

## Critical Technical Awakenings

Maya Malik and Momin M. Malik\*

**Abstract:** Starting with Philip E. Agre’s 1997 essay on “critical technical practice”, we consider examples of writings from computer science where authors describe “waking up” from a previously narrow technical approach to the world, enabling them to recognize how their previous efforts towards social change had been ineffective. We use these examples first to talk about the underlying assumptions of a technology-centric approach to social problems, and second to theorize these awakenings in terms of Paulo Freire’s idea of *critical consciousness*. Specifically, understanding these awakenings among technical practitioners as examples of this more general phenomenon gives guidance for how we might encourage and guide critical awakenings in order to get more technologists working effectively towards positive social change.

**Key words:** critical technical practice; critical consciousness; perspective transformation; education; machine learning

### 1 Introduction

In 1997, then-UCLA professor Philip E. Agre published a remarkable essay, entitled “Towards a Critical Technical Practice: Lessons Learned in Trying to Reform AI”<sup>[1]</sup>. In it, Agre describes his experience as a doctoral student in AI at MIT in the 1980s undergoing a crisis of faith in his discipline and looking to other disciplines for answers. Agre writes (bold emphasis added):

“As an AI practitioner already well immersed in the AI literature, I had incorporated the field’s taste for technical formalization so thoroughly into my own cognitive style that **I literally could not read the literatures of nontechnical fields at anything beyond a popular level**. The problem was not exactly that I could not understand the vocabulary, but that **I insisted**

**on trying to read everything as a narration of the workings of a mechanism**.

“My first intellectual breakthrough came when, for reasons I do not recall, it finally occurred to me to stop translating these strange disciplinary languages into technical schemata, and instead simply to learn **them on their own terms**.

“**I still remember the vertigo I felt during this period**; I was speaking these strange disciplinary languages, in a wobbly fashion at first, without knowing what they meant—without knowing what *sort* of meaning they had... **in retrospect this was the period during which I began to ‘wake up’, breaking out of a technical cognitive style that I now regard as extremely constricting.**”

In this paper, we use Agre’s essay as a foil to discuss what we call *critical technical awakenings*: when people from technical disciplines, previously committed to a narrow technical view of the world, “wake up” from that perspective to what we identify as seeing the world through a critical, constructivist lens.

Other articles in this special issue do a fantastic job of analyzing the political economy of tech ethics<sup>[2, 3]</sup>. While recognizing that structural change at this level is our ultimate goal, our focus here is in taking up a specific slice of how to achieve this: what makes certain technical

• Maya Malik is with the School of Social Work, McGill University, Montréal, H3A 0G4, Canada. E-mail: maya.malik@mail.mcgill.ca.

• Momin M. Malik is with the Institute in Critical Quantitative, Computational, & Mixed Methodologies, Johns Hopkins University, Baltimore, MD 21218, USA. E-mail: momin.malik@gmail.com.

\* To whom correspondence should be addressed.

Manuscript received: 2021-05-20; revised: 2021-12-07; accepted: 2021-12-08

practitioners come to care about understanding this larger context, and how do some individuals become committed to working towards structural change? We do not mean to imply that “ethics” are a problem at the level of individuals; but, as we will argue, individual-level awakenings play a central role in building communities that effectively work towards positive structural change, and so are crucial to consider.

Our goal is not necessarily to convince people purely within a “technical perspective” that they should change (indeed, we argue that rational argumentation alone is insufficient to cause change), but rather to speak to people who are in the process of undergoing, or who have recently undergone, the type of awakening we identify. Awakenings can be a lonely and confusing process, but need not be. By pointing to existing examples and theorizing this process, and by providing guidance about how to productively channel and shape awakenings, we hope to make it less difficult to go through an awakening, and thereby encourage and contribute to growing a community of critical technical practitioners within modern data practice and technology design.

Specifically, we aim to:

- Review the existence of different ways of approaching the world and their different underlying assumptions (in Section 2);
- Identify what is initially compelling about a “technical perspective”, but how and why some of its adherents rightly come to see this perspective as insufficient (in Section 3);
- Draw on Paulo Freire’s idea of *critical consciousness* and subsequent theory from adult education<sup>[4]</sup>, in order to theorize critical awakenings more broadly (in Section 4);
- Present a specific view of ethics and argue that this should be the goal of critical technical awakenings (in Section 5);
- Examine potential shortcomings of existing examples of critical technical awakenings in light of adult education’s prescriptive positions on what makes a “complete” awakening, and by advocating for a care-based ethical code which the examples do not seem to have arrived at (in Section 6).

As a note, the awakenings we discuss are not technical in nature. Perhaps “critical-technical awakenings”, “critical awakenings in tech”, or “critical sociotechnical

awakenings” would be more appropriate; we use the phrase “critical technical awakenings” to emphasize a connection to Agre’s critical technical practice and, in contrast to other examples of people writing about “critical awakenings”<sup>[5, 6]</sup> to emphasize the awakenings in question being experienced by people in *technical* fields.

## 2 Paradigms of Social Research

Training in social research includes, as a basic part of any research methods course, an introduction to different research paradigms. For people who carry out social research from a technical background, this may not be something they have been exposed to; but even if it is, the abstract layout of different paradigms may not be meaningful. To set up the remainder of the discussion, we first present our take on the standard view of the contours of social research in Table 1, with further descriptions in a glossary Appendix, and try to point out how it relates to a technical perspective versus what people might awake to.

The rows correspond to subfields of philosophy, but here more specifically and narrowly represent types of assumptions within that philosophical domain, respectively about the nature of things (ontology), how we *can* know things (epistemology), and how we actually *go about* knowing things (methodology). While not always present in charts like this one, *axiology* is an additional branch of philosophy that contains ethics (what is good) and aesthetics (what is beautiful). Within this, we specifically care about *normative ethics*, which are choices of codes of conduct to which we should adhere (which are how we go about being ethical), as opposed to, say, descriptive ethics (descriptions of what certain people believe to be ethical).

The columns represent different paradigms of social research, and the cells are the assumptions that each paradigm makes. These assumptions are fundamental and foundational, and cannot be debated, justified, or refuted through empirical means (since, among other things, these assumptions are about the very possibility, reliability, and even definition of empirical evidence). In the Appendix, we provide a glossary with extensive descriptions of these columns and some specific terms that appear in the cells.

Neither the rows nor the columns are cleanly separated or singular; positions can bleed into one another, and a

**Table 1 Assumptions of social research paradigms. Based on Guba and Lincoln’s “Basic beliefs (metaphysics) of alternative inquiry paradigms”<sup>[7]</sup>. See Appendix for details.**

Issue	Positivism	Postpositivism	Critical theory	Constructivism	Participatory
Ontology (assumptions about the nature of things)	Naïve realism. Reality is independent of and prior to human conception of it, and apprehensible.	Critical realism: Reality is independent of and prior to human conception of it, but only imperfectly and approximately apprehensible.	Disenchantment theory: there is a reality, shaped by social, political, cultural, economic, ethnic, and gender values and solidified over time, but it is secret/hidden.	Relativism: There are multiple realities and experiences of truth, constructed in history through social processes.	Participative: multiple realities, each co-constructed through interactions between specific people and environments.
Epistemology (assumptions about how can know things)	Reality is knowable through reason and observation. It is possible to have findings that are singular, perspective-independent and neutral, atemporal, and therefore universally true.	Findings are provisionally true; multiple descriptions can be valid but are probably equivalent; findings can be affected/distorted by social and cultural factors.	The truth of findings is mediated by their value; how we come to know something, or who comes to know something, matters for how meaningful it is.	Relativistic: there is no neutral or objective perspective from which to adjudicate competing perspective or truth claims; truth is relative to a given perspective.	We come to know things, and create new understandings that can transform the world, by involving other people in the process of inquiry.
Methodology (how we go about trying to know things)	Experimental/manipulative (hypothetico-deductive); hypotheses can be verified as true. Chiefly quantitative methods, and mathematical representation.	Modified experimental/manipulative; <i>falsification</i> of hypotheses; primacy of quantitative methods, but may include qualitative and mixed methods.	Dialogic (through conversation and debate) or dialectical (through a process of thesis, antithesis, and a synthesis which becomes a new thesis)	Dialectical, or hermeneutical (a process of reading sources “against themselves” to identify inconsistencies, underlying assumptions, or implicit messages, and thereby interpret meaning).	Collaborative, action-focused; flattening researcher/participant hierarchies; engaging in self- and collective reflection; jointly deciding to engage in individual or collective action.
Axiology (ethics; values; who matters, who is important, who has standing)	Knowledge achieved through hypothetico-deductive means is more valuable than other knowledge. The people who can carry out such investigation have privileged access to the truth, and thus have a special role and importance (and potentially a special responsibility).	Knowledge achieved through hypothetico-deductive is more valuable, but can be distorted by social/cultural factors, and this can sometimes only be uncovered by qualitative means and insight. Qualitative methods can provide checks and context, or raw material for quantification.	Marginalization is what is most important; experience of marginalization provides unique insights, and the knowledge of the marginalized is more valuable than the knowledge of dominant/legitimate paradigms.	Understanding the process of construction is what is valuable; value (including valuing understanding the process of construction) is relative to a given perspective.	Everyone is valuable. Reflexivity, co-created knowledge, and non-western ways of knowing are valuable and combat erasure and dehumanization.

single column can cover a variety of irreconcilable different perspectives (for example, *logical* positivism tries to remove the ontological assumptions of realism from positivism’s quantitative empirical commitments, and conversely, mathematical realism often disdains empiricism). We identify the purest form of a “technical perspective” as falling squarely within the “positivism” column, but the perspective we discuss is more

specifically about the power of technology to effect social change.<sup>①</sup>

These columns are not exhaustive or mutually exclusive, but represent useful clusters. But, even

<sup>①</sup> This includes the perspective of *technological determinism*, a position largely rejected in social science that holds that given technology inherently effects certain causal changes, independent of context. See Green’s article in this special issue<sup>[8]</sup> for details. A softer version allows for context as a moderator, but still sees technology as having inherent causal power.

beyond this, as individuals we human beings can be inconsistent or even contradictory in the sets of assumptions we make (crossing multiple columns at different times or even at once), and we may not even be self-aware of the underlying assumptions we are making. Technical disciplines in particular are frequently positivist without realizing that it is a specific position, or that it is not the only way to see the world. Part of undergoing a critical awakening is coming to be aware that a technical perspective is only one way of looking at the world, and starting to recognize its core underlying assumptions—and reject them.

### 3 The Technical Perspective

One piece of Agre’s argument is about the importance of taking AI seriously:

“The central practice of the field of AI, and its central value, was technical formalization. Inasmuch as they regarded technical formalization as the most scientific and the most productive of all known intellectual methods, the field’s most prominent members tended to treat their research as the heir of virtually the whole of intellectual history. I have often heard AI people portray philosophy, for example, as a failed project, and describe the social sciences as intellectually sterile. In each case their diagnosis is the same: lacking the precise and expressive methods of AI, these fields are inherently imprecise, woolly, and vague. **Any attempt at a critical engagement with AI should begin with an appreciation of the experiences that have made these extreme views seem so compelling.**”

The target of Agre’s critique (and the focus of the first half of his essay) is the AI that existed in the 1980s and 1990s, a very specific and peculiar field (seeing itself as seeking to understanding mechanisms of cognition, in contrast to the machine learning of today which is instrumentally focused on achieving specific tasks and effectively unconcerned with cognition; see Ref. [9]). But the same logic remains: we begin a critical engagement with an appreciation of the experiences that make extreme technical views seem so compelling.

The specific “technical perspective” we refer to here is a position around computation and digital technology and has been identified and critiqued under a series of related terms: Morozov’s “tech solutionism”<sup>[10]</sup>; Toyama’s “tech commandments”<sup>[11]</sup>; Broussard’s “tech chauvinism”<sup>[12]</sup>; and Green’s “tech goggles”<sup>[13]</sup>. These labels emphasize something about

the arrogance and absolutism of the technical perspective, and all authors emphasize how adherents are dazzled by the apparent ability of technology (or, if engaging more with the intellectual content than the material artifacts, being dazzled by the apparent power of formalizing goals, operations, and human concepts into mathematical and/or software abstractions) to control and change the world.

As we noted above, at their purest, technical perspectives fall purely within the “positivism” column. We first review the overall appeal of positivism, before focusing specifically on its tech solutionist variety.

A statement by physicists Jean Bricmont and Alan Sokal<sup>[14]</sup> provides a pure expression:

“In the same way that nearly everyone in his or her everyday life disregards solipsism and radical skepticism and spontaneously adopts a ‘realist’ or ‘objectivist’ attitude toward the external world, scientists spontaneously do likewise in their professional work. Indeed, scientists rarely use the word ‘realist,’ because it is taken for granted: *of course* they want to discover (some aspects of) how the world really is! And *of course* they adhere to the so-called correspondence theory of truth (again, a word that is barely used): if someone says that it is true that a given disease is caused by a given virus, she means that, in actual fact, the disease is caused by the virus.

“We would not even call it a ‘theory’; rather, we consider it a *precondition for the intelligibility* of assertions about the world.”

This captures something about the aesthetic appeal of positivism and specifically its realist ontology: the world is fundamentally *knowable*. Furthermore, the technical person experiences the satisfaction of having command of the sole means by which to achieve that knowledge.

While, as suggested in this quote, this perspective is widespread in the natural, mathematical, or “hard” sciences, “positivism” was actually coined as an aspiration for social science in the 19th century (see Appendix). Past that period, Porter<sup>[15]</sup> describes post-WWII behavioralists adopting quantitative methodologies in social science in pursuit of “*liberating essence* of a proper objective methodology” that could “*rise above stubborn tradition and invisible culture*” (emphasis added). That is, they pursued a vision where it is possible to know how the social world “really is”, such that it is possible to have intelligible assertions about it (rather than “stubborn tradition and invisible

culture” getting in the way of intelligibility).

This idea of liberation through science leads to a view where quantification and formalization are not only *practically* superior, but *morally* superior as well. Everything else in the world is anecdotal evidence, naïve heuristics, and armchair philosophy—shackles of ignorance either useless for accomplishing concrete goals and characterized only by failure, or achieving success only through sheer luck or cheap trickery. That is, even if there is a case where technical approaches are not practically superior (like, for example, convincing climate change deniers), there is a view that they are *morally* superior: even if attempting to understand or intervene in the world through means other than abstraction (i.e., through means like through rhetoric, or narrative) may succeed, those alternatives are dishonest, unprincipled, or otherwise somehow ignoble and compromise our moral integrity.

In addition to this intrinsic moral superiority, positivism seems to comport well with a basis for morality. An observer-independent external world also justifies universal morality—a standard which we can hope to define, and then appeal to for solving moral questions. Indeed, in the so-called “science wars” of the 1990s, when some scientists (initially led by Alan Sokal) took up arms against what they saw as the “fashionable nonsense” of science and technology studies (and related areas), those scientists also bemoaned that while they and the “postmodernists” seemed to share progressive political goals of greater justice and equity, the postmodern perspective was undermining the basis for pursuing that goal and the basis of forming coalitions.

Even worse than getting in the way, “postmodern” arguments are in fact deployed in support of<sup>[16, 17]</sup> and by climate change deniers, creationists, and all sorts of religious nationalists and right-wing movements across the world. These reactionary elements of society seek to undermine the legitimacy of science in pursuit of a regressive political agenda, and while they clearly believe in a single reality (corresponding to their own beliefs), they co-opt language around plurality and relativism to prevent critique. One of the more forceful arguments around this is by Nanda<sup>[18]</sup>, who argues how Enlightenment beliefs in universality are what we need to defend against perspectives like those of Hindu nationalists, whose weaponization of science studies she documents.

The computation- and technology-focused variety of positivism discussed by Morozov<sup>[10]</sup>, Toyama<sup>[11]</sup>, Broussard<sup>[12]</sup>, and Green<sup>[13]</sup> is not necessarily about understanding the world, but about acting within it. Toyama discusses (before undergoing what seems like an awakening) thinking technology addresses “real problems”; that both means that the problems are prior to and independent of the perspective of the technologists, and that technology in itself can actually address and solve those problems. Morozov lists examples of Silicon Valley rhetoric about technology changing the world and solving global problems. He summarized the implicit technologist vision of the future in a satirical prediction:

“If Silicon Valley had a designated futurist, her bright vision of the near future... would go something like this: Humanity, equipped with powerful self-tracking devices, finally conquers obesity, insomnia, and global warming as everyone eats less, sleeps better, and emits more appropriately. The fallibility of human memory is conquered too, as the very same tracking devices record and store everything we do. Car keys, faces, factoids: We will never forget them again...

“Politics, finally under the constant and far-reaching gaze of the electorate, is freed from all the sleazy corruption, backroom deals, and inefficient horse trading. Parties are disaggregated and replaced by Groupon-like political campaigns, where users come together—once—to weigh in on issues of direct and immediate relevance to their lives, only to disband shortly afterward. Now that every word—nay, sound—ever uttered by politicians is recorded and stored for posterity, hypocrisy has become obsolete as well. Lobbyists of all stripes have gone extinct as the wealth of data about politicians—their schedules, lunch menus, travel expenses—are posted online for everyone to review...

“Crime is a distant memory, while courts are overstaffed and underworked. Both physical and virtual environments—walls, pavements, doors, and log-in screens—have become ‘smart.’ That is, they have integrated the plethora of data generated by the self-tracking devices and social-networking services so that now they can predict and prevent criminal behavior simply by analyzing their users. And as users don’t even have the chance to commit crimes, prisons are no longer needed either. A triumph of humanism, courtesy of

Silicon Valley.”

This is a synthetic caricature, but we can use it to discuss what might be compelling in the perspective that Morozov identifies and critiques. There is a view that technology is practically superior, in that it will succeed where stubborn tradition and invisible culture have failed. But also, tradition and culture are the cause of social problems in the first place; technology is not compromised by their failings, and thus to approach social problems with technology rather than society is a morally superior and more responsible move.

There is an ignominious aspect of the appeal of this technical perspective as well, which Broussard shows. She argues that technologists, who are frequently white, male, and upper-class, fixate on technology as a way to try and solve social problems traditionally managed by people who are Black, women, and/or poor. These men seek to use technology to avoid engaging with the complex and messy labor and understandings these groups have mobilized to manage and address social problems. That is, part of the appeal to the technical perspective is a chauvinistic one: of providing a means to distance oneself from the knowledge, labor, and even existence of devalued people who are women and/or non-white. If we just invent the right device, formalization, or processes, the thinking goes, we can avoid needing to deal with all the ambiguities, nuances, and emotional labor with which, say, Black women social workers engage.

These are the appeals of a technical perspective. What, then, leads people away from it? In awakenings, a common theme seems to be a precipitating event or moment that put the sleeper into a moment of crisis. For Agre, what he described is fairly abstract and intellectual: when trying to decide on a dissertation topic, he found that “Every topic I investigated seemed driven by its own powerful internal logic into a small number of technical solutions, each of which had already been investigated in the literature”. In his description, it was his search for a novel topic led him to read the literatures of other disciplines.

Agre does allude to a “large and diverse set of historical conditions” beyond what he presents in the essay. But as he does not elaborate on this, we turn to two other examples of described awakenings, respectively from Kentaro Toyama and Phil Rogaway.

First, we consider Kentaro Toyama, who rejected a

technical perspective in a rather “scientific” way. In his book *Geek Heresy: Rescuing Social Change from the Cult of Technology*<sup>[11]</sup>, he describes working after his PhD on “ICT4D”-type projects (Information and Communication Technologies for Development) for Microsoft in India. His position involved expanding technology products’ audiences beyond the educationally advantaged Indian middle class to try and help those in poverty. But he repeatedly found his attempted interventions failing.

“In the course of five years, I oversaw at least ten different technology-for-education projects. We explored video-recorded lessons by master teachers; presentation tools that minimized prep time; learning games customizable through simple text editing; inexpensive clickers to poll and track student understanding; software to convert PowerPoint slides into discs for commonly available DVD players; split screens to allow students to work side by side; and on and on. Each time, we thought we were addressing a real problem. But while the designs varied, in the end it didn’t matter—technology never made up for a lack of good teachers or good principals. Indifferent administrators didn’t suddenly care more because their schools gained clever gadgets; undertrained teachers didn’t improve just because they could use digital content; and school budgets didn’t expand no matter how many ‘cost-saving’ machines the schools purchased. If anything, these problems were exacerbated by the technology, which brought its own burdens.

“These revelations were hard to take. I was a computer scientist, a Microsoft employee, and the head of a group that aimed to find digital solutions for the developing world. I wanted nothing more than to see innovation triumph, just as it always did in the engineering papers I was immersed in. But exactly where the need was greatest, technology seemed unable to make a difference.”

This was “scientific” in the sense that Toyama was open to evidence by which he tested his assumption that technical tools can circumvent the messiness of society. But the fact that he was even able to recognize that he had such foundational assumptions is not a given; Toyama contrasts his insights to the perspective of a prominent technologist, One Laptop Per Child founder Nicholas Negroponte:

“I was once on a panel at MIT with Negroponte where I outlined my hard-won lessons about technology for education. He didn’t like what I said, and he went on the

offensive. But he did it with such confidence and self-assurance that, as I listened, I felt myself wanting to be persuaded: Children *are* naturally curious, aren't they? Why *wouldn't* they teach themselves on a nice, friendly laptop?

“As I heard more of the technology hype, however, I realized that it didn't engage with rigorous evidence. It was empty sloganeering that collapsed under critical thinking.”

That is, many scientists and technologists are not, in this sense, open to a particular type of empirical evidence. This is not inherently bad or even “unscientific”—work in the history, sociology, and philosophy of science points out that interpretations of empirical evidence require layers of theories and assumptions<sup>[19]</sup>, including the idea that evidence can be erroneous due to human error, issues with instrumentation, or natural variability. Indeed, skepticism of evidence that challenges established theory is an important part of science: but this is all to say, evidence alone is not enough to change minds, such as in an awakening. Kuhn<sup>[20]</sup> famously theorized that one-off failures in experimental science seldom affect theory, but strings of failures can precipitate a *crisis*, potentially leading to a *paradigm shift* in understanding and defining basic scientific concepts differently (and, conversely, it takes a crisis and not simply routine failures to produce a paradigm shift).

Second, we look at the account of cryptographer Phil Rogaway in his essay, “The Moral Character of Cryptographic Work”<sup>[21]</sup>. For Rogaway as well, there was a discrete empirical event that led to his identifying and rethinking some fundamental assumptions, but here the challenge posed was a moral one rather than one of assumptions about how the world works not fitting evidence.

“Most academic cryptographers seem to think that our field is a fun, deep, and politically neutral game—a set of puzzles involving communicating parties and notional adversaries. This vision of who we are animates a field whose work is intellectually impressive and rapidly produced, but also quite inbred and divorced from real-world concerns. Is this what cryptography *should* be like? Is it how we *should* expend the bulk of our intellectual capital?

“For me, these questions came to a head with the Snowden disclosures of 2013. If cryptography's most

basic aim is to enable secure communications, how could it not be a colossal failure of our field when ordinary people lack even a modicum of communication privacy when interacting electronically? Yet I soon realized that most cryptographers didn't see it this way. Most seemed to feel that the disclosures didn't even implicate us cryptographers.”

Also noteworthy is how both Rogaway and Toyama (and Agre as well) describe resistance from their peers to their crisis of faith, and how the experience that led to their transformation did not succeed in triggering others. This contrast again emphasizes that evidence, or external triggers, are not sufficient to cause an awakening; they are only catalysts for already-existing potential.

These accounts do not reflect on what made their authors different from their peers. But understanding these accounts through the lens of adult education and specifically work on critical consciousness (see Appendix), below, will help fill in key answers.

We can also contrast these descriptions to others who, while recognizing the limitations of purely technical approaches, remain within a positivist paradigm (or, at most, soften to a post-positivist one).

Physicist and applied mathematician turned sociologist Duncan Watts<sup>[22]</sup> wrote that “many of the ideas and metrics of the ‘new’ science of networks have either been borrowed from, or else rediscovered independently of, a distinguished lineage of work in mathematics, economics, and sociology”, acknowledging sociological contributions but reading them in an essentially positivist light. Another person trained in physics and working in network science, César Hidalgo<sup>[23]</sup>, wrote about realizing why “social and natural scientists fail to see eye to eye”: “Social scientists focus on explaining how context specific social and economic mechanisms drive the structure of networks and on how networks shape social and economic outcomes. By contrast, natural scientists focus primarily on modeling network characteristics that are independent of context, since their focus is to identify universal characteristics of systems instead of context specific mechanisms”. This again positions social science's role by reference to the task of finding universal and objective truths, rather than understanding that (at least some) social science rejects the idea that there could be universal characteristics.

A more personal potential example is Hannah

Wallach’s viewpoint, “Computational Social Science ≠ Computer Science + Social Data”<sup>[24]</sup>. In this she writes, “Despite all the hype, machine learning is not a be-all and end-all solution. We still need social scientists if we are going to use machine learning to study social phenomena in a responsible and ethical manner.” A dilemma was only hinted at:

“When I first started working in computational social science, I kept overhearing conversations between computer scientists and social scientists that involved sentences like, ‘I don’t get it—how is that even research?’ And I could not understand why. But then I found this quote by Gary King and Dan Hopkins—two political scientists—that, I think, really captures the heart of this disconnect: ‘computer scientists may be interested in finding the needle in the haystack—such as... the right Web page to display from a search—but social scientists are more commonly interested in characterizing the haystack.’

“In other words, the conversations I kept overhearing were occurring because the goals typically pursued by computer scientists and social scientists fall into two very different categories... models for prediction are often intended to *replace* human interpretation or reasoning, whereas models for explanation are intended to *inform* or *guide* human reasoning.”

But what she describes overall only goes so far as to recognize the importance of *quantitative* social science—areas of economics like econometrics and game theory, and political science, all of which build formal models for the task of causal understanding. There is no mention of “thick” disciplines that do not use quantitative modeling, such as cultural anthropology, critical sociology, critical race studies, human geography, critical gender studies, media studies, or cultural studies, let alone any mention of experiential ways of knowing outside of academic disciplines.

Like with Agre, from this piece alone it is impossible to know if this encapsulates Wallach’s understandings, or if it is rhetorical strategy (indeed, in a later piece, Wallach<sup>[25]</sup> seems to go beyond post-positivism in recognizing that the notions of “objectivity” are both ill-defined and not desirable, as well as acknowledging *positionality*<sup>②</sup> [see Appendix]). After all, it is much

② “Will these changes of always having a sociotechnical lens make machine learning less fun? Maybe, for some people. But that is their privilege talking about their ethical debt. Machine learning has never been all that fun for people who are involuntarily represented in datasets or subject to uncontested life-altering decisions made by machine learning systems.”

easier to convince computer scientists of the value of the formalism- and data-heavy discipline of economics than of interpretive disciplines like cultural studies, or of knowledge that comes from lived experience.

#### 4 Critical Awakenings

Earlier, we mentioned Kuhn’s idea of paradigm shifts. Recognizing that this may be too simple a model for scientific development<sup>[26]</sup>, Mezirow<sup>[27, 28]</sup> offers a similar model but instead describing individual psychosocial development, which he called *perspective transformation*. More immediately, Mezirow’s idea comes from the work of Paulo Freire and his idea of *critical consciousness* (see Appendix), and has a robust body of follow-up work investigating the idea empirically<sup>[29]</sup> and developing it theoretically<sup>[30–32]</sup>. We will also draw on subsequent work that has noted shortcomings in Mezirow’s theory not going far enough in considering context, other cultural settings, and the significance of interpersonal relationships<sup>[32]</sup>.

Perspective transformation came from Mezirow’s study with women who re-entered college programs mid-life. He identified the ultimate value of such programs as being in the personal transformation that took place among the women, rather than any material outcomes. He theorized 10 stages of this process:

- “(1) A disorienting dilemma;
- “(2) Self-examination with feelings of fear, anger, guilt, or shame;
- “(3) A critical assessment of assumptions and a sense of alienation from taken-for-granted social roles and expectations;
- “(4) Recognition that one’s discontent and the process of transformation are shared and that others have negotiated a similar change;
- “(5) Exploration of options for new roles, relationships, and actions;
- “(6) Planning a course of action;
- “(7) Acquiring knowledge and skills for implementing one’s plans;
- “(8) Provisional trying of new roles;
- “(9) Building competence and self-confidence in new roles and relationships;
- “(10) A reintegration into one’s life on the basis of conditions dictated by one’s new perspective.”

These ten stages are somewhere between descriptive and normative. They are descriptive, insofar as they



describe a process undergone by the subjects of Mezirow's study, but normative, insofar as Mezirow identified perspective transformation as something valuable and possibly aided by knowing about this sequence in advance and following it (following Freire, and the idea of critical consciousness as a normative goal). While this alone does not necessarily shed light on *who* would experience a dilemma as disorienting and change in response (since Mezirow encountered women already pursuing a change), it does point to how this change does not happen in isolation, and indeed how connecting with others who have negotiated a similar change is key for shaping awakenings towards productive ends. But Mezirow<sup>[33]</sup> does provide an answer for the question of what is needed beyond evidence, observing that an additional condition is that a person *reflect* about assumptions and beliefs that structured how they understood an experience (or evidence).

Also noteworthy are the examples of disorienting dilemmas: they included “the death of a husband, a divorce, the loss of a job, a change of city of residence, retirement, an empty nest, a remarriage, the near fatal accident of an only child, or jealousy of a friend who had launched a new career successfully”. In comparison, the *dilemmas* of Agre, Toyama, and Rogaway are decidedly elite and privileged experiences. Still, we can identify critical technical awakenings as a specific form of a much more general phenomenon of critical consciousness, thus making it appropriate to theorize with perspective transformation.

There are several lessons to draw from this connection. The first is how critical technical awakenings may relate to critical consciousness (CC) overall. Jemal<sup>[34]</sup> notes that much work on critical consciousness has deliberately excluded privileged populations, but argues this exclusion “...may inadvertently support the proposition that oppression is a problem for the oppressed to solve. When, in essence, CC is important for members of privileged groups who have greater access to resources and power and can operate as allies privileged by the system of social injustice, unfair distribution of resources and opportunities, and inequity, be able to recognize unjust social processes and acquire the knowledge and skills needed for social change.”

Drawing from Freire, she continues:

“It is imperative that those who may be privileged by

the system of social injustice, unfair distribution of resources and opportunities, and inequity, be able to recognize unjust social processes and acquire the knowledge and skills needed for social change... CC would help individuals understand their role in a system of oppression, as members of either the privileged or stigmatized groups. Liberation requires true solidarity in which the oppressor not only fights at the side of the oppressed, but also takes a radical posture of empathy by ‘entering into the situation of those with whom one is solidary’.<sup>[35]</sup> Thus, CC, with the goal of liberation, has the radical requirement that the oppressor, those who deny others the right to speak their word, and the oppressed, those whose right to speak has been denied, must collaborate to transform the structures that beget oppression.<sup>[35]”</sup>

The second is that all of the descriptions of possible critical technical awakenings do not recognize “that one’s discontent and the process of transformation are shared and that others have negotiated a similar change”. From the perspective of Mezirow’s theory, this means they fall short. Indeed, our article here is an attempt to directly address the fragmentary nature of narratives of critical technical awakenings, and to draw connections between people’s experiences. We can also continue the normative route, and note that in order to fully achieve the potential for social change from critical technical awakenings, we should try to see how to continue past stage (5) and on to stages (6)–(10).

What might new roles (stages (5)–(9)) be, in which technical practitioners should build competence and self-confidence, and make provisional efforts? We suggest that one role might be in opposing gatekeeping. It is rare even for qualitative researchers to have a seat at the table of technological adoption, let alone communities affected by it. But by leveraging the social standing that comes with quantitative legitimacy, and translating concerns into terms that are (more) acceptable for technical audiences as a first step, technical practitioners can help bring others into the processes of technology development—whether to participate, or to oppose development and deployment that does not empower those communities.

The relationships that come with those roles would be with allies outside of technical disciplines and sectors, and particularly through learning from and working with communities affected by technology (whether directly,

by a technology itself, or indirectly, such as in gentrification resulting from real estate expansions by the tech industry or of universities who receive influxes of tech money). These would be new roles not only for the technical practitioners, but indeed new social roles, and would require weathering all the difficulties of negotiating roles outside of recognized categories.

Drawing on the follow-up work to Mezirow, we also draw attention to the importance of looking at perspective transformations outside of frames of self-realization<sup>[31]</sup>, and indeed outside of depicting the process as a deeply *rational* one in molds of western rationality. One example is a study that identifies disorienting dilemmas among women in Botswana that led to questioning assumptions, but with the value of the outcome being oriented towards the spiritual, community responsibility and relationships, and gender roles<sup>[36]</sup>. Indeed, acknowledging other ways of knowing that are not expressed in the language of rationality makes perspective transformation far less novel. Johnson-Bailey<sup>[37]</sup>, coming from the perspective of a Black woman, writes about “transformational learning as the only medium in which we exist, learn, and teach. Since it is the air we breathe, maybe we just take it for granted and didn’t attend to or claim it sufficiently.” This is also an example of a more general issue; in “The Race for Theory”, Barbara Christian<sup>[38]</sup> wrote, “people of color have always theorized—but in forms quite different from the Western form of abstract logic... our theorizing is often in narrative forms, in the stories we create, in riddles and proverb, in the play with language, since dynamic rather than fixed ideas seem more to our liking. How else have we managed to survive with such spiritedness the assault on our bodies, social institutions, countries our very humanity? ...My folk, in other words, have always been a race for theory”.

The third is in looking at recommendations from adult education about how we might encourage perspective transformations. Unfortunately, as Taylor and Snyder<sup>[32]</sup> note, work has focused on support based around assumptions from Mezirow, “such as creating a safe and inclusive learning environment, focusing on the individual learner’s needs, and building on life experiences”. One strand of work that does go beyond Mezirow’s assumptions looks at how the significance of spontaneous action depends on social recognition. That work finds that what would otherwise be a spontaneous

action becomes personally meaningful when others point it out and provide positive feedback about it.

Combining these strands together, we can say: those who have undergone a critical technical awakening should think about relationships with others in which we create safe and inclusive learning environments, facilitate opportunities for experience, serve as guides who can give focus to specific learning needs, and give positive feedback around disorienting dilemmas and other opportunities for reflecting and questioning assumptions.

While these principles were developed in opposition to existing formalized education, there may be opportunities to incorporate them into formal education as well. Trbušić<sup>[39]</sup> argues for integrating critical methods into engineering education as a way of making ethics more than a superficial part of training. She specifically suggests using Augusto Boal’s technique of *Theatre of the Oppressed*<sup>[40]</sup> (itself based on the work of Freire, with whom Boal was friends), using improvisation and role-playing to encourage critical consciousness. Incorporating role-playing with scenarios where engineering students are put into ethical dilemmas could encourage taking an active stance, trying different roles, and stimulating reflection in a way that presenting formal models of ethics would not.

Especially insofar as critical technical awakenings may fall short more than other types of critical consciousness, there is also a task for how to deepen our own awareness and practice. Taylor and Snyder<sup>[32]</sup> identify work about “social accountability”, where a moral underpinning is an outcome of transformative learning. More specifically, “the outcome of transformative learning involves recognizing the reasons why, for what purpose, and for whom a new identity was constructed”, especially as an essential component of trusting relationships<sup>[41]</sup>. Having transformations be ethically grounded for what kind of world we want to see and work towards, and making this a focus of interpersonal relationships and community-building, can also help achieve more complete and powerful transformations.

## 5 Ethics

Earlier, we raised reasons why it seems like positivism is compelling as a basis for ethics. But Rogaway and Toyama’s accounts, in particular, get at how positivism

and technical disciplines are harmful in the consequences of their epistemological assumptions: if quantitative forms of knowledge are superior, then other forms of knowledge are inferior. Consequently, those who do not hold quantitative knowledge do not have anything to offer.

De Sousa Santos<sup>[42, 43]</sup> discusses the interconnection of ecologies of knowledge and how people are valued. When knowledge is put in hierarchies, it also places people into hierarchies. Sylvia Wynter, in her landmark work on “No Humans Involved”<sup>[44]</sup>, has a stark presentation of this idea. Her title refers to a term used by the Los Angeles Police Department to classify police encounters where they enacted violence on young Black men who were jobless in the inner city: by saying that these encounters did not involve “humans”, the department excused themselves from documenting their use of force and gave them a license to continue. The literal, administrative category reflected metaphorical dehumanization: there is no brutality or injustice if the targets are not human.

Critical, constructivist, and participatory paradigms link epistemology and axiology, saying: how do we value people, if we do not value their knowledge? Even post-positivism is insufficient; we can see calls for “Human-Centered AI”, or “Human-Centered Machine Learning”, or “Human-Centered Data Science” as fitting into a post-positivist frame, where we pursue objective knowledge and “real” technology that is focused around the figure of the human and its subjectivity. But human-centeredness does not address dehumanization, who gets recognition as being in the category of “human”, and how exclusion happens (e.g., being “human” is reserved for people who look, talk, think, act, and exist in certain ways). Any form of human-centered computing that takes the category of “human” for granted will not undo the status quo of what Wynter calls “narrative condemnation”. Participatory approaches, in particular, start with the proposition that everyone is valuable, and then derive knowledge from there.

As in the premise of critical theory, the Enlightenment led to or at least did not prevent the atrocities of the Holocaust, to which we can also add the atrocities of indigenous genocides in the Americas and Australia, the brutality of colonialism like in the anthropogenic Bengal famine or the atrocities in Congo Free State, and especially the trans-Atlantic slave trade. Science was a

weapon to dehumanize and make exclusionary standards for moral standing throughout history<sup>[45]</sup>. It was utilized as a tool to control otherized populations, alienate them from the public sphere, and remove them from societal participation. Pretending these things did not happen, or pretending as though they were aberrations from the natural course of science, does nothing to prevent them from happening in the future. Atrocity and oppression cannot happen without devaluing entire groups of people, and excluding them from belonging to the same sort of category of being; this is the only way we can apply different standards, for example, of surveillance or accountability or resource distribution or violence to people based on different labels (e.g., criminal, immigrant, welfare beneficiary, and foreign citizen). Then, instead of making universal morality the basis of our ethics, we should seek to dismantle knowledge hierarchies. We should valorize knowledge creation that resisted and persisted through dehumanization<sup>[46]</sup> through empirical but also artistic, narrative, and cultural means, and see these as no lesser than quantitative forms of knowledge.

We advocate specifically for the *ethics of care* from Black feminist frameworks<sup>[47–50]</sup>. Traditionally, descriptive ethics have linked recognition, belonging, and moral standing: normatively, the way to be ethical, and achieve justice, is to extend recognition, equal standing, and the protection of rights to people who have been marginalized and excluded (such as by bringing marginalized people into full participation in the public sphere, or by policies framed around safeguarding human or civil rights). In contrast, the ethics of care is a normative ethical position that reacts to the ethics of recognition and how it descriptively concedes to “recognition” as being an acceptable basis for treatment. This ethical position is found in a long history of the labor of Black women (including potentially not under the explicit label of “ethics of care”<sup>[51]</sup>), specifically in Black feminist circles and in value-based social services disciplines<sup>[52]</sup> like social work, thinking about how to have ethical and holistic interpersonal relationships, and focusing on care for marginalized people<sup>[53–55]</sup>. Instead of recognition, the basis of these ethics is empathy, love, and connection, coming from non-Eurocentric world-views, and advocating treating every living being with care. Scaling up interpersonal care to systems creates a principle that systems must serve the most

marginalized and disadvantaged, rather than those people needing to fit into systems or gain social capital before they are respected or considered important.

## 6 Traps

A critical technical awakening destabilizes a positivist worldview, opening up the possibility of a perspective transformation that leads to people working with deliberation and awareness towards a better world. But it is not sufficient. In a reflection of the language of Selbst et al.<sup>[56]</sup> who talk about five “traps” of the (positivist) formalisms of computer science, we discuss two traps in critical technical awakenings that reject positivism but may fail to achieve genuine transformation. There are other traps as well, for example co-option, as discussed in other articles in this special issue<sup>[2, 3]</sup>, but here we discuss *incomplete awakenings*, and *technical abandonment*.

The first and most important trap is of *incomplete awakenings*, where one’s perspective only widens somewhat, and specifically does not get past knowledge hierarchies. We have sketched out a particular normative path for an awakening, with this dismantlement as the goal. But none of the critical technical awakenings we identify necessarily get this far. Agre’s characterization of his awakening, for example, seemed more like it was about intellectual fulfillment, and (at least from the description) did not engage with positionality. What he describes is coming to see some other forms of elite knowledge, namely those from the humanities and social sciences, as superior to his former narrow technical worldview.

The blindness Broussard<sup>[12]</sup> identifies of technologists to other forms of knowledge from experience is not ever recognized or addressed in Agre’s work. Again, the work may not reflect the full extent of Agre’s experience, and it may do so in a particular rhetorical strategy of not trying to overturn positivism *and* knowledge hierarchies all at once; but, this is a theme across the other descriptions of awakenings as well. In none of them is there a recognition of the existence and value of other very different forms of knowledge, or the value of the people who hold those other forms of knowledge.

The second trap is a more subjective one: that of *abandonment*. There is a temptation, upon having an awakening and becoming disillusioned, to abandon technical work entirely. We argue this is bad for two

reasons. The first is a strategic one: at the risk of reifying quantification and technology, we believe that there is a role for those trained in these methods to push back and develop critiques in “internal” terms that can be intelligible to those still in a technical mindset (and perhaps even leading others to having their own critical technical awakening). These are some of the potential “new roles”, as in Mezirow’s ten steps, we explore above. This temptation is parallel to how, upon recognizing and becoming disillusioned with privilege, one temptation is to attempt to reject that privilege; but, such attempts do not actually erase the privilege one has benefitted from in the past. Finding ways to engage with and leverage this privilege is the more responsible course.

The second reason we argue against abandonment is more abstract and speculative. Just as modern qualitative research originated in the oppressive project of colonial anthropology but has since worked to reform on grounds of being reflexive and pursuing justice, so too might quantitative research move away from positivism<sup>[57]</sup>.

Given that quantification is about abstraction<sup>[56]</sup>, and abstraction flattens meanings<sup>[15]</sup>, it is difficult to imagine quantitative knowledge that can be reflexive and acknowledge other forms of knowledge, but is worth exploring. Agre’s own suggestion of a “critical technical practice” is itself a call to continue creating technical knowledge, but through a critical lens. What that might mean or how it might look is unclear from Agre’s work or the handful of subsequent works that have taken on that label, but the development of technical knowledge on something other than a realist ontology and a hierarchical axiology can be seen as a worthwhile challenge.

## 7 The Path Forward

Despite being a powerful expression of a profound shared experience, Agre’s call for “critical technical practice” has largely languished for the past two decades. For personal reasons, Agre himself has not been active in academia<sup>[58]</sup> to continue exploring and developing this idea himself. Critical technical practice has been continued by a few people, like Phoebe Sengers<sup>[59]</sup>, but even that has been mostly within design and Human-Computer Interaction<sup>[60, 61]</sup>, rather than in more formal mathematic and technical areas where critical and constructivist approaches are most alien.

As discussed before, one key missing element from

Agre's narrative and those of others is Mezirow's stage (4), "Recognition that one's discontent and process of transformation are shared and that others have negotiated a similar change." While it is hard to say why critical technical practice failed to take hold—Agre no longer being active in academia? The original essay not having any clear statement of what, exactly, critical technical practice is or looks like? Critical technical practice not being a good way to productively channel awakenings? There not being enough awakenings to form a critical mass? Agre simply being ahead of his time<sup>[62]</sup>?—building community and coalitions seems to be a critical missing step.

Some of what we detail in sources of awakening suggest ways that we can try to encourage more people with a technical perspective to undergo critical awakenings: exposure to anti-positivist and anti-realist ideas, putting them in contact with non-technical individuals, and finding ways to attack compartmentalization (as is done in other articles in this collection like those of Green<sup>[8]</sup>, and in the design method that Stark<sup>[63]</sup> offers). Or, if these were integrated in technical education sufficiently early on<sup>[25, 39]</sup>, perhaps people would never develop a distinctly technical perspective and would not need (as abrupt of) an awakening, in a topic that also relates to the article in this special issue by Korn<sup>[64]</sup>. This article (as well as that of Hu<sup>[3]</sup>) also partially take the form of personal reflections, which are central in critical awakenings; while we have chosen, primarily for reasons of length and coherence, to make this essay a primarily informational and analytic one rather than discuss our experiences, we cite these articles as examples of how we should seek to create more opportunities for technical practitioners to, respectively, engage in their own personal reflections as a technical practitioner<sup>[3]</sup> and with the experiences of others<sup>[64]</sup>.

Seeking out perspectives from others, both contemporary and historical, is one way to break through ossified visions. In "Informatics of the Oppressed", Ochigame describes in English for the first time two Latin American informatics projects<sup>[65]</sup>. First, Cuban librarians and computer scientists in the 1980s, facing US embargoes, set up an alternative information indexing and retrieval system whose mathematical model, among other features, adjusted readership-based indexes by the number of librarians in recognition of the

"author-reader social communication that happens in libraries". Second, liberation theologians in Brazil resisting the post-1964 military dictatorship set up a print and mail-based "intercommunication network" to solicit and internationally distribute writings by those most subjected to domination, in a vision of advancing Freire's project past a need for intermediaries and towards "'inter-conscientization' between the oppressed". Ochigame notes that these projects were, like libertarian fantasies coming out of California, overly optimistic in what technology (alone) would achieve; but these visions were still valuable in the alternative they offered to ranking based only on productivity or popularity (in Cuba), and in justifying and structuring dissemination not just in terms of free speech or in the politics of "whether one is free to speak, but whose voices one can hear and which listeners one's voice can reach" (in Brazil). We can take inspiration from these alternative visions, and seek out others that have similarly been silenced and pushed aside (indeed, Ochigame's discovery of these projects came through personal meetings, and not online searches). Those of us trained in technology development and quantitative forms of knowing should try to build on these, and explore alternative visions. We hold that the potential value of quantitative knowledge *outside* of its connection to and role in upholding power, hierarchy, and privileged access to truth have yet to be fully explored.

Another key part of any path forward is to *build community* to encourage, support, and guide critical technical awakenings, and channel those who undergo such awakenings towards developing a critical technical practice. Here, we can point to conference workshops<sup>[60, 61, 63, 66]</sup>, networks like the one formed from the Ethical Tech Working Group that generated this special issue, fellowship cohorts, and mentorship as paths forward. But as a caveat, while community-building aimed at reaching technical practitioners will most likely need to operate within institutional elitism (indeed, like the Ethical Tech Working Group being at Harvard), this should only be one part of larger community-building. After all, during his exile under Brazil's 21-year military dictatorship, Freire also spent a year as a visiting professor at Harvard; but he eventually returned to Brazil and continued to develop both theory and practice, including serving as a municipal Secretary of Education.

But questions remain. What is the value of quantitative approaches outside of knowledge hierarchies? As Bricmont and Sokal suggest<sup>[14]</sup>, are quantitative and technical approaches to the world only valuable if they are getting at a single universal truth? If we reject positivism, and choose participatory paradigms and the ethics of care, must we reject technical approaches? Or even if not, how can we integrate the ethics of care into technology to achieve “doing no unintended harm”, and not further marginalizing resource-deprived communities? What sorts of technical practices might emerge not from an *elite* critical stance, but from a critical *pedagogical* stance?

It seems daunting, but qualitative research also was once positivist and hierarchical, for example, in seeing the role of a colonial anthropologist as providing neutral description about colonized or imperialized peoples to better facilitate control.

Lastly, we hope this article has served as an orientation, encouragement, and guidance for those who are undergoing the kind of vertigo that Agre described. The technical variant of critical consciousness is a profound and important experience, just like critical consciousness in general. But if it happens in isolation, it may be unnecessarily painful, and more importantly may not overcome the most pernicious part of positivism: creating and defending hierarchies of knowledge that structure the ways we approach the world, value ideas, and treat other beings. We hope that this article points to how this experience is not isolated, and gives support towards building community, overcoming knowledge hierarchies, adopting an ethics of care, and taking action towards more liberated ways of being.

## Appendix

### Glossary of Key Terms

**Realism** is the belief in a single underlying reality that exists independent of and prior to human conception of it. A specific form of this was articulated by Plato, where mathematical forms are immutable and that invariance what determines what is “real”. Confusingly but perhaps more appropriately, this is sometimes also called “idealism”, since reality is associated with *ideas* rather than perception.

**Positivism** was coined by Auguste Comte in philosophical writings around 1830–1842. It was an

application of methodology from natural sciences to study human behavior and social phenomena. Comte articulated positivism<sup>[57]</sup> in terms of a premise that universal truths exist for human behavior and social phenomenon (i.e., a realist ontology), and that empirical observations through scientific measurement can discover these universal truths (i.e., an empiricist epistemology and methodology).

Positivism now describes any research paradigm that holds that a singular truth exists and can be uncovered by empirical observation, and covers natural sciences as well as social and behavioral sciences. There are versions of positivism that try to avoid the realist commitment, and there can also be realism without empiricism (such as in pure mathematics) but the key point of either realism or positivism as compared to other sets of assumptions is *belief in an external world that takes primacy over actors’ interpretations and renegotiations of it*<sup>[67]</sup>.

**Post-positivism** is a softening of positivism, and held by people who still find positivism aesthetically compelling, but acknowledge that contingent and malleable (and non-scientifically measurable) history, society, and culture can come in the way of our ability to discover universal truths through observation, and so must be accounted for (potentially through qualitative means).

**Critical theory** is a type of philosophy often viewed as originating from a specific group of European intellectuals based in Frankfurt in the period between the World Wars. Against the prevailing view that the Enlightenment had led to constant social improvement, this Frankfurt School and their successors sought to theorize how the Enlightenment led to, or at least failed to prevent, World War I, the rise of anti-Semitism, and other forms of oppression (eventually leading to the Holocaust) in liberal capitalist societies. Of course, earlier major atrocities—such as the trans-Atlantic slave trade, or colonial genocide of indigenous populations—tellingly did not lead to similar soul-searching among European intellectuals about the consequences of the Enlightenment. Still, the Frankfurt School represented when a major European philosophical school caught up to people in the colonized world in acknowledging marginalization as a central philosophical question. For example, Rabaka<sup>[68]</sup> argues that Martinique-born psychiatrist and

philosopher Frantz Fanon (discussed more below), built on prior work from the colonized world and went far beyond the Frankfurt school in analyzing the nature of the racism and exploitation of settler colonialism.

It is from the Frankfurt School's use of "critical" that the term is applied to theories that dispute prevailing assumptions about social development needing only continue along its current course to eventually result in the end of forms of oppression, e.g., around gender, race, sexuality, disability, etc.

A good definition of what makes a "critical social science" is in Fay's *Critical Social Science: Liberation and its Limits*<sup>[69]</sup>. Fay conceives of critical social science as a type of "estrangement theory". This is a view of the world that holds that there is a manifest/ordinary sphere in which most people live, but this keeps them trapped from what is best in life, which exists in a hidden/extraordinary sphere. Specifically, critical social science is a *humanist* variant of estrangement theory, that locates the hidden/extraordinary sphere not in a religious or spiritual plane (like religious and mystical traditions do), but in the social plane. He additionally theorizes that a complete critical theory includes a *theory of false consciousness* (identifying certain understandings and explaining how they are false and/or incoherent, and how they come to be and are maintained), a *theory of crisis* (how a society is in a crisis from felt dissatisfactions that threaten social cohesion and cannot be resolved within existing social organization and self-understandings), a *theory of education* (the necessary and sufficient conditions for overcoming the false consciousness), and a *theory of transformative action* (identifying what needs to change, and a plan of action for who are "carriers" of anticipated social change and how they will go about achieving it).

Note that positivism (or realism) can have an estrangement aspect as well, where there is a hidden truth that reality is apprehensible through the language of mathematics and/or experimental methods, leading to liberation. Indeed, Plato's parable of the cave, and Platonism (as well as the neo-Platonism of mystic cults throughout the Mediterranean and West Asia centuries after Plato) sees universal abstract mathematical forms as the truth from which the masses are estranged. But the estrangement aspect of positivism need not be present, whereas it is an essential part of any critical theory.

**Relativism** is a stance that potentially spans ontology,

epistemology, and axiology. Ontological (or conceptual) relativism holds that there is no observer-independent reality, and that an observer creates their own reality. Epistemically, relativism holds that there is no neutral frame in which we can arbitrate whether claims are "true" or "false". This can be understood empirically (rather than normatively)<sup>③</sup>: for example, speaking purely empirically, there is no frame of reference to which a Biblical creationist and an evolutionary biologist would agree for arbitrating their competing claims about the origin of biological diversity. Each would insist on their own frame being the "neutral" or superior one, and any logical or empirical basis for deciding between the frames would itself rely on agreement over what counts as logical or empirical. Moral relativism holds that there is no neutral frame in which we can decide what is good or bad. Similar to epistemic relativism, moral relativism may be a descriptive rather than a normative position, built on the observation that people have genuine disagreements about morality that cannot be logically resolved by an appeal to universal underlying principles. That is, a relativist can have their own (non-relativist) normative morality that they believe is correct, alongside a relativist ontology and/or relativist epistemology that they also believe is correct, but they recognize that there is not necessarily any deeper universal principle to which to appeal and *logically* convince others. As a corollary, we can account for people with perspectives we find bizarre or moral codes that we find abhorrent who cannot be convinced through logical means, rather than needing to dismiss them as insane.

Relativism represents a break from a singular truth, and can be deeply uncomfortable and threatening for those accustomed to the pursuit of certainty and finality. Worse, when every possible position and action *can* be critiqued, relativized, destabilized, and once we know how to do this, it can be debilitating. See below for how participatory paradigms provide a way out of this.

**Constructivism** is built on relativism, and describes

<sup>③</sup> Barnes and Bloor<sup>[70]</sup> have a relatively simple response to the frequent initial objection that relativism is paradoxical or self-refuting (i.e., if all perspectives are equally valid, then by its own admission relativism concedes to non-relativism): relativism is not saying we cannot hold our own perspective, or we cannot condemn those of others or say they are wrong (whether morally, or in terms of knowledge); relativism can be just the recognition that others can and will reject our views or condemnations, and that our condemnations alone will not convince them otherwise. Of course, it is possible to interpret relativism in such a way as to defend the right of regressive perspectives to exist, but that treats relativism as a standard to which to aspire, rather than a description of how things are. And, relativism is self-referential and can create paradoxes, but we believe that accepting these paradoxes as axiomatic is enormously insightful.

the process by which multiple “truths” come to exist. It is an idea coming out of the sociology of knowledge that holds that our experiences of the world, and knowledge, are not references to or reflections of an underlying external reality, but are the product of historical, cultural, and material forces that, had they been different, would have built something different. Note that saying something (like scientific knowledge) is “constructed” does not mean that it is not real, or not solid, or not robust; a metaphor used to illustrate this perspective<sup>[71]</sup> is that of a house, which is perfectly “real” but it came to exist at a certain point in time, and was built in one specific way out of specific materials out of many alternatives. We can come to understand this building process without claiming the building is anything other than solid and durable. However, other versions of constructivism stress the fluidity of things like scientific knowledge, rejecting the idea of knowledge as hierarchical structures anchored to, if not a solid underlying reality, then to society and history; these versions of constructivism instead see knowledge as ungrounded webs of mutual reference. Then, the task of inquiry is to understand the construction and maintenance of these webs of mutual reference (with the inquiry being itself a part of the webs it considers).

There is a tension between critical theory and constructivism<sup>[72]</sup> in how *critical* perspectives can end up holding that there is an external world, just that it is something different than what most people think it is. So, for example, Fay offers the Marxist-humanist model of political revolution as an example of a critical theory, where there is a “true nature” that bourgeoisie oppressors derive power from the self-understandings of the oppressed working classes.

However, they frequently appear together. Hacking<sup>[73]</sup> points out how looking at how things are put together also gives people grounds to see how they come apart, and *deconstruct* them. A crucial part of a critical toolbox is in showing the historical construction of ideas, forms of knowledge, institutions, and cultural forms, thereby demonstrating that they are not inevitable, and letting us imagine and advocate for alternatives.

For example, in critical race studies and critical gender studies, there is a “false consciousness” of thinking that the categories by which people are marginalized are based on biological traits or even cultural ones. But there is no such thing as biological race or gender, let alone inferiority by them (and the “value” of cultures, like European culture versus indigenous cultures, come from

how they are valued, and not something intrinsic). Instead, such categories and their value are socially constructed by and maintained through power relationships. Going further, there is a second layer to the false consciousness, of holding marginalized people individually responsible for their suffering and deprivation. Once categories are so constructed, those that fall within the marginalized categories like women of color and others with individual or intersecting marginalized identities are treated as inferior, in ways often enacted on an interpersonal level but structurally and culturally encouraged and permitted. The result is marginalized people face greater mental and physical suffering, and material deprivation, entirely apart from their individual “effort”, yet over which they are held responsible. Even holding those who *enact* the double standards individually responsible (i.e., seeing racism or sexism as an interpersonal problem), rather than seeing the larger structure, is a false consciousness. Only by recognizing the true nature of modern civilization as fundamentally structured on white supremacy, patriarchy, colonialism, and other forms of domination can we effect change and improve human life.

Indeed, Agre’s<sup>[74]</sup> idea of “critical” is actually more about constructivism (and unfortunately he sets it up using ableist language). In one entry from his Red Rock Eater Newsletter (a listserv over which Agre sent out writings that has been cited as a precedent for blogs), he wrote:

“I finally comprehended the difference between critical thinking and its opposite. Technical people are not dumb [sic], quite the contrary, but technical curricula rarely include critical thinking in the sense I have in mind. **Critical thinking means that you can, so to speak, see your glasses. You can look at the world, or you can back up and look at the framework of concepts and assumptions and practices through which you look at the world.**”

Agre continues: “Not that critical thinking makes you omniscient: you’re still wearing glasses even when you’re looking at your glasses.” That is, there is no perspective without any glasses, no “view from nowhere”.<sup>④</sup> The experience of “seeing one’s glasses” is different than just replacing one’s glasses; it opens the

<sup>④</sup> Ludwig Wittgenstein, another figure who underwent a transformation in his basic beliefs and how he saw the world, also used this metaphor much earlier<sup>[75]</sup>: “The ideal, as we think of it, is unshakable. You can never get outside it; you must always turn back. There is no outside; outside you cannot breathe.—Where does this idea come from? It is like a pair of glasses on our nose through which we see whatever we look at. It never occurs to us to take them off.”



path to understanding endless contingency in ideas, structures, institutions, and frameworks.

**Critical consciousness** is a theory that came out of political mobilization and community development, also known as popular education, in the Global South<sup>[76–78]</sup>, and specifically from the work of Brazilian educator, philosopher, and politician Paulo Freire (1921–1997).

Freire worked in the 1960s with populations like marginalized sugarcane harvesters with no access to formal education. He started education programs for political mobilization in conjunction with them, and used that mobilization to get the Brazilian government to financially support the programs they had created. He challenged a “banking” conception of education that assumed he was more of a knowledge holder and knowledge creator than the farmers he worked with, and that placed more value on him as a teacher, because he had access to formal education. He inverted the hierarchy to say that the marginalized are valuable because of their response to marginalization, their resilience, and how the experience of marginalization showed larger societal structures in a way that Freire, with his privilege, had not seen. He theorized how to unseat the teacher or researcher as the expert, and sought to develop a model where we all bring something to the table and learn from each other, and understanding emerges from our interactions.

Another key input, on whose work Freire drew, was Frantz Fanon (1925–1961). Fanon was hired as a psychiatrist by the French colonial government in Algeria to treat mental illness in colonial subjects. There, Fanon realized that his patients were not having mental health crises, but reacting to oppression, and the French government did not understand that their reaction was the most logical response to being otherized, dehumanized, and oppressed. Building on his previous work theorizing his own experience being treated as a French colonial subject<sup>[79]</sup>, in interacting with his patients in Algeria he learned about his own position in a larger oppressive system and how it was causing harm to others<sup>[80, 81]</sup>. From this, he wrote about working with marginalized populations, unlearning harmful frames, and mobilizing for revolution and equity, himself joining the Algerian National Liberation Front to support Algeria’s War of independence from France.

Freire gave the name *conscientização* to the transformative process of interacting with other individuals and other communities<sup>[82, 83]</sup>, translated as

critical consciousness, or more literally as “conscientization”, and sometimes as consciousness-raising<sup>[84]</sup>.<sup>⑤</sup> From there, others have continued to systematically develop tools, strategies, and methods for critical consciousness, including dialogue and critical reflection, reflective questioning, psychosocial support, co-learning, group processes, civic engagement and sociopolitical action, and identity development<sup>[34]</sup>. Critical consciousness has inspired a field within education known as critical pedagogy<sup>[86]</sup> which has been carried forward particularly in adult education<sup>[4]</sup> and has had a large impact on the development of Participatory Action Research<sup>[87]</sup> and Community Based Participatory Research<sup>[88]</sup>.

**Positionality** is awareness and discussion of ones’ social and institutional position with regards to research, particularly of power imbalances, and limitations the researcher may have because of differences in lived experience.

**Reflexivity** is the process of “turning back on” and reflecting on experience and our positionality. For example, in anthropology, this is researchers being explicit about their emotions and how they related to research subjects<sup>[89]</sup>. Positivism, in particular, does not and cannot engage in reflexivity<sup>[90]</sup>, since it holds that knowledge is independent of the knowledge-holder.

**Participatory** paradigms address an important moral aspect lacking in both critical theory and constructivism. Certain streams of critical theory frequently have a condescending aspect to them: that people are unaware of their own oppression, and it is the role of the critical theorist to educate them. On the other hand, constructivism does not account for experiential knowing<sup>[90]</sup>. Building explicitly from the ideas of Freire, participatory paradigms value and highlight experience, following a methodology that challenges hierarchies between teacher and student, or researcher and subject, and seeks to construct knowledge collectively. Its methodology and axiology prioritize understanding and improving the world by changing it through collective, reflexive inquiry<sup>[91]</sup>.

This paradigm has a relativistic component in seeing knowledge as malleable and multiple rather than absolute and singular; by locating value in others and their experiences, rather than seeing the status of

<sup>⑤</sup> Consciousness-raising also appears, without reference to Freire, in US feminist movements in the 1960s<sup>[85]</sup>.

knowledge as the most important thing in life, the instability of knowledge does not become a reason to be nihilistic.

## Acknowledgment

Thanks to two anonymous reviewers for great feedback and pointing us to some relevant literature, and Ben Green both for fantastic continuous comments and for corralling and managing this special issue.

## References

- [1] P. E. Agre, Toward a critical technical practice: Lessons learned in trying to reform AI, in *Bridging the Great Divide: Social Science, Technical Systems, and Cooperative Work*, G. Bowker, L. Star, B. Turner, and L. Gasser, eds. Mahwah, NJ, USA: Lawrence Erlbaum Associates, Inc., 1997, pp. 131–157.
- [2] B. Green, The contestation of tech ethics: A sociotechnical approach to technology ethics in practice, *Journal of Social Computing*, doi: 10.23919/JSC.2021.0018.
- [3] L. Hu, Tech ethics: Speaking ethics to power, or power speaking ethics? *Journal of Social Computing*, doi: 10.23919/JSC.2021.0033.
- [4] S. B. Merriam and L. M. Baumgartner, Transformational learning, in *Learning in Adulthood: A Comprehensive Guide, 4th edition*, S. B. Merriam and L. M. Baumgartner, eds. Hoboken, NJ, USA: Wiley, 2020, pp. 166–195.
- [5] D. A. Henhawk, My critical awakening: A process of struggles and decolonizing hope, *International Review of Qualitative Research*, doi: 10.1525/irqr.2013.6.4.510.
- [6] J. Prada, The critical awakening of a pre-service teacher in a Spanish graduate program: A phenomenology of translanguaging as pedagogy and as content, *International Journal of Bilingual Education and Bilingualism*, doi: 10.1080/13670050.2021.1881945.
- [7] E. G. Guba and Y. S. Lincoln, Paradigmatic controversies, contradictions, and emerging confluences, in *The SAGE Handbook of Qualitative Research*, N. K. Denzin and Y. S. Lincoln, eds. London, UK: SAGE, 2005, pp. 191–215.
- [8] B. Green, Data science as political action: Grounding data science in a politics of justice, *Journal of Social Computing*, doi: 10.23919/JSC.2021.0029.
- [9] M. L. Jones, How we became instrumentalists (again): Data positivism since World War II, *Historical Studies in the Natural Sciences*, doi: 10.1525/hsns.2018.48.5.673.
- [10] E. Morozov, *To Save Everything, Click Here: The Folly of Technological Solutionism*. New York, NY, USA: Public Affairs, 2013.
- [11] K. Toyama, *Geek Heresy: Rescuing Social Change from the Cult of Technology*. New York, NY, USA: Public Affairs, 2015.
- [12] M. Broussard, *Artificial Unintelligence: How Computers Misunderstand the World*. Cambridge, MA, USA: MIT Press, 2018.
- [13] B. Green, *The Smart Enough City: Putting Technology in Its Place to Reclaim Our Urban Future*. Cambridge, MA, USA: MIT Press, 2019.
- [14] J. Bricmont and A. Sokal, Science and sociology of science: Beyond war and peace, in *The One Culture? A Conversation about Science*, J. A. Labinger and H. Collins, eds. Chicago, IL, USA: The University of Chicago Press, 2001, pp. 27–47.
- [15] T. M. Porter, Thin description: Surface and depth in science and science studies, *Osiris*, doi: 10.1086/667828.
- [16] G. Edmond and D. Mercer, Anti-social epistemologies, *Social Studies of Science*, doi: 10.1177/0306312706067900.
- [17] M. Lynch, From ruse to farce, *Social Studies of Science*, doi: 10.1177/0306312706067897.
- [18] M. Nanda, *Prophets Facing Backward: Postmodern Critiques of Science and Hindu Nationalism in India*. New Brunswick, NJ, USA: Rutgers University Press, 2003.
- [19] H. M. Collins, Son of seven sexes: The social destruction of a physical phenomenon, *Social Studies of Science*, doi: 10.1177/030631278101100103.
- [20] T. H. Kuhn, *The Structure of Scientific Revolutions, 3rd ed.* Chicago, IL, USA: The University of Chicago Press, 1996.
- [21] P. Rogaway, The moral character of cryptographic work, <https://web.cs.ucdavis.edu/~rogaway/papers/moral-fn.pdf>, 2015.
- [22] D. J. Watts, The “new” science of networks, *Annual Review of Sociology*, doi: 10.1146/annurev.soc.30.020404.104342.
- [23] C. A. Hidalgo, Disconnected, fragmented, or united? A trans-disciplinary review of network science, *Applied Network Science*, doi: 10.1007/s41109-016-0010-3.
- [24] H. Wallach, Computational social science ≠ computer science + social data, *Communications of the ACM*, doi: 10.1145/3132698.
- [25] H. Wallach, Navigating the broader impacts of machine learning research, Medium, <https://hannawallach.medium.com/navigating-the-broader-impacts-of-machine-learning-research-f2d72a37a5b>, 2021.
- [26] S. E. Toulmin, Does the distinction between normal and revolutionary science hold water? in *Criticism and the Growth of Knowledge*, I. Lakatos and A. Musgrave, eds. Cambridge, UK: Cambridge University Press, 1970, pp. 39–48.
- [27] J. Mezirow, Perspective transformation, *Adult Education Quarterly*, doi: 10.1177/074171367802800202.
- [28] J. Mezirow, Learning to think like an adult: Core concepts of transformation theory, in *The Handbook of Transformative Learning: Theory, Research and Practice*, E. W. Taylor and P. Cranton, eds. San Francisco, CA, USA: Jossey-Bass, 2012, pp. 73–95.
- [29] E. W. Taylor, Building upon the theoretical debate: A critical review of the empirical studies of Mezirow’s transformative learning theory, *Adult Education Quarterly*, doi: 10.1177/074171369704800104.
- [30] V. Sheared, Giving voice: An inclusive model of instruction—a womanist perspective, *New Directions for Adult and Continuing Education*, doi: 10.1002/ace.36719946105.
- [31] E. W. Taylor, Making meaning of the varied and contested perspectives of transformative learning, in *Proceedings of the Sixth International Conference on Transformative Learning*, D. Vlosak, G. Kielbaso, and J. Radford, eds. East Lansing, MI, USA: Michigan State University, 2005,

- pp. 459–464.
- [32] E. W. Taylor and M. J. Snyder, A critical review of research on transformative learning theory, 2006–2010, in *The Handbook of Transformative Learning: Theory, Research and Practice*, E. W. Taylor and P. Cranton, eds. San Francisco, CA, USA: Jossey-Bass, 2012, pp. 37–55.
- [33] J. Mezirow, A critical theory of adult learning and education, *Adult Education Quarterly*, doi: 10.1177/074171368103200101.
- [34] A. Jemal, Critical consciousness: A critique and critical analysis of the literature, *The Urban Review*, doi: 10.1007/s11256-017-0411-3.
- [35] P. Freire, *Pedagogy of the Oppressed*. New York, NY, USA: Continuum, 1970.
- [36] S. B. Merriam and G. Ntseane, Transformational learning in Botswana: How culture shapes the process, *Adult Education Quarterly*, doi: 10.1177/0741713608314087.
- [37] J. Johnson-Bailey, *Sistahs in College: Making a Way Out of No Way*. Malabar, FL, USA: Krieger Publishing, 2001.
- [38] B. Christian, The race for theory, *Cultural Critique*, doi: 10.2307/1354255.
- [39] H. Trbušić, Engineering in the community: Critical consciousness and engineering education, *Interdisciplinary Description of Complex Systems*, doi: 10.7906/indecs.12.2.1.
- [40] A. Boal, *Theatre of the Oppressed*. New York, NY, USA: Theatre Communications Group, 1993.
- [41] K. Jokikokko, The role of significant others in the intercultural learning of teachers, *Journal of Research in International Education*, doi: 10.1177/1475240909105202.
- [42] B. de Sousa Santos, Public sphere and epistemologies of the South, *Africa Development*, vol. 37, no. 1, pp. 43–67, 2012.
- [43] B. de Sousa Santos, *Epistemologies of the South: Justice Against Epistemicide*. New York, NY, USA: Routledge, 2014.
- [44] S. Wynter, No humans involved: An open letter to my colleagues, in *Forum N. H. I. : Knowledge for the 21st Century*. Stanford, CA, USA: Giant Horse, Inc., 1994, pp. 42–71.
- [45] S. Wynter, Unsettling the coloniality of being/power/truth/freedom: Towards the human, after man, its overrepresentation—an argument, *CR: The New Centennial Review*, doi: 10.1353/ncr.2004.0015.
- [46] B. Rusert, *Fugitive Science: Empiricism and Freedom in Early African American Culture*. New York, NY, USA: NYU Press, 2017.
- [47] M. Graham, The ethics of care, Black women and the social professions: Implications of a new analysis, *Ethics and Social Welfare*, doi: 10.1080/17496530701450372.
- [48] L. Bass, Fostering an ethic of care in leadership: A conversation with five African American women, *Advances in Developing Human Resources*, doi: 10.1177/1523422309352075.
- [49] O. Hankivsky, Rethinking care ethics: On the promise and potential of an intersectional analysis, *American Political Science Review*, doi: 10.1017/S0003055414000094.
- [50] P. Raghuram, Race and feminist care ethics: Intersectionality as method, *Gender, Place & Culture*, doi: 10.1080/0966369X.2019.1567471.
- [51] M. Lane, “For real love”: How Black girls benefit from a politicized ethic of care, *International Journal of Education Reform*, doi: 10.1177/105678791802700303.
- [52] L. Bass, When care trumps justice: The operationalization of Black feminist caring in educational leadership, *International Journal of Qualitative Studies in Education*, doi: 10.1080/09518398.2011.647721.
- [53] b. hooks, Love as the practice of freedom, in *Outlaw Culture: Resisting Representations*. b. hooks, ed. New York, NY, USA: Routledge, 1994, pp. 243–250.
- [54] b. hooks, *All About Love: New Visions*. New York, NY, USA: Harper, 2000.
- [55] P. H. Collins, *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment, 2nd ed.* New York, NY, USA: Routledge, 2000.
- [56] A. D. Selbst, d. m. boyd, S. A. Friedler, S. Venkatasubramanian, and J. Vertesi, Fairness and abstraction in sociotechnical systems, in *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT\* '19)*, Atlanta, GA, USA, 2019, pp. 59–68.
- [57] E. J. Dixon-Román, Diffractive possibilities: Cultural studies and quantification, *Transforming Anthropology*, doi: 10.1111/traa.12074.
- [58] J. Masis, Making AI philosophical again: On Philip E. Agre’s legacy, *Continent*, vol. 4, no. 1, pp. 58–70, 2014.
- [59] P. Sengers and G. Hertz, Critical technical practice and critical making: Phoebe Sengers in conversation with Garnet Hertz, in *Conversations in Critical Making*, G. Hertz, ed. CTheory Books, 2015, pp. 9–20.
- [60] K. Boehner, S. David, J. Kaye, and P. Sengers, Critical technical practice as a methodology for values in design, presented at CHI 2005 Workshop on Quality, Values, and Choices, Portland, OR, USA, 2005.
- [61] P. Dourish, J. Finlay, P. Sengers, and P. Wright, Reflective HCI: Towards a critical technical practice, presented at CHI '04 Extended Abstracts on Human Factors in Computing Systems, Vienna, Austria, 2004.
- [62] R. Albergotti, He predicted the dark side of the Internet 30 years ago. Why did no one listen? Philip Agre, a computer scientist turned humanities professor, was prescient about many of the ways technology would impact the world, *Washington Post*, Aug 2021.
- [63] L. Stark, Apologos: A lightweight design method for sociotechnical inquiry, *Journal of Social Computing*, doi: 10.23919/JSC.2021.0028.
- [64] J. U. Korn, Connecting race to ethics related to technology: A call for critical tech ethics, *Journal of Social Computing*, doi: 10.23919/JSC.2021.0026.
- [65] R. Ochigame, Informatics of the oppressed, *Logic*, vol. 11, pp. 53–74, 2020.
- [66] K. Mayer and M. M. Malik, Critical data scientists at work: Summary report of the ICWSM-2019 Workshop on Critical Data Science, presented at the Thirteenth International AAAI Conference on Web and Social Media (ICWSM-2019), Munich, Germany, 2019.
- [67] G. Payne and J. Payne, Positivism and realism, in *Key Concepts in Social Research*. London, UK: SAGE Publications, 2004, pp. 171–174.
- [68] R. Rabaka, *Forms of Fanonism: Frantz Fanon’s Critical Theory and the Dialectics of Decolonization*. Lanham, MD, USA: Lexington Books, 2011.

- [69] B. Fay, *Critical Social Science: Liberation and Its Limits*. Ithaca, NY, USA: Cornell University Press, 1987.
- [70] B. Barnes and D. Bloor, Relativism, rationalism and the sociology of knowledge, in *Rationality and Relativism*, M. Hollis and S. Lukes, eds. Oxford, UK: Basil Blackwell, 1982, pp. 21–47.
- [71] S. Shapin and S. Schaffer, *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*. Princeton, NJ, USA: Princeton University Press, 1985.
- [72] L. S. Clark, Critical theory and constructivism, Media, Culture and Meaning site, Center for Mass Media, School of Journalism and Mass Communication, University of Colorado, <https://web.archive.org/web/20051201184243/http://www.colorado.edu/journalism/mcm/qmr-crit-theory.htm>, 1999.
- [73] I. Hacking, *The Social Construction of What?* Cambridge, MA, USA: Harvard University Press, 1999.
- [74] P. E. Agre, Notes and recommendations for 12 July 2000, Red Rock Eater Newsletter, <https://pages.gseis.ucla.edu/faculty/agre/notes/00-7-12.html>, 2000.
- [75] L. Wittgenstein, *Philosophical Investigations*. Oxford, UK: Wiley-Blackwell, 1953.
- [76] O. H. Jara, Popular education and social change in Latin America, *Community Development Journal*, doi: 10.1093/cdj/bsq022.
- [77] L. Kane, *Popular Education and Social Change in Latin America*. London, UK: Latin American Bureau, 2001.
- [78] L. A. Hadfield, *Liberation and Development: Black Consciousness Community Programs in South Africa*. East Lansing, MI, USA: Michigan State University Press, 2016.
- [79] F. Fanon, *Black Skin, White Masks*. New York, NY, USA: Grove Press, 1952.
- [80] N. C. Gibson, ed., *Decolonizing Madness: The Psychiatric Writings of Frantz Fanon*. New York, NY, USA: Palgrave Macmillan, 2014.
- [81] N. C. Gibson, Decolonizing madness: The psychiatric writings of Frantz Fanon, in *Fanon, Phenomenology, and Psychology*, L. Laubscher, D. Hook, and M. U. Desai, eds. New York, NY, USA: Routledge, 2021.
- [82] P. Freire, *Education for Critical Consciousness*. New York, NY, USA: Seabury Press, 1974.
- [83] P. Freire, *Education, the Practice of Freedom*. London, UK: Writers and Readers Publishing Cooperative, 1976.
- [84] T. J. L. Belle, From consciousness raising to popular education in Latin America and the Caribbean, *Comparative Education Review*, doi: 10.1086/446677.
- [85] K. Sarachild, Consciousness-raising: A radical weapon, in *Feminist evolution*, K. Sarachild, ed. New York, NY, USA: Random House, 1978, pp. 144–150.
- [86] A. Darder, Conscientização: Awakening critical consciousness, in *Freire and Education*, A. Darder, ed. New York, NY, USA: Routledge, 2015, pp. 80–132.
- [87] J. Cammarota and M. Fine, *Revolutionizing Education: Youth Participatory Action Research*. New York, NY, USA: Routledge, 2007.
- [88] N. Wallerstein and B. Duran, The theoretical, historical, and practice roots of CBPR, in *Community-Based Participatory Research for Health: Advancing Social and Health Equity*, N. Wallerstein, B. Duran, J. Oetzel, and M. Minkler, eds. San Francisco, CA, USA: Jossey Bass, 2008, pp. 25–46.
- [89] P. C. Salzman, On reflexivity, *American Anthropologist*, doi: 10.1525/aa.2002.104.3.805.
- [90] J. Heron and P. Reason, A participatory inquiry paradigm, *Qualitative Inquiry*, doi: 10.1177/107780049700300302.
- [91] F. Baum, C. MacDougall, and D. Smith, Participatory action research, *Journal of Epidemiology and Community Health*, doi: 10.1136/jech.2004.028662.



**Maya Malik** received the BA degree in psychology from Warren Wilson College in 2014 and the MS degree in Social Work from Columbia University's School of Social Work in 2017, focusing on international social welfare and rights for immigrants and refugees through program design, research, and evaluation. They are

currently a doctoral student at the School of Social Work, McGill University, where they are researching how to utilize arts-based Youth-Led Participatory Action Research (YPAR) methods to work with Queer Black American youth who have been justice-involved to improve educational intervention programs. They are also a researcher in the Participatory Axis of the McGill Global Child Research Group, and a research contributor at the Youth and Media project at the Berkman Klein Center for Internet & Society at Harvard University. They are also a member of an interdisciplinary public health team to address the possible impact of reparations on quality of life and mortality for Black Americans.



**Momin M. Malik** received the AB degree in History and Science from Harvard University in 2009, the MSc degree in Social Science of the Internet from the University of Oxford in 2012, and the MS degree in machine learning and the PhD degree in societal computing from Carnegie Mellon University both in 2018.

Previously, he was the Data Science Postdoctoral Fellow at the Berkman Klein Center for Internet & Society at Harvard University, and the Director of Data Science at Avant-garde Health. He is currently a Senior Data Scientist for AI Ethics at the Center for Digital Health, Mayo Clinic, a fellow at the Institute in Critical Quantitative, Computational, & Mixed Methodologies, and a lecturer for the School of Social Policy & Practice at the University of Pennsylvania. He is a Senior Program Committee Member for the International Conference for Web and Social Media.