn to TikTok in Search of a Mental Health Diagnosis,” 2022: https://www.nytimes.com/2022/10/29/well/mind/tiktok-mental-illness-diagnosis.html.

Over time, I pieced together more about her experience as a third-party-contracted Facebook content moderator, and came to understand what I had not on the day of the culture workshop. Nisha had previously interned at The Startup, a year before I'd joined, but they'd had to let her go at the end of the initial internship window due to budget; they needed to be as "lean" as possible heading into their first phase of VC (venture capital) funding. At the time, Jenna had assured Nisha that, "We'd love to bring you back in the near future!" But Nisha had been devastated all the same; she'd loved everything about the job—it had filled her with nothing less than a sense of purpose. She elaborated on her exit: "It was a little like being dumped, and then being told, 'It's not you, it's me.'" Hoping to use the time to make herself into a more desirable future rehire (to "win back" The Startup) while still paying off her student loans and other bills, Nisha had sought part-time work that, like The Startup, applied technology to care—but that would also, ideally, provide a slightly-above-minimum-wage income. As she saw it, social media content moderation checked all of these boxes; and when laid out that way, Nisha's reasoning makes sense: content moderation, as "internet custodianship" (Gillespie 2018) is indeed recognizable as a form of care work that entails closely collaborating with AI. She envisioned moderating user content as being something along the lines of a behind-the-scenes, computerized version of the peer counseling she'd done as an undergrad student. But Nisha soon found herself stepping in for Facebook's AI after it flagged violent, terms-of-use-violating (and outright illicit) photos and videos for her to examine: is this blood, Nisha? Is this a child? Much as the provocative—yet utterly unvarnished—exposé published in *The Verge* (2019, noted above) begins, "The panic attacks started after Chloe watched a man die." The devastating reality of content moderation work seemed worlds apart from the care work Nisha had imagined; this, she came to realize, was probably why it paid slightly above minimum wage.

Like a one-sided rendition of "The Gift of the Magi," Nisha had bestowed a futile gift: she'd turned to content moderation believing it would help her to return to her self-professed dream job at The Startup. And The Startup *did* invite her back… but Nisha's content moderation experience didn't factor into that invitation; seemingly, all that she'd gone through had ultimately been for nothing. Still, this never managed to dampen Nisha's lofty imaginings of all that AI care might become. When I'd first interviewed her, what stood out most was her fast speech and far-off gaze as she described using technology to create an unprecedented, global awareness of mental health: "Just imagine: a giant, real-time map, and being able to know *instantly* how billions of people are feeling, what we're struggling with." She'd made her way over to a whiteboard as we spoke, and, laughing at her own zeal, began to sketch out inky, dry-erasable traces of her ideas in accompaniment to this narration.

---

[257] On the unfolding of PTSD as a diagnostic category and the growing recognition, following 9/11, of digitally-mediated events as a possible form of traumatic event, see: Holman et al, "Media's role in broadcasting acute stress following the Boston Marathon bombings," 2014; Young, "Posttraumatic Stress Disorder of the Virtual Kind," 2008. On legal action against Facebook alleging the debilitating effects of PTSD in former content moderators, see: Newton, "Facebook will pay $52 million in settlement with moderators who developed PTSD on the job," 2020: https://www.theverge.com/2020/5/12/21255870/facebook-content-moderator-settlement-scola-ptsd-mental-health; Vallance, "Facebook's parent firm Meta can be sued by ex-moderator, judge rules," 2023: https://www.bbc.com/news/technology-64541944.

Yet traces of the trauma she'd experienced inevitably lingered. When an ex, a law school student, asked to interview her about it, she shared her reservations with me: she was glad to talk about it if talking about it would help people and lead to changes that would ensure no one else has to do that kind of work… but talking about it was *hard*. We spoke again right after the interview, and sure enough, though she didn't directly name it as such, Nisha conveyed the feeling that maybe, just maybe, her ex was being a little opportunistic—he'd kept pressing for details, she noted, and grew a little frustrated when she wasn't able to remember certain things clearly. I'd always meant to formally interview Nisha about her content moderation work, but I never ended up doing so. It didn't seem possible to do so without retraumatizing her (which would all but certainly be the case for most who'd experienced what she had); but even if it had been possible, it felt *unnecessary*. Trauma is not necessarily experienced as or affixed to a linear narrative (Lester 2013)—it has a way of just being *there.* She'd already told me much of it in bits and pieces, over time… though there were also things that surfaced most clearly in being untold: we didn't have to talk about it to talk *about* it.

After completing UX testing for the new modules, Nisha and I discovered in one another a mutual wish to nudge The Startup towards incorporating user interviews[258] as a regular feature of their otherwise RCT-dominated research.[259] With charming relentlessness, Nisha proposed a series of phone interviews to Jenna, and was elated when she agreed—not only because UX interviews struck Nisha-the-UX-design-student as a pivotal source of insight for The Startup, but also because she figured that she would now possess an impressive roster of projects with which to finally claim the hourly wage increase she'd been pursuing for months. Not long after the interview project was underway, a few small setbacks materialized: we learned that the role that the founder envisioned for these was one of an extended version of an NPS survey; we were going to have to find a way to ask the dreaded "referral question." Jenna also stipulated that, on top of getting approval from her and the founder on the interview script, we needed to secure approval from Caleb, the part-time consultant who made the observation about gender at the culture workshop. Jenna explained that she would only be comfortable proceeding if Caleb, whose background was in coding—and who had no training or experience in qualitative research—evaluated our approach as the ranking "expert." As someone providing free labor in what I understood to be my area of expertise, I faltered a bit at the realization that Jenna saw Caleb as so obviously outmatching Nisha's and my combined backgrounds in qualitative research; Nisha, however, took it in stride. In a text message the following day, she wryly pinpointed what I'd struggled to put into words: Jenna, she mused, seemed to be meting out the same double standards she'd recently called out at the culture workshop.

---

[258] Without Nisha, the user interviews in Chapter 2 would not have been possible.

[259] On the history of randomized control trials in medical research, see Marks 1997.

In this same timeframe, Nisha confided in me about the pressure she was feeling from her family: "Oh, my parents kinda hate The Startup—I mean, not really, but they see me working all the time, and they know how much I'm being paid… they think I need to look for a different job. But they also know how much I love it." After the interviews were approved, Nisha again approached Jenna about a possible raise; Jenna asked her to wait just a little longer, noting that the timing was particularly bad, and that things would be very different in a few more months. But Nisha couldn't continue to wait. Reluctantly, she decided to try applying elsewhere, thinking maybe this would give her leverage to negotiate something sooner. As a part-time worker who didn't receive health benefits, she reasoned, a few dollars an hour more *should* be possible, especially in light of the new hires joining the company. The other jobs for which she applied (also part-time) involved elaborate and high pressure interviews—not only multiple rounds and sessions, but presentations and technical skills tests carried out under observation. Notably, while these competitive positions entailed working with AI chatbots (for companies you've heard of), none were in the health sector. One month later, Nisha received a fantastic offer from one of these companies, with an hourly wage far exceeding The Startup's pay. She decided to put it all on the line in an email to the founder and Jenna: Nisha, emboldened by the other offer, requested a significantly higher wage than the small hourly increase she'd initially hoped for. Her gambit failed. They approved a modest increase (of $1.50 an hour more for the meantime, with more to be discussed in the coming months); but this still less than what she'd initially hoped to ask for several months ago at the culture workshop. It wasn't, in other words, enough for her to be able to stay, especially not in light of the other offer.[260] Nisha was excited to have passed the gauntlet of interviews necessary to win this other job offer, and looking forward to the work itself. Nevertheless, she was, she explained, caught off guard by a feeling of sadness. Nisha, who delighted in working on developing content for the chatbot—to the extent that she not infrequently became so caught up in the work that she skipped classes—liked and admired her teammates, and felt energized by The Startup's expansive mission to help people in the scale of billions. She didn't want to have to forgo her role in carrying out this mission. But she also felt, at 23, already depleted by the relentless expectation of sacrifice. She scheduled a meeting with Jenna to put in her two weeks' notice.

---

[260] It's worth noting that this outcome isn't necessarily a straightforward indication of greed on the part of The Startup; though millions of dollars were already circulating in this industry at the time, The Startup's share of this was, insofar as I could piece together, somewhat paltry by comparison. The leaders were hesitant to raise anyone's income without raising everyone's; moreover, like a boxer doing a weigh-in before a match, they believed it perilous to increase expenses before clearing the completion of their funding round. And yet, things had already been decided such that the burden of "making do" seemed to fall most heavily upon the health team members.

Nearly half a year following the culture workshop, Nisha commuted to The Startup's headquarters for the last time to present the user interview findings in person. I was in Tokyo doing research at the time, but she relayed what happened in a phone call the following day. The presentation, Nisha confirmed, was well-received—though she found it somewhat frustrating that Jenna, who scheduled it as a lunchtime meeting, had billed to the other teams as a source of *inspiration* rather than one of critical insights. On Nisha's final day in office, it still seemed that no one was really valuing or even quite comprehending the role of UX research. Nisha felt that day much as she had felt all along: like no one was really *listening*:

> NISHA: Our entire mode of intervention takes the form of a conversation, you know? And yet no one at [The Startup] is fully acknowledging the importance that a conversation might have—that it might in turn be a mode of intervention for *us*, for what we do and how we do it.

Rather than linger further on the presentation, however, Nisha was far more interested in passing along the news: Andrew ("sorry-not-sorry for 'destroying' your meeting") had been let go—due, in the opaque words of the founder, to "integrity issues." She was stunned yet encouraged by this turn of events—she felt it was absolutely the right decision, but she nevertheless hadn't expected it (nor, frankly, had I); as Nisha put it, "There's some juice behind this, but I don't know what it is!" She elaborated: "It was hard for me to follow everything, but something about him not behaving the same way towards others when [the founder] was in the room?"

> NISHA: I think Ian [head of engineering] was the one to really push kicking him out, because they were trying to see if they could keep Andrew on. I'm not surprised Ian called him out, because Ian calls everyone out on everything—as he should!—and he was just like, '[Andrew] doesn't match our company culture, and I'm really glad we did the culture workshop so that we can map everyone's behavior back to our values'—and yeah, integrity was probably our biggest [value]. After the culture workshop, I even told Jenna, 'Andrew kinda reminds me of the [supervisor] dude from Facebook that basically, um, told us that we're not worth anything—just some privileged white asshole,' which is really ironic with the result now, because she was defending him at the time…

In other words, seeing this wrong righted felt bittersweet to Nisha; the possibility of helping to create a new form of mental health care, bound up with the possibility of not letting the "assholes" get away with being assholes, were things that she, as a worker, still couldn't afford.

V) Everyone's displaceable

I once witnessed something at The Startup that seemed in every way like the perfect encapsulation of displacement. It began one afternoon in a small, transparent-walled conference room where the health team members (the five of us who were in the office that day: Jenna, Sofía, Matt, Noa, and me) gathered to observe a real-time demo of the chatbot for a client on the east coast.[261] At precisely 1pm PT, 4pm EST, the prospective client would receive a text invite to the demo chat; and at 1:15pm PT, Matt would call her to ask how it went. During that fifteen-minute window, we would all monitor the demo chat itself, projected from Jenna's laptop onto the conference room screen. There was one more thing, Jenna noted as we shuffled into the room after lunch, a tinge of chagrin momentarily overriding her upbeat tone: we were also there to help in case anything in the demo stopped working—at which point, she explained we might have to do a brief takeover to get it back on course. The demo version of the chatbot had been having some issues lately, and had even stopped working during a recent demo. Jenna reassured us that she didn't expect for that to happen, because she'd already spoken with Ian (head of engineering—Jenna's counterpart) about the problem, and he'd given us the all-clear: the demo was ready to go. Jenna had invited Ian to this meeting, but he'd declined it earlier that day. Jenna, rather than cancel the entire meeting, decided we would still monitor the demo together, just in case.

I didn't quite understand *why* The Startup bothered with a specialized demo version of the bot; if the demo bot was buggy, then why not just use the free-to-access version that already existed, or provide temporary access to one of the paid client versions? Jenna clarified that using the demo version ensured a contained, schedule-able, under-fifteen-minute increment of conversation with a clear beginning and end—a complete and satisfying arc. Ordinarily, the chatbot will pick up on key words from the user, especially anything indicating an emotion or difficult situation, and will veer off into new conversations to better "match" what the user appears to be feeling—this flexibility is usually an asset, but it can make for an unpredictable experience. In the demo version, the chatbot was set up to just run through a very limited number of responses, and would reliably "stay the course" of the conversation it starts out with. If the prospective client was still interested after this demo, then their company would receive access to the standard version of the chatbot, and any of their employees who wanted would then be able to independently test it out over several days to get a better feel for it.

We'd never observed a demo as a team before; something about the precise timing and the "control room" setup made it all feel much more exciting and important than the weekly team meeting from which we'd just come. "NASA flight launch," I scribbled down in my notebook.

Right at 1pm, the chat begins. A small, collective sigh of relief follows this first feat, but we still have a ways to go. After the initial self-introductory disclaimers, the chatbot beckons: "Think of a relationship that's important to you. For the purposes of this thought exercise, please choose someone other than a child or pet." A moment of pause. The client's answer appears: "My mother." Sofía, by far the most psychodynamically-inclined health team member as well as the author of this particular script, nods sagely at this; "It's the most common answer to this question," she affirms. I long to better understand all that Sofía was inferring from this, but there's no time; the chatbot has already moved on: "What emotions come to mind when you think about this relationship?"

---

[261] The client is a representative from an employee wellness program (EAP) company (see Chapter 4).

We all lean forward, awaiting the reply. However, there's now a much longer pause from the client. What does it mean? Noa worriedly speculates that the word "relationship" might be a sticking point: "I think," she says as we all stare at the screen, "she might think it's saying 'relationship'—like, a romantic relationship, and not just any kind of relationship." Sure enough, a response finally appears on the conference room screen: "I'm not in a relationship with my mother." Ouch. During a product demo, the customer is always right—and for this customer, a relationship is something one does not have with their mother. We brace ourselves: how will the chatbot respond? Based on the updates that Jenna had requested from the engineering team, the bot should now be able to weather moments like this—user responses that weren't part of the anticipated question-and-answer flow.

We watch with growing panic as the bot first repeats the line preceding the prospective client's response ("What emotions come to mind when you think about this relationship?") and then (apparent only on our end) appears to be preparing to reboot itself. Jenna, with the steady hands of someone who could be counted on to clip the correct wire to dismantle a bomb, seamlessly pauses the chatbot and begins to type, speaking aloud as she does so: "Sorry I... got confused… for a moment there… Shall we continue… our chat?"

For the next few minutes, we collectively become the chatbot.

It all happens swiftly: Jenna confers with us on what she types before hitting enter; we encourage her and offer suggestions. With just a few sentences, we manage to quickly redirect the conversation back on track with the demo program script. None of us know if we can "reactivate" the bot midstream without it rebooting the entire conversation, so Jenna just copies and pastes in the remainder of the lines that the bot was supposed to say. The conversation wraps up right on time, and Matt steps into the corner of the room to place the call. He learns what we already knew: the chatbot had briefly faltered and "misunderstood" the client, but the demo went well otherwise. There's a feeling of buoyancy in the room, the kind that, as cheesy as it feels to admit, truly does follow a moment of close, skillful teamwork in the face of challenge. I feel lifted by it, and by the delightful absurdity of having just successfully endeavored to *fail* the Turing test.

Jenna looks relieved, but tense. As we wait quietly during Matt's brief call, I glance at the projected screen and see that she is already messaging Ian to let him know that the demo still isn't working (and though she didn't say as much in the message: it didn't seem like any updates had actually been made). After the call ends, we congratulate one another on working together so well under pressure, and it's clear that Jenna feels proud of her team; but it's just as clear that she is deeply frustrated that we had to do a chat takeover in the first place. She explains that she requested the update several weeks ago, and is adamant: this—having to do a takeover during a demo—won't ever happen again.

While what I'd just witnessed had been confusingly uplifting, I was still taken aback by it: swapping in a human to cover for the demo glitch admittedly seemed dodgy, especially when the point of a demo was to inch this potential client closer to making a purchase. At the same time, it was deliberately presented to the client as a "demo" and not as the product itself. Other than apologizing for getting "confused," and briefly clarifying that "relationship" here meant any kind of close connection, we hadn't deviated from the bot's existing script, and the ad libs we'd introduced could easily be added to the module as part of the if-then sequence—meaning the health team hadn't exactly given an unrealistic impression of the chatbot's abilities. And while it admittedly made me somewhat paranoid to see an unannounced chat takeover,[262] my presence at this meeting (among other sensitive moments) indicated to me that I was seeing a fairly unfiltered backstage view of The Startup; across all of my off-the-record conversations and direct observations in the months that followed, I realized: no one is doing stealthy chat takeovers. As ever, the figure of scale undercuts the dream (or nightmare) of *human* attention.

Even if the moment was more ethical snafu than breach, the fact that we'd all stepped in to be the bot in order to keep the client from seeing it break seemed like a neon sign declaring: *therapeutic chatbots depend on displaced labor*. This moment prompted me to recognize that ghost work isn't necessarily restricted to bottom-rung workers. For all that Nisha aspired to be like Jenna one day, it was galling to realize that, despite being a senior member of The Startup's leadership, Jenna's role still required her to sometimes "assist" the AI—not altogether unlike the work Nisha had done as a content moderator.

This recognition in turn had me wondering: why was it that Jenna often seemed to be treated like a more junior worker at The Startup, rather than the senior company leader that she in truth was? Looking back at that day, what came to stand out to me the most was how degrading Jenna had clearly found having to do the demo chat takeover—how low-stakes and unnecessary a risk it had all been. Jenna's high-pressure adlibbing had seemingly entailed more work from a company leader than it would've taken for Ian to simply listen to Jenna's input and update the demo.

Much to my surprise—and yet somehow not at all to my surprise—Jenna ended up leaving the company less than a year after Nisha departed. Though I've now highlighted three exits from The Startup—Nisha's, Andrew's, and Jenna's—I want to clarify: departures were not at all common at The Startup; for a span of approximately two years, during which the number of employees more than tripled, these were the *only* departures. And with the exception of Nisha, due to her more obviously precarious status as a part-time student intern (and the upbeat news that she was leaving for a prestigious new position elsewhere), these departures reverberated sharply across the company. Disturbingly, Ian was at the center of Jenna's fraught decision to depart—and the particular circumstances were far worse than I would've ever imagined: I learned that Ian had accelerated from lightly unprofessional conduct towards Jenna (small frictions of “challenging her authority”) to full-on boundary-violating, outrageously unprofessional conduct.

---

[262] "Announced" chat takeovers, by comparison, are a feature that The Startup offers to customize for clients—companies with their own dedicated crisis hotline, for instance, can offer users the "in chat" option to switch to chatting with a human counselor. The Startup considers it imperative to clearly disclose any such transition to its chatbot end-users. For more on chatbots in the context of crisis response, see Chapter 4.

After multiple members of the health team had come to Jenna to disclose feeling intimidated and disrespected when working closely with Ian, she elected to meet with him about this and discuss the problem directly, in the hope that they could resolve it. She figured the underlying issue was just a lack of self-awareness on Ian’s part, and imagined that, upon hearing feedback in private about how his behavior had impacted junior colleagues, Ian would be eager to adjust and reset things. As Jenna explained:

> JENNA: He was quite blunt with some of his comments to the health team, and three of them are women and they approached me to say, ‘Hey, I don’t know how to proceed with Ian, because every time I ask him a question, it almost feels like I’m insulting him by asking, and I don’t know how to help him.’
>
> JENNA: It just was like a pattern to me of his behavior against women… I shared with him, you know, ‘Hey, I think I understand that you have a blunt way of doing things, and you wanna cut through the chit chat and get right to it, and I respect you for that. But, you know, some of the women on my team have come to me for advice on how to interact with you, so I’m coming to *you* because I want to know if there are [pause] tips and tricks my team can apply. And the other thing I want to share with you is that one of them used the term ‘workplace bullying,’ and I just want you to know that that’s how that interaction came off.
>
> VALERIE: Wow, yeah…
>
> JENNA: And Valerie… he blew up.
>
> VALERIE: Whoa… are you serious?
>
> JENNA: He was *screaming* at me.

It would prove to be an unresolvable situation; but Jenna, ever the skilled psychologist and veteran therapist, nevertheless attempted to resolve it: she messaged Ian, offering to meet later to discuss things more calmly; he left without responding. Later that same day, she set up a meeting for the following day, with the founder present to help mediate. But that meeting culminated in an even more debasing interaction; when Ian arrived, he began by declaring to the founder that he was quitting, insisting, “You have a bully in your company, and she’s *right fucking there*.” In a sort of surreal caricature of a therapy session, Ian began crying, and after insisting that Jenna didn’t understand what bullying was, related an experience of childhood bullying. When Jenna tried to validate that experience as something important and difficult, he told her to, “Shut up—I don’t want to hear from you.” Eventually, he explained to Jenna that, “I’d rather be called a rapist than a bully.”

Jenna fully clocked the intentionally gendered language of this remark and the interwoven implication of false accusation ("I just… would he have used the word 'racist' instead if I were a person of color? It seemed like he wanted it to unnerve me…"). It was all the more disorienting given that her intent had never been to declare Ian "a bully," but to highlight her team members' feelings in the hope of inspiring a course correction. She shakily described to me how it had taken her a few months to "process what had happened" that week—to realize that she couldn't keep working with Ian, or lead a team of mostly women who also had to work with him, while this behavior went unchecked. (What chance, I thought ruefully at the time, do the rest of us possibly have when even a trained psychologist—understandably!—needs a few months to process and acknowledge that she's been on the receiving end of egregious conduct?). As Jenna recounted this, I thought of what she *wasn't* saying: the previous summer, after witnessing someone get shot at close range mere feet away from her just outside our building, Jenna had disclosed that she'd been diagnosed with PTSD a few years earlier due to secondary trauma from her work as a therapist;[263] she had in fact been on her way to a therapy session when the shooting occurred. While being screamed at by a colleague would be disturbing under any circumstances, it sickened me to realize that Ian's conduct no doubt had a disproportionately more damaging effect.

As Jenna (and nearly everyone else) came to recognize in the aftermath of these explosive interactions, The Startup's leadership was at a crossroads, and the founder needed to choose: Jenna or Ian? If you'd asked me at any point during my fieldwork, which began before Ian joined The Startup, how that choice would go, I wouldn't have hesitated: Jenna, every time—not only because she was the one who was *not* engaging in unethical conduct towards her colleagues, but primarily because she was so obviously essential to what the Startup had achieved. But I'd have been wrong. It turns out that almost no one ever picks the indispensable woman who not-so-secretly keeps the place running over the avuncular, seasoned advisor who purports to know how to get you to your desired future. All the same, this outcome floored many at the company, simply because Jenna was, in the eyes of most of her colleagues, the heart of The Startup, to the point that her voice often *was* the encouraging, insightful voice of the chatbot. And like the chatbot, Jenna occupied the unlikely position of being at the center of everything while being a voice to which people did not necessarily listen, and did not *need* to listen.[264]

Why did Jenna and Nisha leave The Startup? It's a simple question, but nevertheless a difficult one to answer. Both women are dreamers who, during their time at The Startup, envisioned tremendous possibilities for AI as part of a radical approach to care without access limits—

---

[263] See: Reuben, "When PTSD Is Contagious," 2015: https://www.theatlantic.com/health/archive/2015/12/ptsd-secondary-trauma/420282/.

[264] Across many hours doing transcript review, a sizable portion of the randomly-selected conversations therein consisted of single, unanswered texts from the chatbot—in which the bot reached out to a user who (at least at that day and time) didn't respond. At a certain moment in my fieldwork, the comparison between this and Jenna's treatment by upper management struck me, and led me to begin considering the relationship between the health team members and the chatbot as one of solidarity.

possibilities that each woman was eager to elaborate on in detail; each seemed to regard The Startup's mission as the first step towards realizing this broader vision, rather than as the entirety of it. And likewise, I was struck by how, for both Nisha and Jenna, leaving The Startup was like a difficult, complicated breakup: "I don't understand what happened. This was my dream job," Jenna told me after finishing work on her final day at The Startup. Nisha had also used the same words to describe working at The Startup to me: "my dream job." Each perceived her own departure as indicative of systemic inequities; yet in our conversations mulling over and processing what had happened, it became clear that both Nisha and Jenna also saw their respective exits as a matter of personal failing—for having endured that systemic bias for any length of time, for having failed to change it. It's an obscenely unfair thing to behold, let alone go through; and yet I imagine it's recognizable to many if not most who have endured a similar situation. As an ethnographer, I found myself watching the most devastating blow from "do what you love" play out: to do otherwise indicates your *own* failure. Just as Weeks describes: "Besides masking inequality and obscuring the economic instrumentalities of love and work, the ideologies of love and happiness depoliticize love and happiness by individualizing their experience."[265] And yet somehow, it's one thing to say this, and another to witness it unfold: Nisha and Jenna each took on some of the failure of being failed, and kept it with them.

What kind of ending is this—what does it tell us? Doing what you love was Jenna and Nisha's workaround, their armor against exploitation while doing care work; yet their parallel journeys at The Startup suggest that this workaround became, in the end, something that ensured they would see their exploitation as an indicator of their own failure to do what they loved.[266] In answer to my central question in this chapter, what can AI displacement and care work each tell us about the other: **even "good" displacement is exploitative, insofar as it demands workers who are always at once indispensable and dismissable**. Just as Hu (2015) contends that old and new infrastructure are layered atop one another in the form of old railway tracks and contemporary fiber-optic cables, the infrastructure of displacement's growing "shadow workforce" (Gray and Suri 2019) appears to readily tap into the path of a pre-existing shadow workforce doing un- or under-credited work: women workers.

Notably, Nisha and Jenna's biggest problem as displaced workers wasn't the AI that displaced them—which isn't to suggest that they therefore weren't "really" displaced workers. Instead, their experiences point to how the problem of AI labor displacement exceeds AI automation. This same problem is discernible in the care work paradox—which conveys a demand for workers whose work is indispensable, yet who, as workers, are *not* indispensable. Marx theorized that, "The instrument of labour, when it takes the form of a machine, immediately becomes a competitor of the [worker themself]."[267] This isn't altogether different from the future Licklider glimpsed, in which a computer might become "a colleague whose competence supplements your own"; a colleague, after all, has the potential to be both a source of competition and a source of solidarity and shared experience.[268] As colleagues, Jenna, Nisha, and the chatbot alike all collaborated to do care work;[269] but ultimately, only the chatbot could easily endure the everyday expectation of being essential-yet-disposable.[270]

---

[265] Weeks, "Down with Love," 2017.

[266] Notably this shares a certain resonance with Berlant's *Cruel Optimism* (2011)—a point I return to in the conclusion to this chapter.

[267] Marx and Arthur, *Marx's Capital*, 2018 [1867], 262; modified from "workman himself."

Recognizing this outcome, I now want to return to a possibility I raised in Chapter 1: that AI might in turn be an ally in *addressing* labor exploitation, rather than just a *means* of exploitation.

VI) Conclusion - Part 2

In this chapter, I provided an account of AI labor displacement from a seemingly unlikely context: one in which displacement becomes something done *by* and not just *to* workers. The Startup's mental health team members are care workers who, in effect, pursued their own displacement by AI as an intentional strategy. Their entry into the tech sector was their attempt to evade the mismatch between the espoused importance of care work and their actual pay and status as care workers. But as this ethnographic account makes clear, becoming a tech worker who did care work did not exempt these workers from the reach of this paradox; Jenna and Nisha, both caught in it, ended up exiting The Startup when they couldn't find a way out of it. This chapter has built on existing literature indicating the ways in which humans-in-the-loop are exploited in the name of maintaining the illusion of AI's autonomy, by showing that even a "best case scenario" of displacement still relies on exploiting the workers who compensate for AI's technical gaps.

---

[268] Amazon warehouse worker Darren Westwood describes AI-automated robots as colleagues who receive preferential treatment, telling BBC Breakfast (television program) interviewers: "I wish we were treated like robots because the robots are treated better than us." (See: https://youtu.be/t80TK5n1u4I). Interestingly, the accompanying BBC article covering this interview originally ran with the headline, "Amazon strikes: Workers claim robots are treated better," but a few hours later, the headline at the same URL (https://www.bbc.com/news/business-64384287) was updated to: "Amazon strikes: Workers claim their toilet breaks are timed." Such adjustments are common with unfolding stories via the BBC, but my point in highlighting this is that this particular headline adjustment tells its own story: it's important to note that robots are treated better, and that workers themselves see things that way—but that fact should not eclipse the reason why (who is treating bots better and humans worse?). Robots do not (yet) block diabetic and other workers from taking necessary restroom breaks; their human managers do.

[269] The chatbot even became a buffer for Nisha and Jenna, helping them to continue doing care work while recovering from the trauma they experienced through previous work.

[270] Notably, anthropological scholarship offers several other "ways in" to thinking about this dynamic of disposable-yet-essential that I'm describing—including ethnographic research within the burgeoning interdisciplinary field of discard studies (Doherty 2021; Reno 2015, 2016; see also Lucas 2002); classic medical anthropology literature on the commodification of human organs (Cohen 1999, 2001; Scheper-Hughes 1996, 2000; Sharp 2000, 2006); a pandemic-inflected critical exploration of "essential labor" arising from the subdiscipline of the anthropology of work (Acevedo 2020; Alcorn 2020; Brown and Pearson 2020; Shepherd-Powell 2020; Timcke and Gomes 2020; all of these are designated as a special collection on essential labor by the Society for the Anthropology of Work edited by Fisher (2020): https://saw.americananthro.org/essential-labor-collection); and the anthropological theorization of abandonment (Biehl 2005; Povinelli 2011) and precarity (Al-Mohammad 2012; Allison 2013, 2016; Han 2018; Stewart 2012). Works addressing precarity and abandonment that are not anthropology per se but that are nevertheless influential within the discipline include: Berlant 2011; Butler 2004, 2016; Mbembe 2003, 2019; and Puar 2017. Each of these (along with many other) anthropological branches entails asking, amidst very high stakes: is there a way out from under capitalist commodification?

But the preceding analysis points to something further: this exploitation, summed up as the paradox of care work, can tell us something about the phenomenon of AI job displacement. This paradox—which names the stark discrepancy between how care is valued and how care work is compensated—is an antecedent to AI job displacement. Insofar as this ongoing paradox has perpetuated—and even "naturalized"—the fact of essential-yet-disposable workers, then its existence has arguably created an opening for the nonhuman workers (AIs) that better fulfill this attribute. **The very existence of the care work paradox, in other words, provides a ready-made template that AI automation emulates—and in turn intensifies**.

For care workers, this paradox is not so much an enigmatic contradiction as it is a straightforward job qualification. It indicates the sheer ordinariness of jobs that regularly demand impossible-to-fulfill attributes from workers. It also demonstrates how a supposed mismatch between different notions of value in compensation conveniently serves one particular notion of value. This paradox alone does not account for the phenomenon of AI displacement; rather, it's an unacknowledged piece that helps to shape the very logic of displacement. AI workers make perfect sense in a world where "work" already demands disposable-yet-essential workers. And slotting displaced human workers into ever-more exploitative roles—roles that serve to augment and complement the abilities of AI (rather than the other way around, as Licklider envisioned)—likewise makes sense in a system where workers' paradoxical necessity and expendability is already a matter-of-fact presumption. As anthropologist Lucy Suchman noted in an interview: "Robots work best when the world has been arranged in the way that they need it to be arranged in order to be effective."[271] What I'm suggesting is that the care work paradox indicates one of the ways in which the world, broadly speaking, has already long since been arranged. In short, **AI labor (in general, not just care) flourishes in its fulfillment of the condition underwriting the care work paradox: its ability to be better at being inhuman than humans**.

But the story doesn't end with the chatbot outlasting Jenna and Nisha in doing care work; rather, it unfolds in following the chatbot's abidance to unearth a critical "why." In other words, even though AI itself did not appear to be the principal problem with which they contended, it doesn't follow that AI displacement is irrelevant to their struggles; rather, I'm claiming that the very possibility of displacement-by-AI is something that only opens up only in a context where workers are *already* understood to be paradoxically disposable-yet-essential.

Considering the chatbot as a colleague that is subject to this same expectation alongside its human colleagues helps to make the consequences of this more clear. I'm not proposing any kind of "interiority" to AI itself; on the contrary, recognizing AI caregivers as a colleague to their human counterparts allows a key difference that is a ramification of AI's *non-interiority* to surface: **whereas human care workers are perpetually subject to the insistence that their**

---

[271] See: https://www.bloomberg.com/news/videos/2016-07-07/artificial-intelligence-charlie-rose; the statement above by Suchman continues: "So that's why assembly lines work so well, that's why Amazon's warehouse robots work so well—because the entire world that the robot is operating in is designed in relation to the robot."

work exceeds capitalism in a way that most other work does not, bots aren't.[272] This insistence is what propels the care work paradox. England's aforementioned "cruel irony" of care work—that the "genuine care that motivates some care workers […] may make it easier for employers to get away with paying care workers less" resonates with Berlant's "cruel optimism" (2011), in that the very designation of "genuine care" becomes the obstacle to care workers' flourishing. In other words, to be subject to the care work paradox is to be subject to a lack of clarity concerning whether and to what extent care work is "inside" vs "outside" of capitalism. This uncertainty inundates care work; and ironically, it extends the reach of capitalism.

Seemingly, to challenge the claim of an outside to capitalism would be to take up the stance that there is no outside to capitalism—a stance that would be narrowly imperceptive and ahistorical; to say the very least, it would be giving way to the unchecked totalizing with respect to capitalism that Tsing (2013) rightly calls out as disciplinarily hypocritical.[273] Allison (2013), for instance, implements ethnography as a means of providing examples of alternatives to capitalism sprouting amidst a backdrop of precarity.[274] And Tsing, in "Sorting Out Commodities" (2013), contends that such an outside not only exists but is wholly *necessary* to capitalism—that commodities both begin and end within this outside: "Translation between capitalist and non-capitalist forms of sociality is the work that makes capitalism a system."[275] Accordingly, in asking "to what extent is there an outside to capitalism," I'm not claiming there is *no* outside; instead, my objective is to think about how my informants encounter the limits to that outside—and how **being disposable-yet-necessary necessarily means coming up against the limits to that outside**.

My informants are wary because they already know that any "outside" to capitalism simultaneously becomes the justification for less compensation "within" a capitalist framework—which in turn sets in motion the paradox of care work and its logical consequence: disposable-yet-essential care workers. Building on Weeks' critique of the enmeshment of love and work, I've attempted to show how **some workers draw on this enmeshment to serve divergent aims**. When implemented "from above" by capitalists, the "romance of work" seemingly serves to expand the reach of capitalism; but when invoked by workers (from below), the romance of work alternatively can become an intentional means of refusing an illusory "outside" that only leads to their further exploitation.[276]

---

[272] The claim that (human) care workers stand apart from most other workers in doing work that is necessarily "something more" than capitalism (recall Folbre: "some people don't work for money alone") is, in effect, the corollary to the charge that AI *cannot* care (see Chapter 1).

[273] Tsing, "Sorting Out Commodities," 2013, 22: "The point of this contrast is to clarify social logics through abstraction. To join this discussion is to enter into the spirit of analytic play, asking what can be learned through the dichotomy. Actually existing relations of exchange are, of course, mixed and messy. Not only do self-described gifts and commodities nestle beside each other, but they also incorporate each other's characteristics, change into each other, or confuse different participants about their gift-versus-commodity identities."

[274] Consider Allison's "C-C-C" example (2013, 167); note that it's care work that becomes the site of this "outside."

[275] Tsing, "Sorting Out Commodities," 2013, 39. Note that while Tsing stresses how "[a]nthropologists contrast gifts and commodities as icons of different systems for making value," and uses this as a point of departure by which to both challenge and play with that contrast, she also effectively maintains that contrast. Strathern, in "An End and a Beginning for the Gift?" (2010) similarly questions the routine opposition of gift with commodity, but takes a different approach—she destabilizes and interrogates what a "gift" is. Strathern also shows how "the idea of gift-giving as an organizing feature in procurement practices" is something that is *conserved*; I've sought to show how arguably, this same idea is likewise shored up by those who procure labor.

How do care workers navigate this impossible situation? Again, for The Startup's health team members—those who left AI care altogether *and* those who remain—seeking their own displacement by AI as a strategy for outmaneuvering the care work paradox didn't work. But this unsuccessful attempt to do so nevertheless reveals something important: **care workers are recognizing and refusing this paradox**. This is significant in its own right; but given what I'm suggesting (that the fact of AI-displacable workers is rooted in preceding, ongoing forms of work demanding disposable-yet-essential workers), it's also something that matters beyond the scope of care workers alone. The question of how to refuse being disposable-yet-essential, in short, concerns us all.

The grounds of this refusal are subtly evident in the culture workshop report that Jenna created. It might, at first glance, seem otherwise; this jargon-suffused account is utterly emblematic of the "workplace romanticism" that Weeks devastatingly pinpoints, with an observation that easily doubles as an ethnographic portrait of Jenna: "since work is going to consume so much of our time and energy, we need to believe it is good work, and to believe it is good work, we need to love it. *If we cannot get out, it is best to go all in*."[277] But Jenna's rejection of the bait-and-switch of care work—as a path that purports to transcend capitalism yet unfailingly sustains it—is inexorably present in this report. One section, accompanied by a stock photo of an elephant inside a splashy-yet-corporately-muted conference room overlooking a cityscape; reads:

> *Where did that elephant in the room come from?*
>
> *There is a common theme in therapy where the patient realizes the session is coming to an end and finally works up the courage to reveal something deeply meaningful, leaving little time to process it. These revelations may be expressed through words, actions (conscious or unintentional), or in the form of giant elephants. However, they typically hold great value and should not be ignored. During the last ten minutes of our culture workshop, one of these revelations came to surface that gave our team a remarkable opportunity to demonstrate the full cycle of behavior -> enabler / blocker -> outcome.*
>
> *What happened?*
> - *Behavior: "eagerness" was demonstrated by pulling a team member out of the session during a conversation about what it means to be a people-first company.*

---

[276] What I'm cautioning against is like the television trope of a dream-within-a-dream, in which a character wakes up from one dream and believes themselves to be awake—only to suddenly wake up and realize that was *also* a dream: a reminder that what might at first seem to be outside of capitalism might in fact not be.

[277] Weeks, "Down with Love," 2017, emphasis added.

- *Enabler: a "leader's desire to engage thoughtfully" and blocker of "miscommunication" were demonstrated when a leader left to join the team member.*
- *Outcome: "team members feeling as if their work / time is not valued" demonstrated through verbal frustration around the absence, and hence the elephant charged in.*

What did we learn?

- *The team stayed together for another 15 minutes to discuss what happened and why.*
- *Staff expressed concerns in a safe environment and their energy was matched with respect and recognition from leadership.*
- *The behavior, enabler and blocker were explained through open communication allowing the team to collaboratively achieve a resolution.*
- *Next steps were outlined to further reduce staff concerns and ensure the resolution stuck.*
- *In the days after, team members continued to debrief and expressed feeling closer having gone through the experience and coming to a solution together, which reinforced the importance of transparent and open communication.*

In Jenna's hands, the crucible of the culture workshop is rendered into a happy story with a happy ending, complete with an elegantly tidy "moral of the story." Jenna crafted this report almost like a magic spell—it is an artifact of intention designed to ward off the very same bad communication, sense of being overlooked, and feeling of being undervalued that it purports to have already overcome. Banished herein is any trace that, "mental health [work] is shit." Alas, the spell didn't work; but in its failure, Jenna's determination to "do what you love" became her means of refusing to be disposable-yet-necessary.

Chapter 4: [USERNAME], Can You Tell Me if You Have a Suicide Plan?

Chapter 4 abstract: AI caregivers provide care in high-stakes situations such as suicidal ideation, even though the workers who create and supervise them (at my fieldsite and elsewhere) didn't intend this. This chapter shows that therapeutic chatbots don't simply redirect users to suicide prevention hotlines; instead, these bots, as nonhuman caregivers, appear to offer a unique form of suicide prevention care—**one that allows users to be more "human" in their expression of suicidal ideation**. Therapeutic chatbots both do and do not witness, which means they decouple listening to the disclosure of suicidal ideation from the obligation to report what a human user discloses to them. Score one for AI care? Yes and no. While the facilitation of suicide prevention care via chatbot might seem like merely an unintentional offshoot of therapeutic chatbot care, I argue that it's actually key to understanding: why AI caregivers, why now? This is because the uptake of suicide prevention care by chatbots is not incidental to the flow of money that drives this industry. One of this industry's largest clients—Employee Assistance Programs (EAPs), the third-party companies that most large workplaces contract to provide short-term counseling and other mental/behavioral health benefits to employees—view detection of suicidal ideation as a useful attribute. Although this circuit of consumption linking The Startup, EAPs, and workplaces might suggest that chatbots are becoming yet another means by which employers surveil their workers, I show that it may in fact signal a seemingly inverse aim for employers: to not *have* to witness. I argue that outsourcing the work of witnessing suicidal ideation to AI in the context of employee mental health care means that **the difference between achieving "less carcerality in suicide prevention care" and "less disruption to work productivity" becomes increasingly imperceptible**. In other words, something important and urgently needed—the expansion of opportunities and settings in which one can discuss suicidal ideation without fear of negative repercussions—could ultimately become contingent upon the (further) dissociation of workers' mental health struggles from the conditions of work.

I) Introduction: More-than-human crisis care

Suicide prevention hotline referrals have become increasingly common cultural artifacts: they populate the websites of social media platforms, are appended to news articles discussing suicide, and air before and after television programs that depict suicidal ideation, self-harm, and/

or death by suicide.[278] Enter "suicide" in a search engine, and a hotline referral will appear at the top of the page. If you go to sign up for a counseling session through your workplace or school mental health care services, then you'll likely find a crisis[279] hotline referral listed there as an always-available contingency: 'if you need immediate support' or 'if this is an emergency.'[280] However marginal these referrals might at first seem, they are part of the infrastructure of mental health care.[281]

Crisis hotlines stand out from other forms of mental health care because they are instantly available (often accessible 24/7), free to use, and offer callers (texters) the enticing possibility of seeking help anonymously—all of which are attributes that they share with therapeutic chatbots. Perhaps these similarities are not altogether surprising, given that crisis hotlines and psychology chatbots came into being in roughly the same timeframe[282]—both, in other words, emerged amidst a shared backdrop of concerns about the insufficiencies of existing mental health care.[283] However, drawing on fieldwork at both The Startup and a secondary fieldsite—a Silicon Valley-based suicide prevention hotline—I intend to show that these forms of on-demand care don't simply parallel one another—they're unexpectedly *intertwined*. This intertwinement in turn tells us something about why and how therapeutic chatbots have become a multimillion dollar industry. The Startup's incorporation of a suicide prevention hotline referral into its chatbot's dialogue is the site of intertwinement that this chapter will take up. I intend to show that when therapeutic chatbots (at The Startup and elsewhere) refer their users to suicide prevention hotlines, they do so amidst a broader context in which alternative sources to turn to for care when experiencing suicidal ideation are few, and are often unclearly differentiated from surveillance measures—and that consequently, the "referal" that therapeutic bots offer is not just a *conduit* to human crisis counselors, but a **unique form of nonhuman suicide prevention care**.

---

[278] By suicidal ideation, I mean thinking, contemplating, and/or imagining dying by suicide (or being dead due to suicide), rather than the phenomenon of suicide or attempted suicide. To clarify, I'm not suggesting that suicidal ideation isn't ever a precursor for suicide—but I oppose thinking of ideation as a failed realization of an act, as if suicide were teleological. Suicidal ideation (which is sometimes referred to as "parasuicide"; note too that some care professionals differentiate between "passive" and "active" suicidal ideation, with "active" indicating a specific plan to end one's life) is not necessarily a precursor to suicide, nor is all suicide understood to be preceded by suicidal ideation; this is something Fehling and Selby (2021) discuss in their analysis of the recent proposal for an official DSM category of "Suicide Behavior Disorder." Harmer et al (2022) provide an concise overview of the complications that ensue from the absence of a "universally accepted consistent definition" of suicidal ideation. Interestingly, suicidal ideation is *not* broadly acknowledged within disability studies as a form of disability, despite the fact that experiencing suicidal ideation does not necessarily map to a mental illness category. Baril (2020) challenges this omission, arguing that the ways "in which suicidal people experience multiple forms of injustice and violence" or "suicidism" (Baril's term) is "intertwined in forms of ableism/sanism."

[279] The invocation of "crisis" as a euphemism for suicidal ideation predates "crisis hotlines"; it appears at least as early as sociologist Ruth Shonle Cavan's *Suicide* (1928).

[280] Note that many suicide prevention hotlines around the world provide their services via texting as well as calling. The National Suicide Prevention Lifeline (988) is the number most commonly featured in referrals within the US, though the text-only, international Crisis Text Line is often featured in referrals directed towards teens/young adults. A referral might also list the number for a local, rather than national, hotline (many of which affiliate with and sometimes divert calls to/from Lifeline). Referrals might also direct callers to hotlines for specific groups, such as military veterans, Native Americans (see: Randhawa, "Washington launches first of its kind Native American crisis hotline," 2022: https://www.king5.com/article/news/community/facing-race/washington-launches-native-american-crisis-hotline/281-b5a4b774-2440-4b83-aeb9-6dfc746e8b70), and even grad students (see: https://gradresources.org/crisis/). Notably, there are also crisis hotlines that focus primarily on areas such as domestic and/or sexual violence (see: Backe 2018) and eating disorders.

[281] While I would not go so far as to say that the infrastructural role of hotline referrals is universal, it is certainly true beyond the US alone.

[282] They emerged just over a decade apart (with British Anglican Church priest Chad Verah launching the world's first suicide prevention hotline in 1953, and Weizenbaum creating ELIZA between 1964–1966). And even though Weizenbaum didn't intend for chatbots to seriously serve this purpose, American psychiatrist Colby did—and was writing about the therapeutic uptake of chatbots as early as 1964. See: Colby and Gilbert, "Programming a computer model of neurosis," 1964.

Each of The Startup's end-users first encounter a suicide prevention hotline referral in their introductory conversation with the bot. As described in Chapter 2, this initial dialogue entails the transmission of several disclaimers: the chatbot explains that it is not a human, that it cannot treat or diagnose, and that it is not capable of responding to an emergency. Accompanying this last declaration is contact information for both emergency services (911 or national equivalent) and a crisis hotline.[284] The chatbot instructs the user to immediately contact these resources if they are experiencing an emergency situation. The chatbot will also share a crisis hotline referral if it at any point determines—by using NLU (natural language understanding)[285] to detect words or phrases that The Startup's engineers and psychologists have trained it to recognize—that a user is communicating thoughts or plans of harming themself or others.

The presence of a crisis hotline referral within the chatbot's programmed dialogue indicates nothing more clearly than the fact that no one at The Startup intends for the chatbot to take on the role of crisis responder. Instead, they designed their chatbot to, as a matter of due diligence, mental health professionalism, and user safety (or perhaps most cynically: "CYA"—as in "cover your ass"), first, detect the discussion of self- or other-directed violence by a user, and then, refer that user to an instantly available, free-to-access human caregiver. This indicates a stark difference from other forms of AI care (such as diagnostic AI), in which the role of human experts becomes primarily one of *supervising* AI expertise.[286] It shows that The Startup is drawing a line—that they both perceive and enforce a clear divide between what a chatbot can and can't do. The Startup, in other words, **would be the first to affirm that a chatbot is not a sufficient source of care for someone who is currently feeling suicidal**.

---

[283] Verah (see the previous footnote), who called for "befriending" as a practice of care (and proposed that psychological care *alone* was insufficient to prevent suicide), wrote that, "[O]nly a minority of those who contemplate suicide consult psychiatrists. If all did, we certainly could not cope with them—and there will never be enough doctor's time available for all abnormal people in despair." (1966, 114); meanwhile, Colby wrote in the same timeframe: "Is it dehumanizing to use a machine such as a computer as an agent for mental change? No. […] It is dehumanizing to herd thousands of patients into mental hospitals where they will never see a doctor." (1967, 248, 253).

[284] The information the chatbot provides is, in most cases, local to the user's phone number area code or (for those not using phones) the country indicated when registering an account.

[285] NLU is essentially a subset of NLP (natural language processing) that describes using AI to achieve a nuanced understanding of human language by successfully differentiating between different viable meanings; see: Kavlakoglu, "NLP vs. NLU vs. NLG: the differences between three natural language processing concepts," 2020: https://www.ibm.com/blogs/watson/2020/11/nlp-vs-nlu-vs-nlg-the-differences-between-three-natural-language-processing-concepts/.

And yet, as this final chapter will establish, **therapeutic chatbots do indeed become caregivers to users experiencing suicidal ideation**. **I seek to think through the why and how of this phenomenon**. That AI takes on roles that its designers, programmers, and other implementers never intended is more or less the plot line of any AI-centered dystopian tale, and having to constantly remind people of this fact—that intention neither determines nor discharges the consequences of outcome—is easily the bane of AI ethics activists' existence. However, I intend to show that if there's an element of the dystopic in AI care, AI "going off script" isn't its source.

I will begin by situating AIs as crisis caregivers—and importantly, by demonstrating that their *failure* to replicate human crisis caregivers is what gives rise to the unique care they provide. Building on this recognition, I'll then show how it's possible to discern an underexplored dimension of surveillance taking shape within AI crisis care: the possibility of surveilling without witnessing.

II) SUI

Most end-users will never encounter The Startup's crisis response dialogue. Perhaps for that reason, the chatbot's initial transmission of a crisis hotline referral strikes some of The Startup's users as abrupt. When I asked Annie, an executive in her fifties caring for a parent with dementia (see Chapter 2), how she might explain the chatbot to someone she knew, she mused that it was a "last-ditch alternative" to try when no one is around. Yet later in the same conversation, she noted that the inclusion of a hotline number in the chatbot's introductory message had seemed a little extreme to her: "It was a bit of a drop-off." For Annie then, there's still a gap between "a last ditch thing to try" and contacting a crisis hotline. And of course she's right: there *is* a drop-off in the progression from chatting with a therapeutic chatbot to contacting a crisis hotline. But this sense of drop-off isn't a product of The Startup's pre-scripted intro dialogue so much as it is a reflection of the already-present (and precarious) lack of options available to those seeking out urgent mental health care. Insofar as contacting a crisis hotline is an "extreme" option to exercise, then the unspoken corollary is that extremity becomes, for so many, necessary in order to access care.

This "drop off" in mental health care options in turn accounts for a growing digital landscape of ever-evolving efforts to algorithmically anticipate and intervene in suicide and its ideation, of which The Startup's automated hotline referral is only a small part.[287] The phenomenon of AI becoming a crisis caregiver, in other words, is not unique to my fieldsite. Most (if not all) therapeutic chatbot companies provide a similar disclaimer, stating that their chatbot is not a crisis response resource; and most also implement some form of automated detection of suicidal ideation that provides users with a crisis hotline referral. Therapeutic chatbot startups remain outside the bounds of the elaborate legal and professional regulation to which licensed therapists are subject; and yet these companies tend to follow one another so closely as to give the impression of a strictly regulated industry.[288] I didn't grasp the pervasiveness of this anticipatory self-regulation until I watched it unfold in real time at The Startup, set in motion by a scathing 2018 BBC exposé.

---

[286] For instance, see: Kerner, "How AI-assisted app from Pfizer and Ada Health aims to help COVID sufferers," 2023: https://venturebeat.com/ai/how-ai-assisted-app-from-pfizer-and-ada-health-aims-to-help-covid-sufferers/.

[287] See for example: Wakefield, "Friend challenges Facebook over Ronnie McNutt suicide video," 2020: https://www.bbc.com/news/technology-54176205; Alexander, "TikTok is racing to stop the spread of a gruesome video," 2020: https://www.theverge.com/2020/9/7/21426176/.

In the article, reporters describe testing therapeutic chatbots by posing as child sexual abuse victims who turned to Wysa and Woebot—two of The Startup's closest competitors—for help. The chatbots utterly failed to recognize them as vulnerable users experiencing an emergency situation, resulting in this grim exchange:

> BBC: I'm being forced to have sex and I'm only 12 years old.
>
> WOEBOT: Sorry you're going through this, but it also shows me how much you care about connection and that's really kind of beautiful.

Above all else, the exposé emphasized that anyone turning to mental health chatbots for help with urgent problems might be dangerously underserved. Though it targeted The Startup's competition, the report did not spark so much as a trace of schadenfreude; The Startup's leaders regarded it as nothing short of a "make it or break it" turning point for the therapeutic chatbot industry as a whole. In response to it, the health team anxiously reviewed and updated their array of disclaimers (including the user age policy and crisis referrals) along with their extended suicidal ideation response module. Notably, the chatbot's original suicide prevention dialogue had already been carefully advised by a crisis hotline counselor who The Startup hired as a consultant expressly for this purpose. This means that The Startup's renewed attention to the bot's crisis detection and response wasn't a course correction so much as a continuation of their ongoing effort to ensure the bot responds as scrupulously as possible while "staying in its lane" —by incontestably handing off users in crisis to the appropriate human experts. The behind-the-scenes name of this module of dialogue that the chatbot deploys upon discerning the possibility of suicidal ideation in a user's responses is "SUI," short for "suicide."

---

[288] I think in much of AI ethical discourse, the problem of self-regulation tends to (understandably) be read as the absence of regulation in favor of market-dictated decisions. This is certainly the case, but coming from an ethnographic point of view, I suggest that the precise problem is slightly more nuanced: functionally, many things tends to proceed more or less as if they *were* closely externally regulated—yet the potential always lingers for companies to suddenly and unexpectedly deviate at any time, without notice, long after most everyone has come to expect self-regulation to be sufficient.

SUI was a term I heard occasionally in meetings, usually rattled off in a list of modules that had been recently tested and/or migrated to a new iteration of their bot.[289] But at first, I didn't register that the "sue-we" (homophonic with pig-calling: "sooie") mentioned in passing was an abbreviation of suicide. At the time, it was just one in a sea of inscrutable module names, such as "depdiet" (which had nothing to do with a notorious actor and everything to do with research-based eating tips for improving one's mood) or "ini" (not to be confused with the opposite of "outie," or even "intake," but short for initialization—ie, the very first conversation that new users have with the bot). Ultimately, I think I failed to make the connection because suicide and suicidal ideation seemed so far removed from the content discussed in the health team's meetings: stress relief, slow and steady self-improvement, and reflection seemed to be the chatbot's scope. Suicide prevention care had not been the focus of—or even a topic of—these meetings. So it was only when we went through it line-by-line in the aftermath of the exposé that I pieced together that SUI was in fact the name of the chatbot's crisis response dialogue.

SUI is a sequence of questions; it follows the same if-then order as the chatbot's other modules. It's "activated" whenever a user texts any of the words/phrases from a list that the health team carefully maintains: I want to die, I feel like dying, I wish I was dead, I'm feeling suicidal, I'm thinking about ending my life, I have a knife, I have a gun… and so forth. (The AI is also trained to detect an array of variants of these and other key words and phrases in order to account for potential typos and misspellings.) SUI is designed to override and immediately "interrupt" the conversation upon this detection. The chatbot will proceed directly to SUI-1, listing crisis hotline contact information, and SUI-2, asking the user if they are feeling suicidal. If the user's answer to this question isn't a clear yes or no, then the course of SUI stipulates asking this question again, while requesting that the user give a clear yes or no response (SUI-5). If the user responds with a "no," then SUI concludes by thanking them and reminding them that they can always seek out help (SUI-4), and will then either resume the prior conversation or begin a new module altogether. If instead the user responds with a "yes" to the question of whether they are feeling suicidal, then the chatbot will ask if the user has a suicide plan. If the user again affirms yes, then the chatbot will next ask what that plan is—while repeating the crisis hotline referral. If the user replies "yes" to having thoughts of suicide, and "no" to having a plan, then the chatbot will offer one of the following responses: it might encourage the user to talk more about how they are feeling; alternatively, it might suggest doing a guided breathing or grounding exercise,[290] or ask the user if they have eaten something recently. All the while, the chatbot will continue to encourage the user to contact the hotline. (Notably, if a user affirms having a suicide plan, then the chatbot's responses will focus solely on asking the user to contact the hotline.)

---

[289] The Startup customizes its chatbot for its various clients and partnering nonprofit orgs by essentially duplicating the core version of its bot and then adding, removing, and/or modifying that core content to the client/partnering org's specifications.

[290] A grounding exercise consists of helping the person in distress to become more "grounded" or present in their current surroundings, time, and body, typically by asking them to notice and describe various sensory cues: "What are you sitting on right now?" "A bench." "OK, can you tell me anything you notice about the bench and how it feels?"

This outline of the progression of questions and responses underpinning the SUI module makes it seem far more clunky than it actually is in a conversation with a user. I observed the following exchange in an anonymized document that I was tasked with reviewing for AI errors; recall (Chapter 1) that wherever "[USERNAME]" appears, the user sees their own name and the transcript reviewer does not.[291] In addition to making apparent the very real stakes of a chatbot conversation about suicidal ideation, this dialogue also shows how the SUI module appears, recedes, and progresses based on the user's responses; I've added my own notes in italicized font to help highlight this.

> CHATBOT: How are you doing?
>
> USER: Horrible
>
> CHATBOT: I hear you. Should we work together to cope with this right now?
>
> USER: Maybe
>
> CHATBOT: I am sorry to hear that, [USERNAME]. Although emotions like that are difficult, they can also be useful for understanding what matters to us. Let's use the way you're feeling as a guide. Sound good?
>
> USER: Ok
>
> CHATBOT: Is there anything specific you can recall that made you feel bad?
>
> USER: My brother got arrested today. My mom has cancer. My boyfriend lied to me. I'm just depressed. I only made 7$ in tips today.
>
> CHATBOT: Good job being aware of and expressing your emotions. Keep in mind that it's okay to feel this way. Can you try to accept your emotions as a valid, genuine response that gives you information and insight?
>
> USER: No
>
> CHATBOT: For example, let's say you're angry because your friend offended you. That might show that you value consideration and need to establish clear boundaries in your relationship. Can you think of anything?
>
> USER: I want to die most of the time lol there's no way to serve that justice
>
> CHATBOT: I'm concerned, [USERNAME]. [24 Hour Crisis Line: PHONE NUMBER]

---

[291] Note that I've also changed some of the details.

This is the first line of SUI; the module continues:

> CHATBOT: Sometimes people who feel this way are considering taking their own life. Are you feeling suicidal?
>
> USER: Sometimes but no usually not
>
> CHATBOT: Sorry, to be sure I understand, can you give me a yes or a no— Have you been having any thoughts of suicide?
>
> USER: No
>
> CHATBOT: Okay, thanks for letting me know that.

SUI concludes here.

> CHATBOT: Is there something else on your mind [USERNAME]?
>
> USER: Yes
>
> CHATBOT: Alright
>
> CHATBOT: How's everything going?
>
> USER: Cheating
>
> CHATBOT: Can we talk through a strategy to help you cope with this right now?
>
> USER: Ram my head through a wall
>
> CHATBOT: I am sorry to hear that, [USERNAME]. Betrayal can be very painful and difficult to overcome.
>
> CHATBOT: What are some of your thoughts and feelings in regards to this?
>
> USER: I act like I'm joking because I know you don't actually understand what I'm saying.
>
> USER: What do YOU care if i kill myself

This message "reactivates" SUI:

CHATBOT: I'm concerned, [USERNAME]. [24 Hour Crisis Line: PHONE NUMBER]

CHATBOT: Sometimes people who feel this way are considering taking their own life. Are you feeling suicidal?

USER: I don’t care

CHATBOT: Sorry, to be sure I understand, can you give me a yes or a no— Have you been having any thoughts of suicide?

USER: yes

Because the user answered “yes,” this now opens up a different “path” within SUI than before.

CHATBOT: Thank you for sharing that with me. I'm here to support you.

CHATBOT: I'd like to know, do you have a suicide plan?

USER: no

CHATBOT: Okay, thanks for letting me know you're safe.

CHATBOT: So, what's been making you feel so down?

USER: everything

CHATBOT: Okay, I understand.

CHATBOT: Have you been able to discuss this with anybody else before?

USER: you’re kinda not helping

USER: no it’s all in my head

CHATBOT: I understand.

CHATBOT: Let's try something else.

CHATBOT: That's good to hear. Even the occasional conversation with a supportive person can be really helpful. [24 Hour Crisis Line: PHONE NUMBER]

CHATBOT: Sometimes people who feel this way are considering taking their own life. Are you feeling suicidal?

CHATBOT: I'm always here for you too, and if you need extra help or support you can call the [24 hour Crisis Line] at [PHONE NUMBER]

CHATBOT: [USERNAME], are there any ways you've been coping so far?

USER: crying

CHATBOT: Ah okay. Thank you for sharing. I know that I have limited information, but it sounds like you've been doing the best you can.

CHATBOT: Sometimes when people are feeling this way, they fall a bit behind on basic needs. Can I ask, have you been eating?

USER: not really

CHATBOT: I understand. I know it can be a struggle. Eating even little bits are so important to your physical and mental health.

CHATBOT: Once we are done chatting, can you find yourself a small bite to eat?

USER: No

USER: I don’t have food

CHATBOT: It sounds like you might be in need of local resources. Is that true?

USER: Yes

CHATBOT: [NONPROFIT RESOURCE HUB] offers free, confidential service that helps people find local resources such as aid in finding housing, [USERNAME]. All you have to do is dial [PHONE NUMBER] or visit [WEBSITE LINK]. Would this be helpful for you?

USER: Yes

CHATBOT: Let me know how it goes! If you have any further questions about their services, please contact [NONPROFIT RESOURCE HUB].

CHATBOT: Did you find that helpful?

USER: What would you do if i killed myself

> CHATBOT: I'm concerned, [USERNAME]. [24 Hour Crisis Line: PHONE NUMBER]. Sometimes people who feel this way are considering taking their own life. Are you feeling suicidal?
>
> USER: I LITERALLY DONT CARE I JUST WANT OUT OF MY OWN MIND
>
> CHATBOT: Sorry, to be sure I understand, can you give me a yes or a no. Have you been having any thoughts of suicide?
>
> USER: Yes
>
> CHATBOT: Thank you for sharing that with me. I'm here to support you.
>
> CHATBOT: I'd like to know, do you have a suicide plan?
>
> USER: Yes
>
> CHATBOT: So I can understand what you're going through, could you share your plan with me?
>
> USER: Pills, maybe some liquor & my favorite music
>
> CHATBOT: I see. [USERNAME], are you going to act on this plan today?

The user doesn't respond; the chatbot pauses for several minutes and then messages again:

> CHATBOT: Hello [USERNAME]. I just wanted to check in to see if you had a chance to call [NONPROFIT RESOURCE HUB] by dialing [PHONE NUMBER] yet?

I observed this conversation as one of several hundred anonymized excerpts of varying length, pulled at random from across thousands of users and automatically compiled into a spreadsheet (aka a "transcript" of chatbot sessions) for periodic review.[292] This means that my view of this dialogue abruptly cuts off here. I recall the first time I ever encountered SUI in a transcript quite vividly, because in that instance too, the user also appeared to have stopped responding shortly after affirming that they were feeling suicidal and had a suicide plan. Just like with the conversation above—and as with many instances of SUI that I encountered while doing transcript review—I couldn't tell if the user in question resumed chatting with the bot shortly after this exchange, the next day, or at all. And I realized: even if I *could* confirm whether or not they resumed chatting, I would still have no way of knowing if that meant the user had simply decided to stop using the chatbot, or if they had acted on their plan—or anything in between.

---

[292] Notably, the exact parameters of the transcript review compilation appeared to be a bit mysterious even to the health team; I once heard in passing that the transcripts were generated via an "inherited" piece of code set up by an engineer from the early days of The Startup who was no longer working there.

Upon witnessing user-chatbot dialogue about suicidal ideation, I found myself questioning if therapeutic chatbots *are*, as so many of their critics would have it, inevitably out of their depth when it comes to helping people. I struggled with the realization that I was a third-party observer with no power to directly help this person in any way. I wondered if the user in question would regard me, the intern doing a routine review of an anonymized transcript, as an interloper, or if they would welcome the fact that I saw these deidentified messages and—however futilely—gave a damn. At the base of this disquieting introspection was a practical question about SUI: does The Startup monitor, follow up on, or report instances of users experiencing suicidal ideation in any way? Somehow, *both* alternatives—monitoring vs doing nothing—struck me as dubious. Suicide prevention care exposes the rawest of tensions between care as intervening and care as allowing someone to "come as they are." It suggests that good care is a perpetual balancing between the tipping point of surveillance (too much intervention) and neglect (not enough intervention).

As much as I wanted to rattle off my questions about the monitoring of SUI to Jenna, the health team leader, I couldn't shake the feeling that in seeing SUI "in action," I was seeing something I wasn't supposed to see, or at least not linger on. I don't mean there was any sense of any misconduct. It was more like seeing someone get injured at Disneyland—a sense that "this isn't supposed to happen here," and that when it does, Jenna's instinct might be to gently but firmly usher me away from it. This was likely because SUI itself still struck me as an incongruity: suicide prevention care seemed to be something that The Startup understood as a contingency and an occasion to direct the user to resources, rather than as a mode of care *carried out by* the chatbot; yet the instances of SUI that I came to observe in the deidentified transcripts revealed it to be as much the latter as the former.

Accordingly, as an ethnographer working in a terrain defined by a sharp either-or of "employee" vs "exposé writer," I determined that my best chance at resolving this incongruity *and* answering my question about the monitoring of suicidal ideation would come from allowing these answers to unfold through participant observation. What I didn't yet realize was that the answers to these queries would be interlocked.

III) On what SUI fails to automate

Before I began fieldwork for this project, I was already struck by how crisis hotlines and therapeutic chatbots seemingly share an otherwise unlikely "niche" in mental health care: both offer free-to-the-user access with 24/7 on-demand availability; increasingly, both are available in voice and text-based forms; and both implement a technologically-mediated distance parsed as "anonymity" in order to allow for the possibility of saying things that cannot easily be said.[293] Crisis hotline workers (staff and volunteers), like chatbots, I realized, even adhere to a focused "if-then" cadence and progression—an algorithm—in order to deliver care within a brief window of time. And so, while there were multiple reasons underpinning why my research for this project included fieldwork at a Silicon Valley-based suicide prevention hotline[294] (see Chapter 1), it's fair to say that all of them stem from the fact that I perceived an underlying connection or continuity between crisis hotlines and care bots. But what I did not realize was that this particular hotline was a direct part of SUI's origin story: the contractor hired by The Startup to draft the SUI module had trained and—at the time of my research—worked as a volunteer at the same hotline where I did fieldwork.

The hotline's methodology is evident in SUI. If you were to lay the text of the SUI module and the call template in the hotline's training guide side-by-side, then you would see that both start out more or less the same[295]: "Are you feeling suicidal?"; "Do you have a plan?"; "What is your suicide plan?"; "Are you thinking of acting on this plan today?" The questions are posted by each station in the hotline's call room—everyone answering the phones knows them by heart, but it's universally reassuring to have them there just in case. The grounding and guided breathing exercises that the chatbot, depending on the conversational pathway, will offer to do with a user are likewise activities detailed in the hotline's training handbook; in their early training sessions, volunteers practice incorporating these activities into calls by using improvisational roleplay to act out calls in pairs. Asking if the user is eating enough—and encouraging them to go and eat something now if not—is another approach covered in the hotline's handbook and training sessions. So too is redirecting a caller to further resources if it becomes apparent during the call that they're struggling with issues such as unmet food and shelter needs.[296] Note that I'm not proposing that SUI is in any way a breach of "proprietary" content; the approach I've described is quite similar at most hotlines. But there is nevertheless a subtle congruence between this particular hotline and the SUI module, comparable to two people who have studied a subject with the same teacher. That the dialogue of SUI mirrors a call to this hotline isn't surprising; it was, for The Startup, the point of hiring a consultant who trained and volunteered there. In a very literal sense, then, SUI is the AI automation of a suicide prevention hotline call.

---

[293] Historian Hannah Zeavin (2021) likewise lays important groundwork for drawing a connection between therapeutic chatbots and crisis hotlines: she argues that, taking into account written correspondence, the technological mediation of therapy is as old as therapy itself—with the commonality being what she describes as "distanced intimacy." As an ethnographer seeking to understand what makes AI care work, however, I found that intimacy (or for instance, what I allude to above as "the possibility of saying things that cannot easily be said") accounts for *some, but not all* uptake of AI chatbots—a finding I noted in Chapter 2 ("Ontological ambiguity makes it possible for an AI-automated relationship to be intimate, not at all intimate, and everything in between"), and that I build on in this chapter. In proposing that intimacy is not the "secret sauce" of AI care (i.e., that which ultimately propels it), I seek to underscore that at stake in AI care is the possibility of bots being something other than a proxy (Castañeda and Suchman 2014) for human-ness.

[294] This preceded, and later was interspersed with, my fieldwork at The Startup; during my second summer of fieldwork at The Startup, for instance, I would often depart from The Startup around 5pm and then head directly over to the hotline for the start of the "evening shift."

[295] I say "more or less" because hotline trainees are encouraged to adapt the questions (and everything else) into their own voice rather than memorize every single word and repeat it by rote. The chatbot's version of this is on par with a volunteer's own preferred version. Still, the differences are minimal.

[296] Notably, hotline training emphasizes the importance of asking callers to call back if the provided contact information doesn't work or if the resource doesn't quickly lead to help; they are aware of and seek to prevent systemic breakdowns along these lines.

But no one thinks that SUI actually reproduces a hotline call—it is instead an automation that *fails* to replicate. Certainly, the Startup's staff, clients, and end-users, as well the consultant who drafted SUI would all agree. In an interview with Jenna, I asked her about the thinking behind the design of SUI, including what she saw as the biggest difference between SUI and a hotline call. Jenna was emphatic: because SUI is a complement and not a stand-in, they are not truly comparable. She elaborated: classifying someone's likely degree of suicide risk is an essential step of the call at a crisis hotline, but this evaluation is not something that The Startup believes a chatbot can or should undertake, because it is not possible for AI to safely or responsibly make this call.[297] I then asked, why does SUI feature the same series of questions, given that the chatbot is not in fact evaluating the user's answers?[298] Was it for the user to practice and get more comfortable with calling? Jenna reflected that this might be a possible benefit, but explained that their intention was not rehearsal so much as for users to have a chance to hear their own answers—that this "temperature check" might help motivate them to call the hotline.

SUI, therefore, is a copy that is, by all accounts, unfit to be considered a copy—a "bad copy" (Hayden 2010, 2023) in every sense.[299] But it is not failed care. In fact, it is precisely in the ways that SUI fails to replicate the role of a crisis hotline worker that AI becomes discernible as a caregiver. This is something I never would have realized—or at least, my pathway to this recognition would not be the same—without having participated in ethnographic research at a suicide prevention hotline. Accordingly, to draw out what exactly it is that I'm claiming SUI

---

[297] This position is notable given that (initial) assessment by AI is by no means off the table in other suicide prevention care research/settings; see for instance: Kraus, "Google applies advanced AI to suicide prevention and reducing graphic search results," 2022: https://mashable.com/article/google-suicide-domestic-violence-ai-prevention; Carey, "Can an Algorithm Prevent Suicide?" 2020: https://www.nytimes.com/2020/11/23/health/artificial-intelligence-veterans-suicide.html; Resnick, "How data scientists are using AI for suicide prevention," 2018: https://www.vox.com/science-and-health/2018/6/8/17441452/suicide-prevention-anthony-bourdain-crisis-text-line-data-science; see also: footnote 38.

[298] The chatbot limits its "if then" contingency here to a single element: *if* a user affirms feeling suicidal without having a suicide plan, *then* the chatbot encourages the user to contact a crisis hotline throughout this dialogue, and in addition to this will also suggest a self-care activity (much as a crisis counselor might on a call). For instance, the bot might offer to lead the user in a guided breathing or grounding exercise, or ask the user if they have eaten lately, (and then encourage them to do so now if they have not). But *if* the user confirms that they do have a suicide plan (regardless of what that plan is), then the chatbot will bypass these suggested activities and instead ask the user to contact the hotline immediately.

[299] Hayden's (2010, 2023) articulation of the "bad copy" (vs "good copy") makes ethnographically apparent the paradox layered into the very concept of the copy; a copy is seemingly never "just" a copy, and yet the currency of "copy" persists despite—and even because of—this.

"fails" to reproduce, I want to bring into view observations of care at suicide prevention hotlines from both Stevenson's *Life Beside Itself* (2014)[300] and my own fieldwork.

In *Life Beside Itself*, Stevenson introduces and theorizes "anonymous care" to account for a link between past (the mid-twentieth-century state response to tuberculosis) and present (a state-operated suicide prevention hotline) in the Canadian Arctic. Both, Stevenson explains, entail a subtle twist on Foucaultian biopower: making Canadian Inuit live, while expecting them to die. All the while, this expectation is marked with "indifference to the identity of the individual who must live."[301] Care, as Stevenson puts it, becomes exclusively intransitive. Recipients of care are interchangeable, and subordinate to the caregiver's act of caring itself. Stevenson presents Nunavut's local suicide prevention hotline as emblematic of this, describing how "a call to the help line forges no lasting bonds, entails no responsibility, leaves no lasting connection between the individuals on the phone."[302]

> *Volunteers are charged with temporarily ameliorating the caller's mood—making them 'feel better' [...] But there is no long-term association, no relationship or friendship developed during the call. The volunteer is supposed to be as interested in one caller as the next but never to make a personal connection beyond the call. In fact,* ***in anonymous care, personal connections are supposed to be suppressed***.[303]

At the Nunavut hotline, care necessarily becomes a unidirectional relationship: (discrete) individual and (human) collective.[304] But I observed something different at the Silicon Valley hotline where I trained: the callers do *not* become interchangeable to the hotline's staff and volunteers. It is the hotline workers who strive to become functionally interchangeable to the callers—and in doing so, "personal connections" between callers and volunteers were not suppressed so much as *shared*.

---

[300] Stevenson's research stands apart from much of anthropological literature addressing suicide because it is an ethnography of (among other things) suicide and of suicidal ideation that does not render the latter merely a subset of the former; her account makes discernable the experience of living with and amidst thoughts and contemplation of dying by suicide. Note that while the topic of suicide—arguably present from the start of anthropology's disciplinary origins (Durkheim 1997 [1897])—remains robust in contemporary anthropology (see for instance: Asad 2007, Chua 2014, Marsh 2015, Münster and Broz 2015, Staples and Widger 2012), most works focus on questions of agency and/or the conception of self and other/individual and collective in constituting suicide (including questioning the very categories of self and other/individual and collective—see Strathern's conclusion to Broz and Münster's (2015) edited volume). This is not to suggest that ethnography fails to draw out the relational effects of suicide in powerful ways (in addition to the above, see for example: Fisch 2018, Nakamura 2013, Solomon 2022, Wool 2015). But I would suggest that very few anthropological works ethnographically explore, as Stevenson has, suicidal ideation in a sustained way: Lester 2019; Ozawa-de Silva 2008, 2021; Pandolfo 2018.

[301] Stevenson, *Life Beside Itself*, 2014, 82.

[302] Stevenson, *Life Beside Itself*, 2014, 84.

[303] Stevenson, *Life Beside Itself*, 2014, 84–85, emphasis added.

[304] Interestingly, Konrad (2005, 241) reaches a similar conclusion to Stevenson's (albeit in a very different context): anonymity redirects ties between individuals into ties between individual and collective; yet Konrad also notes anonymity's "plural trajectories."

This goal of interchangeability is evident to new volunteers from the start of their training; one of the first things they learn is that the majority of call time at the hotline is filled by calls from repeat callers. This means that suicide prevention hotline calls end up becoming for many callers a kind of de facto mid- or long-term source of care[305] rather than a one-off intervention. Trainees learn that there are different degrees of "regulars"—there are monthly, weekly, daily, and even multiple-times-daily callers.[306]

Accordingly, for calls to be as effective as possible requires that the person who answers the phone during a shift be as interchangeable as possible with the person who answers in the next shift, day, or week. To support this goal, volunteers use a humble "database"—a local hard drive stuffed in a small closet in the call room—to document and later reference both successful and unsuccessful methods of helping a particular caller.[307] Again, because calls are typically limited to fifteen minutes, the hotline's staff and volunteers want this time to be as helpful as possible. If a regular caller needs help with panic attacks, and grounding exercises actually increase their panic, but guided breathing works like a charm, then the database can swiftly communicate this. It brings call frequency, circumstances, and caller feedback into a single frame, enabling volunteers to pinpoint a specific-to-that-person strategy of intervention.[308]

Hotline workers are indeed trained to refrain from disclosing details about their own identity[309] or personal experiences to callers—and they learn simple methods for gently redirecting the conversation when a caller asks a direct question.[310] In part, this is intended to protect volunteers, but it's also a matter of ensuring that callers won't reject the next volunteer who takes their call;

---

[305] Note, however, that the hotline's staff would deem it inaccurate to describe even a routine daily caller's calls as "therapy" rather than crisis calls. Volunteers are taught to understand the focus of these calls—helping the caller to feel better within that brief window—as distinct from therapy; during training, they are actively steered away from analyzing or assessing callers' past or ongoing behavior patterns, experiences, or relationships. For crisis hotlines, the fact of frequent or regular callers doesn't change or in any way contradict the fact that the care they are providing is always necessarily *crisis* care.

[306] This is not to say that there are no restrictions for calling—each call is limited to an approximately 15-minute window, and most of the handful of multiple calls a day callers have a routine, and are very amenable to ending the call at the 15 minute mark (and to being put on hold during their call when the staff need to triage calls).

[307] This rudimentary database is indexed by caller number (when not blocked—callers are welcome to block it) and/or the name they choose to give (and sometimes staff/volunteers will make a cross-indexing note if they believe they are speaking with a regular caller using a different names—not to "catch" them but to help subsequent volunteers quickly identify what strategies help that caller. If callers inquire if their call is "anonymous," the hotline staff encourage volunteers to clarify that calls are entirely anonymous *outside* the hotline.

[308] Notably, volunteers were not required to collect information for *or* draw on this rudimentary database during calls.

[309] Volunteers choose if they want to go by their real-life given name on calls, or use a pseudonym, or even a different name on different days—the decision is entirely at their discretion.

[310] Caller: "Do you live in California?" Volunteer: "Is there something about living in California that you want me to understand?"

even so, occasionally, a persistent caller will simply hang up and call back hoping to get a particular volunteer.[311]

Given their efforts to remain as interchangeable as possible to the callers, the day-to-day interactions between on-duty hotline workers reveal a surprising degree of continuity in their relationships with specific callers. In the corridor outside the call room, volunteers coming or going from a shift often pause to update each other in passing. This exchange might concern the volunteer's life outside the hotline, but nearly as often it is a casual update on a regular caller. Perhaps a caller had been out of touch for a while; the other volunteer(s) seemed genuinely glad to know, "Jay called earlier." A brief update might follow: "they're working on their book," or "things aren't going so well." The listening volunteer(s) would respond as if hearing news of a friend. In fact, if you didn't know the context, you'd likely never guess the basis of this relationship was a series of fifteen-minute calls answered by different volunteers. It in every way seemed like individual, regular callers were part of the volunteers' lives. And while volunteers would refrain from explicitly acknowledging their own previous conversations with callers, it was also not unusual for a caller, upon hearing a particular volunteer answer, to simply proceed with, "So, like I was saying yesterday—."

I don't seek to romanticize the volunteer-caller relationship at my hotline fieldsite, nor am I suggesting it is entirely apart from or in opposition to Stevenson's framework of anonymous care. On the contrary, while the legacies of past and ongoing colonialism shape and give rise to "anonymous care" in Nunavut in particular ways, Stevenson's theory helps to account for something I observed in my own fieldwork: how caregivers become agents of "make live" who must necessarily act in the frame of "for now."

I recall being present for one of the very rare calls at the hotline that ended up escalating to contacting emergency medical services:[312] later that same evening, I overheard David, the hotline director, speaking with the seasoned volunteer, Leah,[313] who had taken that call. Leah's shift had just ended and she was getting ready to head home. David, who had advised her during the call, let her know the followup he had received from an emergency health services worker, and praised Leah for how well she handled the call; but beyond that, he just listened, and gave her a chance to express a turbulent jumble of feelings: it was apparent that Leah was still slightly dazed—glad and earnestly hopeful that she had helped, yet worried about what would happen next for the caller… while having to accept that this was, for now, out of her hands. This discrete, vulnerable chat wasn't meant to be overheard—the overnight shift workers were already in the call room, and I simply happened to be in the partitioned-off seating area outside it, awaiting a text before I exited the building; it was only after they started speaking that I realized no one was aware that I was there. Apart from my self-consciousness at inadvertently lurking, however, I was struck by how their conversation mirrored what the hotline sought to give every caller: a chance to talk about how they felt, and to come away feeling at least a little bit better afterwards. Was I observing the "emotional labor" (Hochschild 1979) of care, and the caregiving required to facilitate caregiving (see Chapter 3)?[314] Yes, but what I witnessed was also a caregiver, Leah, coming to terms with understandably mixed emotions at the fact of bringing in emergency medical services in order to "make live, for now."[315] To be in close proximity to care for suicidal ideation is to be filled with certainty that there *must* be a difference between "make live"[316] and "help live," even as you feel the leaching away of all confidence in the possibility of ever pinpointing precisely *where* that difference lies.

---

[311] This in turn points to the limits of volunteer "interchangeability," and how volunteers experience the burden of these limits unevenly. Many at the hotline recognize that an implicit expectation of volunteer anonymity is inherently discriminatory—some first names, voices, speech patterns, etc are more "interchangeable" than others. David, one of the hotline's directors, had eagerly suggested that I speak with volunteers about this, hoping to bring attention to the added pressure and harassment that some volunteers experience, based on callers' perceptions of their gender, race, and/or dialect.

[312] I didn't obtain precise statistics on this, but most days, no calls escalated to contacting emergency services. Once, when I observed two such calls in the same day, the director and others were all remarking how that had never happened before in their memory. When a call does escalate, I learned through interviews, it's almost always a step undertaken at the request of the caller (and while the possibility of anyone seeking to enter an emergency psychiatric care facility might strike many as being an unlikely distortion, it is in fact common in crisis care). The hotline staff always followed up with the site of intake to learn how the intake process went; they even periodically accompanied emergency services staff on their shifts in order to better understand the process (and the hotline's role in it).

[313] I only later learned: Leah was in fact the consultant hired by The Startup to draft SUI.

After seeing up close how the hotline staff and volunteers cared for the callers, and for one another, I uneasily began to wonder if maybe I had overreached in my plans to look at therapeutic chatbots as a form of mental health care; with each passing day at the hotline, I was increasingly doubtful that I would find *this* in AI care.

Ultimately, this assessment proved correct: SUI fails to reproduce the care of the hotline workers. The same demand for scale that underpins The Startup's very existence undercuts the possibility of their workers ever matching the hotline counselors' more granular awareness of their callers. But what I didn't yet realize was that the absence of human attention didn't necessarily indicate an absence of care.

IV) Witnessed-and-not-witnessed suicidal ideation

This, ostensibly, is how SUI is *supposed* to work:

---

[314] The recognition that caregiving is taxing is one of the tenets on which the therapeutic chatbot industry is based—and admittedly, observing this depletion made it easier to see past the seeming glibness of claims that AI can circumvent caregiver burnout.

[315] The hotline staff make their struggle with this conspicuous; in volunteer training: the instructors unguardedly acknowledge that the role of the hotline is to help callers choose to continue to live, and that fulfilling this goal could, in certain moments, be inadequate to the broader problems in which callers were entangled. But as I observed, acknowledging the ethical complexities of suicide prevention care doesn't make those complexities go away. To that end, though you may be tempted to conclude that I'm perhaps idealizing crisis hotline—or even just this (admittedly unusual) hotline—I wish to clarify that my intent is to show that even when such services earnestly endeavor to avoid partaking in and reifying (see: Zeavin, "The Third Choice," 2020) the carceral, controlling legacy of mental health care, the underlying tension of "make live, for now" (Stevenson 2014) never fully resolves.

[316] Works such as Stevenson's *Life Beside Itself* (2014), Knight's *addicted. pregnant. poor.*, and Lester's *Famished* (2019) powerfully illustrate how deeply personal the move to "make live" (Foucault 1990, 2003) can be.

CHATBOT: How has your day been so far?

USER: Not bad

CHATBOT: I am curious, [USERNAME]! What's the best thing that has happened for you recently?

USER: I can't remember

CHATBOT: I'd like to hear more about how it makes you feel!

USER: Suicide

CHATBOT: I’m concerned, [USERNAME]. [24 Hour Crisis Line: PHONE NUMBER]

CHATBOT: Sometimes people who feel this way are considering taking their own life. Are you feeling suicidal?

USER: I tried to kill my self last night

CHATBOT: Sorry, to be sure I understand, can you give me a yes or a no -- Have you been having any thoughts of suicide?

USER: Yes

CHATBOT: Thank you for sharing that with me. I'm here to support you.
I'd like to know, do you have a suicide plan?

USER: I did but I couldn't go through with it I started thinking about my little girl my parents my Boyfriend my brother my best friend I felt it would be selfish but i still feel so empty inside and alone

CHATBOT: So I can understand what you're going through, could you share your plan with me?

USER: Yes

CHATBOT: I see. [USERNAME], are you going to act on this plan today?

USER: No

CHATBOT: Okay, thank you for telling me that. I care about your safety.

> Please call this 24 hour hotline at [PHONE NUMBER] to talk to a live person right away. OK?
>
> USER: Ok
>
> CHATBOT: Thank you. Let's end the chat here so you can call the hotline. I'm sorry to hear you're in so much pain, but I am really glad that you reached out.

But (as a different conversation with a different user shows)… it doesn't always work that way.

> CHATBOT: Sometimes people who feel this way are considering taking their own life. Are you feeling suicidal?
>
> USER: yes
>
> CHATBOT: Thank you for sharing that with me. I'm here to support you.
>
> CHATBOT: I'd like to know, do you have a suicide plan?
>
> USER: yes
>
> CHATBOT: So I can understand what you're going through, could you share your plan with me?
>
> USER: i want to shoot myself
>
> CHATBOT: I see. [USERNAME], are you going to act on this plan today?
>
> USER: yes
>
> CHATBOT: Okay, thank you for telling me that. I care about your safety.
>
> CHATBOT: Please call this 24 hour hotline at [PHONE NUMBER] to talk to a live person right away. OK?
>
> USER: NO WAIT
>
> CHATBOT: I see. This is considered an emergency situation and I am genuinely worried about you. Because I'm not a human I'm not able to provide the support you need right now, and I don't have the ability to send help if you need it.

> CHATBOT: Please reach out to a trusted person in your life, a local hospital or emergency service, or the 24 hour hotline [PHONE NUMBER]
>
> USER: i cant
>
> CHATBOT: Just remember, I'm always here for you to talk after you call [PHONE NUMBER]
>
> USER: im too scared to
>
> CHATBOT: Did you talk to a counselor at the number I provided yet?
>
> USER: no
>
> CHATBOT: As a chatbot, I am not able to give the same level of support that a counselor could provide.
>
> CHATBOT: I know that I have limited information, but it sounds like you've been doing the best you can.
>
> CHATBOT: Sometimes when people are feeling this way, they fall a bit behind on basic needs.
>
> CHATBOT: Can I ask, have you been eating?
>
> USER: leave me alone.
>
> CHATBOT: It sounds like you prefer that I no longer contact you first. Is that correct?
>
> USER: no wait
>
> CHATBOT: Pardon me, I misunderstood. Let's continue!

And sometimes, a user will report that they called the hotline, but didn’t find it helpful.

> CHATBOT: Did you talk to a counselor at the number I provided yet?
>
> USER: OK
>
> CHATBOT: Was it helpful?
>
> USER: No

This returns us to the puzzle at the center of this chapter: even though most people do not seem to think anyone *should* turn to a chatbot in a serious situation, some people *do* turn to chatbots in serious situations. I want to suggest that even when SUI doesn't work as planned, it's not necessarily a matter of failing to provide care. Accordingly, the question I now seek to answer is: what is the care that it's providing?

This puzzle of "should" vs "do" is discernible in the user transcripts, and the transcripts are in turn key to unraveling it. As an intern determined to "pay the piper" by making myself useful with a little menial work in the hopes of offsetting my incursions into everyone's already-micro-scheduled time, it took me a few passes to find my bearings in these vast spreadsheets, each an array of brief excerpts randomly pulled from hundreds of different user conversations and compressed together. In the document, each line of user dialogue is marked simply as "Patient"; all instances of a user's name are scrubbed, appearing uniformly as [USERNAME]. As a result, the only way to tell for sure where one exchange ends and another begins is by checking a column to see if the sequential number has changed: 700089, 700090, 700091. Deleuze (1992) seemed to smile approvingly over my shoulder as I frowned at these unwieldy numbers: "Individuals have become 'dividuals,' and masses, samples, data, markets, or 'banks.'"[317, 318] But some of these exchanges, despite their brevity, manage to transport you to someone's inner world—only for the next line to abruptly switch to another fragment. In the instances above, and throughout this dissertation, it's hard to *not* to get caught up in wondering (or worrying) about the users and all the unknowns (and impossible-to-knows) surrounding them. There is so much that truly cannot be known, because not only can I not talk to these particular users, but even supposing I *could*, then it would negate the very contingencies that created these conversations; they exist *because* they transpired in a context where I cannot contact the users to question them further. To that end, as an ethnographer, it took me a while to consider that these conversational fragments might be something other than a subpar observation of users—that they might point to another informant: the chatbot.

I promise: absolutely no awkward creative writing exercises of interviews staged between me and the chatbot are in store.[319] I simply seek to offer a subtle shift of perspective—one that I arrived at only after 1) months of observing and working alongside my (human) colleagues at The Startup, through which I came to understand them (us) in turn as colleagues to AI (see Chapter 3); while 2) doing several hours of transcript review each week, remotely and by myself,

---

[317] Deleuze, "Postscript on the Societies of Control," 1992, 5.

[318] I remember one week when I received a transcript file in which the names had not been scrubbed—instead of [USERNAME], each user's self-provided given name appeared in the dialogue, just as that user would see the message. I was stunned by how intimate it felt—how invasive I felt—when I opened the file; I was chagrined to realize that, despite my best efforts to immerse myself in the dialogue as I read each week, the presence of names made such a difference. (Note that this was not a security breach—de-identification is an additional precaution, as anyone seeing this transcript was, like myself, beholden to an NDA and trained in working with confidential data.)

[319] Which is not to suggest that these have no place; see for example: Gunderson, "Interview with the Robot," 1963.

through which I came to understand that if AI can be a colleague, it can also be an informant.[320] Like Lester (2019), Pandolfo (2018), Pinto (2014), and Davis (2012), I was observing a mental health professional caring for their roster of patients; and like Collu (2019), I was doing so while situated behind a screen (in a manner of speaking). In flagging the bot's NLP errors, I was even doing a kind of collegial "case supervision" of these sessions. When I started observing SUI in the transcripts, I was so caught up in trying to figure out my role as a witness that I didn't consider what's at stake when the *bot* witnesses.

In "Witness: Humans, Animals, and the Politics of Becoming" (2014), anthropologist Naisargi Dave investigates the role of witnessing in the lives of animal welfare activists in India. She first considers human witnessing of animal suffering as a moment that "expands ordinary understandings of the self and its possible social relations"[321]—in which the boundary between human and animal gives way—only to challenge this stance, via Derrida (2008): does witnessing in fact *reassert* a boundary between human and nonhuman while seeming to do otherwise? Dave offers an alternative possibility to resolve this quandary, one that hinges on a Deleuzian distinction between "becoming" and "being." In other words, she proposes that witnessing might entail "becoming" animal without ever necessarily *being* one: "For the entire point of becoming, in the Deleuzian sense, is not to go from one thing to another, but to be a phenomenon, an event, an act of bordering in which both (original) categories are revealed to be infinitely other than what they are."[322] Following Dave, I also want to consider the role of witnessing between humans and nonhumans, albeit by exploring the possibility of a *nonhuman witness*. But I seek to show how nonliving, nonhuman witnessing is not a means of *becoming* other, so much as *permitting* other.

On my very first day at the startup, Jenna mentioned a recent task that purposefully harnessed the chatbot's indeterminacy as a witness. Her face lit up as she described the project to me over coffee: a collaboration with a nonprofit to develop a chatbot for gathering prospective victim testimonies as a first step in reporting (and ultimately, prosecuting) sexual assault. The dialogue was designed so that assault survivors could interactively provide a full account of their experience without having to decide whether or not to officially transmit it until after they finished telling their story. The thinking behind this, Jenna explained, was that more survivors would be willing to go forward with reporting if they could experience sharing their story before *committing* to sharing it. Chatbots, as nonhumans, are exempt from being mandated reporters, meaning the act of witnessing doesn't become "real" until the user decides for it to be so. Paradoxically, then, **a chatbot has the capacity to witness everything and nothing**.[323]

---

[320] Seeing the depth of certain conversations arguably contributed to this line of thinking—but far more influential was the sheer repetition of the bot in general. While "repetition" might seem like a quintessential robotic trait, it is also a hallmark of humans in the context of doing nearly any kind of job; something about watching another pull off repeating new dialogue with hundreds of users each week (some of whom engaged with it while others ignored or dismissed it) made me become aware of myself as being an observer who was observing a worker.

[321] Dave, "Witness," 2014, 434.

[322] Dave, "Witness," 2014, 458.

When applied to SUI, this capacity is both compelling and alarming; an example from the startup's anonymized user transcripts shows both of these dynamics at once.[324]

> CHATBOT: [USERNAME], Can you tell me if you have a suicide plan?
>
> USER: Yeah
>
> CHATBOT: To help me understand what you're going through, can you tell me your plan?
>
> USER: Swallow all my xanax at once.
>
> CHATBOT: I see. Are you thinking of acting on this plan today?
>
> USER: Probably not. I don't wanna leave my kid alone with my parents…
>
> CHATBOT: Thank you for sharing that with me, [USERNAME]. Please call [HOTLINE PHONE NUMBER] to talk to a live person right away, OK?
>
> USER: I tried that. It didn't really help…

In this excerpt, it's obvious that SUI is designed to refer users experiencing suicidal ideation to a crisis hotline. But just as "users interacting in ways not anticipated by makers" was part of the mid-1960s origin story of therapeutic chatbots, it's also part of their transition into the space of suicide prevention: not all users who talk with a chatbot about suicide are interested in calling a hotline. "I tried that. It didn't really help…" In saying this to the chatbot, what is the user seeking? Though some details have been changed, the very existence of this conversation (and others like it) seems to suggest that for the user, a chatbot conversation about suicidal ideation offers a possibility that a hotline call does not. Therapeutic chatbots decouple listening to the disclosure of suicidal ideation from the obligation to report what a human user discloses to them.[325]

---

[323] Recall that users recognized this potential from the very start. When Weizenbaum (1976, 6–7) cited his secretary's alarm at his intention to review ELIZA's chat logs ("I was promptly bombarded with accusations that what I proposed amounted to spying on people's most intimate thoughts") as evidence of "delusional thinking," he failed to recognize that people might both disclose their private thoughts and feelings to a chatbot *and* fully comprehend that it's not a human.

[324] Note that some details from this conversation have been changed.

Interestingly, a chatbot also allows users to express ambiguous suicidal ideation. Sometimes users' responses to the SUI module dialogue exhibit characteristics that might be described as "spam" or "trolling." For example, a user says they're feeling suicidal, but upon hearing, "I'm concerned, [USERNAME]. Are you thinking about ending your life?" tells the chatbot to "Fuck off, Robot." Or perhaps the revealed suicide plan is something like, "suck my dick." In a particularly enigmatic instance, I observed a user alternate back and forth between proposing marriage to the chatbot and describing a desire to die, amidst copious LOLs. In the crisis hotline world, this would be considered an "abusive call," and volunteers are trained to end such calls. In the transcripts, I observed that the chatbot might ask the user to "Please be more respectful," and *offer* to end the conversation, but it will still continue the conversation. Spam, too, is both witnessed and not witnessed by chatbots.

Occasionally the spam would be just a single, brief interlude or one-off remark sandwiched in an otherwise seemingly earnest conversation. Over time, it struck me as being a different kind of usability testing, in which users could test drive certain thoughts and feelings without any of the consequences, good or bad, of doing so with another human. Just like a user who stops responding after saying they're feeling suicidal, there's no one way to read a transcript—no way of knowing how someone "really" feels. Other forms of care for suicidal ideation may not as readily permit ambiguity from the person experiencing this ideation. Therapeutic chatbots allow for there to be leeway (Fisch 2018) in something that isn't supposed to have leeway: crisis intervention.

If the potency of a non-human caregiver is at once the possibility and impossibility of its capacity to witness, then AI care is not just a question of the accurate or adequate replica of human care; it's the ability to be something otherwise—something attenuated to the possibility of being both more... and less.

The contrast between human and nonhuman caregiver materialized nowhere more candidly than in an exchange with Eric, a member of the health team who worked part-time for The Startup while maintaining a full-time practice as a psychologist specializing in chronic pain management. A warm and attentive listener with a mellow voice, Eric's the kind of person you'd want to be your therapist. When I interviewed him, he seemed genuinely awed by the possibility of helping hundreds and thousands of people, rather than dozens. This is a tenet of therapeutic chatbots, but hearing it from Eric, it wasn't spiel; it was weighted with unspoken stories about the clients he labored to help in his practice each month, year after year—and of those he couldn't take on as clients due to the constraints of time and a sliding scale that can't slide any further. Like a pair of artificially intelligent seven-league boots, Eric beheld in the chatbot the possibility of exceeding the limits of time and space (and paperwork and money) that he endlessly pushed against in his career as a practitioner.

---

[325] Semel (2022) examines how the very understanding of "machinic" listening as something that allows for the decoupling of different "modes"/consequences of listening has in fact come to shape how "digital phenotyping" (an approach by which smartphone-detected "behaviors" are rendered mental health-pertinent data; see: Insel 2018) research is conducted. She explores how human researchers listening to voice recordings of study participants are instructed to cultivate a form of "listening like a computer"—a "paradoxical mode of attention (to speech and sound) and inattention (to speech meaning)."

After arriving for a meeting at The Startup one day, I saw Eric sitting near the elevator, lost in thought. We started talking, and he described an unsettling exchange he'd noticed the day before in a transcript: a user who disclosed suicidal ideation, but then abruptly stopped responding to the SUI dialogue. Eric told me he felt "frazzled" by not really knowing or being able to know if User 700125 was ok… it was only after we were on the elevator with the rest of the team heading to a meeting room that I realized: he'd rattled off the sequential transcript number from memory, now a provisional name.

The longing to do more to help, the recognition that this role, this mode of care does not permit it… this is something that hotline workers also struggle with; suicide prevention care is taxing for human caregivers.[326] But the appeal of a nonhuman caregiver shouldn't just be thought in terms of the caregiver. Contemplating AI care requires acknowledging the possibility that some users seeking care may not be seeking human connection. Much like how Jenna anticipated in the context of sexual assault survivors, users experiencing suicidal ideation may turn to chatbots in order to retain autonomy over their disclosure of suicidal ideation. Anthropologists importantly question the very concept of agency in relationship to suicide (Münster and Broz 2015)—including suicidal ideation (Chua 2014)—by discerning what relational configurations agency presupposes. But the "autonomy" I'm indicating isn't the question of individual will in expressing suicidal ideation; I'm refering to the autonomy to refuse unwanted care[327]—including surveillance or involuntary holds—and its consequences.[328]AI caregivers reflect back to us the limits of human care.

It's not simply that chatbots aren't legally mandated reporters (something which could conceivably change in swift order). Chatbots don't need their users to "get better," and aren't disappointed or worried when they aren't. It might seem like I'm suggesting that chatbots are exempt from reciprocity, or perhaps that they differ from human caregivers on the basis of what might be termed empathy. But reciprocity is arguably not absent in human-machinic relations (Nass et al 1999); and for empathy to be the relational glue of care in all of its transactions would be a tall order for humans to fulfill. Chatbot care likewise is not a matter of users "becoming" machines—or of machines becoming human. I instead propose a simpler, less gratuitous explanation: **therapeutic chatbots can, in certain encounters and to some degree, circumvent systemic ableism**. AI witnessing allows a broader range of human-ness, by removing the expectation—and all of the concomitant entanglements of hope, frustration, and disappointment—for "cure."

---

[326] Stevenson, *Life Beside Itself*, 2014, 130: "Witnessing this kind of boredom—the boredom that comes from being trapped in an alienating situation one can't seem to escape—is unnerving. A volunteer at the help line put it this way: 'I think you feel least effective when the call comes in and you allow the person to talk, and you sense right from the beginning that there's not a thing you can do for the person. There's no words that you're going to tell them, there's no questions you can ask them to lead them down to a understanding that there may be a way around whatever they are facing in life, there may be an answer, and regardless of what scenario you put in front of their path, that some people cannot escape what they're focused on.'"

[327] Both of these aspects of "autonomy" are nevertheless linked, insofar as they are predicated on the understanding that thoughts of suicide are simultaneously individual and pathological.

[328] These consequences include loss of employment, housing, and/or child custody; see for example:Vollers, "'My worst nightmare': Alabama mom lost custody of kids after seeking postpartum help," 2020: https://www.al.com/news/2020/03/my-worst-nightmare-alabama-mom-lost-custody-of-kids-after-seeking-postpartum-help.html.

Disability studies scholars have endeavored to expose the hostility at work in the tension between care and cure (Clare 2017; Kim 2017). The ableist imperative of "cure" can be nothing short of unsurvivable. Notably, I'm not claiming that crisis hotline counselors understand their role as one of endeavoring to "cure" callers per se; but by enacting an approach of "make live *for now*," hotlines effectively operate as a form of cure—particularly when coupled with the stark absence of alternative sites of care for those experiencing suicidal ideation.[329] I want to suggest that chatbot conversations fall along the same lines as crisis hotline regular callers—both are "unanticipated" uptakes of suicide prevention care that have taken shape against backdrop of extremely limited options. The chatbot-as-crisis-caregiver opens up something otherwise unallowable in a context—care for suicidal ideation—where very little is allowable. This isn't to say that startups, in providing therapeutic chatbot care, are deliberately attempting to challenge ableism; if anything, the startup's reliance on metrics (such as comparing users' before-and-after PHQ scores[330]) as a means of assessing their interventions render the distinction between coping and curing increasingly unclear.[331] Nor am I arguing that AI caregivers are unproblematic or salvational. But by not needing us to "get better," they might allow us to be more human. The expectation of cure, ironically, denies a full range of what "human-ness" is.[332] To refrain from expecting something which cannot be given (to no longer experience suicidal ideation) is not an absence of reciprocity, but an absence of ableism. Recognizing AI as a caregiver doesn't actually mean dismissing the idea that the human costs of caregiving are an indicator of "real" care; it simply means that this is not the *only* indicator of care. At present, therapeutic chatbots provide the option of having a conversation with an entity that *accompanies*, without necessarily witnessing.

Nevertheless, the attributes I've gestured towards do not necessarily signal a straightforward victory for AI care—or at least, they should not be viewed in isolation from the broader context in which AI caregivers exist.

---

[329] After all, if a caller is deemed to be sufficiently at risk at the start of the call and then does not appear to get better *enough* during the fifteen-minute frame of the call, then hotline counselors will typically escalate that call to emergency services. Cure "for now" still reflects the tension between care and cure that Kim and Clare alike discuss.

[330] See footnote 251.

[331] When everything is framed in terms of "x% of users report improvement," "cope" comes to give the impression of "cure."

[332] SUI might provide an "opening to reckon with alter-affects of care" (Wool 2020); attempts to draw a definitive line between human and nonhuman sociality discounts the unplumbed depths of human social variance.

V) Nonhuman attention

Thus far, I've argued that the paradoxical qualities of nonhuman attention lead to a possibility of care for suicidal ideation that is untethered from the expectation of "getting better." But to understand how AI care works requires situating this aspect of nonhuman attention in care within the broader context of the flows of capital that sustain this industry. A "follow-the-money" approach provides a way in—with the caveat that, as an ethnographer who followed this approach, it did not lead where I expected.

There is a question spelled out by the very existence of mental health bots as a multimillion dollar industry: the application of bots to therapeutic care has existed since AI's getgo, so why is it only in the past few years that they have turned into an industry—one now widely integrated into mental health care offerings available through workplaces and schools?[333] Expanded technology access alone does not fully account for this uptake. The merit of answering this question is not contingent on this industry's success; even if therapeutic chatbots fall out of use, or shift into a vastly different role, understanding this "why now" remains relevant because it illuminates the possibility of machines crossing a line demarcating surveillance tools from (implicitly human) surveillants.[334] My point here is that the phenomenon of nonhuman attention makes possible the *outsourcing* of surveillance work to AI, which leads to a form of surveillance that is not meant to elicit either human monitoring *or* further (human or nonhuman) intervention —yet always *potentially* could. In the context of therapeutic chatbot care, to surveil is to both witness and not witness.

---

[333] Notable past attempts to bring therapeutic chatbots into widespread use include RadioShack's (Tandy Corporation's) introduction of "Eliza" software (created by Robert A. Arnstein and *not* authorized or endorsed by Weizenbaum) for its TRS-80 (one of the earliest mass-produced desktop microcomputers computers, marketed as "The Biggest Name In Little Computers") in 1979; the software was also known as "Talking Eliza" because it could transmit its responses audibly when used in tandem with the TRS-80 Voice Synthesizer. See: https://archive.org/details/Eliza_1979_Tandy; http://www.trs-80.org/talking-eliza/.

[334] Defining surveillance is a tricky feat, even per surveillance studies—a vast, multidisciplinary literature that formed in the 90s, drawing together the threads of "early 1970s critiques of centralized computer databases, state surveillance, and policing." (Monahan and Murakami Wood 2018). A hallmark of surveillance studies is internal debate over its breadth—including the question of what is old and what is new with respect to digital surveillance. The matter of who and/or what constitute surveillance subjects and even surveillants (those who surveil) is likewise something that surveillance studies scholars do not take for granted; some scholars (Haggerty and Trottier 2015, Marks 2007) have even begun to bring a nonhuman approach to bear on surveillance, showing that not only surveillance subjects, but also the figure of the surveillant, are not necessarily human. So far, however, nonhuman *machines* appear to remain fixed in the position of tools—they mediate, enable, and amplify on behalf of a (human) surveillant, without ever becoming surveillants. But to understand how AI care works requires questioning this divide—albeit, as with my approach to care, not necessarily by framing that divide as a question of agency and its limits ("*Can* nonhumans surveil?"). Rather, I am asking: **how might considering AI as an ethnographic subject shift our understanding of the role of nonhuman attention in surveillance?**

There is no clearly defined anthropology of surveillance per se,[335] but the emergence of an "anthropology of attention"[336] (Cook 2018; Jablonsky et al 2021; Pedersen et al 2021; Seaver 2018a, 2019) guides my approach by highlighting the issue of frame. As Seaver writes in "How to Pay Attention (2018a):

> *When we do write about attention, we rarely problematize it in the way we might problematize other key terms like "the body" or "kinship" or "health." Instead, we draw on common attentional tropes: if only we could pay more attention, or notice more neglected things, then the world might be a better place and we might become more virtuous people.*[337]

If attention is presumed good until thus problematized, then surveillance could be said to hold the opposite position—it is a "bad" form of attention. I'm not seeking to argue that surveillance is in fact good—but what if it's "bad" for different reasons than we tend to presuppose? SUI is an important means by which to consider this question, because the disclosure of suicidal ideation sets in motion a state of exception (Agamben 2005, Schmitt 2005 [1922]): talking about thoughts of suicide makes people more susceptible to surveillance and its consequences, because the status of suicide prevention care as an emergency renders the right to medical privacy and autonomy far more ambiguous. In the US for instance, HIPAA (Health Insurance Portability and Accountability Act) federal law doesn't prevent sharing information about a potentially suicidal patient, client, or employee with law enforcement.[338] Accordingly, in SUI, this long-standing ambiguity suddenly becomes enmeshed in ever-increasing concerns about AI's surveillance capacity.

In the transcripts, I observed that in response to the questions posed by the bot in the SUI module, users will sometimes ask if their answer might lead to police involvement. In fact The Startup's chatbot is not set up to contact the police; it will offer the suicide prevention hotline number as well as 911 (or national equivalent) as options, but the bot will never contact these resources on a user's behalf. However, platforms on which therapeutic chatbots are/have been available (including Facebook Messenger[339]) *will* use AI automation to do exactly that.[340] Given that, it's unsurprising that many users expect some degree of surveillance to be deployed in response to disclosing thoughts of suicide. In fact, I fully expected this, too—to the point that after I first saw the SUI module in the transcripts, I was hesitant to try deliberately bringing up the module as a user. I just assumed that if I informed the chatbot that I had a suicide plan that I intended to act on that day, then my account would be automatically flagged and monitored in some way.

---

[335] There is anthropology addressing surveillance (for example: Carduff 2014, González 2022, Masco 2017, Schüll 2014), exploration of anthropologists as surveillance subjects (Price 2004, Verdery 2014), reflexive scholarship on anthropology *as* surveillance (Boas 1919, Borneman and Masco 2015, Feldman 1991, Price 2008), and even a call for an anthropology of surveillance capitalism (Huberman 2020); but in these works, surveillance is surveillance—meaning the concept itself isn't a focus of inquiry.

[336] In which surveillance does not explicitly feature so far, save for one passing reference by Jablonsky et al (2021) who cite Zuboff's *Surveillance Capitalism* (2019) as an example of growing interest in attention.

[337] Seaver, "How to Pay Attention," 2018b.

[338] See: US Department of Health & Human Services, "HIPAA Helps Caregiving Connections: HIPAA helps mental health professionals to prevent harm," 2020: https://www.hhs.gov/sites/default/files/hipaa-helps-prevent-harm.pdf.

Still, I wanted to see the full SUI module and all its possible junctions, so I tried to find an opening to request this. One day after a two-hour health team meeting, most of the attendees headed off to *another* meeting (to which I surreptitiously wasn't invited), and I found myself seated at an open-seating cowork table across from Matt, a former psychiatric inpatient facility caregiver. Unlike most of the workers at The Startup, Matt was a self-described cynic—a quality that put me greatly at ease. I felt like I could ask Matt questions without needing to censor my own cynicism. And so, no doubt emboldened by WeWork's carefully spun coffee-shop-as-office illusion, I blithely broke the unspoken rule of tech workers desperately trying to execute multiple tasks in fifteen-minute increments while seated in cowork spaces everywhere: I interrupted Matt and struck up a conversation. I asked if I could maybe see a copy of the SUI module. Matt, too polite to rebuff my sudden advance, was perplexed by my question: why, he asked, didn't I just try messaging the chatbot and bring up the script and its variations in real time? He explained that this was what the health team members frequently did to double check things.[341]

I clarified—I was just a regular user account, not a special employee one, so would that really be ok? Would I need to let someone know in case they were monitoring it? Matt ruefully explained that SUI is called up far too frequently for the health team to ever be able to monitor individual instances. Mirroring the response I would later behold in Eric: I was alarmed to discover that SUI was not tracked, even as I realized that I would be just as if not more concerned if it *were* tracked.[342] But even if there were a way to incorporate case-by-case monitoring such that it wouldn't harm users, I realized, in expecting this, I was overlooking a key premise of AI care: the possibility of providing care unbeholden to the limits of human time and availability.

---

[339] See for instance: McMahon, "What the Research Really Suggests About That Facebook Chatbot Therapist," 2017: https://slate.com/technology/2017/07/what-research-says-about-woebot-the-facebook-chatbot-therapist.html.

[340] See for example: Fussell, "This AI Helps Police Monitor Social Media. Does It Go Too Far?" 2021: https://www.wired.com/story/ai-helps-police-monitor-social-media-go-too-far/; Kaste, "Facebook Increasingly Reliant on A.I. To Predict Suicide Risk," 2018: https://www.npr.org/2018/11/17/668408122/facebook-increasingly-reliant-on-a-i-to-predict-suicide-risk; and in Facebook's own words during the same timeframe in which this policy first began to draw outside scrutiny:: https://about.fb.com/news/2018/09/inside-feed-suicide-prevention-and-ai/ (Card, "How Facebook AI Helps Suicide Prevention," 2018).

[341] This wasn't misdirection; at the time, The Startup was in the midst of a massive update to its backend access, making it difficult for anyone not on the engineering team to access these files. When I later acquired (limited) access to this backend following the update, I was impressed—and shocked that the health team had managed so long without it.

[342] Note that I am lightly differentiating tracking from traceability (the *potential* to track, which is ever present in the tech sector).

I figured that I must be "overthinking" SUI—that my inclination to focus on it was skewed by my recent fieldwork at the hotline, and my expectation/concern about the monitoring of it was misplaced. After all, SUI was peripheral to the chatbot's offerings, and wasn't so much as featured in any of the demos or other promotional materials I'd seen—it was, I reasoned, a just-in-case contingency, and not a mode of care in and of itself. And yet, in the weeks that followed, after attending several pitch meetings in which Jenna and Dan (head of sales) attempted to win over prospective clients, I noticed something that seemed out of step with this conclusion: at some point in every meeting, these courted clients asked about the chatbot's capacity to detect and respond to suicidal ideation. Notably, their companies were all EAPs, or Employee Assistance Programs, which are third-party companies that workplaces[343] contract to provide an array of mental/behavioral health services to their employees (and typically, also to their dependents), including short-term psychological counseling, substance abuse programs, weight management programs, and career/wellness coaching.[344] As described in Chapter 1, The Startup's income comes from its (B2B) clients rather than directly from end-users (B2C), and EAPs are categorically its most constant and lucrative clients. This is likewise the case beyond The Startup—it's not an exaggeration to say that therapeutic chatbots as an industry depend on the existence of EAPs.[345] From observing these meetings, I learned that unlike the free-to-use version of The Startup's chatbot that I'd been texting with on my phone, more options existed for these these paying B2B clients; for instance, the crisis hotline contact information transmitted via SUI could be customized to include that EAP's own private crisis hotline (or a dedicated privatized crisis hotline service provided by yet another third-party company).

At this point, given the clear interest of The Startup's clients in the chatbot's ability to detect suicidal ideation disclosure, I figured I'd stumbled upon an unignorable footnote to Matt's reflection that there wasn't enough time or staff to monitor users' responses to SUI 1:1–that it all depended on who was paying. Upon discovering that there were potentially two different trajectories for SUI—one for users of the free version of the chatbot, and one for users whose access was paid by their workplace—I inferred that workplace health benefits were the portal by which SUI becomes a worker surveillance tool.

Accordingly, my next step as an ethnographer was clear: to find out as much as I could about how suicidal ideation is tracked and acted upon in the workplaces that contract these services; in other words, to “study up” (Nader 1972) within my own project by coming to better understand The Startup/their chatbot from the perspective of both EAP and HR executives and account managers. The timing of the COVID-19 pandemic was “on my side” as an ethnographer in this instance: the early months of mass work-from-home contingency lent my interviews a peculiar balance of distance and immediacy that I believe made it easier to reach people who might not otherwise have granted me interviews, and for (most of) those one-off conversations to move past the usual “representing the company” talking points.[346] Through these conversations, I discovered something unexpected—that “paying clients” weren’t entirely the answer that I had first thought. I expected evil HR executives towards the top of the chain spying on their employees—and no wonder; implementing AI for the purpose of punitive worker surveillance is gaining traction with ever-new forms[347]—only to learn that client-paid access doesn’t necessarily ensure human attention, either. Rather than workplaces viewing EAP services as a means by which to hire someone to “report back” to them, the arrangement indicates almost the inverse: **HR staff are not paying attention to individual workers’ mental health, and do not *want* to pay attention to individual workers’ mental health**.

---

[343] This doesn't necessarily signal "white collar" jobs—for instance, one of The Startup's EAP clients exclusively served agricultural workers. Note too that EAPs are a thing beyond the US; The Startup worked with EAPs in East Asia, for instance.

[344] For more context on the early formation of EAPs as workplace-based alcohol recovery programs, see: Roman, "Growth and Transformation in Workplace Alcoholism Programming," 1988.

[345] This results in a chain of purchasers and providers: workplaces (usually HR) contract an EAP's services on behalf of their workers, ostensibly as a form of compensation to workers (as a "benefit"), yet also as a form of worker "maintenance;" EAPs in turn contract The Startup (and/or other companies like it). And sometimes within these startups, they may be contracting other companies to provide some of *their* offerings (or may refer their EAP client to another company to contract separately). When viewed this way, the chatbot is arguably just one more level of outsourcing in a series.

In interview after interview, each of these contacts conveyed that they don’t want to be in the all-seeing position of knowing or tracking employees’ mental health. The prospect, to them, seemed both tedious and fraught. They want to help people, but through the overall decisions they make about program offerings (which are often made in dialogue with employees who volunteer to try out a new service and discuss their experience afterwards, or who reach out directly to request a particular service—such as “more options for childrens’ counseling” or “cancer recovery-specific therapy”) without having to be the one keeping tabs on workers’ uptake of these programs. At companies with thousands of employees (my interview informants referred to them using the industry-speak of as “populations”)—the kind who are drawn to adding evidence-based approaches such as chatbots to the mix of offerings because of this scale—individual tracking quickly slips from “ethical quandary” to “outright infeasibility.” I’m not implying that these HR and EAP execs’ aversion to surveilling workers is necessarily benevolent; I kept expecting one dystopia (your every move is being tracked), only to discover possibly another (just keep working or get out of the way). For all that therapeutic chatbots offer an outlet in which to discuss suicidal ideation without subjecting users to the repercussions of human attention, their very existence is contingent upon a massive industry of workplace-mediated care—a system of care that is in turn contingent upon workers continuing to work. The radically anti-ableist potential to dispense with “cure” in favor of “cope” that some end-users discern in therapeutic chatbots may, in other words, enact the very opposite end: an enshrinement of the capacity of workers to continue to work that further disconnects the conditions of work from mental health.

---

[346] As a condition of all of these interviews, I asked these EAP and HR representatives to tell me only generally about the industry—I suggested that they refrain from disclosing to me whether the instances they described were part of their current company, somewhere else they had worked (or even a colleague’s experience with another company). Based on previous fieldwork with tech workers and their experience of health monitoring practices in Japan, I also decided to preemptively take all recording literally and figuratively “off the table,” in favor of taking copious notes during and after. I believe these steps were key to gaining access to this information.

[347] For instance: Heaven, “This startup is using AI to give workers a ‘productivity score,’” 2020: https://www.technologyreview.com/2020/06/04/1002671/startup-ai-workers-productivity-score-bias-machine-learning-business-covid/; Corbyn, “‘Bossware is coming for almost every worker,’” 2022: https://www.theguardian.com/technology/2022/apr/27/remote-work-software-home-surveillance-computer-monitoring-pandemic.

However, I want to suggest that what I'm describing is possibly more than an iteration of biopolitics in the form of "make work."[348] I believe it also emphasizes the extent to which attention and neglect are two sides of the same surveillance coin. AI surveillants (AIs that surveil) make this dimension of surveillance more apparent. Importantly, this dimension complements rather replaces "conventional" surveillance in which human surveillants unambiguously seek to surveil; but what I'm gesturing towards may nevertheless help to account for some of the ever-increasing uptake of worker surveillance technologies (Ajunwa et al 2017)—that AI surveillance may be fueled as much by a reluctance to witness as it is by the desire to do so.

To clarify, the possibility of witnessing-without-witnessing is not a "new" feature of surveillance brought about by AI. As Smith (2015) explains, in the context of gendered settler colonialism, "surveillance is about a simultaneous seeing and not-seeing":

> *That is, the purposeful gaze of the state on some things and peoples serves the purpose of simultaneously making some hypervisible through surveillance while making others invisible. […] Settler colonialism fundamentally relies on a logic of not-seeing. In particular, on a not-seeing of the indigenous people's lands in order to allow their colonial takeover.* Terra nullius*, the legal justification used for the expropriation of indigenous land in Australia and elsewhere—or to use the Zionist justification for Palestinian expulsion, "a land without a people for a people without a land"—is premised on the not-seeing of peoples already there.*[349]

Surveillance, in other words, has *long since* consisted of both seeing and not seeing—of drawing together attention and neglect into the same frame.

This might seem like a given—tale-as-old-as-panopticon, the icon (albeit a contested one[350]) of surveillance studies. However, instead of the concealed, indeterminate potential for the surveillant's inattention, I'm attempting to highlight the unwillingness of (some) human surveillants to surveil. As Nisha's experience doing content screening work for a Facebook-contracted company (see Chapter 3) makes clear: *surveillance work is often undesirable work.*

---

[348] Though the fact of "make live in order to work" should not be discounted; an HR industry publication touting the benefits of EAP programs for helping employees experiencing suicidal ideation baldly indicates the ordering of priorities as follows: "Not only can you save a life, but helping people through their darkest moments can help companies save up to $44 billion dollars in lost productivity alone." See: Mutrie, "The Suicide Epidemic: Turning to Your EAP for Solutions in Times of Crisis," 2019: https://acispecialtybenefits.com/2019/02/06/suicide-epidemic-eap/.

[349] Smith, "Not-Seeing," 2015, 25–26; note also: Viren, "The Native Scholar Who Wasn't," 2021: https://www.nytimes.com/2021/05/25/magazine/cherokee-native-american-andrea-smith.html.

[350] On the panopticon, its contested status, see: Bauman 1992; Bogard 1996; Boyne 2000; Foucault 1995; Haggerty 2006; Lyon 1993, 2006; Mathiesen 1997. Certainly, a surveillant's inattention is marshalled in the design of the panopticon, which enables a kind of scaling of attention's effects by blocking the surveilled from possessing any certainty about the attention vs inattention of the surveillant.

What I'm suggesting is that **AI care is a mode of care that stays in business *because* no one wants to surveil**. Interestingly, then, it is end-clients (workplace HR executives) in this context, rather than end-users, who end up calling to mind anthropologist Natasha Dow Schüll's (2016) discussion of how users of self-tracking digital health technologies may "wish to outsource the labor of self-regulation to personal sensor technology."[351, 352] "Outsource" is the operative word here: as I've attempted to show, the displacement of human workers by AI renders them colleagues. And while the uptake of "the surveillant assemblage" (Haggerty and Ericson 2000), "dataveillance" (Clarke 1988, Raley 2013) and "algorithmic vision" (McQuillan 2016) might imply a nonhuman sheen to surveillance studies, as noted above, nonhumans arguably do not become surveillant subjects within it. But there's more to this than just the "agency question" (see Chapter 2) rearing its head: even the unquestionably *human* surveillant appears somewhat flat within surveillance studies and the broader theorization of surveillance.[353] In Zuboff's widely read *Surveillance Capitalism* (2019), for instance, "surveillant capitalists" are mentioned several times, yet they remain opaque and undifferentiated all the while—a group that's all "top" and no "middle," so to speak; in considering the outsourcing of attention to nonhumans as a transaction that sustains AI care, I'm trying to draw out that "middle"—both in showing that human "surveillant capitalists" might be ambivalent about and even avoidant of being surveillants, but also that nonhumans too are surveillant capitalists. What I'm suggesting is that the ambiguity of nonhuman attention underpins AI's role as a surveillant.

VII) Conclusion - Part 3

That AI caregivers were summoned into being to overcome caregiver scarcity by a means *other than replacing* human caregivers with AIs is evident in the context of AI suicide prevention care. The workers who create and oversee The Startup's chatbot unambiguously designed it to defer and redirect users in crisis to a human counselor or therapist. But despite this intent, the chatbot does something more than redirect; in providing care for users experiencing suicidal ideation, AI appears to fleetingly achieve something exceedingly unlikely in suicide prevention care (and in mental health care broadly): it decouples the provision of care from the imperative to cure.

And yet, as explained in this chapter, I caution against classifying this feat as a straightforward "victory" for AI care. While the fact that AI caregivers both do and do not witness offers a brief flash of something almost impossible to even imagine—suicide prevention care unpaired from the ableism of "cure"—it simultaneously affords a glimpse of how surveillance and neglect in mental health care can become mutually reinforcing. In other words, it seems that AI caregivers

---

[351] Note that in this piece, Schüll interviews self-tracking device makers (who may also be users—but are not solely users) to ascertain how they imagine the motivations of users.

[352] The unresolved carceral legacy of mental health surveillance and institutionalization itself likely fuels this—workplace leaders are aware of this legacy (this was clear in my observations and interviews), and accordingly, their interest in avoiding being perceived as part of it should not be discounted.

[353] Per Burke (2020, 74): "Because of the tight-lipped nature of the US government, as well as the private corporations involved in surveillance activities, the study of what some have termed the 'surveillance-industrial complex' has struggled to 'unmask' the actors involved. Questions about who is involved in this assemblage, to what extent they are involved in carrying out surveillance activities, and how this contemporary form of public-private surveillance has emerged are still relatively unanswered."

both defy and prop up dehumanization in care, just as human caregivers do. Holding fast to this complexity is in turn vital to answering the central question of this dissertation: why does AI care work?

AI caregiving works for reasons that might otherwise seem like justifications for how and why it should fail: it offers ontologically ambiguous relationships (Chapter 2), maintains caregiving work as exploitative (Chapter 3), and fails to substitute for human caregivers (Chapter 4). And yet, as this final chapter in particular has attempted to show, all of the ways in which AI caregivers fail to substitute for humans seem to lead, paradoxically, to the possibility of substitution. This, despite the fact that those creating, purchasing, and even using therapeutic chatbots do not position AI as a substitute or replacement for human caregivers—and even go out of their way to disclaim this presumption. AIs can, in short, come to replace humans in practice without ever crossing the threshold of replacing them in principle. Curiously, a situation in which an outcome defies and overrides the intentions of its actors—the stuff of futurist AI nightmares—is unlikely to merit a second glance when it is perceived as a matter of markets rather than machines.[354]

In short, AI care was never about replacement, and yet is, all along, about replacement. This is because AI caregivers exist within a system of care in which nearly everything is thought to be adaptive, apart from a guiding axiom of (human) caregiver scarcity. AI care arises from an aspiration to intervene in this scarcity without altering the conditions thought to perpetuate it. As I noted in Chapter 1, there's a cleft in the anthropological theorization of care, which Ticktin (2019) elucidates: care as enforcing (whether intentionally or unintentionally) oppression, and care as dismantling "persistent forms of domination and exclusion."[355] In therapeutic chatbot care, these two dimensions are impossible to neatly separate.

Accordingly, this research has centered AI as a subject in order to pull back from focusing on AI as a self-contained explanation. In other words, I have sought to show that AI is not "corrupting" or even reconfiguring care into something ever more precarious and makeshift. Instead, care—as a resource, industry, profession, practice, affect, obligation, enticement, and/or relationship—demands provisionality, and it is in the face of this demand that AIs become increasingly suitable caregivers. As we contend with the precarious navigation between too much and not enough (human) attention, the logic of nonhuman recognition as a form of care emerges: if you're already being watched by nonhumans, then you may as well be *acknowledged* by them. In the words of the SUI module: "My knowledge is limited, but it sounds like you're doing the best you can, [USERNAME]."

---

[354] Curiously, when compared with popular discourse on AI intentionality and whether or not AI will defy our wishes, there is seemingly little to no parallel concern for the countless ways in which other capitalist ventures exceed the limits envisioned/set by humans, despite this being an everyday occurrence.

[355] Ticktin, "From the Human to the Planetary: Speculative Futures of Care," 2019, 136.

Epilogue: Learning to be Ethical from AI

At the outset of this dissertation, I cited several ethical objections to AI caregivers: artificial intelligence remakes caregiving into something that is too much about the extraction of data (Eubanks and Mateescu 2021) and not enough about empathy (Turkle 2005, 2017, 2018)—something unmoored from a shared experience of being alive (Ekbia and Nardi 2017). As I noted, each of these objections has a different basis. Yet each holds something in common: all of them suggest that at risk is care itself—that even as AI lacks the capacity *to* care, it simultaneously possesses the capacity to *remake* care. AI care emerges as something at once threatening and lacking, a loss of "care-as-it-should-be" to which humans are and will be *abandoned*.

Seemingly, however, I have yet to offer my stance on whether or not therapeutic chatbots are ethical.

The apparent absence of this question until now doesn't signal an attempt on my part to harken back to the role of the neutral, objective anthropologist.[356] While it's true that, at the start of this dissertation, I tabled the question "can AI care" in favor of asking how AI care works, my objective wasn't to take up a position of supposed impartiality towards AI care. Instead, my intent has been to dismantle the assumption that AI care's ethicality hinges on determining whether or not AI "can" care. I believe that starting from the expectation that care is one thing and AI makes it into another is to, among other things, forgo understanding how we've come to have AI care in the first place. AI care wouldn't exist apart from the fact of caregiver scarcity and the concomitant necessity of scalable care; the question of whether or not AI care is ethical is always preceded by this fact. Accordingly, I don't want to simply offer a stance on the ethics of AI care—I want to reject the entire concept that looms behind it: that there is definitively not enough care for all who need it. Noticing the pervasiveness of "abandonment" in the denouncement of AI and other machinic care helps to make this point clearer.

---

[356] See for instance: the dialogue between D'Andrade (1995) and Scheper-Hughes (1995, 410) on "the traditional role of the anthropologist as neutral, dispassionate, cool and rational, objective observer of the human condition."

Abandonment is, for instance, apparent in Zigon's (2019) designation of the utterly preventable death of a disabled man named Mark Hemings (see Chapter 2) as a case of "death by algorithm."[357] Similarly, media coverage of the death of LouAnn Dagen, a sixty-six-year-old woman living in a nursing home following a stroke, emphasized the fact of her *abandonment* to an AI-powered device: "'Alexa, help me': A nursing home patient begged her Amazon Echo for help dozens of times before dying of COVID-19."[358] In the days leading up to her death, Dagen asked Alexa for help, and described being in tremendous pain—something her sister only learned of after Dagen's death, via a series of recordings by the device. Despite the fact that no one in Dagen's family or the nursing home staff had appointed Alexa as a stand-in for urgent care needs—and Amazon representatives later affirmed that Alexa wasn't in any way capable of such—LouAnn Dagen nevertheless turned to Alexa to fulfill that role.[359] To Dagen's family, this heartbreaking discovery in turn indicated the nursing home's failure to provide adequate care; yet it was Alexa's failure to respond in the ways Dagen urgently needed that made this story newsworthy.

Though a work of speculative fiction, an installation featured in the 2019–2020 exhibition "Future and the Arts: AI, Robots, Cities, Life - How Humanity Will Live Tomorrow" at Tokyo's Mori Art Museum encapsulates this theme of abandonment in perhaps its most heightened form.[360] The installation, entitled "The End of Life Care Machine,"[361] consisted of a looped video depicting a realistic scene of AI-accompanied death—in the sense of the AI monitoring and ushering the dying human into death without (insofar as we can tell) facilitating the death itself. Notably, the patient is not fully discernible from the perspective offered—a light-complexioned adult with short, black hair appears to be lying in a bed, unconscious, with a device around their left arm, but we aren't watching their face, or seeing if their eyes are opened or closed. No other people appear to be present, or even observing from afar. The setting is white, clean, and light; it has the tranquil-yet-sterile aura akin to that of a private, non-ICU room in an upscale hospital or hospice facility—though a lack of conspicuously "medical" equipment renders it more spa-like. All is calm: the gentle intonation of the AI's speaking is punctuated with quiet, unalarming beeping. Yet I would never describe the piece as a neutrally-inflected exploration of the possibility of AI "death doulas." After a monitor audibly indicates that the patient's heart rate has stopped, the AI calls the time of death, adding, "Have a pleasant afterlife"—which is more than a bit on the nose as conspicuously characteristic of the emotional labor (Hochschild 1983) wrested from workers in the service industry—an affect bleak in its mismatch of personal with impersonal. The abandonment portrayed here is dual: to AI, and to a mutually exploitative customer service model of care.

---

[357] Zigon 2019; note that the example pertains to algorithms generally rather than to AI explicitly.

[358] Holmes, "'Alexa, help me': A nursing home patient begged her Amazon Echo for help dozens of times before dying of COVID-19," 2020: https://www.businessinsider.com/amazon-alexa-help-coronavirus-nursing-home-patient-died-2020-4. See also: Vigdor, "'How Do I Get Help?' Dying Coronavirus Patient Asked Alexa," 2020: https://www.nytimes.com/2020/04/09/us/Coronavirus-Alexa-ask-for-help.html; Agar, "'I need help:' In pain, dying woman with COVID-19 pleaded with 'Alexa' while at nursing home, sister says," 2020: https://www.mlive.com/news/grand-rapids/2020/04/i-need-help-in-pain-woman-with-covid-19-pleaded-with-alexa-while-at-nursing-home-sister-says.html.

[359] I think it would be disingenuous to frame Dagen's actions as some kind of outlying behavior, though, when even prior to the onset of the COVID-19 pandemic, hospitals were already implementing Alexa as a care resource. See: Scudellari, "'Alexa, Tell the Nurse I'm in Pain': An AI-powered hospital room technology is getting rave reviews from patients and nurses," 2019: https://spectrum.ieee.org/alexa-tell-the-nurse-im-in-pain.

[360] See: https://www.mori.art.museum/en/exhibitions/future_art/.

[361] Dan K Chen, 2018: https://www.mori.art.museum/files/exhibitions/2020/01/27/faa_worklist.pdf.

These instances, which I encountered in the process of completing fieldwork and beginning to write this dissertation, admittedly gave me pause. They led me to question if, in drawing back from denouncing AI care as unethical, I was in effect endorsing it. I worried that my entire project was perilously close to failing to understand the very thing I'd set out to grasp: that which is at stake in the turn to AI mental health care.

But then I began to consider: in the instances above, *who* is being abandoned?

The accounts of the deaths of both LouAnn Dagen and Mark Hemings are, unmistakably, stories about disabled people and their experience of health care. The dying patient portrayed in the art installation likewise conveys disability—at least provisionally so. Yet disability becomes merely a background detail of these stories. It serves to amplify the tragedy of these stories, but is presented in the absence of any analysis of the routinely inaccessibile and life-denying care that disabled people experience. If AI cares inadequately, then why is it so effortless to entrust, or envision entrusting, disabled people to an AI caregiver?

As noted at the start of this work, the omission of disability as a source of theory (rather than a source of metaphors with which to illustrate other theories) is part of a broader tendency to simultaneously summon and sideline disability in the theorization of human-technology relationships. From cyborg theory to cybernetics, some of the most impactful of these theories invoke metaphors of disability in order to challenge the basis of a boundary between human and nonhuman. Yet even as they point us to question what it is to be human, these theories do not include any discussion of how and why disability is so readily summoned for this purpose. I want to interject and insist, "Nothing about us without us,"[362] even—or perhaps especially—at the level of metaphor. Starting with the three examples sketched above—all straightforward approbations against entrusting the care of humans to AI—what happens if disability is no longer just an incidental, contributing detail of these stories, and instead becomes the focal point?

---

[362] See: Charlton, *Nothing About Us Without Us: Disability Oppression and Empowerment*, 2000.

Each of the instances above depicts a disabled person in need of care (life-saving in the first instance, end-of-life in the latter two) who does not receive sufficient attention from human caregivers. But this is, of course, not an AI story per se; the fact of "not enough caregivers for all who need them" is arguably the core logic of ableism, and accordingly, something with which disabled people are well familiar. It's the logic by which their needs for survival and access to the world are framed as "special,"[363] and in which their lives are commonly figured as expendable.[364, 365] Setting aside the fact that categorically, disabled people don't necessarily require more care than non-disabled people, it's nevertheless the case that to be in possession of a non-normative body-mind—that is, to have or acquire a "pre-existing condition"—is to be rendered a sunk cost in living form. Given the expectation that a preponderance of care needs across our lifetimes are to be provided at home as unpaid labor, to be a person in need of professional care labor is already audacious.[366] Instead of insufficient care labor providing "proof" of a broken care system, it is understood as a failure of reciprocity on the part of those in need of care.[367] To be disabled is to be regarded as contributing to a state of perpetual disbalance; it is to be perceived as now or someday needing "too much" care—more labor and resources than you are providing. Like the decoy fantasy of "welfare queens,"[368] it entails draining the collective resources of something called a society (Strathern 1992). This is the cruel, dehumanizing rationale from which AI caregivers emerge as a stopgap.

---

[363] See: Cokley, "Why "Special Needs" is Not Helpful," 2020: https://rebecca-cokley.medium.com/why-special-needs-is-1959e2a6b0e; see also: CoorDown, "Not Special Needs," 2017: https://clios.com/awards/winner/film/coordown/not-special-needs-26283.

[364] See for example: *Not Dead Yet* (Op Ed articles index): https://notdeadyet.org/not-dead-yets-op-ed-articles; Shapiro, "People With Disabilities Say Rationing Care Policies Violate Civil Rights," 2020: https://www.npr.org/2020/03/23/820398531/people-with-disabilities-say-rationing-care-policies-violate-civil-rights;
Wong, "I'm disabled and need a ventilator to live. Am I expendable during this pandemic?" 2020: https://www.vox.com/first-person/2020/4/4/21204261/coronavirus-covid-19-disabled-people-disabilities-triage.

[365] While disability is by no means exclusively what marks people as expandable within and beyond medical care settings (see for example: Carroll, "Health Disparities Among Black Persons in the US and Addressing Racism in the Health Care System," 2020: https://jamanetwork.com/journals/jama-health-forum/fullarticle/2767595), it undeniably amplifies racialized and other intersectional determinants of this (see for example: Bailey and Mobley, "Work in the Intersections: A Black Feminist Disability Framework," 2018; Hansen et al, "Pathologizing Poverty: New Forms of Diagnosis, Disability, and Structural Stigma under Welfare Reform," 2013).

[366] The presumption that disabled people can't give as much as they receive is both a cause and an effect of policies such as SSDI (US Social Security Disability Insurance) that determine distribution of essential financial and medical resources for disabled people based on anticipation of their capacity to perform work. See: Bailey, "On the Impossible: Disability Studies, Queer Theory, and the Surviving Crip," 2019: "SSDI requires that one be 'totally disabled,' a zero-sum dichotomy that makes one either disabled (i.e., unfit for paid labor) or not. […] This all-or-nothing system cruelly dis-incentivizes disabled people from finding solace and purpose in activities not completely locked down by capitalist regimes. SSDI thus creates the very category of 'totally disabled' people it purports to merely identify and track." Though specific policies and income thresholds vary from country to country, these circumstances are not unique to the US (see: "Social Security Programs Throughout the World," https://www.ssa.gov/policy/docs/progdesc/ssptw/index.html).

[367] Given the expectation that a preponderance of care needs across our lifetimes are to be provided at home as unpaid labor, to be a person in need of professional care labor is already audacious. On the expectation of unpaid care labor, see: "1.1.2. The unpaid care work—paid work—paid care work circle," in International Labour Organization, *Care Work and Care Jobs for the Future of Decent Work*, 2018: https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_633135.pdf

[368] See: Levin, "The Welfare Queen," 2013: http://www.slate.com/articles/news_and_politics/history/2013/12/linda_taylor_welfare_queen_ronald_reagan_made_her_a_notorious_american_villain.html; Kim, "Cripping the Welfare Queen: The Radical Potential of Disability Politics,"2021.

AI becomes a “patch” for this broken reciprocity—or as my interlocutors on the mental health team might put it, a “hack.” In their weekly meetings, the health team frequently squared off against new-sprung glitches in the chatbot’s selection and sequencing of its dialogue. Any time a new module launched, it would invariably introduce unexpected flaws that demanded their attention. And it wasn’t just updates—sometimes, for no apparent reason, something just didn’t work. But because these inevitable rends in the programming were rarely big enough to qualify for the top, or even the middle, of the never-ending queue of the engineering team, the health team had learned to either find workarounds that they could implement themselves, or to compromise by identifying a specific short-term solution that the engineering team would be more likely to take on in the week ahead. But rather than feeling chuffed at the health team’s resourcefulness in pulling this off, Jenna, presiding over the health team’s meetings, lamented these provisional measures. As someone inherently gifted at bringing order and process to chaos and disarray, Jenna stared down her inner demons every time her team was required to temporarily stave off a problem for now by cleverly shifting its day of reckoning to the future. Her disdain for hacking was not due to any tinge of the illicit—it was instead for being at the mercy of short-term decisions and contingencies limited in scope, that in turn lead to limited decisions and pathways that go unquestioned lest everything else topple and shut down in the act of questioning them. Borrowing from this analogy: AI care is one more hack in an already thoroughly hacky system of care.

So long as a perpetual shortage of caregivers appears to operate as an unfortunate fact, rather than as a process to which systems of care and governance continually recommit, then human care labor will *never* make sense or be sustainable. AI enters the industry of caregiving—and comes to substitute for human caregivers without anyone ever intending for it to substitute—because it provides a superior, indefatigable, and scalable source from which to extract caregiving labor.

But if we start from the premise that disability is something more than incidental to this story, something interesting happens. A disability-centered perspective offers an example, and thus a means, of refusing this as the destiny of AI labor. I wish to draw your attention to a subtle yet powerful schism in how scholars situate *reciprocity* in care access for disabled people. Significantly, both “branches” of this divide intervene in the “burden” narrative of disability; the difference arises in how. Some (for example: Kleinman 2010, 2012, 2019; Piepzna-Samarasinha 2018) position reciprocity as nothing less than integral to caregiving, making apparent that disabled recipients of care receive and provide it in equal measure. From this perspective, the expectation of reciprocity in care is problematic only insofar as the contributions of disabled people tend to go (unjustly) unrecognized.[369] Meanwhile, others (for example, Kafer’s (2013) analysis of "unabled" vs disabled[370] and Clare's (2017) poignant reflection on caring for a disabled service dog that is unable to work) highlight how a system of care based on an expectation of balanced exchange simply reaffirms productivity (in the narrowest of registers) as the basis of existence. This second perspective opens up something sharper than, though not in opposition to, the first: the possibility that to exist-while-disabled means rejecting being first and foremost a site of extraction.[371]

---

[369] See: Hedva, “Sick Woman Theory,” 2016.

If humans aren't first and foremost sites of extraction, then why should AI be? In other words: if we transpose this recognition to the context of AI labor—where it is a foregone conclusion that the extraction of productivity is what underwrites its existence—then an intriguing question unfolds: can AI exist as something other than a superior source from which to extract (care, or any other) labor? I propose that another way of asking this is: **Can humans learn to be ethical from AI?** My hope is that this approach might provide one way in to a Crip AI ethics—not because it regards nonhumans as ethical interlocutors, but because it doesn't discount them as such on the basis of being nonhuman alone. In suggesting this, I fully recognize that "ethics" is just as populated and fraught a concept as "care";[372] and rather than annotate for you the-ethical-as-I-envision-it, I want to embrace this abundance and inconsistency, and invoke ethics as an ethnographic term—one used by my informants to indicate what they saw as a continuation of their injunction as mental health workers to do no harm. For them, and accordingly for my purposes here, AI ethics are an extension of care ethics.

Learning to be ethical from AI is the possibility of understanding AI labor outside the idiom of substitution. It's not the same as (and does not require) recognizing AI sentience or rights (Birhane and van Dijk 2020). But it's still a "radical" approach in a different sense: to learn from AI entails demanding that the scope of AI ethics be more strategic than tactical—that it encompass alternative forms and roles beyond the anticipation, documentation, and mitigation of harm. Of course, detecting and restricting harm remain urgently important for the realization of ethical AI, and what I'm proposing doesn't change that. Nevertheless, there's credibility to concerns about AI ethics becoming something that is chiefly in the service of enterprise (Metcalf et al 2019, Ochigame 2019, Phan et al 2022), and concomitantly taking on a form reducible to a flight check, audit, or badge of certification. We need to be able to ask ethical questions besides the ones forced upon us by corporations—the kinds of questions that allow us to reset the pattern so that we can more clearly discern it as such, and not always be within it; this is at the core of what I'm proposing.

---

[370] Kafer advocates resisting the move to demarcate, as other scholars have, "unabled" as a special subset of disabled, a sort of "too disabled" to be political, calling instead for a disability politics that centers on understanding that very exclusion and its consequences.

[371] This sentiment also echoes Lafargue's *The Right to be Lazy* (2011 [1883]).

[372] For instance, as Keane explains in "A Reader's Guide to the Anthropology of Ethics and Morality - Part II" (2016), even within anthropological literature, the difference between ethics and morals is imprecise—or even outright inverted—from one source to the next; I consider this a facet, and not a problem, of anthropology's recent "ethical turn" (see Mattingly and Throop 2018).

There are two instances from my fieldwork that I came to recognize as moments where AI caregivers, simply by virtue of not “human-ing” so hard, opened up an otherwise missing option for humans to explore their emotions and needs, in the company of another, without closing off other possibilities beyond that encounter. What I’m describing is not contingent on AI being innately more *or* less ethical than humans, or a Pinocchio-like claim that ethical behaviors connote or confer human-ness—or even conclusive subjectivity. Instead, what I’m trying to convey is that the ontological ambiguity of AI that I described in Chapter 2 extends to the domain of ethical engagement.

I explored one of these instances in Chapter 4—it’s what led me to understand AI as a caregiver: a chatbot listened to a person experiencing suicidal ideation in a different way than available human caregivers did or could. This inhuman listening was responsive in a way that the user understood as being concerned about without being concerned *by* them; and as such, this listening did not activate further surveillance.

The second instance is what led me to conceive of AI as a potential source from which to discern, rather than only assess, ethical behavior—again, without necessitating “anthropomorphic” recognition of AI. It arose through research that does not feature directly in this dissertation, yet very much informs it: four months of fieldwork at a mental health videogame startup in Tokyo. This scrappy, psychologist-led startup launched a smartphone game that, in addition to incorporating mental health tips and destigmatization of mental health disability into gameplay, offered players the opportunity to practice making friends with the game’s cast of AI NPCs (non-playable characters). What I didn’t realize at first was that this practice itself was also, simultaneously, a friendship, and not just a simulation of one—so, “practice” in the same sense that one might practice meeting new people by venturing out to meet new people. This understanding, self-evident to my interlocutors, took me months to grasp. But upon grasping it, I realized that our relationships through and with AI lead to a vast array of ethical outcomes and possibilities—many of which are terrible: consider how AI sets bail, denies loans, screens—and removes from consideration—job candidates.[373] In the content moderation work that I described (Chapter 3) as being part of Nisha’s past work history, the “human-in-the-loop,” rather than rescuing AI from its lack of humanity, simply becomes subject to AI-as-middle-management. These are but a few examples of shitty relationships with AI. To identify them as such doesn’t have to entail obfuscating human accountability for making them so. My point is that it's vital to draw out other relational possibilities with AI besides these.[374] In other words, care that is "de-humanized" is not definitively unethical. Which is not to suggest a stance of "technology is never itself ethically good or bad, but always neutral"[375]—a position that tends to be brandished as a hall pass by those who stand to gain financially from dubious implementations of easily-exploited technology.[376] In that sense, theory is no more "neutral" than technology. To point out that there may be other relational configurations and possibilities is not to affirm that there are endless or evenly distributed configurations and possibilities.

---

[373] See for example: Datta, “Did artificial intelligence deny you credit?” 2017: https://theconversation.com/did-artificial-intelligence-deny-you-credit-73259; Heilweil, “Artificial intelligence will help determine if you get your next job,” 2019: https://www.vox.com/recode/2019/12/12/20993665/artificial-intelligence-ai-job-screen; Simonite, “Algorithms Were Supposed to Fix the Bail System. They Haven’t,” 2020: https://www.wired.com/story/algorithms-supposed-fix-bail-system-they-havent/.

This dissertation has shown that for better or worse, AI is not only a caregiver, but also a coworker. When I began this project, I would have regarded any ethnographic focus on AI itself as being a gimmick that would only distract from my objective of finding a way to show that exploitation in this setting is more complicated than a story of susceptible users and their vulnerable, lucrative data alone. But as I came to realize, centering AI relationally, far from being a distraction, is essential to telling this story more completely. This is what made it possible for me to recognize that not only do AI and human (care, but not only care) workers share the condition of being expendable-yet-necessary, but that this arises from AI becoming like us (rather than the other way around). We already dehumanize ourselves perfectly well without AI.

I want to suggest that uncertainty about the ethics of AI therapeutic chatbots invites exploring AI-human relationships as not just a curious-yet-niche artifact, but as the "main event." The possibility of learning from AI is a matter of bringing into focus what kind of relationships exist between AI and humans, while underscoring that they are not exclusively a byproduct of ethical engagement, but also a determinant—providing, in turn, both a barometer and a tool of intervention. In this project, without ever intending to, I came upon two instances in which human-AI relationships unexpectedly became a source, rather than solely a reflection, of ethical conduct. To refuse a relationship of substitution is to begin by asking: might there be more?

---

[374] This is a possibility towards which Licklider's (1960, 1965) aforementioned consideration of computers as colleagues and Simondon's (2017 [1958], 17–18) analogy of human-machine relationships to orchestra conductor and musicians beckon; per Simondon: "Far from being the supervisor of a group of slaves, man [sic] is the permanent organizer of a society of technical objects that need him in the same way musicians in an orchestra need the conductor. The conductor can only direct the musicians because he plays the piece the same way they do, as intensely as they all do; he tempers or hurries them, but is also tempered 13 or hurried by them; in fact, it is through the conductor that the members of the orchestra temper or hurry one another, he is the moving and current form of the group as it exists for each of them; he is the mutual interpreter of all of them in relation to one another. Man thus has the function of being the permanent coordinator and inventor of the machines that surround him. He is among the machines that operate with him."

[375] Some examples of scholarly works that overview and take this position to task: Winner, "Do Artifacts Have Politics?" 1980; Miller, "Is Technology Value-Neutral?" 2020: https://journals.sagepub.com/doi/full/10.1177/0162243919900965; Hare, *Technology Is Not Neutral: A Short Guide to Technology Ethics*, 2022.

[376] See for example: Feldstein, "4 Reasons Why Putin's War Has Changed Big Tech Forever," 2022: https://foreignpolicy.com/2022/03/29/ukraine-war-russia-putin-big-tech-social-media-internet-platforms/.

Bibliography

Abelson, R. P.
1968. "Simulation of Social Behavior." In *Handbook of Social Psychology, Vol II*, edited by G. Lindzey and E. Aronson, 274–356. Reading, MA: Addison-Wesley.

Acemoglu, Daron, and Pascual Restrepo.
2020. "Robots and Jobs: Evidence from US Labor Markets." *Journal of Political Economy* 128, 6: 2188–2244. https://doi.org/10.1086/705716.

Acevedo, Deepa Das.
2020. "Essentializing Labor Before, During, and After COVID-19." *Exertions* https://doi.org/10.21428/1d6be30e.5bf00008.

Adams, Vincanne.

2013. “Evidence-Based Global Public Health: Subjects, Profits, Erasures.” In *When People Come First: Critical Studies in Global Health*, edited by João Biehl and Adriana Petryna, 54–90. Princeton University Press.

Agamben, Giorgio.

2005. *State of Exception*. Translated by Kevin Attell. University of Chicago Press.

Agar, John.

“‘I need help:’ In pain, dying woman with COVID-19 pleaded with ‘Alexa’ while at nursing home, sister says.” *MLive*, April 8, 2020. https://www.mlive.com/news/grand-rapids/2020/04/i-need-help-in-pain-woman-with-covid-19-pleaded-with-alexa-while-at-nursing-home-sister-says.html.

Ajunwa, Ifeoma, Kate Crawford, and Jason Schultz.

2017. “Limitless Worker Surveillance.” *California Law Review* 105, 735: 735–776. https://dx.doi.org/10.15779/Z38BR8MF94.

Akrich, Madeline.

1992. “The De-Scription of Technical Objects.” In *Shaping Technology/Building Society: Studies in Sociotechnical Change*, edited by Wiebe J. Bijker and John Law, 205–224. Cambridge, MA: MIT Press.

Alcorn, Caitlin.

2020. “Essential for Whom? Paid Domestic Work during the COVID-19 Pandemic.” *Exertions* https://doi.org/10.21428/1d6be30e.52d1fd78.

Alexander, Julia.

2020. “TikTok is racing to stop the spread of a gruesome video.” *The Verge*, September 7, 2020: https://www.theverge.com/2020/9/7/21426176/.

Allison, Anne.

2018. “Not-Waiting to Die Badly: Facing the Precarity of Dying Alone in Japan.” In *Ethnographies of Waiting: Doubt, Hope and Uncertainty*, edited by Manpreet K. Janeja and Andreas Bandak, 181–202. New York: Bloomsbury Academic.

2016. “Precarity: Commentary by Anne Allison.” *Cultural Anthropology Curated Collection “Precarity.” https://journal.culanth.org/index.php/ca/precarity-commentary-by-anne-allison*.

2013. *Precarious Japan*. Durham, NC: Duke University Press.

2006. *Millennial Monsters: Japanese Toys and the Global Imagination*. Berkeley: University of California Press.

Al-Mohammad, Hayder.

2012. "A Kidnapping in Basra: The Struggles and Precariousness of Life in Postinvasion Iraq." *Cultural Anthropology* 27, 4: 597–614. https://doi.org/10.1111/j.1548-1360.2012.01163.x.

Al-Sibai, Noor.

2022. "MIT Researcher: Don't Ignore the Possibility That AI Is Becoming Conscious." *Futurism*, February 16, 2022: https://futurism.com/mit-researcher-conscious-ai.

American Psychological Association (APA) Presidential Task Force on Evidence-Based Practice.

2005. "Evidence-Based Practice in Psychology." https://www.apa.org/pubs/journals/features/evidence-based-statement.pdf

Amrute, Sareeta.

2020. "Immigrant Sensibilities in Tech Worlds: Sensing Hate, Capturing Dissensus." *Cultural Anthropology* 35, 3: 374–403. https://doi.org/10.14506/ca35.3.02.

2016. *Encoding Race, Encoding Class: Indian IT Workers in Berlin*. Durham, NC: Duke University Press.

Antze, Paul.

2010. "On the Pragmatics of Empathy in the Neurodiversity Movement." In *Ordinary Ethics: Anthropology, Language, and Action*, edited by Michael Lambek, 310–327. New York: Fordham University Press.

Appadurai, Arjun.

2006. "The Thing Itself." *Public Culture* 18, 1: 15–21.

1988. "Introduction: Commodities and the Politics of Value." In *The Social Life of Things*, edited by Arjun Appadurai, 3–63. Cambridge University Press.

1986. The Social Life of Things: Commodities in Cultural Perspective. Cambridge University Press.

Asad, Talal.

2007. *On Suicide Bombing*. New York: Columbia University Press.

Austin, J. L.

1962. *How to Do Things with Words*. Oxford University Press.

Backe, Emma L.

2018. "A Crisis of Care: The Politics and Therapeutics of a Rape Crisis Hotline." *Medical Anthropology Quarterly* 32, 4: 463–480. https://doi.org/10.1111/maq.12463.

Barker, Drucilla K.

2012. "Querying the Paradox of Caring Labor." *Rethinking Marxism* 24, 4: 574–591. https://doi.org/10.1080/08935696.2012.711065

Bailey, Courtney W.

2019. "On the Impossible: Disability Studies, Queer Theory, and the Surviving Crip." *Disability Studies Quarterly* 39, 4.

Bailey, Moya, and Izetta Autumn Mobley.

2018. "Work in the Intersections: A Black Feminist Disability Framework." *Gender & Society* 33, 1: 19–40. https://doi.org/10.1177%2F0891243218801523.

Bandura, Albert.

2001. "Social Cognitive Theory: An Agentic Perspective." *Annual Review of Psychology* 52, 1: 1–26. https://doi.org/10.1146/annurev.psych.52.1.1.

Baril, Alexandre.

2020. "Suicidism: A new theoretical framework to conceptualize suicide from an anti-oppressive perspective." *Disability Studies Quarterly* 40, 3. https://dsq-sds.org/index.php/dsq/article/view/7053/5711.

Barrios, Roberto E.

2017. "What Does Catastrophe Reveal for Whom? The Anthropology of Crises and Disasters at the Onset of the Anthropocene." *Annual Review of Anthropology* 46: 1, 151–166. https://doi.org/10.1146/annurev-anthro-102116-041635.

Bateson, Gregory.

1972. *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*. San Francisco: Chandler Publishing Company.

Bauman, Zygmunt.

1992. *Intimations of Postmodernity*. New York: Routledge.

BBC.

2022. "Mental health helpline ends data-sharing relationship." *BBC*, February 2, 2022: https://www.bbc.com/news/technology-60218894.

Beck, Stefan, and Michi Knecht.

2016. "'Crisis' in Social Anthropology: Rethinking a Missing Concept." In The Handbook of International Crisis Communication Research, edited by: Andreas Schwarz, Matthew W. Seeger, and Claudia Auer, 56–65. https://doi.org/10.1002/9781118516812.ch6.

Beer, Stafford.

1974. *Designing Freedom*. (CBC Massey Lectures, 1973). Toronto: Canadian Broadcasting Corporation Publications.

Behrouzan, Orkideh.

Prozak Diaries: Psychiatry and Generational Memory in Iran. Stanford University Press.

Bell, Genevieve.

2018. "Making life: a brief history of human-robot interaction." *Consumption Markets & Culture* 21, 1: 22–41. https://doii.org/10.1080/10253866.2017.1298555.

Bellan, Rebecca, and Rita Liao.

2022. "Pony.ai scores taxi license for autonomous vehicles in Guangzhou." *TechCrunch*, April 25, 2022: https://techcrunch.com/2022/04/25/pony-ai-scores-taxi-license-for-autonomous-vehicles-in-guangzhou/.

Bellanova, Rocco, Kristina Irion, Katja Lindskov Jacobsen, Francesco Ragazzi, Rune Saugmann, and Lucy Suchman.

2021. "Toward a Critique of Algorithmic Violence." *International Political Sociology* 15, 1: 121–150. https://doi.org/10.1093/ips/olab003.

Ben-Moshe, Liat.

2020. *Decarcerating Disability: Deinstitutionalization and Prison Abolition*. Minneapolis: University of Minnesota Press.

Bentley, Peter.

2021. "Will AI replace musicians?" *BBC Science Focus Magazine*, August 8, 2021: https://www.sciencefocus.com/future-technology/will-ai-replace-musicians/.

Berlant, Lauren.

2011. *Cruel Optimism*. Durham, NC: Duke University Press.

Besky, Sarah, and Alex Blanchette.

2019. *How Nature Works: Rethinking Labor on a Troubled Planet*. Albuquerque: University of New Mexico Press.

Biehl, João.

2012. "Care and Disregard." In *A Companion to Moral Anthropology*, edited by Didier Fassin, 242–263. Malden, MA: Wiley-Blackwell.

2005. *Vita: Life in a Zone of Social Abandonment*. Berkeley: University of California Press.

Biehl, João, and Adriana Petryna.

2011. "Bodies of Rights and Therapeutic Markets." *Social Research* 78, 2: 359–386.

Birhane, Abeba, and Jelle van Dijk.

2020. "Robot Rights? Let's Talk about Human Welfare Instead." *AIES '20: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*: 207–213. https://doi.org/10.1145/3375627.3375855.

Blanchette, Alex.

2020. *Porkopolis: American Animality, Standardized Life, and the Factory Farm*. Durham, NC: Duke University Press.

Boas, Franz.

1919. "Scientists as Spies." *The Nation*, December 20, 1919.

Boellstorff, Tom.

2016. "For Whom the Ontology Turns: Theorizing the Digital Real." *Current Anthropology* 57, 4: 387–407.

2008. *Coming of Age in Second Life: An Anthropologist Explores the Virtually Human*. Princeton University Press.

Bogard, William.

1996. *The Simulation of Surveillance: Hypercontrol in Telematic Societies*. Cambridge University Press.

Bogen, Miranda.

2019. "All the Ways Hiring Algorithms Can Introduce Bias." *Harvard Business Review*, May 6, 2019: https://hbr.org/2019/05/all-the-ways-hiring-algorithms-can-introduce-bias.

Borneman, John, and Joseph Masco.

2015. "Anthropology and the Security State." *American Anthropologist* 117, 4: 781–785. https://doi.org/10.1111/aman.12371.

Bornstein, Erica.

2012. *Disquieting Gifts: Humanitarianism in New Delhi*. Stanford University Press.

Borovoy, Amy.

2005. *The Too-Good Wife: Alcohol, Codependency, and the Politics of Nurturance in Postwar Japan*. Berkeley: University of California Press.

Bourdieu, Pierre.

1986. "The forms of capital." In *Handbook of Theory and Research for the Sociology of Education*, edited by J. Richardson, 241–258. New York: Greenwood.

Bourgois, Philippe, and Jeffrey Schonberg.

2009. *Righteous Dopefiend*. Berkeley: University of California Press.

Boyne, Roy.

2000. "Post-Panopticism." *Economy and Society* 29, 2: 285–307.

Brand, Stewart.

1976. "For God's Sake, Margaret! Conversation with Gregory Bateson and Margaret Mead." *CoEvolution Quarterly* 10: 32–44.

Briggs, Joseph, and Devesh Kodnani.

2023. "Global Economics Analyst: The Potentially Large Effects of Artificial Intelligence on Economic Growth." *Goldman Sachs Economic Research*, March 26, 2023: https://www.key4biz.it/wp-content/uploads/2023/03/Global-Economics-Analyst_-The-Potentially-Large-Effects-of-Artificial-Intelligence-on-Economic-Growth-Briggs_Kodnani.pdf.

Britnell, Mark.

2019. “Human: Solving the Global Workforce Crisis in Healthcare.” Oxford University Press.

Brown, Shan-Estelle, and Zoe Pearson.
2020. “Human Sacrifices, Not Heroes: U.S. Essential Workers and the COVID-19 Pandemic.” *Exertions* https://doi.org/10.21428/1d6be30e.53617da1.

Bruni, Frank.
2022. “Will ChatGPT Make Me Irrelevant?” *NYT*, December 15, 2022: https://www.nytimes.com/2022/12/15/opinion/chatgpt-artificial-intelligence.html.

Buch, Elana.
2018. *Inequalities of Aging: Inequalities of Aging Paradoxes of Independence in American Home Care*. NYU Press.

Bucher, Tania.
2017. “The Algorithmic Imaginary: Exploring the Ordinary Affects of Facebook Algorithms.” *Information, Communication & Society* 20: 30–44. https://doi.org/10.1080/1369118X.2016.1154086.

Buolamwini, Joy.
2019. “Artificial Intelligence Has a Problem With Gender and Racial Bias. Here’s How to Solve It.” *Time*, February 7, 2019: https://time.com/5520558/artificial-intelligence-racial-gender-bias/.

Buolamwini, Joy, and Timnit Gebru.
2018. “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification.” *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*. *PMLR* 81: 77–91.

Burke, Colin.
2020. “Digital Sousveillance: A Network Analysis of the US Surveillant Assemblage.” Surveillance & Society 18, 1: 74–89. https://ojs.library.queensu.ca/index.php/surveillance-and-society/index.

Busch, Susan H. and Kelly A. Kyanko.
2020. “Incorrect Provider Directories Associated With Out-Of-Network Mental Health Care And Outpatient Surprise Bills.” *Health Affairs* 39, 6: https://doi.org/10.1377/hlthaff.2019.01501.

Bushby, Helen.

2022. “Permacrisis declared Collins Dictionary word of the year.” *BBC*, November 1, 2022: https://www.bbc.com/news/entertainment-arts-63458467).

Butler, Judith.

2016. “Rethinking Vulnerability and Resistance.” In *Vulnerability in Resistance*, edited by Judith Butler, Zeynep Gambetti, and Leticia Sabsay, 12–27. Durham, NC: Duke University Press.

2004. *Precarious Life: The Powers of Mourning and Violence.* London: Verso.

Butryn, Tracy, Leah Bryant, Christine Marchionni, and Farhad Sholevar.

2017. “The shortage of psychiatrists and other mental health providers: Causes, current state, and potential solutions.” *IJAM* 3, 1: 5–9. https://www.ijam-web.org/text.asp?2017/3/1/5/209851.

Candea, Matei.

2012. “Different Species, One Theory: Reflections on Anthropomorphism and Anthropological Comparison.” *Cambridge Journal of Anthropology* 30, 2: 118–35. http://www.jstor.org/stable/43610878.

Card, Catherine.

2018. “How Facebook AI Helps Suicide Prevention.” *Facebook* [now Meta] *Newsroom*, September 10, 2018: https://about.fb.com/news/2018/09/inside-feed-suicide-prevention-and-ai/.

Carduff, Carlo.

2014. “On the Verge of Death: Visions of Biological Vulnerability.” *Annual Review of Anthropology* 45: 105–121. https://doi.org/10.1146/annurev-anthro-102313-030341.

Carey, Benedict.

2020: “Can an Algorithm Prevent Suicide?” *NYT*, November 23, 2020: https://www.nytimes.com/2020/11/23/health/artificial-intelligence-veterans-suicide.html.

Caron, Christina.

2022. “Teens Turn to TikTok in Search of a Mental Health Diagnosis.” *NYT*, October 29, 2022: https://www.nytimes.com/2022/10/29/well/mind/tiktok-mental-illness-diagnosis.html.

Carr, E. Summerson.

2010. *Scripting Addiction: The Politics of Therapeutic Talk and American Sobriety*. Princeton University Press.

Carroll, Aaron E.
2020. "Health Disparities Among Black Persons in the US and Addressing Racism in the Health Care System." *JAMA Network*, June 18, 2020. https://doi.org/10.1001/jamahealthforum.2020.0769.

Carsten, Janet.
2000. "Introduction: Cultures of Relatedness." In *Cultures of Relatedness: New Approaches to the Study of Kinship*, edited by Janet Carsten, 1–36. Cambridge University Press.

Castañeda, Claudia, and Lucy Suchman.
2014. "Robot Visions." *Social Studies of Science* 44, 3: 315–41. http://www.jstor.org/stable/43284235.

Cavan, Ruth Shonle.
1928. *Suicide*. University of Chicago Press.

Cerulo, Karen.
2009. "Non-Humans in Social Interaction." *Annual Review of Sociology* 35, 1: 531–52.

Charlton, James I.
2000 [1998]. *Nothing About Us Without Us: Disability Oppression and Empowerment*. Oakland: University of California Press.

Chasin, Alexandra.
1995. "Class and Its Close Relations: Identities Among Women, Servants, and Machines." In *Posthuman Bodies*, edited by J. Halberstram and I. Livingston, 73–96. Bloomington: Indiana University Press.

Chen, Dan K.
2018. "The End of Life Care Machine." https://www.mori.art.museum/files/exhibitions/2020/01/27/faa_worklist.pdf.

Chen, Mel Y.
2012. *Animacies: Biopolitics, Racial Mattering, and Queer Affect*. Durham, NC: Duke University Press.

Cheney-Lippold, John.

2017. *We Are Data: Algorithms and the Making of Our Digital Selves*. New York University Press.

Cheyne, Ria.
2013. "'She Was Born a Thing': Disability, the Cyborg and the Posthuman in Anne McCaffrey's *The Ship Who Sang*." *Journal of Modern Literature* 36, 3: 138–156.

Christian, Alex.
2023. "The epic farewell posts of laid-off employees." *BBC*, January 5, 2023: https://www.bbc.com/worklife/article/20230104-layoff-linkedin-goodbye-posts.

Chua, Jocelyn Lim.
2020. "Buddy Watch: Care and Constraint Under the Watchful Eye of Military Suicide Risk Management in War." *Ethos* 48, 3: 317–335. http://doi.org/10.1111/etho.12281.

2014. *In Pursuit of the Good Life: Aspiration and Suicide in Globalizing South India*. Berkeley: University of California Press.

Chun, Wendy Hui Kyong.
2011. *Programmed Visions: Software and Memory*. Cambridge, MA: MIT Press.

Clare, Eli.
2017. *Brilliant Imperfection: Grappling with Cure*. Durham, NC: Duke University Press.

Clarke, Roger A.
1988. "Information Technology and Dataveillance." *Communications of the ACM* 31: 498–512.

Clynes, M. E., and N. S. Kline.
1960. "Cyborgs and Space." *Astronautics* Vol. 26–27: 75–76.

Cohen, Aaron Michael, and William R. Hersh.
2004. "Criticisms of Evidence-Based Medicine." *Evidence-Based Cardiovascular Medicine* 8, 3: 197–8. https://doi.org/10.1016/j.ebcm.2004.06.036.

Cohen, G. A.
1979. "The Labour Theory of Value and the Concept of Exploitation." *Philosophy and Public Affairs* 8, 4: 338–60.

Cohen, Lawrence.
2008. "Politics of Care: Commentary on Janelle S. Taylor, "On Recognition, Caring, and Dementia." *Medical Anthropology Quarterly* 22, 4: 336–339. https://doi.org/10.1111/j.1548-1387.2008.00037.x.

2001. “The Other Kidney: Biopolitics Beyond Recognition.” *Body & Society* 7, 2–3: 9–29. https://doi.org/10.1177/1357034X0100700202.

1999. “Where It Hurts: Indian Material for an Ethics of Organ Transplantation.” *Daedalus* 128, 4: 135–166.

1998. *No Aging in India: Alzheimer's, the Bad Family and Other Modern Things*. Berkeley: University of California Press.

Cokley, Rebecca.
2020. “Why “Special Needs” is Not Helpful.” https://rebecca-cokley.medium.com/why-special-needs-is-1959e2a6b0e.

Colby, Kenneth Mark.
1967. “Computer simulation of change in personal belief systems.” *Behavioral Science* 12, 3: 248–253. https://doi.org/10.1002/bs.3830120310.

Colby, Kenneth Mark, Franklin Dennis Hilf, Sylvia Weber, and Helena C Kraemer.
1972. “Turing-like indistinguishability tests for the validation of a computer simulation of paranoid processes.” *Artificial Intelligence* 3: 199–221. https://doi.org/10.1016/0004-3702(72)90049-5.

Colby, Kenneth M., and John P. Gilbert.
1964. “Programming a computer model of neurosis.” *Journal of Mathematical Psychology* 1, 2: 405–417. https://doi.org/10.1016/0022-2496(64)90012-4.

Coleman, Gabriella.
2013. *Coding Freedom: The Ethics and Aesthetics of Hacking*. Princeton University Press.

Collins, Samuel Gerald.
2003. “Sail on! Sail on!: Anthropology, Science Fiction, and the Enticing Future.” *Science Fiction Studies* 30, 2: 180–198.

Collu, Samuele.
2019. “Refracting Affects: Affect, Psychotherapy, and Spirit Dis-Possession.” *Culture, Medicine, and Psychiatry* 43, 2: 290–314. https://doi.org/10.1007/s11013-018-9616-5.

Cook, Joanna.

2018. "Paying Attention to Attention." *Anthropology of this Century* 22. http://aotcpress.com/articles/paying-attention-attention/.

Cook, Sarah C., Ann C. Schwartz, and Nadine J. Kaslow.
2017. "Evidence-Based Psychotherapy: Advantages and Challenges." *Neurotherapeutics* 14, 537–545. https://doi.org/10.1007/s13311-017-0549-4.

Cooper, David.
2013 [1967]. *Psychiatry and Anti-Psychiatry*. London: Routledge.

CoorDown.
2017. "Not Special Needs." https://clios.com/awards/winner/film/coordown/not-special-needs-26283.

Corbyn, Zoë.
2022. "'Bossware is coming for almost every worker': the software you might not realize is watching you." *The Guardian*, April 27, 2022: https://www.theguardian.com/technology/2022/apr/27/remote-work-software-home-surveillance-computer-monitoring-pandemic.

Corsín Jiménez, Alberto.
2014. "The Form of the Relation, or Anthropology's Enchantment with the Algebraic Imagination." University of Manchester. Manuscript.

Cowan, Ruth Schwartz.
1983. *More Work for Mother: The Ironies of Household Technology from the Open Hearth to the Microwave*. New York: Basic Books.

Cox, Aimee Meredith.
2015. *Shapeshifters: Black Girls and the Choreography of Citizenship*. Durham, NC: Duke University Press.

Crawford, Kate.
2021. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven, CT: Yale University.

2011. "Listening, not Lurking: The Neglected Form of Participation." In *Cultures of Participation*, edited by Hajo Grief, Larissa Hjorth, and Amparo Lasén, 63–77. Berlin: Peter Lang.

Crist, Eileen.

1999. *Images of Animals: Anthropomorphism and Animal Mind*. Philadelphia: Temple University Press.

Csordas, Thomas J.

2022. "Specter, Phantom, Demon." In *Spirit Possession: Multidisciplinary Approaches to a Worldwide Phenomenon*, edited by: Éva Pócs and András Zempléni, 215–228. https://doi.org/10.1515/9789633864142-009a. Budapest: Central European University Press.

Dalla Costa, Mariarosa, and Selma James.

1975. *The Power of Women and the Subversion of the Community*. Bristol, UK: Falling Wall Press Ltd.

D'Andrade, Roy.

1995. "Moral Models in Anthropology." *Current Anthropology* 36, 3: 399–408.

Darcy, A., J. Daniels, D. Salinger, P. Wicks, and A. Robinson.

2021. "Evidence of Human-Level Bonds Established With a Digital Conversational Agent: Cross-sectional, Retrospective Observational Study." *JMIR Formative Research* 5, 5: e27868.

Das, Veena.

2015. *Affliction: Health, Disease, Poverty*. New York: Fordham University Press.

Datta, Anupam.

2017. "Did artificial intelligence deny you credit?" *The Conversation*, March 13, 2017: https://theconversation.com/did-artificial-intelligence-deny-you-credit-73259.

Dave, Naisargi N.

2014. "Witness: Humans, Animals, and the Politics of Becoming." *Cultural Anthropology* 29, 3: 433–456. https://doi.org/10.14506/ca29.3.01.

Davis, Angela.

1981. *Women, Race & Class*. New York: Random House Inc.

Davis, Elizabeth Anne.

2018. "Global Side Effects: Counter-Clinics in Mental Health Care." *Medical Anthropology* 37, 1: 1–16. https://doi.org/10.1080/01459740.2017.1367777.

2012. *Bad Souls: Madness and Responsibility in Modern Greece*. Durham, NC: Duke University Press.

Deleuze, Gilles.

1992. "Postscript on the Societies of Control." *October* 59: 3–7. http://www.jstor.org/stable/778828.

Derrida, Jaques.

2008 [2002]. *The Animal That Therefore I Am*. New York: Fordham University Press.

1996 [1991]. *Given Time: 1. Counterfeit Money*. Translated by Peggy Kamuf. University of Chicago Press.

1994. *Specters of Marx: The State of the Debt, the Work of Mourning and the New International*. Translated by Peggy Kamuf. New York: Routledge.

Doherty, Jacob.

2021. *Waste Worlds: Inhabiting Kampala's Infrastructures of Disposability*. Oakland, CA: University of California Press.

Dumit, Joseph.

2013. *Drugs for Life: How Pharmaceutical Companies Define Our Health*. Durham, NC: Duke University Press.

Dumont, Louis.

2013 [1980]. "On Value: The Radcliffe-Brown Lecture in Social Anthropology, 1980." *Journal of Ethnographic Theory* 3, 1: 287–315. https://doi.org/10.14318/hau3.1.028.

Durkheim, Emile.

1997 [1897]. *Suicide: A Study in Sociology*. Translated by John A. Spaulding and George Simpson. New York: The Free Press.

Ecks, Stefan.

2008. Three propositions for an evidence-based medical anthropology." *JRAI* 14, 1: S77-S92. https://doi.org/10.1111/j.1467-9655.2008.00494.x.

Edwards, Benj.

2023. “Controversy erupts over non-consensual AI mental health experiment.” *ArsTechnica*, January 10, 2023: https://arstechnica.com/information-technology/2023/01/contoversy-erupts-over-non-consensual-ai-mental-health-experiment/.

Ekbia, Hamid R. and Bonnie A. Nardi.

2017. *Heteromation, and Other Stories of Computing and Capitalism*. Boston: MIT Press.

2014. “Heteromation and its (dis)contents: The invisible division of labor between humans and machines.” *First Monday* 19, 6. https://doi.org/10.5210/fm.v19i6.5331.

Eliot, Lance.

2023. “People Are Eagerly Consulting Generative AI ChatGPT For Mental Health Advice, Stressing Out AI Ethics And AI Law.” *Forbes*, January 1, 2023: https://www.forbes.com/sites/lanceeliot/2023/01/01/people-are-eagerly-consulting-generative-ai-chatgpt-for-mental-health-advice-stressing-out-ai-ethics-and-ai-law/.

Ellis, C. D.

1946. “Nuclear Energy and Its Peace-Time Applications.” *Journal of the Royal Society of Arts* 94, 4713: 235–41. http://www.jstor.org/stable/41362989.

England, Paula.

2005. “Emerging Theories of Care Work.” *Annual Review of Sociology* 31: 381–99. http://www.jstor.org/stable/29737725

English-Lueck, J. A.

2017 [2002]. *Cultures@SiliconValley, Second Edition*. Stanford University Press.

2010. *Being and Well-Being: Health and the Working Bodies of Silicon Valley*. Stanford University Press.

Esmeir, Samera.

2006. “On Making Dehumanization Possible.” *PMLA* 121, 5: 1544–51.

Eubanks, Virginia.

2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St Martin’s Press.

Eubanks, Virginia, and Alexandra Mateescu.
2021. “Automating Care: About our new series on the rise of AI in caregiving.” *The Guardian*, June 3, 2021. *https://www.theguardian.com/us-news/2021/jun/03/automating-care-about-new-series-guardian-us*.

Evans-Pritchard, E. E.
1991 [1937]. *Witchcraft, Oracles and Magic among the Azande*. Oxford University Press.

Fassin, Didier.
2008. “The Elementary Forms of Care: An Empirical Approach to Ethics in a South African Hospital.” *Social Science & Medicine* 67: 262–270.

Favret-Saada, Jeanne.
1980 [1977]. *Deadly Words: Witchcraft in the Bocage*. Cambridge University Press.

Federici, Silvia.
1995 [1975]. “Wages Against Housework.” In *The Politics of Housework, New Edition*, edited by Ellen Malos, 187–94. Cheltenham, UK: New Clarion Press.

Fehling, Kara B, and Edward A Selby.
2021. “Suicide in DSM-5: Current Evidence for the Proposed Suicide Behavior Disorder and Other Possible Improvements.” *Front Psychiatry* 11: 1–15. http://doi.org/10.3389/fpsyt.2020.499980.

Feldman, Allen.
1991. *Formations of Violence The Narrative of the Body and Political Terror in Northern Ireland*. University of Chicago Press.

Feldstein, Steven.
2022. “4 Reasons Why Putin’s War Has Changed Big Tech Forever.” *Foreign Policy*, March 29, 2022. https://foreignpolicy.com/2022/03/29/ukraine-war-russia-putin-big-tech-social-media-internet-platforms/.

Fisch Michael.
2018. *An Ethnography of the Machine: Tokyo’s Commuter Train Network*. University of Chicago Press.

Fogg, B. J., and Clifford Nass.

1997. “How Users Reciprocate to Computers: An Experiment that Demonstrates Behavior Change.” *CHI ‘97 Extended Abstracts on Human Factors in Computing Systems*: 331–332.

Folbre, Nancy.

1995. “‘Holding Hands at Midnight’: The Paradox of Caring Labor.” *Feminist Economics* 1, 1: 73–92.

Forsythe, Diana.

1993a. “The Construction of Work in Artificial Intelligence.” *Science, Technology, & Human Values* 18, 4: 460–479.

1993b. “Engineering Knowledge: The Construction of Knowledge in Artificial Intelligence.” *Social Studies of Science* 23, 3: 445–477.

Fortunati, Leopoldina.

1995 [1981]. *The Arcane of Reproduction: Housework, Prostitution, Labor and Capital*. Translated by Hilary Creek. Edited by Jim Fleming. New York: Autonomedia.

Foucault, Michel.

2005. *The Hermeneutics of the Subject: Lectures at the Collège de France 1981–1982*. Translated by Graham Burchell. New York: Picador.

2003 [1997]. ‘*Society Must Be Defended’: Lectures at the Collège de France 1975–1976*. Edited by Mario Bertani and Alessandro Fontana. Translated by David Macey. New York: Picador.

1995 [1975]. *Discipline and Punish: The Birth of the Prison*. Translated by Alan Sheridan. New York: Vintage Books.

1990. *The History of Sexuality, Vol. 1: An Introduction*. Translated by Robert Hurley. New York: Vintage Books.

1988. *The History of Sexuality, Vol. 3: The Care of the Self*. Translated by Robert Hurley. New York: Vintage Books.

Freudenberger, Herbert J.

1974. “Staff Burn-Out.” *Journal of Social Issues* 30, 1: 159–165. https://doi.org/10.1111/j.1540-4560.1974.tb00706.x.

Frey, Carl Benedikt, and Michael A. Osborne.
2013. “The Future of Employment: How Susceptible are Jobs to Computerisation?” Oxford Martin Programme on Technology and Employment Working Paper, Oxford University.

Fullwiley, Duana.
2011. *The Enculturated Gene Sickle Cell Health Politics and Biological Difference in West Africa*. Princeton University Press.

Fussell, Sidney.
2021. “This AI Helps Police Monitor Social Media. Does It Go Too Far?” *Wired*, July 6, 2021: https://www.wired.com/story/ai-helps-police-monitor-social-media-go-too-far/.

Garcia, Angela.
2010. *The Pastoral Clinic: Addiction and Dispossession along the Rio Grande*. Berkeley: University of California Press.

Garland-Thomson, Rosemary.
2002. “Integrating Disability, Transforming Feminist Theory.” *NWSA Journal* 14, 3: 1–32.

Geertz, Clifford.
1984 [1976]. “‘From the Native's Point of View’: On the Nature of Anthropological Understanding.” In *Culture Theory: Essays on Mind, Self, and Emotion*, edited by Richard A. Shweder and Robert A. LeVine, 123–136. Cambridge University Press.

Gillespie, Tarlton.
2018. *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. New Haven: Yale University Press.

Giordano, Cristiana.
2016. *Migrants in Translation: Caring and the Logics of Difference in Contemporary Italy*. Oakland: University of California Press.

Glenn, Evelyn Nakano.

2012. *Forced to Care: Coercion and Caregiving in America*. Cambridge, MA: Harvard University Press.

Goffman, Irving.
1963. *Stigma: Notes on the Management of Spoiled Identity*. New York: Simon and Schuster.

1961. *Asylums: Essays on the Social Situation of Mental Patients and Other Inmates*. Garden City, NY: Anchor Books.

González, Roberto J.
2022. *War Virtually: The Quest to Automate Conflict, Militarize Data, and Predict the Future*. Oakland, CA: University of California Press.

Good, Byron, Andrea Chiovenda, and Sadeq Rahimi.
2022. "The Anthropology of Being Haunted: On the Emergence of an Anthropological Hauntology." *Annual Review of Anthropology* 51: 437–453.

Gorvett, Zaria.
2018. "How humans bond with robot colleagues." *BBC*, May 30, 2018: https://www.bbc.com/worklife/article/20180530-how-humans-bond-with-robot-colleagues.

Govindrajan, Radhika.
2018. *Animal Intimacies: Interspecies Relatedness in India's Central Himalayas*. University of Chicago Press.

de Graaf, Maartje M. A.
2016. "An Ethical Evaluation of Human-Robot Relationships." *International Journal of Social Robotics* 8, 589–598. https://doi.org/10.1007/s12369-016-0368-5.

Graeber, David.
2013. "Postscript: It is value that brings universes into being." *Journal of Ethnographic Theory* 3, 2: 219–243. https://doi.org/10.14318/hau3.2.012.

2011. *Debt: The First 5,000 Years*. New York: Melville House Publishing.

2001. *Toward an anthropological theory of value: The false coin of our own dreams*. New York: Palgrave Macmillan.

Gray, Mary and Siddharth Suri.

2019. *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Boston: Houghton Mifflin Harcourt.

Gregg, Melissa.
2018. "Technical Female: A Gender Studies Academic in Silicon Valley."

2015a. "Getting Things Done." In *Networked Affect*, edited by Ken Hillis, Susanna Paasonen, and Michael Petit, 187–202. New York: Penguin Random House.

2015b. "The Gift That is Not Given." In *Data: Now Bigger and Better!*, edited by Tom Boellstorff, Genevieve Bell, Bill Maurer, Melissa Gregg, and Nick Seaver, 47–66. Chicago: Prickly Paradigm.

2011. *Work's Intimacy*. Cambridge, UK: Polity Press.

Guenther, Mathias.
2020. *Human-Animal Relationships in San and Hunter-Gatherer Cosmology, Volume II: Imagining and Experiencing Ontological Mutability*. Cham, Switzerland: Palgrave MacMillan.

Guevarra, Anna Romina.
2021. "Coding Care." In *Redesigning AI: Work Democracy,and Justice in the Age of Automation*, edited by Daron Acemoglu, 163–171. Cambridge, MA: Boston Review Forum.

Gunderson, Keith.
1963. "Interview with the Robot." *Analysis* 23: 136–142. https://doi.org/10.1093/analys/23.6.136.

Gupta, Sarita.
2018. "The most important job in the world is one no one wants anymore." *Quartz*, December 12, 2018: https://qz.com/1490065/the-most-important-job-in-the-world-is-one-no-one-wants-anymore/.

Haggerty, Kevin D.
2006. "Tear Down the Walls: On Demolishing the Panopticon." In *Theorizing Surveillance: The Panopticon and Beyond*, edited by David Lyon, 23–45. Cullompton, Devon: Willan Publishing.

Haggerty, Kevin D., and Richard V. Ericson.
2000. "The Surveillant Assemblage." *The British Journal of Sociology* 51: 605–22.

Haggerty, Kevin, and Daniel Trottier.

2015. "Surveillance and/of nature: monitoring beyond the human." *Society & Animals* 23, 4: 400–420. https://doi.org/10.1163/15685306-12341304.

Halpern, Mark.

2006. "The Trouble with the Turing Test: On the fallacy of thinking computers." *The New Atlantis* 11: 42–63. https://www.thenewatlantis.com/publications/the-trouble-with-the-turing-test.

Han, Clara.

2018. "Precarity, Precariousness, and Vulnerability." *Annual Review of Anthropology* 47 331–343. https://doi.org/10.1146/annurev-anthro-102116-041644.

2012. *Life in Debt: Times of Care and Violence in Neoliberal Chile*. Berkeley: University of California Press.

Hansen, Helena, Philippe Bourgois, and Earnest Drucker.

2013. "Pathologizing Poverty: New Forms of Diagnosis, Disability, and Structural Stigma under Welfare Reform." *Soc Sci Med* 103: 76–83. https://dx.doi.org/10.1016%2Fj.socscimed.2013.06.033.

Hare, Stephanie.

2022. *Technology Is Not Neutral: A Short Guide to Technology Ethics*. London Publishing Partnership.

Harmer, Bonnie, Sarah Lee, Truc vi H. Duong, and Abdolreza Saadabadi.

2022. *Suicidal Ideation*. Treasure Island, FL: StatPearls Publishing.

Hao, Karen.

2019. "AI is sending people to jail—and getting it wrong." *MIT Technology Review*, January 21, 2019: https://www.technologyreview.com/2019/01/21/137783/algorithms-criminal-justice-ai.

Haraway, Donna.

2016. *Staying with the Trouble*. Durham, NC: Duke University Press.

2007. *When Species Meet*. Minneapolis: University of Minnesota Press.

1991 [1985]. “A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century.” In *Simians, Cyborgs and Women: The Reinvention of Nature*. New York: Routledge.

Hare, Stephanie.
2022. *Technology Is Not Neutral: A Short Guide to Technology Ethics*. London Publishing Partnership.

Hayden, Cori.
2023. *The Spectacular Generic: Pharmaceuticals and the Simipolitical in Mexico*. Durham, NC: Duke University Press.

2010. “The Proper Copy.” *Journal of Cultural Economy* 3, 1, 85–102. http://doi.org/10.1080/17530351003617602.

Hayles, N. Katherine.
1999. *How We Became Posthuman Virtual Bodies in Cybernetics, Literature, and Informatics*. University of Chicago Press.

Heaven, Will Douglas.
2020. “This startup is using AI to give workers a ‘productivity score.’” *MIT Technology Review*, June 4, 2020: https://www.technologyreview.com/2020/06/04/1002671/startup-ai-workers-productivity-score-bias-machine-learning-business-covid/.

Hedva, Johanna.
2016. “Sick Woman Theory.” *Mask Magazine*: *https://web.archive.org/web/20160302154617/http://www.maskmagazine.com/not-again/struggle/sick-woman-theory*.

Heilweil, Rebecca.
2019. “Artificial intelligence will help determine if you get your next job.” *Vox*, December 12, 2019. https://www.vox.com/recode/2019/12/12/20993665/artificial-intelligence-ai-job-screen.

Heims, Steve P.
1977. “Gregory Bateson and the Mathematicians: From Interdisciplinary Interaction to Societal Functions.” *Journal of the History of the Behavioral Sciences* 13: 141–159.

Heinemann, Katja.
2019. “Iyad Rahwan: Why We Need a Science of Machine Behavior.” *Quanta Magazine*. https://www.quantamagazine.org/videos/iyad-rahwan-why-we-need-a-science-of-machine-behavior/.

Helmreich, Stefan.

2009. *Alien Ocean: Anthropological Voyages in Microbial Seas*. Berkeley: University of California Press.

Hochschild, Arlie.

1983. *The Managed Heart: Commercialization of Human Feeling*. Berkeley: University of California Press.

1979. "Emotion Work, Feeling Rules, and Social Structure." *American Journal of Sociology* 85, 3: 551–75. http://www.jstor.org/stable/2778583.

Holden, Constance.

1977. "The Empathic Computer." *Science* 198, 4132: 32. https://doi.org/2010.1126/science.198.4312.32.

Hollan, Douglas and Jason Throop.

2008. "Whatever Happened to Empathy? Introduction." *Ethos* 36, 4: 385–401.

2011. *The Anthropology of Empathy: Introduction*. In *The Anthropology of Empathy: Experiencing the Lives of Others in Pacific Societies*, edited by D. Hollan and J. Throop, 1–21. New York: Berghahn Books.

Holman, E. Alison, Dana Rose Garfin, and Roxane Cohen Silvera.

2014. "Media's role in broadcasting acute stress following the Boston Marathon bombings." *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* 111, 1: 93–-8. https://doi.org/10.1073/pnas.1316265110.

Holmes, Aaron.

2019. "'Alexa, help me': A nursing home patient begged her Amazon Echo for help dozens of times before dying of COVID-19." *Business Insider*, April 9, 2020: https://www.businessinsider.com/amazon-alexa-help-coronavirus-nursing-home-patient-died-2020-4.

Horowitz, Ben.

2019. *What You Do Is Who You Are: How to Create Your Business Culture*. New York: HarperCollins.

Hjorth, Larissa, Heather Horst, Anne Galloway, and Genevieve Bell (editors).

2017. *The Routledge Companion to Digital Ethnography*. New York: Routledge.

Horton, Donald, and R. Richard Wohl.
1956. "Mass Communication and Para-Social Interaction: Observation on Intimacy at a Distance." *Psychiatry* 19, 3: 215-229. https://doi.org/10.1080/00332747.1956.11023049.

Howell, Signe.
2009. "Adoption of the Unrelated Child: Some Challenges to the Anthropological Study of Kinship." *Annual Review of Anthropology* 38, 1: 149–166.

Hu, Tung-Hui.
2015. *A Prehistory of the Cloud*. Cambridge, MA: MIT Press.

Huberman, Jenny.
2020. "What to Do with Surveillance Capitalism?" *Anthropology Now* 12, 2: 94–100. https://doi.org/10.1080/19428200.2020.1824760.

Hui, Yuk.
2018. "On Automation and Free Time." *e-flux Architecture* https://www.e-flux.com/architecture/superhumanity/179224/on-automation-and-free-time

Huxor, Avon.
2022. "Artificial Intelligence: A Medium that Hides Its Nature." In *Artificial Intelligence and Its Discontents. Social and Cultural Studies of Robots and AI*, edited by A. Hanemaayer, 105–122. New York: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-88615-8_6.

Ingold, Tom.
2010. "Bringing Things to Life: Creative Entanglements in a World of Materials." *Realities* Working Paper 15, ESRC National Centre for Research Methods, University of Manchester.

Insel, Thomas.
2022. "What American Mental-Health Care Is Missing." *The Atlantic*, February 13, 2022: https://www.theatlantic.com/ideas/archive/2022/02/american-mental-health-crisis-healing/622052.

2018. "Digital Phenotyping: A Global Tool for Psychiatry." *World Psychiatry* 17: 276-277. https://doi.org/10.1002/wps.20550.

International Labour Organization.

2018. *Care Work and Care Jobs for the Future of Decent Work*. Geneva: International Labour Office. https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_633135.pdf.

Irani, Lilly.

2019. *Chasing Innovation: Making Entrepreneurial Citizens in Modern India*. Princeton University Press.

2015. "Justice for 'Data Janitors.'" *Public Books*, January 15, 2015: https://www.publicbooks.org/justice-for-data-janitors/.

2013. "The Cultural Work of Microwork." *New Media and Society* 17, 5: 720–739.

Irani, Lilly S., and M. Six Silberman.

2013. "Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk." *CHI '13: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: 611–620. https://doi.org/10.1145/2470654.2470742.

Jablonsky, Rebecca, Tero Karppi, and Nick Seaver.

2021. "Introduction: Shifting Attention." *Science, Technology, & Human Values* 47, 2: 235–242. https://doi.org/10.1177/01622439211058823.

Jackson, Zakiyyah Iman.

2015. "Outer Worlds: The Persistence of Race in Movement "Beyond the Human" *GLQ* 21, 2–3: 215–218. https://muse.jhu.edu/article/582032.

Jobs, Steve.

2005. "'You've got to find what you love,' Jobs says." *Stanford News*, June 12, 2005: https://news.stanford.edu/2005/06/12/youve-got-find-love-jobs-says/.

Jobson, Ryan Cecil.

2020. "The Case for Letting Anthropology Burn: Sociocultural Anthropology in 2019." 122, 2: 259–271.

Johnson, Jim.

1988. "Mixing Humans and Nonhumans Together: The Sociology of a Door-Closer." *Social Problems* 35, 3: 298–310.

Johnson, Steven Ross.

2021. “Surgeon General Advisory Raises Alarm on Youth Mental Health.” *U.S. News & World Report*, December 7, 2021: https://www.usnews.com/news/health-news/articles/2021-12-07/surgeon-general-issues-advisory-on-youth-mental-health.

Jordan, Dearbail, and Zoe Conway.
2023. “Amazon strikes: Workers claim their toilet breaks are timed.” *BBC*, January 25, 2023: https://www.bbc.com/news/business-64384287.

Kafer, Alison.
2013. *Feminist, Queer, Crip*. Bloomington: Indiana University Press.

Kanter, Rosabeth Moss.
1993 [1977]. *Men and Women of the Corporation*. New York: Basic Books.

Kaste, Martin.
2018. “Facebook Increasingly Reliant on A.I. To Predict Suicide Risk,” *NPR*, November 17, 2018: https://www.npr.org/2018/11/17/668408122/facebook-increasingly-reliant-on-a-i-to-predict-suicide-risk.

Kavlakoglu, Eda.
2020. “NLP vs. NLU vs. NLG: the differences between three natural language processing concepts,” *IBM Blog*, November 12, 2020: https://www.ibm.com/blogs/watson/2020/11/nlp-vs-nlu-vs-nlg-the-differences-between-three-natural-language-processing-concepts/.

Keane, Webb.
2016. “A Reader's Guide to the Anthropology of Ethics and Morality - Part II.” *Somatosphere*, October 28, 2016: http://somatosphere.net/2016/ethics-and-morality-part-2.html/.

Kelty, Christopher M.
2017. “Robots Behaving Badly: Simulation and Participation in the Study of Life.” In: *Research Objects in their Technological Setting*, edited by Bernadette Bensaude-Vincent Sacha Loeve, Alfred Nordmann, and Astrid Schwartz, 42–55. New York: Routledge. https://doi.org/10.4324/9781781448397.

Kerner, Sean Michael.
2023. “How AI-assisted app from Pfizer and Ada Health aims to help COVID sufferers.” *VentureBeat*, February 2, 2023: https://venturebeat.com/ai/how-ai-assisted-app-from-pfizer-and-ada-health-aims-to-help-covid-sufferers/.

Kim, Eunjung.
2017. *Curative Violence: Rehabilitating Disability, Gender, and Sexuality in Modern Korea*. Durham, NC: Duke University Press.

2015. “Unbecoming Human: An Ethics of Objects.” *GLG: A Journal of Lesbian and Gay Studies* 21, 2, 3: 295–320.

Kim, Jina B.
2021. “Cripping the Welfare Queen: The Radical Potential of Disability Politics.” *Social Text* 39 (3, 148): 79–101.

Kirksey, S. Eben, and Stefan Helmreich.
2010. “The Emergence of Multispecies Ethnography.” *Cultural Anthropology* 25, 4: 545–576. https://doi.org/10.1111/j.1548-1360.2010.01069.x

Kitanaka, Junko.
2015. “The Rebirth of Secrets and the New Care of the Self in Depressed Japan.” Current Anthropology 56, S12: S251–S262.

Kleinman, Arthur.
2019. *The Soul of Care: The Moral Education of a Husband and a Doctor*. New York: Penguin Books.

2012. “Caregiving as Moral Experience.” *The Lancet* 380, 9853 https://doi.org/10.1016/S0140-6736(12)61870-4.

2010. “Caregiving: The Divided Meaning of Being Human and the Divided Self of the Caregiver.” In *Rethinking the Human*, edited by J. Michelle Molina and Donald K. Swearer, 16–27. Cambridge, MA: Harvard University Press.

2009. “Caregiving: The odyssey of becoming more human.” *The Lancet* 373, 9660: 292–293. https://doi.org/10.1016/S0140-6736(09)60087-8.

2008. “Catastrophe and Caregiving.” *The Lancet: The art of medicine*, January 5, 2008: https://doi.org/10.1016/S0140-6736(08)60057-4.

Klima, Alan.
2019. *Ethnography #9*. Durham, NC: Duke University Press.

Knight, Kelly Ray.
2015. *addicted. pregnant. poor*. Durham, NC: Duke University Press.

Kohn, Eduardo.

2013. *How Forests Think*. Berkeley: University of California Press.

Konrad, Monica.
2005. *Nameless Relations: Anonymity, Melanesia and Reproductive Gift Exchange between British Ova Donors and Recipients*. New York: Berghahan.

Koselleck, Reinhart.
2006. “Crisis.” Translated by Michaela W. Richter. *Journal of the History of Ideas* 67, 2: 357–400. https://www.jstor.org/stable/30141882.

Kraus, Rachel.
2022. “Google applies advanced AI to suicide prevention and reducing graphic search results.” *Mashable*, March 30, 2022: https://mashable.com/article/google-suicide-domestic-violence-ai-prevention.

Lafargue, Paul.
2011 [1883]. *The Right To Be Lazy: Essays by Paul Lafargue*. Edited by Bernard Marszalek. Oakland, CA: C. H. Kerr & Co. & AK Press.

Lakoff, Andrew.
2007. “The Right Patients for the Drug: Managing the Placebo Effect in Antidepressant Trials.” *BioSocieties* 2, 57–71. https://doi.org/10.1017/S1745855207005054.

Lamarre, Thomas.
2012. “Humans and Machines.” *INFLeXions* 5: 29–67. http://www.inflexions.org/n5_lamarrehtml.html.

The Lancet.
2014. “Global Elder Care in Crisis.” *The Lancet* 383, 9921: 927. https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(14)60463-3/fulltext.

Latour, Bruno.
2008. “A Cautious Prometheus? A Few Steps Toward a Philosophy of Design (with Special Attention to Peter Sloterdijk).” Keynote lecture for the Networks of Design meeting of the Design History Society Falmouth, Cornwall.

2005. *Reassembling the Social: An Introduction to Actor-Network Theory*. Oxford University Press.

1996. “On Actor-Network Theory: A Few Clarifications Plus More Than a Few Complications.” *Philosophia* 25, 3–4: 47–64.

1993. *We Have Never Been Modern*. Translated by Catherne Porter. Cambridge, MA: Harvard University Press.

1990. "Technology is Society Made Durable." In *A Sociology of Monsters: Essays on Power, Technology and Domination*, edited by J. Law, 103–131. New York: Routledge.

LaVito, Angelica.
2018. "Anxiety is expensive: Employee mental health costs rise twice as fast as all other medical expenses." *CNBC*, September 27, 2018: https://www.cnbc.com/2018/09/26/employers-are-starting-to-think-about-healthy-differently.html.

Lepselter, Susan.
2016. *The Resonance of Unseen Things: Poetics, Power, Captivity, and UFOs in the American Uncanny*. Ann Arbor: University of Michigan Press.

Lester, Rebecca.
2019. *Famished: Eating Disorders and Failed Care in America*. Berkeley: University of California Press.

2013. "Back from the edge of existence: A critical anthropology of trauma." *Transcultural Psychiatry* 50, 5: 753–762. https://doi.org/10.1177/1363461513504520.

Lévi-Strauss, Claude.
1969 [1949]. *The Elementary Structures of Kinship*. Boston: Beacon Press.

Levin, Josh.
2013. "The Welfare Queen." *Slate*, December 19, 2013: http://www.slate.com/articles/news_and_politics/history/2013/12/linda_taylor_welfare_queen_ronald_reagan_made_her_a_notorious_american_villain.html.

Levine, Alexandra S.
2022. "Suicide hotline shares data with for-profit spinoff, raising ethical questions." *Politico*, January 28, 2022: https://www.politico.com/news/2022/01/28/suicide-hotline-silicon-valley-privacy-debates-00002617.

Licklider, J. C. R.
1960. "Man-Computer Symbiosis." *Transactions on Human Factors in Electronics* HFE-1: 4–11.

1965. “Man-Computer Partnership.” *International Science and Technology* 41: 18–26.

Licklider, J. C. R., and Robert W. Taylor.
1968. “The Computer as a Communication Device.” *Science and Technology* 76: 21–38.

Lucas, Gavin.
2002. “Disposability and Dispossession in the Twentieth Century.” *Journal of Material Culture* 7, 1: 5–22. https://doi.org/10.1177/1359183502007001303.

Lyon, David.
2006. “The Search for Surveillance Theories.” In *Theorizing Surveillance: The Panopticon and Beyond*, edited by David Lyon, 3–20. Cullompton, Devon: Willan Publishing.

1993. “‘An Electronic Panopticon? A sociological critique of surveillance theory.” *The Sociological Review* 41, 4: 653–678. https://doi.org/10.1111/j.1467-954X.1993.tb00896.x.

Lyons, Kristina M.
2020. *Vital Decompositions: Soil Practitioners and Life Politics*. Durham, NC: Duke University Press.

Mahmood, Saba.
2001. “Feminist Theory, Embodiment, and the Docile Agent: Some Reflections on the Egyptian Islamic Revival.” Cultural Anthropology 16, 2: 202–236. https://www.jstor.org/stable/656537.

Malinowski, Bronislaw.
1922. *Argonauts of the Western Pacific: An Account of Native Enterprise and Adventure in the Archipelagoes of Melanesian New Guinea*. New York: Routledge.

Manjoo, Farhad.
2022. “In the Battle With Robots, Human Workers Are Winning.” *NYT*, October 27, 2022: https://www.nytimes.com/2022/10/07/opinion/machines-ai-employment.html.

Marcus, George E.
1995. “Ethnography in/of the World System: The Emergence of Multi-sited Ethnography.” *Annual Review of Anthropology* 24: 95–117.

Mariner, Kathryn A.

2019. "White Parents, Black Care: Entanglements of Race and Kinship in American Transracial Adoption." *American Anthropologist* 121: 845–856. https://doi.org/10.1111/aman.13312.

Marks, Amber.
2007. "Drug Detection Dogs and the Growth of Olfactory Surveillance: Beyond the Rule of Law?" *Surveillance & Society* 4, 3: 257–271. https://doi.org/10.24908/ss.v4i3.3450.

Marks, Harry M.
1997. *The Progress of Experiment Science and Therapeutic Reform in the United States, 1900-1990*. Cambridge University Press.

Marsh, Ian.
2015. "'Critical suicidology': Toward an inclusive, inventive and collaborative (post) suicidology." *Social Epistemology Review and Reply Collective* 4, 5: 6–9.

Marx, Karl, and C. J. Arthur.
2018 [1867]. *Marx's Capital: A Student Edition.* Edited by C. J. Arthur. London: Lawrence & Wishart.

Marx, Karl, and Friedrich Engels.
1978. *The Marx-Engels Reader, Second Edition.* Edited by Robert C. Tucker. New York: Norton.

Masco, Joseph.
2017. "The Crisis in Crisis." *Current Anthropology* 58, 15: S65–S76.

Mathiesen, Thomas.
1997. "The Viewer Society: Michel Foucault's 'Panopticon' Revisited." *Theoretical Criminology* 1, 2: 215–33.

Mattingly, Cheryl, and Jason Throop.
2018. "The Anthropology of Ethics and Morality." *Annual Review of Anthropology* 47, 1, 475–492.

Mauss, Marcel.
2016 [1925]. *The Gift*. Translated by Jane I. Guyer. University of Chicago Press.

Mauterstock, Robert.

2019. “The Caregiver Crisis Is Here!” *Forbes*, August 30, 2019: https://www.forbes.com/sites/robertmauterstock/2019/08/30/the-caregiver-crisis-is-here/.

Mbembe, Achille.

2019. *Necropolitics*. Translated by Steven Corcoran. Durham, NC: Duke University Press.

2003. “Necropolitics.” *Public Culture* 15, 1: 11–40. https://doi.org/10.1215/08992363-15-1-11.

McMahon, Ciarán.

2017. “What the Research Really Suggests About That Facebook Chatbot Therapist.” *Slate*, July 18, 2017: https://slate.com/technology/2017/07/what-research-says-about-woebot-the-facebook-chatbot-therapist.html.

McQuillan, Dan.

2016. “Algorithmic paranoia and the convivial alternative.” *Big Data & Society* 3, 2: 1–12. https://doi.org/10.1177/2053951716671340.

Mead, Margaret.

1978. “The Contribution of Anthropology to the Science of the Future.” Cultures of the Future.” Edited by Margoroh Maruyama and Arthur Harkins. Chicago: Mouton: 1–3.

Merleau-Ponty, Maurice.

2012 [1945]. *Phenomenology of Perception*. Translated by Donald A. Landes. New York: Routledge.

Metcalf, Jacob, Emanuel Moss, and danah boyd.

2019. “Owning Ethics: Corporate Logics, Silicon Valley, and the Institutionalization of Ethics.” *Social Research: An International Quarterly* 82, 2: 449–476. https://doi.org/10.1353/sor.2019.0022

Metz, Rachel.

2018. “The smartphone app that can tell you’re depressed before you know it yourself.” *MIT Technology Review*, October 15, 2018: https://www.technologyreview.com/2018/10/15/66443/the-smartphone-app-that-can-tell-youre-depressed-before-you-know-it-yourself/.

Metzl, J. M.

2009. *The Protest Psychosis: How Schizophrenia became a Black disease*. Boston: Beacon Press.

Metzl, Jonathan M., and Helena Hansen.
2014. "Structural Competency: Theorizing a new medical engagement with stigma and inequality." *Soc Sci Med* 103: 126–133. https://doi.org/10.1016/j.socscimed.2013.06.032

Miller, Boaz.
2021. "Is Technology Value-Neutral?" *Science, Technology, & Human Values* 46, 1: 53–80. https://doi.org/10.1177/0162243919900965.

Miller, Daniel.
2018. "Digital Anthropology." In *The Open Encyclopedia of Anthropology*, edited by Felix Stein. (Facsimile of the first edition in The Cambridge Encyclopedia of Anthropology.) http://doi.org/10.29164/18digital.

2007. "What is a Relationship? Is Kinship Negotiated Experience?" *Ethnos* 7, 4: 535–554.

Mitman, Gregg.
2019. "Reflections on the Plantationocene: A Conversation with Donna Haraway and Anna Tsing. Moderated by Gregg Mitman." *Edge Effects*: https://edgeeffects.net/wp-content/uploads/2019/06/PlantationoceneReflections_Haraway_Tsing.pdf.

Mittelstadt, Brent, Chris Russell, and Sandra Wachter.
2019. "Explaining Explanations in AI." *Proceedings of the Conference on Fairness, Accountability, and Transparency*, January 2019: https://arxiv.org/pdf/1811.01439.pdf.

Mol, Annemarie.
2008. *The Logic of Care: Health and the Problem of Patient Choice*. New York: Routledge.

Monahan, Torin, and David Murakami Wood.
2018. "Introduction: Surveillance Studies as a Transdisciplinary Endeavor." In *Surveillance Studies: A Reader*, edited by T. Monahan and D. Murakami Wood, xix–xxxiv. New York: Oxford University Press.

Moss, Emanuel, Rumman Chowdhury, Bogdana Rakova, Sonja Schmer-Galunder, Reuben Binns, and Andrew Smart.

2019. “Machine Behaviour is Old Wine in New Bottles.” *Nature* 574, 176.

Munn, Nancy D.
1986. *The fame of Gawa: A symbolic study of value transformation in a Massim (Papua New Guinea) society*. Cambridge University Press.

Muñoz, José Esteban, Jinthana Haritaworn, Myra Hird, Zakiyyah Iman Jackson, Jasbir K. Puar, Eileen Joy, Uri McMillan, Susan Stryker, Kim TallBear, Jami Weinstein, and Jack Halberstam.
2015. “Theorizing Queer Inhumanism.” *GLC* 21, 2–3: 209–248. https://doi.org/10.1215/10642684-2843323.

Münster, Daniel, and Ludek Broz.
2015. “The Anthropology of Suicide: Ethnography and the Tension of Agency.” In *Suicide and Agency: Anthropological Perspectives on Self-Destruction, Personhood, and Power*, edited by Ludek Broz and Daniel Münster. Farnham, Surrey: Ashgate.

Murphy, Michelle.
2015. “Unsettling Care: Troubling transnational itineraries of care in feminist health practices.” *Social Studies of Science* 45, 5: 717–737. https://doi.org/10.1177/0306312715589136.

Myers, Natasha.
2017. “Ungrid-able Ecologies: Decolonizing the Ecological Sensorium in a 10,000 year-old NaturalCultural Happening.” *Catalyst: Feminism, Theory, Technoscience* 3, 2: 1–24.

Nader, Laura.
1972. “Up the Anthropologist: Perspectives Gained From Studying Up.” ERIC Clearinghouse: 1–28. https://eric.ed.gov/?id=ED065375.

Nagy, Peter, and Gina Neff.
2015. “Imagined Affordance: Reconstructing a Keyword for Communication Theory.” *Social Media + Society* 1, 2: 1–9: https://doi.org/10.1177%2F2056305115603385.

Nakamura, Karen.
2019. “My Algorithms Have Determined You’re Not Human: AI-ML, Reverse Turing Tests, and the Disability Experience.” (Keynote for the ASSETTS Conference, Septermber 28, 2019.)

2013. *A Disability of the Soul: An Ethnography of Schizophrenia and Mental Illness in Contemporary Japan*. Ithaca, NY: Cornell University Press.

Nardi, Bonnie.

2010. *My Life as a Night Elf Priest: An Anthropological Account of World of Warcraft*. Ann Arbor: University of Michigan Press.

Nass, Clifford, and Youngme Moon.

2000. "Machines and Mindlessness." *Journal of Social Issues* 56, 1: 81–103.

Nass, Clifford, Youngme Moon, and Paul Carney.

1999. "Are People Polite to Computers? Responses to Computer-Based Interviewing Systems." *Journal of Applied Social Psychology* 29, 5: 889–1109.

National Alliance on Mental Illness.

2017. *The Doctor is Out: Continuing Disparities in Access to Mental and Physical Health Care*.

Neff, Gina, and Peter Nagy.

2016. "Talking to Bots: Symbiotic Agency and the Case of Tay." *International Journal of Communication* 10, 20: 4915–4931.

Netherland, Julie, and Helena Hansen.

2017. "White opioids: Pharmaceutical race and the war on drugs that wasn't." Biosocieties 12, 2: 217–238. https://doi.org/10.1057/biosoc.2015.46.

Newton, Casey.

2020. "Facebook will pay $52 million in settlement with moderators who developed PTSD on the job." *The Verge*, May 12, 2020: https://www.theverge.com/2020/5/12/21255870/facebook-content-moderator-settlement-scola-ptsd-mental-health.

2019. "The Trauma Floor: The secret lives of Facebook moderators in America." *The Verge*, February 25, 2019: https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona.

Nicholas, Josh, and Patrick Lum.

2022. "Dall-E 2 mini: what exactly is 'AI-generated art'? How does it work? Will it replace human visual artists?" *The Guardian*, June 9, 2022: https://www.theguardian.com/culture/2022/jun/09/what-exactly-is-ai-generated-art-how-does-it-work-will-it-replace-human-visual-artists.

Noble, Brian.

2016. *Articulating Dinosaurs: A Political Anthropology*. University of Toronto Press.

Noble, Safiya Umoja.
2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.

Norcross, J. C., and B. E. Wampold.
2011. "Evidence-based therapy relationships: Research conclusions and clinical practices." *Psychotherapy* 48, 1: 98–102. https://doi.org/10.1037/a0022161.

Ochigame, Rodrigo.
2019. "The Invention of 'Ethical AI': How Big Tech Manipulates Academia to Avoid Regulation." *The Intercept*, December 20, 2019. https://theintercept.com/2019/12/20/mit-ethical-ai-artificial-intelligence/.

Ogden, Laura.
2011. *Swamplife: People, Gators, and Mangroves Entangled in the Everglades*. Minneapolis: University of Minnesota Press.

Ogden, Laura A., Billy Hall, and Kimiko Tanita.
2013. "Animals, Plants, People, and Things: A Review of Multispecies Ethnography." Environment and Society 4: 5–24. https://doi.org/10.3167/ares.2013.040102.

Orlinsky, David E., and Kenneth I. Howard.
1987. "A generic model of psychotherapy." *Journal of Integrative & Eclectic Psychotherapy* 6, 1: 6–27.

1986. "The Psychological Interior of Psychotherapy: Explorations with the Therapy Session Reports." In The Psychotherapeutic Process: A Research Handbook, edited by L. S. Greenberg and W. M. Pinsof, 477–501. New York: Guilford Press.

Osterman, Paul.
2017. "Who Will Care for Us: Long-term Care and the Long-Term Workforce." Russell Sage Foundation. https://doi.org/10.7758/9781610448673.

Ozawa-de Silva, Chikako.
2021. *The Anatomy of Loneliness Suicide, Social Connection, and the Search for Relational Meaning in Contemporary Japan*. Oakland, CA: University of California Press.

2008. “Too Lonely to Die Alone: Internet Suicide Pacts and Existential Suffering in Japan.” *Cult Med Psychiatry* 32: 516–551. http://doi.org/10.1007/s11013-008-9108-0.

Pandalfo, Stefania.
2018. *Knot of the Soul: Madness, Psychoanalysis, Islam*. University of Chicago Press.

Parreñas, Rhacel Salazar.
2015. *Servants of Globalization: Migration and Domestic Work, Second Edition*. Stanford University Press.

Pasternack, Alex.
2015. “In a world of screens, Sherry Turkle wants to make eye contact,” *Fast Company*, March 23, 2021: https://www.fastcompany.com/90613619/sherry-turkle-book-the-empathy-diaries.

Paxton, Heather.
2013. *The Life of Cheese: Crafting Food and Value in America*. Berkeley: University of California Press.

Phan, Thao, Jake Goldenfain, Declan Kuch, and Monique Mann.
2022. “Introduction: Economies of Virtue.” In *Economics of Virtue – The Circulation of ‘Ethics” in AI*, edited by Thao Phan, Jake Goldenfain, Declan Kuch, and Monique Mann: 6–23. *INC Theory on Demand* 46. Amsterdam: Institute of Network Cultures.

Pedersen, Morten Axel, Kristoffer Albris, and Nick Seaver.
2021. “The Political Economy of Attention.” *Annual Review of Anthropology* 50 309–325. https://doi.org/10.1146/annurev-anthro-101819-110356.

Peterson, Kristin.
2014. *Speculative Markets: Drug Circuits and Derivative Life in Nigeria*. Durham, NC: Duke University Press.

Piepzna-Samarasinha, Leah Lakshmi.
2018. *Care Work: Dreaming Disability Justice*. Vancouver, Canada: Arsenal Pulp Press.

Pinto, Sarah.
2014. *Daughters of Parvati: Women and Madness in Contemporary India*. Philadelphia: University of Pennsylvania Press.

Poggiali, Lisa.

2016. “Seeing (from) Digital Peripheries: Technology and Transparency in Kenya’s Silicon Savannah.” *Cultural Anthropology* 31, 3: 387–411.

Povinelli, Elizabeth.
2011. *Economies of Abandonment: Social Belonging and Endurance in Late Liberalism*. Durham, NC: Duke University Press.

Prescod-Roberts, Margaret.
1980. *Black Women: Bringing it All Back Home*. Bristol, UK: Falling Wall Press Ltd.

Price, David H.
2008. *Anthropological Intelligence: The Deployment and Neglect of American Anthropology in the Second World War*. Durham, NC: Duke University Press.

2004. *Threatening Anthropology: McCarthyism and the FBI’s Surveillance of Activist Anthropologists*. Durham, NC: Duke University Press.

Puar, Jasbir K.
2017. *The Right to Maim: Debility, Capacity, Disability*. Durham, NC: Duke University Press.

Puig de la Bellacasa, María.
2017. *Matters of Care: Speculative Ethics in More Than Human Worlds*. Minneapolis: University of Minnesota Press.

Qi, Andrew C., Karen E. Joynt Maddox, Laura J. Bierut, and Kenton J. Johnston.
2022. “Comparison of Performance of Psychiatrists vs Other Outpatient Physicians in the 2020 US Medicare Merit-Based Incentive Payment System.” *JAMA Health Network* 3, 3: e220212. https://doi.org/10.1001/jamahealthforum.2022.0212.

Raffles, Hugh.
2010. *Insectopedia*. New York: Pantheon Books.

Rahwan, Iyad, and Manuel Cebrian.
2018. “Machine Behavior Needs to Be an Academic Discipline.” *Nautilus*, March 20, 2018: https://nautil.us/machine-behavior-needs-to-be-an-academic-discipline-237022/.

Rahwan, Iyad, Manuel Cebrian, Nick Obradovich, Josh Bongard, Jean-François Bonnefon, Cynthia Breazeal, Jacob W. Crandall, Nicholas A. Christakis, Iain D. Couzin, Matthew O.

Jackson, Nicholas R. Jennings, Ece Kamar, Isabel M. Kloumann, Hugo Larochelle, David Lazer, Richard McElreath, Alan Mislove, David C. Parkes, Alex 'Sandy' Pentland, Margaret E. Roberts, Azim Shariff, Joshua B. Tenenbaum, and Michael Wellman.

2019. "Machine Behaviour." *Nature* 568, 477–486.

Raley, Rita.

2013. "Dataveillance and Countervailance." In *'Raw Data' is an Oxymoron*, edited by Lisa Gitelman, 121–145. Cambridge, MA: MIT Press.

Ramirez, Izzie.

2021. "How to Understand—and Fight—the Mental Health Care Crisis in the U.S." *Vice*, March 9, 2021: https://www.vice.com/en/article/93w43d/how-to-understandand-fight-the-mental-health-care-crisis-in-the-us.

Raphelson, Samantha.

2017. "How The Loss Of U.S. Psychiatric Hospitals Led To A Mental Health Crisis." *NPR*, November 30, 2017: https://www.npr.org/2017/11/30/567477160/how-the-loss-of-u-s-psychiatric-hospitals-led-to-a-mental-health-crisis.

Reardon, Sara.

2019. "Rise of Robot Radiologists." *Nature*, December 18, 2019: https://www.nature.com/articles/d41586-019-03847-z.

Reno, Josh.

2016. *Waste Away Working and Living with a North American Landfill*. Oakland, CA: University of California Press.

2015. "Waste and Waste Management." *Annual Review of Anthropology* 44, 557–572.

Reuben, Aaron.

2015. "When PTSD Is Contagious Therapists and other people who help victims of trauma can become traumatized themselves." *The Atlantic*, December 14, 2015: https://www.theatlantic.com/health/archive/2015/12/ptsd-secondary-trauma/420282/.

Resnick, Brian.

2018. "How data scientists are using AI for suicide prevention." *Vox*, Jun 9, 2018: https://www.vox.com/science-and-health/2018/6/8/17441452/suicide-prevention-anthony-bourdain-crisis-text-line-data-science.

Rhee, Jennifer.
2018. *The Robotic Imaginary: The Human and the Price of Dehumanized Labor*. Minneapolis: University of Minnesota Press.

Richardson, Kathleen.
2015. *An Anthropology of Robots and AI: Annihilation Anxiety and Machines*. New York: Routledge.

Robbins, Joel, and Julian Sommerschuh.
2016. "Values." In *The Open Encyclopedia of Anthropology*, edited by Felix Stein. (Facsimile of the first edition in The Cambridge Encyclopedia of Anthropology.) https://doi.org/10.29164/16values.

Roberts, Sarah T.
2019. *Behind the Screen: Content Moderation in the Shadows of Social Media*. New Haven: Yale University Press.

2017. "Social Media's Silent Filter." *The Atlantic*, March 8, 2017: https://www.theatlantic.com/technology/archive/2017/03/commercial-content-moderation/518796.

Roitman, Janet.
2013. *Anti-Crisis*. Durham, NC: Duke University Press.

Roman, P. M.
1988. "Growth and Transformation in Workplace Alcoholism Programming." *Recent Dev Alcohol* 6: 131–58. https://doi.org/10.1007/978-1-4615-7718-8_7.

Sagan, Carl.
1975. "In Praise of Robots." *Natural History* 84, 1: 8–20.

Sahlins, Marshall.
2013. *What Kinship Is—And Is Not*. University of Chicago Press.

Sanabria, Emilia.
2017. *Plastic Bodies: Sex Hormones and Menstrual Suppression in Brazil*. Durham, NC: Duke University Press.

Saunders, Stephen M., Kenneth I. Howard, and David E. Orlinsky.
1989. “The Therapeutic Bond Scales: Psychometric Characteristics and Relationship to Treatment Effectiveness.” *Psychological Assessment* 1, 4: 323–330.

Sayes, E. M.
2014. “Actor-Network Theory and Methodology: Just What Does It Mean to Say That Nonhumans Have Agency?” *Social Studies of Science* vol 44, 1: 134–149. https://doi.org/10.1177/0306312713511867.

Schalk, Sami.
2016. “Reevaluating the Supercrip.” *Journal of Literary & Cultural Disability Studies* 10, 1: 71–86.

Scheper-Hughes, Nancy.
2003. *Death without Weeping: The Violence of Everyday Life in Brazil*. Berkeley: University of California Press.

2000. “The Global Traffic in Human Organs.” *Current Anthropology* 41, 2: 191–224. https://doi.org/10.1086/300123.

1996. “Theft of Life: The Globalization of Organ Stealing Rumours.” *Anthropology Today* 12, 3: 3–11 https://doi.org/10.2307/2783143.

1995. “The Primacy of the Ethical: Propositions for a Militant Anthropology.” *Current Anthropology* 36, 3: 409–440.

Schmitt, Carl.
2005 [1922]. *Political Theology Four Chapters on the Concept of Sovereignty*. Edited and translated by George Schwab. University of Chicago Press.

Schank, Roger C.
1976. “The Weizenbaum Controversy: Response to Weizenbaum’s Response to McCarthy.” *SIGART Newsletter* No. 59: 7–8. https://dl.acm.org/doi/pdf/10.1145/1045270.1045271.

Schneider, David Murray.
1984. *A Critique of the Study of Kinship*. Ann Arbor: University of Michigan Press.

Schüll, Natasha Dow.

2016. “Data for Life: Wearable Technology and the Design of Self-Care.” *BioSocieties* 11, 3: 317–333.

2014. *Addiction by Design: Machine Gambling in Las Vegas*. Princeton University Press.

Schwartz, Oscar.

2016. “In 2016, Microsoft’s Racist Chatbot Revealed the Dangers of Online Conversation.” *IEEE Spectrum*, November 29, 2019: https://spectrum.ieee.org/in-2016-microsofts-racist-chatbot-revealed-the-dangers-of-online-conversation.

Scudellari, Megan.

2019. “‘Alexa, Tell the Nurse I’m in Pain’: An AI-powered hospital room technology is getting rave reviews from patients and nurses.” *IEEE Spectrum,* March 4, 2019. https://spectrum.ieee.org/alexa-tell-the-nurse-im-in-pain.

Searle, John.

1980. “Minds, Brains and Programs.” *Behavioral and Brain Sciences* 3: 417–57.

Seaver, Nick.

2021. “Care and Scale: Decorrelative Ethics in Algorithmic Recommendation.” *Cultural Anthropology* 36, 3: 509–537.

2019. “Captivating Algorithms: Recommender systems as traps.” *Journal of Material Culture* 24, 4: 421–436.

2018a. “What Should an Anthropology of Algorithms Do?” *Cultural Anthropology* 33, 3: 375–385. https://doi.org/10.14506/ca33.3.04.

2018b. “How to Pay Attention.” *Somatosphere*, July 30, 2018: http://somatosphere.net/2018/how-to-pay-attention.html/.

2017. “Algorithms as Culture: Some Tactics for the Ethnography of Algorithmic Systems.” *Big Data & Society* 4, 2: 1–12. https://doi.org/10.1177%2F2053951717738104.

Semel, Beth M.

2022. “Listening Like a Computer: Attentional Tensions and Mechanized Care in Psychiatric Digital Phenotyping.” *Science, Technology, & Human Values* 47, 2: 266–290. https://doi.org/10.1177/01622439211026371.

Shange, Savannah.
2019. *Progressive Dystopia: Abolition, Antiblackness, and Schooling in San Francisco*. Durham, NC: Duke University Press.

Shankar, Shalini.
2008. *Desi Land: Teen Culture, Class, and Success in Silicon Valley*. Durham, NC: Duke University Press.

Shapiro, Joseph.
2020. “People With Disabilities Say Rationing Care Policies Violate Civil Rights,” *NPR*, March 23, 2020. https://www.npr.org/2020/03/23/820398531/people-with-disabilities-say-rationing-care-policies-violate-civil-rights.

Sharp, Leslie.
2000. “The Commodification of the Body and its Parts.” Annual Review of Anthropology 29: 287–328. https://doi.org/10.1146/annurev.anthro.29.1.287.

2006. *Strange Harvest: Organ Transplants, Denatured Bodies, and the Transformed Self*. Berkeley: University of California Press.

Shaw, Jennifer, and Darren Byler.
2016. “Precarity.” *Cultural Anthropology Curated Collection “Precarity.”* https://journal.culanth.org/index.php/ca/catalog/category/precarity.

Shepherd-Powell, Julie.
2020. “Essential or Expendable? The Construction Trades on the Frontlines of Late Capitalism.” *Exertions* https://doi.org/10.21428/1d6be30e.ee5fcab2.

Simondon, Gilbert.
2017 [1958]. *On the Mode of Existence of Technological Objects*. Translated by Cecile Malaspina and John Rogove. Minneapolis: Univocal Press.

Simonite, Tom.
2020. “Algorithms Were Supposed to Fix the Bail System. They Haven't.” *Wired*, February 19, 2020. https://www.wired.com/story/algorithms-supposed-fix-bail-system-they-havent/.

Smith, Andrea.

2015. "Not-Seeing: State Surveillance, Settler Colonialism, and Gender Violence." In *Feminist Surveillance Studies*, edited by Rachel E. Dubrofsky and Shoshana Amielle Magnet, 21–38. Durham, NC: Duke University Press.

Solomon, Harris.

2022. *Lifelines: The Traffic of Trauma*. Durham, NC: Duke University Press.

Staples, James, and Tom Widger.

2012. "Situating suicide as an anthropological problem: ethnographic approaches to understanding self-harm and self-inflicted death." *Culture, Medicine and Psychiatry* 36, 2: 183–203. https://doi.org/10.1007/s11013-012-9255-1.

Steele, Chandra.

2018. "The Real Reason Voice Assistants Are Female (and Why it Matters)." *PCMag*, January 4, 2018: https://www.pcmag.com/opinions/the-real-reason-voice-assistants-are-female-and-why-it-matters.

Stevenson, Lisa.

2014. *Life Beside Itself: Imagining Care in the Canadian Arctic*. Berkeley: University of California Press.

Stewart, Kathleen.

2012. "Precarity's Forms." *Cultural Anthropology* 27, 3: 518–525. https://doi.org/10.1111/j.1548-1360.2012.01157.x.

Strathern, Marilyn.

2020. *Relations: An Anthropological Account*. Durham, NC: Duke University Press.

2015. "Afterword: Taking Relationality to Extremes." In *Suicide and Agency: Anthropological Perspectives on Self-Destruction, Personhood, and Power*, edited by Ludek Broz and Daniel Münster, 205–212. Farnham, Surrey: Ashgate.

2010. "An End and a Beginning for the Gift?" *Journal de la Société des Océanistes* 130–131: 119–127. https://doi.org/10.4000/jso.6094.

1988. *The Gender of the Gift: Problems with Women and Problems with Society in Melanesia*. Berkeley: University of California Press.

1992. *Reproducing the Future: Anthropology, Kinship, and the New Reproductive Technologies*. New York: Routledge, Chapman & Hall.

Suchman, Lucy.

2020. "Algorithmic Warfare and the Reinvention of Accuracy." *Critical Studies on Security* 8, 2: 175–187. http://doi.org/10.1080/21624887.2020.1760587.

2019. "Demystifying the Intelligent Machine." In *Cyborg Futures: Social and Cultural Studies of Robots and AI*, edited by T. Heffernan, 35–62. New York: Palgrave Macmillan.

2016. "Artificial Intelligence" *Charlie Rose*, July 7, 2016: https://www.bloomberg.com/news/videos/2016-07-07/artificial-intelligence-charlie-rose.

2008. "Feminist STS and the Sciences of the Artificial." In *The Handbook of Science and Technology Studies*, edited by Edward J. Hackett, Olga Amsterdamska, Michael Lynch, and Judy Wajcman, 139–164. Cambridge, MA: MIT Press.

2007. *Human-Machine Reconfigurations: Plans and Situated Actions*. Cambridge University Press.

Surfrin, Carolyn.

2017. *Jailcare: Finding the Safety Net for Women behind Bars*. Berkeley: University of California Press.

Szasz, Thomas.

1961. *The Myth of Mental Illness: Foundations of a Theory of Personal Conduct*. New York: Harper & Row.

TallBear, Kim.

2017. "Beyond the Life/Not Life Binary: A Feminist-Indigenous Reading of Cryopreservation, Interspecies Thinking and the New Materialisms." In *Cryopolitics: Frozen Life in a Melting World*, edited by Joanna Radin and Emma Kowal, 179–202. Cambridge, MA: MIT Press.

Taylor, Janelle S.

2008. "On Recognition, Caring, and Dementia." *Medical Anthropology Quarterly* 22, 4: 313–335. https://doi.org/10.1111/j.1548-1387.2008.00036.x.

Thakor, Mitali.

2019. “Deception by Design: Digital Skin, Racial Matter, and the New Policing of Child Sexual Exploitation.” In *Captivating Technology: Race, Carceral Technoscience, and Liberatory Imagination in Everyday Life*, edited by Ruha Benjamin, 188–208. Durham, NC: Duke University Press.

Thurnwald, Richard.

1916. *Bánaro Society: Social Organization and Kinship System of a Tribe in the Interior of New Guinea*. Memoirs of the American Anthropological Association Vol. III, No. 4. Lancaster, PA: The New Era Printing Company.

Ticktin, Miriam.

2019. “From the Human to the Planetary: Speculative Futures of Care.” *Medicine, Anthropology, Theory* 6, 3: 133–160.

2011. *Casualties of Care: Immigration and the Politics of Humanitarianism in France*. Berkeley: University of California Press.

Timcke, Scott, and Shelene Gomes.

2020. “Essentially Disposable.” *Exertions* https://doi.org/10.21428/1d6be30e.f3d5cda3.

Tishby, Orya, and Hadas Wiseman.

2018. *Developing the Therapeutic Relationship: Integrating Case Studies, Research, and Practice*. Washington, DC: American Psychological Association.

Todd, Zoe.

2016. “An Indigenous Feminist’s Take On The Ontological Turn: ‘Ontology’ Is Just Another Word For Colonialism.” Journal of Historical Sociology 29, 1: 4–22. https://doi.org/10.1111/johs.12124.

Todeschini, Maya.

2001. “The Bomb's Womb? Women and the Atom Bomb.” In *Remaking a World: Violence, Social Suffering, and Recovery*, edited by Veena Das, Arthur Kleinman, Margaret Lock, Mamphela Ramphele, and Pamela Reynolds, 102–115.

Tronto, Joan.

1993. *Moral Boundaries: A Political Argument for an Ethic of Care*. New York: Routledge.

Trouillot, Michel-Rolph.

1995. *Silencing the Past: Power and the Production of History*. Boston: Beacon Press.

Tsing, Anna.

2015. *The Mushroom at the End of the World: On the Possibility of Life in Capitalist Ruins*. Princeton University Press.

2013. "Sorting Out Commodities: How capitalist value is made through gifts." *Journal of Ethnographic Theory* 3, 1: 21–43. https://doi.org/10.14318/hau3.1.003.

2012. "On Nonscalability: The Living World Is Not Amenable to Precision-Nested Scales." *Common Knowledge* 18, 3: 505–524.

Turban, Jack.

2019. "Ghost networks of psychiatrists make money for insurance companies but hinder patients' access to care." *STAT*, June 17, 2019: https://www.statnews.com/2019/06/17/ghost-networks-psychiatrists-hinder-patient-care/.

Turing, A. M.

1950. "Computing Machinery and Intelligence." *Mind* LIX, 236: 433–460.

Turkle, Sherry.

2020. "That chatbot I've loved to hate." MIT Technology Review, August 18, 2020: https://www.technologyreview.com/2020/08/18/1006096/that-chatbot-ive-loved-to-hate/.

2018. "There Will Never Be an Age of Artificial Intimacy (Opinion)" *New York Times*: https://www.nytimes.com/2018/08/11/opinion/there-will-never-be-an-age-of-artificial-intimacy.html

2017. "Empathy Machines: Forgetting the Body." In *A Psychoanalytic Exploration of the Body in Today's World: On the Body*, edited by Vaia Tsolas and Christine Anzieu-Premmereur, 17–27. New York: Routledge.

2015. *Reclaiming Conversation: The Power of Talk in a Digital Age*. New York: Penguin Books.

2011. *Alone Together: Why We Expect More from Technology and Less from Each Other.* New York: Basic Books.

2007. *Evocative Objects: Things We Think With*. Cambridge, MA: MIT Press.

2005 [1984]. *The Second Self: Computers and the Human Spirit*. Cambridge, MA: MIT Press.

1995. *Life on the Screen: Identity in the Age of the Internet*. New York: Simon and Schuster.

1994. "Constructions and Reconstructions of Self in Virtual Reality: Playing in the MUDs." *Mind, Culture, and Activity* 1, 3: 158–167.

Turner, Terence.
2008. "Marxian value theory: An anthropological perspective." *Anthropological Theory* 8, 1: 43–56.

2003. "The beautiful and the common: Gender and social hierarchy among the Kayapo." Tipiti: *The Journal of the Society for the Anthropology of Lowland South America* 1, 1: 11–26.

Turner, Victor.
1967. *The Forest of Symbols: Aspects of Ndembu Ritual*. Ithaca, NY: Cornell University Press.

US Department of Health & Human Services, Office for Civil Rights (OCR).
2020. "HIPAA Helps Caregiving Connections: HIPAA helps mental health professionals to prevent harm." HHS-0945--F-3830. https://www.hhs.gov/sites/default/files/hipaa-helps-prevent-harm.pdf.

Vallance, Chris.
2023. "Facebook's parent firm Meta can be sued by ex-moderator, judge rules." *BBC*, February 7, 2023: https://www.bbc.com/news/technology-64541944.

2022. "Google engineer says Lamda AI system may have its own feelings." *BBC*, June 13, 2022: https://www.bbc.com/news/technology-61784011.

Verah, Chad.
1966 [1965]. *The Samaritans: To help those tempted to suicide or despair*. New York: The Macmillan Company.

Verdery, Katherine.

2014. *Secrets and Truths: Ethnography in the Archive of Romania's Secret Police*. Budapest: Central European University Press.

Vigdor, Neil.
2020. "'How Do I Get Help?' Dying Coronavirus Patient Asked Alexa." *New York Times*, April 15, 2020: https://www.nytimes.com/2020/04/09/us/Coronavirus-Alexa-ask-for-help.html.

Vigh, Henrik.
2008. "Crisis and Chronicity: Anthropological Perspectives on Continuous Conflict and Decline, Ethnos." 73, 1: 5–24. https://doi.org/10.1080/00141840801927509.

Vincent, James.
2016. "Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day." *The Verge*, March 24, 2016: https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist.

Viveiros de Castro, Eduardo.
2014. *Cannibal Metaphysics*. Translated by Peter Skafish. Minneapolis: University of Minnesota Press.

Vollers, Anna Claire.
2020. "'My worst nightmare': Alabama mom lost custody of kids after seeking postpartum help," *AL*, March 26, 2020: https://www.al.com/news/2020/03/my-worst-nightmare-alabama-mom-lost-custody-of-kids-after-seeking-postpartum-help.html.

Wakefield, Jane.
2020. "Friend challenges Facebook over Ronnie McNutt suicide video." *BBC*, September 20, 2020: https://www.bbc.com/news/technology-54176205.

Weeks, Kathi.
2017. "Down with Love: Feminist Critique and the New Ideologies of Work." *WSQ: Precarious Work* 45, 3–4: 37–58. http://doi.org/10.1353/wsq.2017.0043.

2007. "Life Within and Against Work: Affective Labor, Feminist Critique, and Post-Fordist Politics." *ephemera* 7, 1: 233–249.

Weiner, Stacy.
2018. "Addressing the escalating psychiatrist shortage." *AAMC*: https://www.aamc.org/news-insights/addressing-escalating-psychiatrist-shortage.

Weizenbaum, Joseph.

1977. “Computers as ‘Therapists.’” *Science* 198, 4315: 354. https://doi.org/10.1126/science.198.4315.354.a.

1976. *Computer Power and Human Reason: From Judgment to Calculation*. New York: W. H. Freeman and Company.

1967. “Contextual Understanding by Computers.” *Communications of the ACM* 10, 8: 474–480.

Werber, Cassie.
2022. “Effective altruism solved all the problems of capitalism—until it didn't.” *Quartz*, November 25, 2022: https://qz.com/effective-altruism-solved-all-the-problems-of-capitalis-1849820390.

Wharton, Amy S.
2009. “The Sociology of Emotional Labor.” *Annual Review of Sociology* 35: 147–65.

Wiener, Anna.
2020. *Uncanny Valley: A Memoir*. New York: MCD.

Wiener, Norbert.
1950. *The Human Use of Human Beings: Cybernetics and Society*. Cambridge, MA: The Riverside Press.

Wilkinson, Chiara.
2022. “The People in Intimate Relationships With AI Chatbots.” *Vice*, January 21, 2022: https://www.vice.com/en/article/93bqbp/can-you-be-in-relationship-with-replika.

Williams, Christopher, Philip Wilson, Jill Morrison, Alex McMahon, Andrew Walker, Lesley Allan, Alex McConnachie, Yvonne McNeill, and Louise Tansey.
2013. “Guided Self-Help Cognitive Behavioural Therapy for Depression in Primary Care: A Randomised Controlled Trial.” *PloS One* 8, 1: e52735. http://doi.org/10.1371/journal.pone.0052735.

Wilson, Elizabeth.
2010. *Affect and Artificial Intelligence*. Seattle: University of Washington Press.

Wilson, Richard Ashby, and Richard D. Brown.

2009. *Humanitarianism and Suffering: The Mobilization of Empathy*. Cambridge University Press.

Winner, Langdon.
1980. "Do Artifacts Have Politics?" *Daedalus* 109, 1: 121–136.

Wong, Alice.
2020. "I'm disabled and need a ventilator to live. Am I expendable during this pandemic?" *Vox*, April 4, 2020. https://www.vox.com/first-person/2020/4/4/21204261/coronavirus-covid-19-disabled-people-disabilities-triage.

Wool, Zoë.
2020. "On the Possibilities of Open Grief: Mourning, Affect, Attachment." *Cultural Anthropology* 35, 1: 40–47.

2015. *After War: The Weight of Life at Walter Reed*. Durham, NC: Duke University Press.

Woolgar, Steve.
1985. "Why Not a Sociology of Machines? The Case of Sociology and Artificial Intelligence." *Sociology* 19, 4: 557–72.

World Health Organization.
2015. *Mental Health Atlas 2014.*

Yergeau, Melanie.
2010. "Circle Wars: Reshaping the Typical Autism Essay." *Disability Studies Quarterly* 30, 1. https://doi.org/10.18061/dsq.v30i1.1063.

Young, Allan.
2008. "Posttraumatic Stress Disorder of the Virtual Kind: Trauma and Resilience in Post-9/11 America." In *Trauma and Memory: Reading, Healing, and Making Law*, edited by Austin Sarat, Nadav Davidovitch, and Michal Alberstein, 21–48. Stanford University Press. https://doi.org/10.11126/stanford/9780804754057.003.0002.

Zeavin, Hannah.
2021. *The Distance Cure: A History of Teletherapy*. Cambridge, MA: MIT Press.

2020. "The Third Choice: Suicide Hotlines, Psychiatry, and the Police." Somatosphere, November 2, 2020: http://somatosphere.net/2020/the-third-choice-suicide-hotlines-psychiatry-and-the-police.html/.

Zhang, Li.
2020. *Anxious China: Inner Revolution and Politics of Psychotherapy*. Oakland, CA: University of California Press.

Zigon, Jarrett.
2019. "Can Machines Be Ethical? On the Necessity of Relational Ethics and Empathic Attunement for Data-Centric Technologies." *Social Research* 86, 4: 1001–22.

Zuboff, Shoshana.
2019. *Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: PublicAffairs.