THE UNIVERSITY OF CHICAGO

SOCIAL SERVICE LABORATORIES: THE POLITICS OF REFORM IN AN OVERDOSE PREVENTION CLINICAL TRIAL

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE CROWN FAMILY

SCHOOL OF SOCIAL WORK, POLICY, AND PRACTICE

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

BY

EMILY JANE CLAYPOOL

CHICAGO, ILLINOIS

AUGUST 2024

Table of Contents

List of Figures

List of Tables

Acknowledgements

Many people made this dissertation possible. First, I have my committee members to thank. My chair, Zhiying Ma, was an immense source of support both intellectually and emotionally as I conducted and wrote this ethnography. She pressured me to scale my arguments and to think about the broader, sociological implications of what I was examining. At the same time, she taught me how to do ethnography in a way that is sensitive, and responsible while also holding onto my own political accountabilities. Importantly, she was kind and reliable. I thank Curtis McMillen, who always patiently listened to my ideas, took them seriously and helped me to bring them down to earth, while also, somehow, simultaneously inspiring creativity. He always provided me encouragement during critical times in my life and in the program and is one of the primary reasons I am a doctoral student at the University of Chicago. I felt honored to have Victoria Stanhope on my committee. She was always willing to deliver thoughtful and critical feedback and offer her insights in ways that made this dissertation sharper.

Many other faculty helped me to develop as a scholar and navigate academic life. At the University of Chicago, Jennifer Mosley provided me with the training and intellectual space to think organizationally and politically about intervention and implementation science. Jennifer was always willing to be a reliable source of support throughout my time at UC. Leyla Ismayilova served as a research advisor during my early years in the program. I learned so much about conducting intervention research under her mentorship. I benefited tremendously from being in her orbit as she launched impressively designed large-scale, international clinical trials. I would also like to thank Roman Kislov, at Manchester Metropolitan University, who was willing to meet with me, engage my work and provide written feedback that made this dissertation more relevant and meaningful to implementation science scholarship. I also would

like to thank Daniel Listoe. He was able to turn lots of convoluted pages into meaningful arguments. I'm certain this dissertation would not be complete without his editorial work.

Then, I must thank my friends from the University of Chicago. Each, in their own way, contributed to how I think in significant ways and made my time at UC more special: Danielle Adams, Zoe Berman, Cameron Day, Jesse Self, Katie Gibson, Kit Ginzky, Julian Thompson, and Tadeo Weiner-Davis. Hannah Norwood was so generous in reading chapters and talking through the trials and tribulations of dissertation writing with me, particularly in the final year. I thank her for her kindness and generosity.

I want to thank Katie Berringer at the University of Michigan for being such a good interlocutor to my project, a reliable source of intellectual support, and a friend. I am also lucky to call Erin Daly a friend and I thank her for the time she put into my project, and the support she provided as I finished it. Finally, I want to express my appreciation to Rishi Arora for his support on this dissertation, from its inception to its completion. Rishi was consistently there for me to talk through fieldwork and other ethnographic challenges, assisted in framing the project, and meticulously copy edited the manuscript during the final stages. I am deeply grateful for his guidance and contributions, without which this work would not have been possible.

También quiero agradecer a mis amigos y fuentes de apoyo en España, donde escribí la mayoría de mi tesis. Quiero agradecer a Manuel Jiménez, Lola Vélez y Magdalena Mannigel por hacerme sentir como en casa en un lugar lejano.

It is hard to encapsulate the support I received from my parents without which I would not be where I am. They have always believed in me and encouraged me to pursue my interests while also keeping me grounded in what really matters, friends and family and trying to make some difference in the world, however small. I love them and thank them.

Abstract

Researchers have made increasing efforts to consider the contexts where evidence is implemented, and more meaningfully include forms of public participation that make science more accountable and responsive to the people and places meant to take it up. The rise of implementation science and hybrid clinical trials are both examples of this contextualization. As the production of evidence moves toward its contextualization and begins to more meaningfully consider the role of the social environment in knowledge making and uptake, novel partnerships and political questions emerge. This dissertation is an ethnographic study of a complex harm reduction and peer recovery-based case management intervention (HaRP). Instead of studying whether and how HaRP works, however, this dissertation examines how researchers and various implementation actors negotiate the tension between conducting experimental science to produce a standardized, scalable intervention, while also being attuned to the "real world" of social service practice. Drawing on over two years of ethnographic field work of a federally funded, multi-sited clinical trial, I ask: how, and with what social effects, do researchers negotiate the principles of scientific inquiry with the exigencies of the "real world," and the diverse sets of institutional commitments that make it up?

Throughout this dissertation, I highlight how local context and scale, objectivity and activism, and standardization and tacit, embodied forms of care are negotiated with experimental logics like standardization, fidelity and objectivity. I argue that the imperatives of the clinical trial, coupled with the institutional obligations to which it is tethered, operate against the ethical and political commitments of practitioners and researchers themselves regarding the most effective strategies for preventing overdose. I show how care is refashioned (Chapter Three), and activism is deferred (Chapter One). However, no one actor had an unwavering commitment to

the scientific integrity of the clinical trial and this dissertation set out to reveal how researchers themselves grapple with these tensions. Indeed, Chapter two demonstrates how the imminent needs of study participants led trial actors to compromise model stability and feasibility (Chapter Two). Indeed, experimental concepts like feasibility and fidelity were contested throughout HaRP's implementation, by PIs and practitioners alike, and were actively redefined to meet the diverse, and often competing nature of hybrid clinical trials. The findings of my dissertation will help researchers, and implementation actors negotiate locally embedded social and professional values and growing activist commitments with the obligations of scientific inquiry during socio-behavioral clinical trials.

Introduction

Since the Center for Disease and Control (CDC) recognized opioid overdose as an "epidemic" in 2010, over one million people have died a preventable death (Center for Disease Control and Prevention [CDC], n.d.; Mann, 2021). During what has felt like an intractable crisis, made ever worse by the COVID19 pandemic—federal administrations have mobilized resources, expertise, and concerted efforts to try to stem the tide of overdose deaths. Overdose is a prime example of a "wicked" problem (Rittel & Webber, 1973). The causes are complex and hard to define, data is difficult to procure and interpret, and fully grasping its nature requires the assembly of many forms of expertise and decision makers, many of whom possess conflicting values, motivations, and ideas for how to address it. This diversity is due to the fact that overdose, like many complex issues, stems from a dynamic and interactive set of factors, including criminal-legal policies, local drug use practices, volatile drug supply markets, and healthcare systems, among others (Campbell, 2020).

Scholars of policy and politics have protested the idea that social problems should necessarily or at least uniformly be addressed in the "scientific idiom" (Jasanoff, 2011). Compared to other countries, American policy processes have tended to lean on the findings of science and the experts who produce it to assist in the governance and rationalization of policy-making decisions (Jasanoff, 2006; Porter, 1995). Depending on the extent to which technical decisions justify governmental decisions, some might call this the "scientization of politics." In the context of scientization, policy deliberations are steered by technical criteria and insulated from overarching social norms (Hess, 2016). Moral and political questions are reframed in scientific terms, which both limit who is able to participate in decision-making as well as what

types of questions and methods are viable. Other types of methods and questions, including what counts as effective and for who, tend to be metaphorically left "off the table."

In contemporary sociological scholarship, randomized controlled trials (RCTs) are used as canonical examples of the present-day manifestation of the scientization of social policy.[1] The basic idea is that RCTs, by randomly assigning study participants to either treatment or control group, ensures that other relevant causal factors are distributed equally, thereby isolating a causal pathway between the treatment and the outcome. This provides an ostensibly objective and reliable truth about a cause-effect relationship (Cartwright, 2013). Through their "mechanical objectivity," the removal of personal bias by appealing to formal rules, they tend to be perceived as a relatively frictionless method to arbitrate social policy (Eyal, 2019; Porter, 1996).[2] The RCT is often therefore understood as a tool that can help to overcome accusations of ideological or political bias, human error, or simply subjectivity. To be sure, in polarized societies and ideologically heated policy domains, the cultural authority of science is relied upon for its capacity to appear politically neutral (Frickel & Moore, 2005; Hess, 2016). Meanwhile, in moments of relative austerity, science can also be used to determine what is "cost-effective," which has tended to become blurred with questions of efficacy in contemporary policy (Carr, 2022). Indeed, both efficacy and cost-effectiveness have tended to be hailed as "objective" indicators of a given practice or policy's value. Yet, we have also long known from science and

[1] As defined formally in the literature: "Randomized controlled trials (RCT) are prospective studies that measure the effectiveness of a new intervention or treatment. Although no study is likely on its own to prove causality, randomization reduces bias and provides a rigorous tool to examine cause-effect relationships between an intervention and outcome. This is because the act of randomization balances participant characteristics (both observed and unobserved) between the groups allowing attribution of any differences in outcome to the study intervention" (Hariton & Locascio, 2018, p 1716).

[2] Mechanical objectivity, according to Porter (1996) "implies personal restraint. It means following the rules. Rules are a check of subjectivity: they should make it impossible for personal biases or preferences to affect the outcome of an investigation" (p. 4).

technology studies, as well as feminist epistemologies, that while framed as value-neutral and objective, all knowledge is contextually situated (Haraway, 1988). Any clinical trial is suffused with the politics and interests of the institutions that field and fund it, as well as the sociological and cultural conventions of the period (Gauchat, 2012; Shapin, 1994). Yet, to maintain both credibility and social status, researchers are pressured to produce evidence that appears objective, context-free and standpoint independent; ostensibly, this creates less political baggage for policymakers who are meant to take up the evidence once it's produced (Hess, 2016).

Framing social problems in the technical language of science, or more specifically deploying the scientific method to resolve policy issues, is often thought to limit who can participate, and de-legitimize or circumscribe other ways of attending to social issues, including those grounded in human rights or social justice (Brown, 2004; Guidry & Sawyer, 2003). Importantly, the primary methods used to produce and evaluate complex social problems such as overdose, like the RCT, struggle to fully comprehend their complexity (Sanabria, 2016; Campbell, 2020). In fact, the very concept of a *wicked problem* was coined to argue that researchers and policymakers must separate the processes of resolving scientific and technical disputes with that of social problems (Rittel & Webber, 1973). This is because the construction of social problems, Rittel and Webber (1973) argued, are imbued with values, morals and diverse perspectives that are integral to successful policy yet elided in more technical, scientific analyses. As researchers and policymakers, we must learn to understand and attend not only to such diverse perspectives, but also the structural conditions as well as the processes and institutions which develop, diffuse, and prioritize forms of evidence.

Critiques abound regarding the relative merits of the RCT and its (in)capacity to produce scalable or universal evidence, a point that will be elaborated further on in this introduction.

Simultaneously, as these critiques have emerged, efforts have been made to more intimately consider the contexts where evidence is implemented, and more meaningfully include forms of public participation that might make science more accountable and responsive to the people and places meant to take it up. While an ideal type, this transformation in the development of the sciences has been referred to as *Mode 2*, which describes a process by which science has increasingly become contextualized, interdisciplinary, dialogic, and more attuned to the social impact of its practices and outputs (Jasanoff, 2003; Nowotny et al., 2001). Unlike *Mode 1* knowledge, which necessitates "translation" for application, *Mode 2* knowledge is inherently integrated into the application context from its inception (Greenhalgh & Wieringa, 2011). This integration also means that it is even *more* difficult to conceptualize science as a closed system, distinguishable from society and the funders, policymakers and service users who help to evaluate and produce it.

The rise of implementation science, as well as hybrid effectiveness-implementation clinical trials, both of which will be addressed more in depth in the following sections, are examples of recent movements toward the localization and contextualization of science. In its own terms, implementation science is "the scientific study of methods to promote the systematic uptake of research findings and other EBPs into routine practice, and, hence, to improve the quality and effectiveness of health services" (Eccles & Mitman, 2006, p. 1). Hybrid clinical trials, which combine questions of efficacy with implementation, are an attempt to condense the translational timeline or "pipeline" between scientific development and scale by examining whether a given intervention works under "real world" conditions.[3] This makes it incumbent

[3] Curran et al. (2013) defined hybrid effectiveness-implementation trials as "A study design that takes a dual focus in assessing clinical effectiveness and implementation."

upon researchers to develop and test a scalable intervention, while, at the same time, trying to make interventions attuned to the needs and desires of service users, providers and the social service organizations that host them. In hybrid clinical trials, researchers are tasked with balancing the maintenance of internal validity, aimed at establishing causal inference regarding the intervention under evaluation, with promoting external validity and generalizability. Blurring the lines between (1) science and routine practice and (2) clinical versus experimental goals, hybrid trials are meant to produce generalizable evidence, while also remaining real world relevant. All that being said, while conducting a clinical trial, researchers and practitioners must balance the imperatives of scalability and experimental science with the norms and values of social service practices. It has been under-examined, however, how the experimental imperatives of the RCT, which are meant to create scalable, universal evidence, might work alongside efforts at its contextualization through implementation science and effectiveness-implementation hybrid trials. As the production of evidence moves toward its contextualization and begins to consider the role of the social environment more heartily in knowledge making and uptake, novel partnerships, and political questions, and thus tensions emerge.

This dissertation intervenes here to pursue the question: How, and with what effects, do researchers negotiate the principles of scientific inquiry (e.g., the epistemologies, theoretical models, and experimental methods of the RCT) with the exigencies of the "real world," and the diverse sets of institutional commitments that make it up? As the field of implementation science has emerged to manage and attend to context, how do implementation actors negotiate the tension of local context and scale, science and activism, and care and control? In pursuing these questions, this dissertation centers on examining the social processes of fielding a complex, multi-sited clinical trial, and more broadly the social processes of scientific reform with a focus

on the ethical, political and epistemological challenges involved in negotiating the imperatives of scientific inquiry with the increasing pressures of social and contextual relevance. I answer these questions through analyzing the case of the Harm Reduction Peer Recovery (HaRP) clinical trial, an ongoing hybrid effectiveness-implementation clinical trial that brings together public health, criminal-legal, and harm reduction actors working to stem the tide of overdose in their local communities.[4] As a hybrid trial of a complex, community-based intervention, The HaRP clinical trial serves as a key site for understanding how the commitments of scientific inquiry are negotiated with social service goals.

While health services research may have borrowed the language and formal methodologies of the natural sciences, characterizing research groups as labs, the "field" of community-based programming like HaRP cannot be so strictly controlled and delineated, and deciding how to do so requires us to apprehend certain moral and political choices. As jails and social services became embedded within and subject to research, trial actors were forced to negotiate broader social and professional values with the commitments of scientific inquiry. Normative accounts of implementation tend to obscure important political and social processes that are essential to apprehend in order to successfully implement a complex intervention like HaRP. That is, the contexts where HaRP was taken up are not simply "sites" for research but also the grounds for its contestation and production. My aim is to highlight the political dynamics and consequences of scientific reform, both micro and macro, through the case of the development of an evidence-based program.

[4] Pseudonyms like "HaRP" are used throughout this dissertation. All names are also pseudonyms.

In this dissertation I argue that the imperatives of the clinical trial, and the institutional obligations to which it was tethered, operated against the ethical and political commitments of practitioners and researchers themselves related to how best to prevent overdose. In other words, what would become recognized as the intervention itself was openly contested throughout implementation. Ultimately, however, the experimental apparatus tended to take precedence over the "real world" imperatives of clinical practice. The pursuit of objectivity, generalizability, and control were met with the vagaries of the real world: local activism, contextual idiosyncrasies and versions of care and social intervention that failed to match up to the logics inherent to the clinical trial and the carceral contexts where it took place. However, no one actor had an unwavering commitment to the scientific integrity of the clinical trial. Experimental concepts like feasibility, fidelity and objectivity were all contested throughout, by PIs and practitioners alike, and were actively redefined to meet the diverse, and often competing nature of hybrid clinical trials.

Harm Reduction and the Criminal-legal System

The questions I pose above are particularly important to apprehend given that HaRP represents a relatively unprecedented federal investment in both rerouting substance use treatment into American jails and prisons and making harm reduction and peer recovery "evidence-based" through a network of large-scale clinical trials. As I will show in the following section, this is because both require fundamentally political processes. A complex, 12-month harm reduction and peer recovery-based case management program, HaRP follows a "hub and spoke model," where training is provided at a central "hub." Peer specialists and case managers then work to deliver the intervention at six "spoke" sites at jails across a midwestern state. Launched in 2021, the primary goal of the trial was to connect people exiting jail to medication

for opioid use disorder (MOUD). The study aims to both test the effectiveness of HaRP *as well as* facilitate its fit within the "real world" contexts of jails and the array of social and medical services that have grown in and around them.

In the federal response to the national opioid crisis, jails and prisons have come to be seen as a "critical target." This is not surprising given that the carceral system has long been a primary mechanism to manage substance use disorder in American policymaking (Campbell, 2020). In late 2018, the National Institute on Drug Abuse (NIDA) published a Request for Applications (RFA) as part of a special call for researchers to build evidence-based solutions to the opioid crisis at the nexus of the country's carceral and public health systems. NIDA budgeted around 25 million dollars to fund the various projects for a given year. This special call prompted health services researchers from a variety of disciplines to submit grant applications for developing and testing strategies to reduce overdose among people who are incarcerated or recently released from criminal-legal settings. Spanning almost 40 states, the funding supported a collection of studies that were designed to support the translation of substance use research into criminal-legal settings. HaRP, one of the trials funded under this mechanism, is a clinical trial that brings together novel partnerships, particularly between harm reductionists, academic researchers and criminal-legal actors.

The federal request for proposals was developed after the publication of a widely circulated study that showed the primary cause of death among people formerly incarcerated was drug overdose, with the majority of overdoses occurring two weeks post-release. Cited over 1,000 times, this article published in the *New England Journal of Medicine* titled "Release from Prison — A High Risk of Death for Former Inmates" found that individuals faced a significantly elevated risk of overdose death in the two weeks following release from jail or prison, compared

to the general population (Bingswanger et al., 2007). This study suspected that a period of abstinence during incarceration resulted in reduced physiological tolerance to drugs, thereby increasing the risk of drug overdose upon release. (Binswanger et al., 2007). The request for proposals therefore understood the rising rate of opioid overdose among populations released from jail or prison due to a loss of tolerance, rather than the criminalization of drug use and enforced abstinence itself. In other words, this framing of the problem (enforced abstinence → loss of tolerance → overdose) meant federal resources would be spent trying to improve availability of medication for opioid use disorder (MOUD) in and around American jails and prisons. This framing would mean that researchers would need to secure buy-in with jails and prisons in ways that compromised their relationship to the practitioners who delivered HaRP, particularly those invested in harm reduction.

With its tethering to the evidence-based policy movement, harm reduction has become a recent focus of federal funding and attention, with HaRP being one harbinger of this progress (Des Jarlais, 2017). During the Spring of 2021, the Biden administration unveiled plans to allocate unprecedented levels of funding towards peer recovery and harm reduction services under its "Unity Agenda." Concurrently, in 2021, the Substance Abuse and Mental Health Services Administration (SAMHSA) launched a grant program focused on funding state-level harm reduction efforts. While these programs demonstrate growing support of harm reduction, it was only recently that harm reduction has become acceptable by federal agencies, philanthropic organizations, and academic institutions alike. Historically, as a form of credible knowledge, harm reduction tended to be resisted, a "dirty word" even, on political and ideological grounds (Campbell, 2020). Harm reduction's growing acceptance, as signaled by increased federal and state investment has also meant, as in the case of HaRP, that harm reduction has expanded into

novel terrains where logics might fundamentally clash, particularly the criminal-legal system. Actors such as police, sheriffs, and state's attorneys, groups often characterized as resistant to institutional change, have been asked to adapt and implement novel approaches outside their typical purview. Sociological research has demonstrated how, and with what consequences, criminal-legal actors have taken on the practices and responsibilities of fields otherwise designated to medicine or other helping professions (Lara-Millan, 2020). Less have shown how fields and social movements that might actively contest these processes operate alongside such transformations.

With the support of the burgeoning HIV/AIDS movement, the harm reduction movement was developed by peer-led grassroots activists amidst the state abandonment of drug users during the 1980s. Together, they focused on housing-first initiatives, meeting the basic needs of drug users, and expanding needle exchange programs, all of which tended to require acts of civil disobedience (McLean, 2011). Importantly, these user-led groups challenged conventional drug policy, particularly the role of the criminal legal system in *perpetuating harm* through mass incarceration, sentencing disparities and over-policing (Campbell, 2020). The goal of harm reduction, per harm reductionists, was not about control, criminalization, or "managing risk," but instead they promoted the radical acceptance of drug users and challenged the criminalization of drugs and drug users.

Of course, the movement's need to appeal to evidence-based policy in recent years has also meant shedding some of its radical roots. For example, many public health-oriented harm reduction organizations tend to focus on syringe exchange and other more "medicalized" strategies of harm reduction, without the underlying adversarial and social justice orientation. Even so, certain elements of harm reduction are only integrated uneasily into formalized public

health and criminal-legal venues, where actors hold different conceptions of the causes of harm, ideal behavioral change goals, the moral worth of drug users, as well as the role of the criminal-legal system and direct action in addressing the causes and consequences of drug overdose. Public health, harm reduction and criminal-legal actors tend to aspire to contrasting solutions to prevent overdose, abide by different ethical norms as well as professional and institutional commitments. Campbell (2020) argued, in referencing the collaboration of harm reduction actors with scientists that, "when governments get scientists to deploy technocratic risk assessment, efficacy assessment, and cost-benefit analysis, all seemingly politically neutral ways of analyzing technology's effects upon populations, the terms of engagement are closed to question" (p. 192) That is, larger political questions about what counts as harm in the first place, for example, tends to be off the table. However, I show that, even in the aseptic, seemingly politically neutral fielding of a clinical trial, these questions are not closed but still very much open to debate and reveal important tensions about the collaboration between harm reduction, academic and carceral actors and the institutional obligations they attempt to fulfill. The following section will attend to randomized controlled trials, providing context to how these political processes might be both elided and produced during their implementation.

Randomized Controlled Trials and Evidence in Social Policy

Evidence-based policy is used as an umbrella concept that refers to the use of a range of empirical methods, from cost-benefit analysis to randomized controlled trials that are used to ground policy-making decisions. By its proponents, it is often used to invalidate other policy-making methods, primarily those regarded as biased, subjective or grounded in personal experience. Randomized controlled trials are often referred to as the *gold standard*, though most

commonly as a rhetorical setup by scholars who then follow with their critiques.[5] Indeed, it should be noted that the majority of practitioners of the RCT are often the first to diagnose their limitations (See: Steeger et al., 2021). Nonetheless, RCTs are ultimately recognized as the principal method to legitimize an intervention within contemporary social service provision (Mosley, Marwell & Ybarra, 2019; Tanenbaum, 2005). Philanthropic organizations and U.S. federal agencies often use RCTs to determine which programs should receive funding or not, including in the evaluation of reentry, education and social service programs.[6] They are also used in program evaluation to aid in Congressional decision-making (Brass, 2006). Recently, federal legislation like child welfare's Families First Prevention Act (2018) linked federal funding to pre-approved evidence-based programs, or those which have shown efficacy through an RCT (*Title IV-E Prevention Program*, 2023). Perhaps because of the RCT's prominence in contemporary policy making, much ink has been spilled over the past decade surrounding its relative merits. These conversations roughly follow three primary themes: 1. The political economy of RCTs, 2. Epistemic critiques about how context is attended to (or not), and 3. The social and organizational effects of the RCT protocol. A brief history of the RCTs' emergence helps to elucidate the stakes of these debates.

Following the "Thalidomide disaster," which showed catastrophic side effects of the drug among pregnant women, legislation passed in 1962 mandated the Food and Drug Administration's (FDA) use of RCTs to demand proof of the safety and efficacy of pharmaceuticals before they could be brought to market. The regulatory mandates of the RCT

[5] Cartwright would argue that there is no such thing as a "gold standard," and that the value of the method relies on the question at hand: "Gold methods are whatever methods will provide the (a) information you need, (b) reliably, (c) from what you can do and from what you can know on the occasion." (Cartwright, 2007, p. 11)

[6] Each year, thousands of RCTs are being fielded. 12,987 NIH-funded clinical trials were registered between 2005 and 2015 (Greshem et al., 2018)

became the *de facto* strategy to evaluate drugs, taking over all other modes of assessment including the clinical reports of physicians who prescribed them (Carpenter, 2014; Eyal, 2019). The RCT was conceived as an objective mechanism that would ostensibly curtail the whims of the American Medical Association (AMA), given its cozy relationship to the pharmaceutical industry and establish "incontrovertible facts" about a drug's safety and efficacy profile (Eyal, 2019, p. 117).[7] The FDA would thus assume full regulatory control while the AMA, namely doctors, or those with more practice-based experience and little statistical training, took to the sidelines. The rise of pharmacology and new statistical techniques promised that this adherence to standards would eliminate subjectivity, and bias, and with it, resolve the FDA's apparent crisis of legitimacy.

Randomized controlled trials were eventually adopted in the field of social welfare. The development and expansion of social-welfare research following World War II prompted an increase in federal funding for social program evaluation across service domains (Brown, 2019). Beginning in the 1970s, large-scale social programs (e.g., Negative Income Tax) began to be subjected to scientific evaluation to curtail what were considered the ideological whims of administrators interested in pursuing their own political interests (Berk et. al, 1985; Campbell, 1969; de Souza Leão & Eyal, 2019). However, by the 1970s the experimental method—as applied to social policy—was eventually deemed unethical, politically contentious, expensive, and altogether too complex to implement. The ethical implications, for example, of randomly assigning people to a control group so as not to receive life altering cash assistance was

[7] Eyal (2020) describes how the RCT would help to bolster the legitimacy of the FDA and reduce the perceived bias of earlier methods of evaluating drugs: "adherence to the RCT protocol–prospective study designs, detailed procedures, centralized administration of experimental assignment, randomization, control of placebo effects–would eliminate subjectivity. 'Winners and losers' would not be picked by the whim of state bureaucrats, nor through collusion between self-interested corporations and venal physicians, but by an objective mechanism that will determine incontrovertible facts about safety and efficacy (p. 117).

perceived as politically fraught. Moreover, it was difficult, during decades long experiments, to know whether the control group had accessed newly introduced programs (e.g., Aid to Families with Dependent Children), thus impacting the perceived purity of the control group (Berk et al., 1985). In response, RCTs underwent a significant transformation in ensuing decades. The interventions tested became smaller (e.g., evidence-based programs that focus on behavioral changes), the time frame shorter (and therefore less expensive), and funding sources more diverse and reliant on philanthropy as well as government support (de Souza Leão & Eyal, 2019).

Meanwhile, American social work programs were also looking for more effective approaches, largely grounded in cognitive and behavioral theories and problem-oriented approaches. Both offered discrete outputs and outcomes and were amenable to quantitative evaluation research and could also ground social work more wholeheartedly in empirical research (Kirk & Reid, 2002; Okpych & Lu, 2014). These new models also were warmly welcomed during the era of what is now referred to as *new public management*, which introduced private sector management techniques into the welfare sector with the goal of making public agencies more effective, and most importantly, efficient. Nonprofit contracting, for instance, became a source of competition among organizations for government bids, and new methods of performance management arose to ensure putatively proper and efficient use of government funds (Smith & Lipsky, 1993). EBPs, with their aura of scientific credibility, were helpful in gaining public trust during times of perceived scarcity (Tanenbaum, 2005). A new type of social work research infrastructure, focused on innovation, and the scaling of standardized programs, thus emerged. It is the narrowing of interventions, political context in which it

emerged and thrived, and its status as an objective arbiter of policy decisions, that lay the foundation for contemporary critiques of the method.

Social scientists have labored to point out how RCTs are fundamentally political, both in their emergence and in their effects. Importantly, RCTs proliferated alongside welfare state retrenchment and neoliberal policies (Biehl and Petryna, 2013; Brives et al., 2016, Adams 2016; Ramachandran, 2020), particularly in the human services where they are used to determine not just effectiveness, but importantly, cost-effectiveness (Carr, 2023). That is, questions of cost have become ideologically "bundled" with questions of efficacy whereby promises to reduce cost are also defined as "effective" (Carr & Obertino-Norwood, 2022). RCTs are therefore deeply connected to the rationalization of the public sector brought into relief by contemporary audit culture and movements toward new public management.

Across the social sciences, RCTs are often therefore tied to a political "narrowing" in two senses. First, their projected authority has constricted the way complex social problems are framed, constricting space away from framing social problems more expansively (e.g., on human rights or social justice grounds). Second, the very method allows for only certain types of interventions to be studied (namely those centered around individual action).[8] By trying to find "what works," a delimited framing that tends to focus on effectiveness and cost-effectiveness, only certain forms of intervening become tenable (Ben Menachem, 2019). That is, focusing on what is measurable, they argue, limits the types of interventions to those more easily put through the sieve of scientific inquiry. Other ways of knowing, and other forms of legitimizing social policy become sidelined and seemingly impossible to put forth as valid policy ideas (Gottschalk,

[8] Scholars have focused on the influence of microeconomics on RCTs proliferation, and how nudges and behavioral change, the primary mechanism of enacting social change within RCTs, as quintessentially neoliberal (Stein et al., 2021; Ramanchandran, 2020).

2016). It is perhaps this inattention to politics and structural change that allows for its putative objectivity.

The Treatment of Context in RCTs

Across the social sciences, researchers have lamented the perceived erasure of context, or the relegation of political and cultural context in an effort to standardize, control for extraneous variables, and give findings their "universal status" (Brives et al., 2016, p. 370). In these accounts, ethnographic, qualitative, and community-based participatory approaches that could help to pinpoint crucial causal mechanisms as well as local cultural, historical and political contexts that might better inform health services, are sidelined to the quantifying and universalizing impulse of the RCT. Here, critics also suggest that the local concerns of practitioners are supplanted by the technocratic strategies of experts (Adams, 2016). It is this presumed erasure of context that critics argue makes them less able to generalize beyond the trial apparatus.

Within traditional clinical trials, context is often treated as an inconvenience to securing a universal cause-effect relationship. Although RCTs are frequently lauded for their robust internal validity, their need to control contextual variables poses challenges regarding external validity.[9] Much has been written sociologically about how the contexts of RCTs are often ignored by researchers at the expense of their transferability. The resources tethered to RCTs often make their interventions more likely to show positive results and less likely to survive the "real world." RCTs are highly resourced, procure highly trained professionals, introduce novel and at times expensive technologies, and follow procedural manuals to ensure staff reliably adhere to

[9] By internal validity, I am referring to the capacity to ascertain a cause-and-effect relationship through randomization, wherein potential confounding variables are evenly distributed between the intervention and control groups (Krauss, 2021).

intervention protocols. In what she refers to as "material standardization" Montgomery (2017) argues that the drive to standardize in clinical trials, though methodologically necessary, may render new innovations relatively unsustainable, while also helping clinics that conduct research more likely to be a vehicle for successful intervention outcomes.

Rather than treating context as a backdrop for the cause-effect relationship to emerge, other researchers argue that context *interacts* with a given intervention to bring about said changes. Failing to appreciate the interactive relationship, it has been argued, oversimplifies how the cause-effect relationship manifests within a particular context. Instead, this line of scholarship highlights the situated practices, and relationships that help make interventions efficacious, rendering any given intervention difficult to analytically separate from its context (Hardon & Sanabria 2017; Brives, 2016). Reidpath, et al. (2022) have characterized this phenomenon as "shielding." This concept elucidates that while an RCT may provide insights into whether a specific treatment caused an outcome, these findings do not extend beyond the parameters of the trial without appropriate "shields" or support structures. An intervention's success is therefore dependent on a range of interactive processes and, for that reason, are contextually bound (Reidpath et al., 2022).

Few studies, with some exceptions, have shown how these tensions, including problems of generalizability, the politics of their objectivity, and the effects of the protocol are actively negotiated during the day-to-day process of running a clinical trial. None have actively highlighted how researchers *themselves* experience, grapple with and describe these very tensions, as this dissertation sets out to do. Studies that do examine the social effects of standards within the context of RCTs tend to focus on the participants of clinical trials, often through the lens of social control and surveillance (Brives, 2016; Winther & Hillersdale, 2020). And, while I

am often sympathetic to the political critiques of RCTs both on a micro and macro scale, they are rhetorically less effective when the reasons critics dismiss RCTs are the very reasons practitioners of RCTs themselves abide by and champion them (e.g., their objectivity). What seems necessary today, is highlighting how scientific ideas become contextualized, or how objectivity, standardization and other scientific processes are negotiated with the complexity of the "real world" in order to produce a putatively scalable intervention. In doing so, we uncover the very political processes that RCTs otherwise might obscure. By focusing on researchers and implementation actors, I am able to show how seemingly stable assumptions about the RCT, for example, what the intervention is, and how social change should occur, were contested and constantly negotiated throughout implementation.

Lastly and importantly, across these critical analyses of RCTs, there has been a lack of research on the emerging field of implementation science and broader trends in health services research toward "pragmatic" and "hybrid" clinical trials. In the field of implementation science, context enters the purview of scientific analysis to deal with some of the very methodological limitations of RCTs described above (e.g., their lack of external validity), thus making social interventions more attuned to "the real world."

Implementation Science

Although randomized controlled trials (RCTs) hold epistemic authority within health and social policy, the evidence-based interventions they produce often fail to translate into practice settings (Cartwright, 2013). In the context of clinical trials, this dilemma has effectively been outsourced to implementation science, a field that has emerged to manage, in a rather technical fashion, the complexity of the "real world" of clinical practice to help facilitate the uptake of evidence-based practices into routine care (Rhodes & Lancaster, 2019; Boulton et al., 2020). In

doing so, the field has tended to congeal around evidence-based practices, and with it, the prioritization of certain forms of evidence and the economic and political imperatives of federal institutions most often invested in clinical and cost effectiveness.[10] Less a problem with the production of evidence itself, the general diagnosis within implementation science is that the findings of "basic" science fail to match the needs, desires and local contexts of clinical practice. In other words, because evidence-based programs failed to show the right effects beyond the clinical trial, the health services field turned their attention to the *context* of intervention adoption. Critically, social service organizations and society are shored up to blame for the null findings of scientific results rather than the intervention or the idea that a given program can be abstracted and universally deployed in the first place.

In recent decades, the field has gained traction. The *Journal of Implementation Science*, established in 2006, brought together a diverse array of researchers and academics to establish consensus across researchers doing similar work in disparate corners of the behavioral health and health services fields (Boulton et al., 2020). Meanwhile, a growing proportion of funding from the National Institute of Health (NIH), UK Medical Research Council, and the Canadian Institutes of Health Research has been allocated to implementation science (See: Chambers & Emmon, 2024 for a review). Most funding agencies now require an implementation science component to clinical trials, understanding the integral nature of context to an intervention's success and future diffusion. This is because implementation science has been recognized as a field of research that can attend to the 'black box' of experiments, identifying the factors within the environment and the intervention itself that contribute to successful implementation and thus

[10] Recently, the field has called for de-tethering itself to evidence-based practices, and instead studying the implementation of various kinds of interventions, policies and practices, particularly those championed by communities of practice themselves (Chambers, 2024; Martinez, Callejas, Hernandez, 2010).

intervention outcomes. Due to its role as a largely ancillary component to clinical trials, implementation science is embedded within the evidence-based policy paradigm, and with it, the values, and epistemic traditions of health services research and funding mechanisms more broadly.

The Scientific Roots of Implementation Science

Rogers' *Diffusion of Innovations* (1962) is the most widely recognized theoretical framework to inform the contemporary field of implementation science. Diffusion of Innovations offers a generalized theory for how new innovations spread and identifies factors most likely to influence adoption. First developed in rural sociology to understand and promote the uptake of agricultural innovations by farmers, the diffusion model came to influence a variety of disciplines including education, communications, and public health. To explain and predict the diffusion of innovations, Rogers categorized aspects of context, the stakeholders taking up the innovation, as well as characteristics of the intervention itself into different ideal types. Based on a synthesis of the existing diffusion literature, Rogers established what he called the *four elements of diffusion.* One could roughly summarize these elements as characteristics of the innovation (relative advantage, complexity, trialability), the types of communication channels at the site of diffusion, the characteristics of the adopters themselves, and the social system (relationship hierarchies, presence of opinion leaders), each of which affects diffusion trajectories. Importantly for the matters at hand, *Diffusion of Innovations* theory provided a model for studying context in a scientific way and formed a foundational model for present day frameworks in implementation science.

While a variety of frameworks exist, each with their own theoretical and conceptual genealogy, most frameworks systematize ways of (1) describing implementation, (2) identifying

barriers and facilitators to an intervention's success, and (3) evaluating the implementation process (Nilsen, 2015). Generally, within these frameworks, interventions are created independently of the context for which they are intended, and movement is "unidirectional, and predominantly linear and proceeds in a stepwise fashion. These steps include awareness of the new evidence, followed by learning, trialing and deciding to adopt, reject or modify that evidence and finally reinforcing the decision" (Dryden-Palmer, 2020, p. 2). Context thus tends to be treated as something to manage, control, or foster and is conceptualized as *analytically distinct* from the intervention as well as the research process. This treatment of context is likely because of its epistemic commitments to evidence-based practice and policy, and its historical roots in the traditions of clinical research, which tends to privilege quantitative research and positivist epistemologies.

Within implementation science, context tends to be treated as "representational," whereby issues like culture, context and climate are quantified and abstracted into discrete categories and framed as either facilitative or as a barrier to intervention efficacy (Dourish, 2004). Similar depictions of how context is conceptualized have come from internal critiques of the field (May et al., 2016; Dryden-Palmer et al., 2020; Kislov, 2019; Wensing & Grol, 2019). This approach to context is rooted in the experimental model, which aims to identify and control for 'extraneous' variables in order to distill causal mechanisms and measure outcomes (May et. al, 2016). Additionally, to make generalizable claims and find patterns across local sites of practice, context needs to paradoxically be made commensurate across place, which requires the simplification of a range of practices and interactions. Though important for developing evidence, this representational approach often limits researchers from apprehending the interactive, and processual nature of implementation.

By virtue of the frameworks and methods used within implementation science, the social processes upon which translation depends are often obscured in accounts of evidence-based intervention development and implementation. As has been argued elsewhere, implementation research has often neglected the heterogeneity of implementation actors, and how these interactions may reveal important political processes and inevitably affect the implementation process (Kislov, 2019). For example, implementation scientists must help to elicit buy-in, study and manage relationships between institutions with different levels of power and, most formidably, attempt to standardize and regiment highly intimate forms of social service labor. Moreover, this representation of context fails to account for the ways in which context is actively produced. By this I mean, if we are to understand the relationship between intervention and context as interactional, work must be done to disentangle them. This dissertation shows the complex negotiations that are required to distinguish an intervention from its context, and by extension, creating a stabilized, scalable intervention. In doing so, I uncover the moral and political processes that are necessary though often omitted in both doing implementation science and fielding complex RCTs.

Extending the Scale of Context

In recent years, the field of implementation science has started to extend the scale of its contextual focus. Still defining the scope of its boundaries, implementation science has started to grapple with how it might address "outer" context or "macro level factors," those obstinate social structures that are known to create and intensify health disparities and impact implementation outcomes of the evidence-based interventions practitioners have been tasked to carry out (Brownson, 2021; Shelton et al., 2021). There is now broad acknowledgement from within the field that it needs to facilitate comprehensive, structural change, and an overwhelming sense that

it has historically elided structural elements of context that might serve as essential "factors" to intervention success – from history and political economy to, more recently, structural racism (Gustafson et al., 2023). Implementation science, and the implementation researchers I worked with, have called for a more explicit consideration of the political, social, and economic conditions that lead to inequitable outcomes in the first place. Yet, pursuing such change, as I demonstrate in this dissertation, might place the field in a tenuous relationship with the very institutions and methods of research that would help it to become a more formal and coherent science.

Recreating "The Gap"

While implementation science has made strides to establish itself as a field of science, it also tends to understand itself as motivated by the pragmatic and ethical concerns of democratizing evidence, or at least making interventions more responsive to local communities of practice. The "gap" is a term deployed by implementation science to describe the, on average, 17 years it takes for the findings of basic science to move into the real world of clinical practice (Rubin, 2023). As such, it is driven by an ethical and moral imperative of ensuring the findings of science are taken up. Importantly, it has done so in an attempt to make what is developed by *other* scientific disciplines useful to patients and service users, namely evidence-based practices and programs. In historical accounts focused here, implementation scientists describe their efforts as being motivated by local practice communities and the people they serve, desiring evidence, strategies and practices that are better aligned with the daily work, cultures, and constraints of social service and health delivery.

Contemporary implementation scholars echo this historiography of the field.[11] Yet, researchers and practitioners have started to lament the growing chasm between implementation research and practice (Harvey et al., 2023). Specifically, researchers and implementation practitioners have expressed frustration over the growing number of complex frameworks that have been developed without the practice context in mind. In other words, the field's attempts to cohere a science have made it into something it has developed in reaction to, a field that produces even more "gaps" between practice and research. I argue that the very issues the field has identified as causes of its misguidance and stagnation–the side effects of seeking legitimacy through science, the re-creation of the research to practice gap, the development of models and theories that fail to serve community partners–all point to a principle tension that exists within the field, and one I address in this dissertation: how do we conduct science while also making it socially and contextually relevant? (See: Beidas et al., 2022).

Throughout this dissertation, I illuminate a paradox within the field of implementation science that researchers and practitioners alike must contend with and negotiate: On the one hand, identifying implementation as an obstacle to desirable health and social service outcomes immediately focuses attention on local conditions and the need to translate and adapt to them. On the other hand, codifying implementation practices into a formal science, means lifting this practice of adaptation outside the specific context and examining it from the point of view of yielding generalizable results (e.g., to build theory and models). Within the context of a clinical

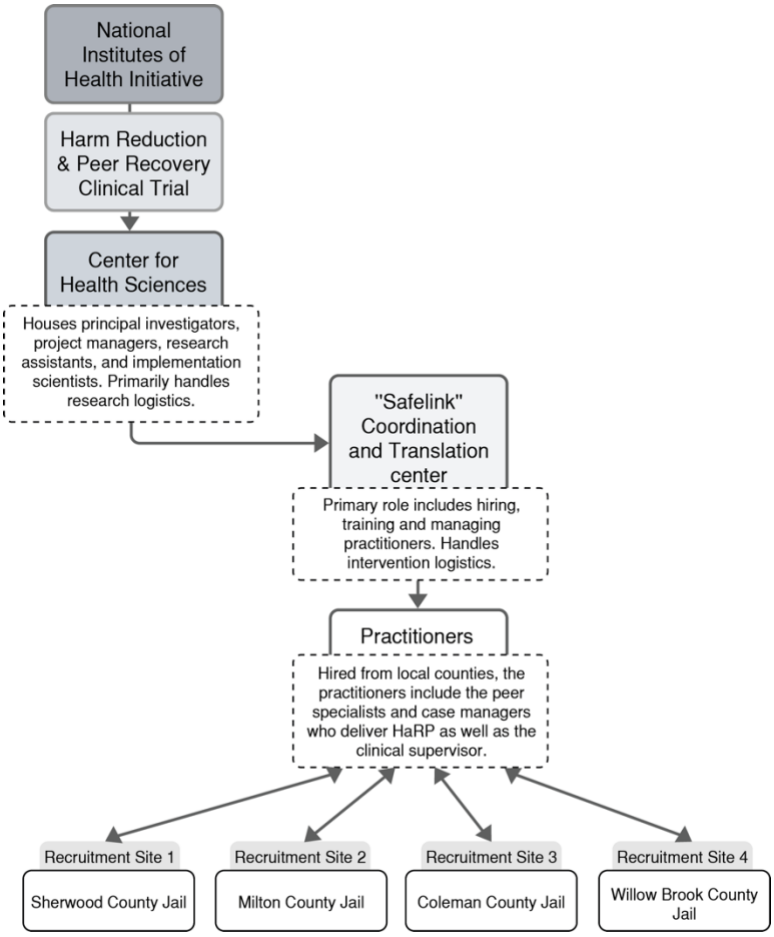
[11] Consider how Beidas et al., 2022, in their post-mortem of the field, describes the concerns of implementation science during the early years of its emergence: "We came to implementation science frustrated because patients and community members did not routinely receive evidence-based practices (EBPs) and because policies were not aligned with high-quality research evidence. We became aware that resources spent developing EBPs were not translating into their routine delivery outside of research contexts, and recognized that racially, ethnically, and socioeconomically diverse communities, populations, and settings that would benefit most from EBPs were not equitably reached" (Beidas et al., 2022).

trial like HaRP, researchers must therefore manage the imperatives of standardization and creating generalizable findings, while also remaining true to the immediate contexts of implementation and the day-to-day realities of implementers and the public they're attempting to serve.

Study Site & Actors

The primary ethnographic site of this dissertation is the HaRP clinical trial, where researchers are collaborating alongside criminal-legal, public health and academic research institutions to create and evaluate the HaRP intervention (See Figure 1). The project was spearheaded by the Center for Health Sciences (CHS), a university-adjacent research center situated in an urban, Midwestern city. Importantly, CHS had experience fielding clinical trials throughout the state, as well as longstanding partnerships with state prisons and jails.

Figure 1: HaRP Organizational Map



The HaRP intervention was rolled out at four jail recruitment sites across the state—one urban, one suburban, and two rural (See Table 1). HaRP researchers were particularly motivated to understand barriers to treatment in rural contexts. In rural contexts, there is thought to be stigma toward people who use drugs, resistance toward harm reduction, a lack of MOUD providers, minimal public transportation and other important barriers that frustrate peoples' ability to access treatment. Due to pandemic-related lockdowns, many of the sites did not begin recruiting until after 2022, meaning that much of my ethnographic research took place during the design and the implementation of HaRP at Sherwood County Jail, recruitment site one.

Table 1 : Clinical Trial Recruitment Sites

County & Jail beds	County Description
Sherwood County Jail	Rural county with a high incidence of fatal overdose. No discharge or re-entry services, though harm reduction and community-based mental health and substance use providers exist. Drug use is highly criminalized through prosecutorial and police officer discretion.
Coleman County Jail	Rural county with a rapidly increasing rate of fatal opioid and meth-related overdose with few re-entry and discharge services. Drug use is criminalized through prosecutorial and police officer discretion.
Willowbrook County Jail	Wealthy suburban county with strong re-entry, harm reduction, substance use treatment, and case management services. Drug use is less criminalized, and people are less likely to be held in jail for possession and individual use.
Milton County Jail	Urban county with strong discharge, re-entry, substance use and harm reduction services. Drug use is less criminalized, and people are less likely to be held in jail for possession and individual use.

HaRP Actors

The HaRP clinical trial provides an important case for studying collaboration and negotiation because it brings together novel, if sometimes unprecedented, partnerships, including between criminal-legal actors, harm reductionists and the academic research community (See Table 2). Within the trial, such actors operate with competing epistemologies, ethical norms, and at times, notions about the very purpose of the trial they are participating in. Yet, it is incumbent on study leadership to manage the tensions of ensuring success of the experiment while also

building trust and credibility with actors performing and hosting the intervention. The following passages will detail the main actors of the HaRP trial.

The majority of clinical trials begin with the *principal investigators*, who, with the help of a cadre of research assistants, and project managers, write the grants, design the study, analyze the data, and then pen the papers and reports. Administratively, PIs must secure data-use agreements, negotiate with funders, establish a team of other researchers, and forge, as well as maintain, relationships with community partners. HaRP was spearheaded by two researchers, Robert, a public health researcher, and Trevor, an infectious disease researcher, both from the Center for Health Sciences (CHS).

To field the HaRP clinical trial, the PIs also require the expertise of *implementation researchers*. A key component of the research team during a hybrid clinical trial, implementation scientists are charged with guiding, documenting, and evaluating the implementation process as well as studying how context meaningfully varies across clinical trial sites. The implementation team assesses and ensures fidelity, which means documenting whether practitioners adhere to the HaRP program model. A key aim guiding the implementation science team is to identify barriers and facilitators to HaRP's implementation success. In doing so, they examine how the various contexts represented in the trial shape intervention success. Lastly, they are also tasked with identifying implementation strategies to foster intervention adoption and implementation.

As part of the NIDA grant agreement, CHS selected SafeLink as their "Translation Center" (See Figure 1), which would provide the social service infrastructure to field the HaRP intervention, including clinical supervision, training, technical assistance, and management of the HaRP practitioners. SafeLink is an urban community-based organization also affiliated with a public university, which conducts both epidemiological research and provides street outreach,

counseling, testing services, and support groups for people who use drugs (PWUD). SafeLink trains peers, or people with lived experience of drug use and incarceration, to conduct outreach and research, some of whom were hired onto the HaRP study.

Hired by the study from local harm reduction and community-based organizations are the *practitioners* who deliver the HaRP intervention at different study sites across the state. The work practitioners is overseen by the clinical supervisor who works with the implementation team to troubleshoot with the case managers and peer specialists in order to ensure they properly employ the intervention through weekly supervision and monthly case consultations. Practitioners exist in a liminal space within the experiment—they are part of the study team in the sense that they are the people performing the intervention, but their activities are also under examination by the research team to ensure standardization of the intervention across sites as well as fidelity to the model.

HaRP employs two types of practitioners: (1) The case manager provides needs assessments, service planning, and linkage to services for intervention participants in the RCT, and (2) the peer specialist or "peer," who is intended to serve as a mentor and advocate to study participants, inspiring hope on their recovery journey. I refer to peer specialists and case managers, when described together, as *practitioners*. When I describe implementation researchers, administrators, managers and practitioners together, I will refer to them as *implementation actors*. While all practitioners received the same training in order to deliver HaRP, they also arrive at the study with diverse ideological relationships to the criminal-legal system, hail from different types of harm reduction service systems, and maintain different relationships to research. While some practitioners from SafeLink had experience conducting research, those hired from community organizations had virtually no experience conducting

clinical research. Importantly, none of the practitioners hired for HaRP had ever been involved in an RCT.

Table 2: HaRP Trial Roles

Trial Actor Title	Trial Actor Role
Principal Investigators	With the support of research assistants and project managers, the PIs write the research question, design the experiment, analyze the data, and publish the academic papers. The PIs must also establish a wide and diverse network of researchers who will make up the research team. Administratively, they must secure data-use agreements, negotiate with funders, and establish relationships with community partners. Also, part of the research team includes the implementation scientists.
Implementation Scientists	A key component of the research team are the implementation scientists, who are charged with guiding, documenting, and evaluating the implementation process. They examine and document how context facilitates or frustrates implementation success. In the HaRP trial, they are also tasked with identifying implementation strategies to foster intervention adoption and implementation.
Practitioners	Hired by the study from local harm reduction and community-based organizations, are the practitioners who deliver the HaRP intervention at different study sites across the state. Their activities are evaluated and managed by the clinical supervisor and research managers. HaRP employs two types of practitioners who work in collaboration: (1) The case manager provides needs assessments, service planning, and linkage to services for HaRP participants. (2) The peer worker, often hired from a local harm reduction organization, serves to inspire hope and advocate for study participants.
Jail Partners	Jail partners are the administrators who work at county jails that serve as sites for HaRP recruitment. Jail partners' level of engagement in the study varied across sites, but most commonly they helped to coordinate space and the "movement" of participants to support recruitment into the HaRP Trial.

Methods

My findings draw on two and a half years of ethnographic research, including participant observation of the HaRP clinical trial. I began as a research assistant on the project in the Winter of 2019 and received institutional review board (IRB) approval for my study of the trial in the summer of 2020.[12] Unlike many researchers who experienced a disruption to their fieldwork during this period, my ethnographic project continued because of the mandates of the clinical trial, and because, as a research study, my interlocutors also adapted to the pandemic by moving most research processes and practices to a virtual format. Even before the pandemic was announced, many meetings took place over Zoom largely out of convenience given that HaRP researchers and other implementation actors were scattered across the city. After the pandemic, however, the practices and processes I followed primarily shifted to a virtual format.

Ethnographic research is grounded in specific contexts, often characterized by extensive immersion. Spending extended time in a given context has empirical value: it allows ethnographers to gain a nuanced understanding of their interlocutors and the everyday institutional constraints they face *as* they encounter them. It also helps researchers to build trust with interlocutors, bringing increased validity to the data. The majority of qualitative research that examines the implementation of evidence-based practices and other health and social service innovations tends to look at variables related to the success or failure of a given intervention, including feasibility and acceptability.[13] Yet, I was not so interested in how or whether HaRP worked, though this is important. Instead, I wanted to understand the way researchers managed

[12] The HaRP study is an ongoing clinical trial that has been substantially delayed by the COVID-19 pandemic as well as other factors, some of which are described in this dissertation. Therefore, clinical trial outcomes are not available at the time of the publication of this dissertation. Therefore, analysis, and other arguments stem from the time I participated in the trial as a participant observer, between July of 2020 and November 2022. Since November of 2022, I have continued on the study in volunteer fashion as a research assistant.

[13] This is of course due to the priorities of funders as well as the more immediate, pragmatic concerns of health and social service providers.

the scientific commitments of the clinical trial with "real world" context, including how feasibility is imagined or with what social effects fidelity is employed. This ethnographic dissertation widens the scope of implementation analysis to understand the research process, funding structures and epistemological frameworks that are often otherwise elided in accounts of intervention and implementation science. Ethnographic analysis is particularly appropriate for observing and analyzing processes, power dynamics as well as the structural conditions that help to shape service delivery. While meeting notes and interviews might explain what decisions are made, they may be less adept at describing how these decisions are made, including whose ideas are listened to and validated, and whose thoughts and concerns are ignored or reinterpreted.

I attended daily meetings for over two years, including formal intervention trainings, clinical case consultations, community advisory board meetings, implementation research team meetings, and "all team" meetings, where directors, managers, and supervisors of SafeLink would convene with the rest of the academic research staff on the HaRP project. Throughout this period, I followed up on points of interest with semi-structured, key-informant interviews with researchers, practitioners, as well as jail and community stakeholders (See Table 2). In interviews, I tried to reflect back to my interlocutors what they were communicating to me, reiterating the subtext underlying their responses that pointed to institutional constraints, as well as the political and ethical tensions that occurred during HaRP's rollout.

All meetings, when provided proper permissions, were recorded, transcribed and coded. Field notes, which I took during the initial two years of my project, were also coded. A research assistant helped me in 2022 with organizing the meeting recordings I had accumulated, as well as with transcribing. I also read as many documents and technical texts as were available. I had access to all documents related to the study, which were located on a secure server using Box

Cloud Storage. I studied implementation models, implementation instruments, PowerPoints, meeting notes and graphics used to guide both the study and process of implementation. Lastly, I conducted a detailed analysis of texts on implementation science concepts and methodologies related to this dissertation, including literature on complex interventions, context, implementation strategies, fidelity, and health equity.

Fieldnotes, memos, transcripts of meetings and interview transcripts were transcribed and analyzed throughout the period of data collection in a process of constant comparison for similarities and differences, in order to refine future collection and organization practices based on emergent themes (Corbin & Strauss 2008). Literature reviews and current theories related to my research questions guided my original analysis, though I looked for theoretically surprising cases and tentative explanations, comparing existing theory against my empirical observations. As themes were produced, I coded pertinent meeting transcripts and memos, establishing a provisional codebook, which I used in subsequent transcripts (Tavory & Timmermans, 2014). Throughout, I engaged in an iterative process where I tested the adequacy of categories against the data (constantly turning between codes and data) and then moved across cases (comparing data to data). In the final year of data-collection, I conducted member checking with the implementation science team (Strauss, 1987). Here, I would present findings and emergent themes from my dissertation to gather how these findings matched up to their reported realities, clarify my understanding of the situation, producing new insights about the data.

Table 3: Organizational Roles

Center for Health Sciences		**SafeLink & "Spoke" Sites****	
Organizational Role	**n**	**Organizational Role**	**n**
Program Officers*	2	SafeLink Director	2
		Clinical Supervisor	1
		SafeLink Service Coordinators	2
Principal Investigator	2	Jail Administrator	4
Implementation Scientist	2	Jail Staff	4
Project Managers	3	CBO directors	3
Research Assistants	3	Practitioners	9

*Program officers were not able to conduct "on the record" interviews due to organizational policy. However, I was provided relevant texts, and meeting recordings to answer specific questions about the development of the request for proposals.

**SafeLink is the training and supervision "hub" and the local counties. The jails and community-based organization (CBO) partners serve as the "spoke" sites.

Participation

While ethnographers and interviewers are arguably always "in the data themselves," I was also a participant on the HaRP clinical trial, which meant my involvement was more than an observer and analyst, but also supporting the project as a research assistant (Small & Calarco, 2022, p. 12). So, layered on top of this knowledge making project (HaRP) was my own engagement in the clinical trial, producing knowledge about and alongside my interlocutors. As a participant and research assistant on the HaRP trial, I helped to develop study instruments including surveys, interview guides, and manuals of procedure. I also conducted interviews with stakeholders, helped with data analysis, and participated on teams that wrote-up implementation

findings. This mode of participation allowed me to observe and participate in the reflexive processes of researchers and implementation actors as they negotiated the ethical, epistemological, and political tensions of their work.

Participation in this mode was critical for obtaining access to the clinical trial, and allowed me to "give back" to my interlocutors for granting me access to study the project. Being involved as a researcher for the Center for Health Sciences, however, also meant that I was formerly considered part of the research project, meaning this may have affected how I was received by community partners and practitioners on the trial, who, as I will show in this dissertation, were at times critical of the researcher's efforts. To mitigate this, I tried to describe my role as an ethnographer of the study, rather than just part of it. Also, most critiques that practitioners voiced were also openly reported to the clinical supervisor and study team for purposes of implementation. Meanwhile, critiques voiced by implementation scientists and the clinical supervisor were also readily relayed to the larger research team.

While I never specifically called myself one, I became known as an implementation scientist by the research team because I too was invested in studying "context." As a burgeoning field, many who come to do "implementation science" do not necessarily recognize themselves as such. Even some of the implementation scientists I worked with on HaRP were not eager to attach the title to themselves, even though it was their defined role on the study. Many who take on the role of implementation scientists on a given trial are often only learning the concepts themselves and may be critical of the very frameworks and theories they are often pressured to deploy. While implementation science tends to, often out of necessity, view context as a physical site where the intervention takes place, I was specifically interested in the research context and practices, which tends to be grayed out or "black boxed" in these types of projects.

Positionality & Para-ethnography

The last thing that was said to me, before I hit record for the first time to launch my dissertation, was "don't do a hack job, Emily." While the researcher who said this to me was, I think, only half serious, I couldn't help but revisit these words throughout my dissertation research and writing. Perhaps because I was "doing an ethnography," I also came with some epistemological and political baggage. While the "paradigm wars" are mostly of the past, there still exists a sense within social service research that those doing fieldwork are there to make categorical, structural critiques of the "positivistic" social sciences (Small & Calarco, 2022). It's true that I did have a more simple and distant critique in the early days of launching this project. I arrived as an ethnographer of HaRP with practice experience as a social worker asked to perform EBPs, I worked for a nonprofit that disseminated one, and engaged in qualitative research that examined the heterogeneous effects of RCTs on human service organizations and service workers. Therefore, I arrived at this dissertation project with my own political critiques of them. And, at the end of the day, my accountabilities and political beliefs were that we should not build up or invest in jails and prisons, positions I shared with and were shared by many of my interlocutors on the HaRP study.

But, the more time I spent in the field, the more this critique blended with empathy, awareness of the institutional pressures that guide any large-scale knowledge making project, and appreciation for the expertise, hard work and knowledge of those I studied and worked with. An inability to suspend judgment, as we are often taught to do in ethnographic methods, could be read as self-indulgent or worse, nihilistically critical of efforts to make a difference in a space where the problems of overdose are imminent and severe. This dissertation attempts to sympathetically examine the limitations of a particular method to develop substance use policies

and practices, many of which the actors themselves grew frustrated by. I try to show the institutional constraints that researchers and implementation actors in this field face, and the types of pressures they experience to perform their roles and duties, rather than focusing on the individual motivations that underlie any given decision.

As Sunder Rajan writes in his book *Multisituated Ethnography*, ethnographers "rarely know things that their interlocutors do not." Instead, he argues, ethnography is about attention, and "attending to things that her interlocutors might attend to differently" (2022, p.17). While I'm not convinced that I *always* attended to things differently, I did try to bring my "lens" to meetings. Through what Marcus and Holmes have coined "Para-ethnography," I attempted to create space for conversations with the implementation science team on tensions and problems that will be outlined in this dissertation. Para-ethnography is a type of ethnographic collaboration in which, rather than seeing our interlocutors as "data" we also understand them as companions in learning. Hopefully, through these conversations we all learned how to approach the challenges we came across or witnessed and approach them in a more radical or transformative way.

I was in the privileged position to be able to study HaRP without having to necessarily think, at least immediately, about how we might resolve the tensions and challenges I identify through science. Instead, I approached the HaRP project in a way that allowed me to understand and "sit with" the tensions that we are otherwise quick to try and resolve through more methods, frameworks, and adaptations. I do understand that it was a luxury to have time and space to reflect. I ask that readers take this dissertation as a moment to think about the "bird's eye view" of the labor it takes, and political processes implicated in making social services into laboratories, or sites for research. Many of the researchers I worked with were frustrated by the

limitations of implementation and intervention science, even if they were actively using the newest models and latest frameworks to tame the complexities of context. Attending to how social work and health services researchers deal with context in their pursuit of science is particularly important given the rather unprecedented social context of HaRP's implementation.

The trial at the center of this dissertation took place during the heart of the pandemic. Overdoses across the country were on the rise due to increased isolation and stress, changes in patterns of drug use and purity, and a loss of formal and informal support (Ghose, Forati & Mantsch, 2022). Meanwhile, social movements to decarcerate and end police brutality against Black life gained more durability and power, and debates were renewed about the importance of the welfare state. These issues were top of mind for those in the social sciences, social work, and other helping professions, and became a frequent topic of discussion during implementation meetings and amongst study actors.

It was common for us to stray into conversations on these topics, though often only in a conceptual register, not only because we had time due to COVID related delays, but we had also grown to rely on each other to make sense of the social effects of the pandemic, which despite its horrific effects, also brought about new horizons for social change. Clinicians and researchers alike were frustrated and inspired to do more than the bounded nature of the experimental apparatus normally allowed us to. We started to ask ourselves: What should constitute the intervention under experimentation, and what type of public impact is being imagined and enacted? How do the institutional constraints of the project affect our ability to advocate and pursue social change in a way that is meaningful? It was these types of questions that seemed generative of important insights rather than a blanket critique of the research endeavor itself.

While these conversations originally occurred organically during work group meetings, we started to make time for intentional conversations that centered the conditions under which we were working, and the challenges of fitting harm reduction and other modalities into the scientific paradigm of the RCT. In some ways, these conversations provided the time and space to reflect during the otherwise rapid mode of knowledge production that characterizes intervention research. The focus is so often on the next publication or deliverable in an effort to appease funders or meet tenure targets. That said, productive conversations took place about the larger sociological questions about the role of experiments in the context of welfare state retrenchment, the politics of knowledge production and the extraction required, and the very premises of an intervention's scalability. Each of the chapters of my dissertation deals with topics that came out of these conversations: the-depoliticizing effects of the RCT, the crisis of transferability, and the existential question of what the experiment was meant to produce, as well as the ways in which the experimental apparatus reshaped social service provision. So, I have my interlocutors to thank too.

Chapter Summaries

Chapter One: Evidence-based Optimism and Activism: Competing Ideas of Social Change

Chapter one describes the divergent political visions represented by trial actors and how they conceptualized the best way to prevent overdose via the HaRP trial. I show how trial actors, from practitioners to PIs, maintained different ideas for how broad, and under what timeline, social change should occur stemming from different professional, disciplinary and activist commitments. Principal investigators on the trial expressed their desire to pursue objectivity and to produce research that would establish generalizable, politically neutral findings for policymakers. Meanwhile, practitioners and harm reductionists on the clinical trial prioritized

advocacy and other activist strategies to reduce overdose deaths in their own communities, including strategies that could involve direct confrontation with the carceral systems that the research team worked hard to win over. Practitioners' concerns were echoed by implementation scientists on the trial, who saw it as their role to deal with the contexts that may frustrate or facilitate the implementation of evidence-based programs like HaRP. Ultimately, the scale at which practitioners and implementation scientists could intervene was circumscribed by PIs to ensure objectivity and maintain model integrity.

Chapter Two: Surviving the Trial: Negotiating "Real world" & Experimental Contexts

Chapter two examines the practices involved in negotiating "real world" concerns like feasibility and fit, with the experimental concerns of stability, standardization and protecting "the model." Importantly, the more control that was imposed on the model, the less "responsive" it became to the needs expressed by trial participants. That said, I show how everyday decisions regarding the model tended to prioritize completing the trial rather than considering HaRP's long-term future. Practitioners and implementation scientists were dismayed by the widespread lack of investment in community-based social services, and the relative disparity between HaRP trial resources as compared to the "real world" of social service delivery. Both researchers and practitioners began to view the experiment and trial resources as an intervention itself: trial resources were allocated not only to establish a stabilized intervention, but also to address the immediate needs of trial participants (e.g., through transportation or material support). Secondly, the resolutely unfeasible nature of some of HaRP's elements represented a political intervention highlighting the underfunded nature of peer work, and other forms of community-based care.

Chapter 3: On Being Faithful: The Politics of Intervention Fidelity

Chapter three describes researchers' efforts at building and implementing the fidelity

protocol, as well as practitioners' experience with its implementation. Despite aspirations from all levels of the research team toward flexibility and ensuring the protocol remained contextually relevant, the practitioners' sense of how to best build a therapeutic relationship with study participants was compromised by the carceral context, the exigencies of scientific inquiry, including treatment goal and recruitment targets, both of which were mediated by the fidelity system itself. This emphasis on flexibility, I argue, might very well obscure fundamental tensions in social service provision, particularly as it relates to self-determination, coercion, and service engagement, made more acute in the context of a clinical trial which took place within carceral systems. Finally, I argue that the fidelity protocol, and the evidence-based practices it promoted, did more than attempt to stabilize practitioners' practices, it also helped to assuage their concerns about how their relational practices were being fundamentally reconfigured–and at times compromised–by the clinical trial protocol.

Chapter 1: Evidence-based Optimism vs Activism

The primary outcome of the HaRP clinical trial was to connect participants to medications for opioid use disorder (MOUD), including methadone and buprenorphine. And, while the primary target for the HaRP trial was engagement with MOUD providers, the broader goal of the trial was to prevent opioid overdose among those recently released from jail. The diverse set of professional actors involved in the trial, however, did not share a cohesive vision for how best to achieve this goal. Harm reductionists, peer workers, and various types of academic researchers (e.g., public health, epidemiology, implementation science) were working "together" and in "collaboration," yet they aspired to employ different, and sometimes competing, strategies to prevent overdose. While some wished for immediate wholesale change—for example, rejecting the criminalization of drug use in the first place—others had their eye toward building evidence for future policy reform. In other words, there were those who understood the trial as a technical effort to produce evidence, and those who understood it as a mechanism for advocacy and immediate, broader social change. This chapter pursues the question: How are the demands of scientific inquiry negotiated with activist commitments during a socio-behavioral clinical trial?

Harm Reduction and the Criminal-legal System

As described in the introduction to this dissertation, the HaRP clinical trial is one study in a broader portfolio of federally funded projects which represented an unprecedented social scientific investment in both the carceral system as well as in building evidence for harm reduction. Historically, harm reduction was ideologically resisted by federal agencies as well as the academic-research community due to its seemingly permissive view on drugs as well as its activist orientation that pit it against the very scientific and medical establishments tasked with

building putatively credible evidence. Meanwhile, jails and prisons, tend to be considered a fundamental source of harm by harm reductionists, and, within the health services research community, tend to be described as reluctant adopters of medication for opioid use disorder (MOUD) and other evidence-based practices. For both health service researchers and harm reductionists then, a source of conflict derives from the carceral systems' tendency to criminalize drug use and therefore punish those who use them. That said, the HaRP project brought together not just a diverse set of institutions, but a set of actors whose professional commitments have historically stood in tension: from grassroots harm reduction activists to carceral stakeholders and health services researchers.

Science and Activism

There is an extensive tradition of social science scholarship that examines the relationship between science and activism. Activists have long adopted a confrontational position toward standard biomedical paradigms of medical knowledge (Brown et al., 2004). Historically, they have done so to center the ways in which structural inequalities and the uneven distribution of social power is responsible for the "fundamental cause" of disease (Brown & Fee, 2014; Phelan et al., 2010). Liberatory harm reduction organizations, for example, were developed by peer-led grassroots activists amidst the relative state abandonment of drug users. They concentrated their efforts on challenging the role of law enforcement on policing people who use drugs, as well as counteracting the detrimental effects of sentencing disparities and mass incarceration on Black communities. As part of mutual aid syringe exchange groups, early harm reduction activists prioritized securing housing and healthcare for drug users, helping to address root factors that made overdose more likely (Des Jarlais, 2017). As these activist movements emerged, social and health services were increasingly adopting an "evidence-based" approach (Cambell, 2020).

Social movement actors have tended to critique randomized controlled trials (RCTs), not only on ethical grounds, but also on political and ideological ones. Political scientists as well as others have argued that evidence-based policy, which usually take RCTs as its 'gold standard,' might constrict political space to pose questions about the carceral state on social justice or human rights grounds (Gottschalk, 2016). A number of social scientists across disciplines have argued that RCTs have the tendency to turn complex and "wicked" social problems into technical ones, making them more amenable for technocratic solutions (Campbell, 2020; Adams, 2016). In doing so, they argue that RCTs fail to account for the social and infrastructural conditions that actually determine one's health (Biehl & Petryna, 2016; Montgomery, 2017).

We also know from social movement scholarship that activists have worked successfully with scientists to transform knowledge production itself. Historically, grassroots activists have been obligated to build an alternative basis of expertise to get their claims onto the agenda of governmental, philanthropic, and community-based organizations (Epstein, 1996). This has often meant partnering with researchers and medical providers to show proof for their claims (Epstein, 1996). Through what they term "evidence-based activism" Rabeharisoa et al. (2014) show how activist groups, rather than adopting a purely confrontational stance toward research institutions and the medical world instead become "reformers," leveraging their experiences to destabilize existing understandings of social problems and health conditions to make research more relevant to service users. Critically, activists have also intervened on questions of method, particularly the exigencies of "purity" in clinical research, arguing for experimental methods that are more pragmatic, contextually relevant and immediate in their effects (Epstein, 1995). Often, however, these relationships are marked by ambivalence; neither researchers nor activists may get precisely what they want from the other in the collaboration. Researchers face risks when

working with activists. They may be persuaded to pursue research that is low in academic relevance because, instead of pursuing innovative topics in the field, they are pulled toward the pragmatic demands of activists (Hess, 2016).

We know less, however, about the interactive processes that take place when social-movement based interventions, like harm reduction, become the subject of a clinical trial. More generally, few studies have examined how fields and social movements that may challenge specific policy decisions might operate alongside these transformations (e.g., evidence-based policy), and with what scientific and practical consequences. As argued by Campbell (2020), RCTs have tended to foreclose alliances with social movements because the "terms of engagement are closed to question," these terms might include who frames the research question, as well as the types of evidence and forms of intervening that are considered credible (Campbell, 2020, p. 46). The following chapter shows how, even within the context of a clinical trial, these questions are still very much open to debate and contestation.

This chapter adds to scholarly conversations about the relationship between activism and research in two ways. My work sheds light on how the commitments of advocacy and science are negotiated as harm reduction-based interventions have become more acceptable to federal grant-making agencies, and therefore increasingly embedded in socio-behavioral clinical trials like HaRP. Within this trial, the harm reduction-based intervention must be subject to the experimental demands of stabilization, standardization, and objectivity but must *also* be attuned to the complex lives of trial participants and the carceral and medical contexts in which they are embedded. Second, I address emerging challenges as clinical trialists attempt to manage "context" via the burgeoning field of implementation science, including how researchers toggle the activist concerns of practitioners with the scientific commitments of the clinical trial. While

implementation science has emerged to ostensibly address issues of context in clinical trials, we know little about how they negotiate the competing commitments of social service delivery with the scientific commitments of the clinical trial.

In this Chapter, I argue that the imperatives of the clinical trial, and the institutional obligations to which it was tethered, compromised efforts at advocacy, and practitioners' ideas for how to reduce overdose deaths most effectively and immediately. This is because, I argue, trial actors operated within different temporal and scalar logics. That is, they had different ideas for how broad, and under what timeline, social change should occur stemming from different professional, disciplinary or activist commitments. I will show how principal investigators on the trial expressed their desire to pursue objectivity and to produce research that would establish generalizable findings for policymakers. Evidence therefore had to be produced in a way that was perceived as politically neutral, isolated from moral and political values, including concerns about social justice brought to the fore by practitioners. The trial, for PIs, was about answering a delimited set of technical research questions that could be used to inform the future decisions of policymakers. At the same time, researchers, and PIs in particular, demonstrated their ideological commitment to the criminal-legal system to gain buy-in and accomplish the trial.

Meanwhile, practitioners and harm reductionists on the clinical trial prioritized advocacy and other activist strategies to reduce overdose deaths in their own communities, including strategies that meant direct confrontation with the carceral system that the research team had worked so hard to win over. Practitioners' concerns were echoed by implementation scientists on the trial, who saw it as their role to deal with the contexts that may frustrate or facilitate the implementation of evidence-based programs like HaRP. Extending the scale at which the field has historically operated, implementation science, and the implementation researchers I worked

with, called for a more explicit consideration of the political, social, and economic conditions (e.g., structural racism, wealth, and income inequality) that lead to inequitable outcomes. Yet, these commitments to advocacy and transforming carceral contexts, were deferred by the academic objectives of producing science. Critically, advocacy would also threaten not only the perceived scientific objectivity of the trial but also the fragile relationships forged over time with stakeholders and administrators within the carceral system who are the very arbiters of access for the research team.

In what follows, I begin with conversations I had with two primary principal investigators on the study, Robert and Trevor, where they described the institutional pressures they faced to pursue objectivity as well as maintain buy-in with fickle criminal-legal stakeholders during HaRP's roll out. Next, I will discuss some of the inherent difficulties of fielding a social intervention in the criminal-legal system, and how this inspired various implementation actors toward advocacy against the criminal-legal system. I will then move on to discuss how implementation science tends to deal with structural problems, and the complexities of their role as both a science and a field invested in social context. Lastly, I describe how this aspired advocacy was deferred out of concern for maintaining objectivity and sustaining relationships with criminal-legal stakeholders.

Gaining buy-in: Objectivity and Evidence-based Optimism

In early 2022, with the clinical trial underway for almost a year, I interviewed the principal investigators of the study, Trevor and Robert, separately to learn more about how the clinical trial was set up to reduce opioid related mortality and how they collaborated with stakeholders during this difficult process of reform. Both emphasized the bipartisan nature of opioids, and relatedly, the extensive federal funding aimed at targeting them. For Robert, as he

saw it, HaRP was one of the "least political" research projects that he had worked on. By this he meant that most stakeholders he worked with did not take too much convincing that reducing overdose was a vital goal, the challenge for researchers, in terms of influencing stakeholders, was how we might go about doing it.

Robert was keenly interested in discussing the centrality of relationships with stakeholders during a clinical trial, and how they are best established through strategies of political neutrality and objectivity. After all, built into the project were mandates for each trial to forge transdisciplinary collaborations that could create timely and translatable evidence for carceral institutions and the web of services that surround them. That said, the PIs were incentivized, if not mandated to build and sustain such partnerships as outlined by the grant guidelines. Importantly, trials like HaRP rely on state institutions for much of their data, particularly as efforts are made for trials to be more pragmatic. Leveraging administrative data helps to relieve the burden of data-collection and allows the research processes to be better integrated into service as usual. While having access to such data is critical for a project like HaRP, it is not guaranteed. Gaining buy-in with community partners, particularly carceral ones, is a difficult process, yet central to accomplishing a clinical trial like HaRP. Gaining access to research sites, and the attending administrative data is likely why Robert stressed the centrality of relationships. As he stated, obtaining data from administrators is one of the more "political" aspects of a given research project.

As Robert explained, inviting researchers to their jails could be risky, especially for sheriffs who are usually elected officials and rely on the public's acceptance of how they manage not just crime, but also safety in the jail. Moreover, jail administrators are already in a constant struggle to meet the pressing need for health and social services among people who are

incarcerated, often with very limited local and state resources (Lara-Millan, 2021). Ultimately, therefore, American jails are acutely concerned with preventing the legal liability and associated public relations crises that stem from inadequate care, a strong motivator for jails to invest in medical services and collaborate with health services programming (Lara-Millan, 2021).[1] That said, to overcome skepticism, Robert felt they would have to frame HaRP in a way that held salience for carceral stakeholders. Deciding to partner with HaRP could also bring additional resources to the jail in the form of material resources, personnel, and new technologies, a fact that will be attended to more in depth in Chapter Two. Indeed, the HaRP intervention was frequently referred to as a form of additional "help" for jail partners rather than a clinical trial where the results were unknown. As a case management intervention that supported individuals as they are discharged, it would make sense that jail administrators would expect HaRP to reduce the recidivism rate, a problem that jail administrators tend to be keenly interested in. So, the research project could serve as a relatively free addition of resources to support discharge, potentially reducing recidivism while offering supplementary services for people with opioid use disorder.

It was also *how* the researchers framed the scientific process of the HaRP trial that would help them to secure buy-in with jail administrators. In the early stages of the project, researchers held buy-in meetings with stakeholders. During these meetings, a significant portion of the research team would attend to deliver a formal PowerPoint presentation over Zoom, discussing the aims and trial design of the HaRP trial. Robert describes how the conversations to gain buy-

[1] Jails are in a constant struggle to avoid litigation for wrongful death, medical malpractice or intentional infliction of emotional distress, but they are frequently unsuccessful. The American Jail Association, in a 2021 issue of their publication *American Jails*, provides guidance on how to carefully avoid such jail-related liabilities. Campbell, 2019 states: "Jails are guided by constitutional mandates and case law, and thus, can be a focus for litigation for liability lawsuits and civil rights claims. Litigation is *costly* and *time-consuming*, but can be reduced with well-designed policies, training, accountability, and diligent operational oversight."

in proceed, and how the RCT provides unique credibility in enrolling organizations that serve as community partners on clinical trials this way:

> The RCT gives us credibility with some of these stakeholders where they say, "we know that their preferences [scientists] didn't pollute that." The fact that I can explain in one PowerPoint slide, "here's the treatment group, here's the control group and here's the difference," that transparency and clarity is incredibly valuable. [...] I think if you come in and you say, you know, "there's a lot of issues with this population. Here's what we randomized, here's what we did. It's rigorous. And we found that they were arrested 40%, less or whatever." People will believe that. Now, whether they believe it can be applied in their context is another question.

The clinical trial also provides, at least in Robert's view, objective evidence that could convince jail stakeholders of their research findings, and importantly, produce evidence not clouded by their moral and political views on jails and the activities that take place there. While not the primary outcome of the clinical trial, Robert highlights recidivism, a salient metric for jails looking to reduce cost. Reference to the technical language of the scientific method, he suggests, also helps to convince stakeholders that their efforts to produce evidence were not "polluted" by the perceived political preferences of researchers. Randomization, itself a mechanized process, reduces various sources of biases in ways that help to distance researchers from the findings they produce.

According to Robert, the average carceral stakeholder also tends to perceive researchers as more politically progressive and therefore more likely to be critical of the punitive nature of jails. The structured nature of the clinical trial, with its predetermined set of outcomes and research questions, quells concerns about political bias. As Robert explains, "outcomes are pre-registered" through databases like clinicaltrials.gov, which assures stakeholders that the research questions and outcomes outlined in the beginning of a trial won't stray from the original trial design once implemented. There's a sense that, as Robert put it, you "can't screw around with the

data" and "run eight regressions" to find a significant outcome, one that could potentially make stakeholders "look bad." Trevor, on the other hand, focused on the institutional demands for political neutrality. He relayed to me: "I think one of the tenets we are taught when we do federal research is that we're not supposed to be lobbying, right? we're supposed to be collecting and providing data to inform policy decisions." In other words, from Trevor's understanding, certain types of social intervention could be restricted by federal grant makers.

It wasn't just jail administrators or grantmakers who might desire objectivity. Robert, for example, also had his eye toward how the public, policymakers and other decision-makers would evaluate the RCT evidence once produced. After all, the way evidence is framed dictates the possibilities of who and when it can be taken up. In his words:

> The other dilemma is that we're trying to make a political case for humane policies that we ourselves are motivated by on the basis of social justice and core values of public health. But we also do not believe that those values are necessarily the most effective sales pitch in the political market. Right?

Evidence was produced for future decision-makers, and thus needed to be produced with them in mind. Robert's sense was that policymakers were acutely concerned, for example, with reducing cost, curbing crime, and other outcomes and modes of evaluation that might attract political actors through promises of cost savings and a reduced community burden. While Robert himself might be motivated by appeals to social justice and human rights, he argued that cost and other more putatively objective indicators might sway individuals who otherwise have a "tenuous" commitment to people who use drugs.

Relatedly, Robert argues that the causal mechanisms of social change tested through a clinical trial like HaRP are more convincing to policymakers and more feasible than other types of social change, producing what he calls "evidence-based optimism." Interventions like HaRP, with actionable, short-term outcomes provide the possibility of "genuinely mov[ing] the needle,"

he states. This is compared to the "profound" structural issues that seem insurmountable. This optimism, he argues, serves as an antidote to what he understands as "scaling pessimism" from the progressive left, who he describes as suspicious of criminal-legal reform if it does not include a plan for structural investment, critique, or abolition. Robert describes how a typical leftist might critique the RCT:

> The structural issues that are so profound, and that's really what's driving things. And you're putting a Band-Aid on the problem with your incremental solution. And I'll (the leftist) be interested when you come back with a serious attack on racism or whatever, and that might be something I totally agree with, but I'm like, that's the 2050 plan. You know, and, and the RCT shows people "no, actually we have not found the cure for racism, but we still got fewer kids arrested."

Robert understood clinical trials as a more pragmatic strategy to create social change as well as more feasible and immediate than social-structural change, "the 2050 plan," as he put it. For Robert, while the clinical trial is not a "cure for racism," it could generate change more swiftly, even if incrementally. For Robert, while this type of change might be gradual it also is appealing both to the administrators fielding the intervention, and policymakers looking for more immediate, technical solutions.

The "causal story," endemic to RCTs—ones that individualize social problems—as least for Robert, tends to be more palatable to organizations and funders asked to efficiently explain why they should pursue a particular policy. As he relayed to me, almost all stories in economics tend to be about individual behavior. In our conversation, he provided one figurative example: "I got the guy into treatment, so he had less need for money for his drug habit. So, he stole less." The underlying theory is that it is possible to isolate the effect of a program from complicating factors through randomization and by controlling for perceived extraneous variables. Importantly, certain factors including structural racism and over-policing often trouble one's ability to "make good," yet tend to not be accounted for when recidivism is deployed as an

outcome measure.[2] This serves to elide the role of the criminal-legal system in causal explanations of recidivism while inflating the role of individual level 'deviant' behaviors. The very context that is elided here, as I will later show, becomes a point of contention for harm reductionists and implementation scientists. That said, the HaRP trial, for Robert, exemplified a specific vision for how social change might occur: change should be feasible to promote "evidence-based optimism," objectively pursued in ways that do not provoke carceral stakeholders, and with an easily digestible causal pathway that can scientifically control for context.

The experimental method, for researchers like Robert, provides what Porter (1996) might call 'mechanical objectivity', the removal of personal bias by appealing to formal rules. This is important for two related reasons. Clinical trials tend to be oriented toward producing evidence for the future, and specifically to support the decision-making processes of policymakers (Donovan, 2018). As such, researchers are pressured to produce evidence that appears objective; ostensibly, this creates less friction for policymakers who are meant to take up the evidence once it's produced (Hess, 2016). Often more vulnerable to accusations of political bias, applied social scientists tend to pursue objectivity for its ability to lend an increased aura of credibility, or at least to quell accusations of self-interest or political bias (Breslau, 1998; Porter, 1996). Meanwhile, narrowly defined, short-term outcomes that follow an economic "style of reasoning" tend to command more attention than outcomes that are long-term (Jasanoff, 2003; Berman; 2022). One could argue that clinical trials even provide "epistemic doxa," or a naturalized set of

[2] Prins & Reich (2021) offered a meta-review and critical analysis of evidence-based rehabilitative programs in the criminal-legal system. Recidivism was rarely operationalized despite it capturing new offenses *and* technical violations, the majority of which do not arise from illegal behavior. Often, technical violations and recidivism are a result of discretionary practices on the part of the probation or parole officers. Altogether, this served to elide the role of the criminal-legal system in causal explanations of recidivism while inflating the role of individual level 'deviant' behaviors.

indicators which are considered the criteria or basis for informing social policy (e.g., reducing recidivism) (Hess, 2016). For example, it would seem much more acceptable to evaluate an intervention on the basis that it can reduce cost than on the basis of broader social values like social justice, even if researchers might be motivated by both.

Aligning with a more economic "style of reasoning" also makes sense given broader changes in the welfare state and carceral funding. Organizations across the human services field have become increasingly obliged to frame complex social problems as achievable within the constraints of short funding cycles, and easily operationalized metrics, all of which were conditionally tied to program effectiveness. The rising cost of corrections amid declining state resources has prompted sentencing reformers to focus on increasingly austere or "smart" and "evidence-based" criminal-legal strategies. These include those interventions that not only 'work' in reducing recidivism, but to simultaneously reduce cost when compared to status quo carceral practices. Within frameworks of evidence-based policymaking, researchers tend to study interventions that are more readily lifted from the local particularities of their research context and scaled to new places and with different populations. Indeed, only certain interventions can be tenably reproduced and scaled without apparent attunement to context–those that are more individualized in nature. This approach toward building evidence, however, would place the research project into epistemological and political tension with the burgeoning imperatives of implementation science and the practitioners whose concerns the field is meant to champion.

Aspirations for Advocacy

HaRP began its recruitment in the summer of 2021, yet was stifled by ongoing pandemic related delays. Partner jails were struggling to contain COVID-19 outbreaks, and administrators were rushing to release the incarcerated people they deemed less of a putative threat to public

safety. In the context of a health crisis and potential journalistic spectacle, jail administrators were reluctant to allow researchers inside the jail to meet with participants for recruitment. At the same time, jails administrators were interested in participating in the trial as it was often perceived as an intervention in itself: it promised discharge support as well as the potential for a reduced recidivism rate, saving money amidst the fiscal crisis faced by many jails (Lara-Millan, 2021). However, strict safety measures needed to be put in place before any research assistants could enter. The resulting paucity of eligible participants expected by trial PIs as well as pandemic-related releases and lockdowns, meant that, early on, the study was persistently behind on meeting its needed recruitment numbers. In 2022, less than 50 participants had been recruited. It would not be until 2023 that the study would reach 100 out of the 1,000 needed to make statistical significance. With recruitment only trickling in and few participants to work with, practitioners and clinical supervisors grew frustrated and eager to begin their efforts to stem opioid overdose deaths around the state.

Importantly, during the heart of the COVID19 pandemic, overdoses across the country were on the rise due to increased isolation and stress, changes in patterns of drug use and purity, and a loss of formal and informal support (Ghose, Forati & Mantsch, 2022). We know it to have accelerated overdose deaths, particularly among Black, Latine and Native American people, as mediated by increased exposure to policing and incarceration (Hansen et al., 2022). As mentioned in the introduction to this dissertation, social movements to decarcerate and end police brutality against Black life gained more durability and power, and debates were renewed about the strength and importance of the welfare state. These issues were top of mind during implementation science meetings. It seemed to me, clinicians and researchers alike were frustrated and inspired to do more than the bounded nature of the experimental apparatus

normally allowed us to. Likewise, they aspired to make change immediately, and at scales inconvenient to the clinical trial design.

Out of all the HaRP trial actors, practitioners were the ones who worked most closely with jail staff. The practitioners trained to deliver the HaRP intervention were often hired from SafeLink, or local harm reduction and community-based organizations. Practitioners needed to enter the jail weekly to meet participants, build relationships with administrators and service providers, and troubleshoot with carceral stakeholders to find ways to get participants into treatment once released. Peers, within the HaRP trial thus found themselves in a difficult position —behavior previously criminalized was now valued as "expertise," and they had to navigate this tension throughout the trial. Taking on a newfound professionalized role that required engaging with jail staff was an emotional burden, but their own personal experiences also provoked a sense of duty to advocate for the abysmal conditions of participants being jailed. This tension will be more thoroughly outlined in Chapter Three.

Many harm reductionists hired for the trial experienced incarceration themselves and had families still caught in the criminal-legal system. Practitioners in the HaRP trial tended to oppose the criminalization of drug use, understanding policing and jail as themselves "hazards to public health" (Levenson et al., 2023 p. 2). The harm reduction movement has actively pursued, and in many cases successfully so, the decriminalization of drug use (e.g., decriminalization of cannabis, Good Samaritan laws).[3] Yet in their efforts to prevent overdose by serving on the HaRP trial, practitioners found themselves having to work *alongside* carceral stakeholders, often

[3] While ethnographies of carceral drug treatment have shown that clinical knowledge of substance use have been mobilized for crime control and prevention, harm reduction has tended to reject this tendency (Garriott, 2011; McKim, 2017).

in their very own county jails. These were sometimes the very county jails they themselves had been incarcerated in.

Heather, a peer on the intervention team, grew up in Sherwood County and had extensive experience dealing with the carceral system. She was recruited to be part of the study because she worked with a local harm reduction organization in Sherwood County called Love, Inc. After she lost her job during the initial wave of COVID in 2020, she volunteered for Meals on Wheels and was then hired part-time through a state grant. Longtime friends with the executive director, Heather had been recruited because of, as he put it, her "way with people." We talked together, early in the trial, about working with the jail system in Sherwood, a place she has intimate knowledge of, having had family in and out of jail since childhood. She spoke to me about the collateral consequences of incarceration and the punitive practices of the Sherwood County jail and police drug task force, explaining how the cost of incarceration propels many into debt:

> Our average is a million-dollar bond. So, $100,000 is what it costs to get somebody out for any minor offense. [...] If it's a DUI, that affects getting your driver's license back, yeah, so you can't let that go to collections. But it's almost like a medical bill, if you get so many medical bills, at some point, you're just like, I'm not even going to try and get this up. Right. At some point, I'll just file bankruptcy. I know that's kind of the impression of a lot of the folks that don't have money and get big fines, and like, I'm never gonna be able to pay this, but it just goes to collections.

As Heather describes it, the arrest and incarceration associated with drug possession can catapult someone into poverty by preventing job attainment, loan qualification, or rejection from public housing (Wang, 2018). Once people are booked into Sherwood County, people remain there, on average, for about a year, even though they are often charged with minor drug possession for amounts that reflect that they are simply "supporting their own use," as one research assistant on the study put it. Incarcerated individuals, and even jail administrators, rarely have clarity on release dates. These decisions are largely determined on the whims of the county court and local

prosecutors. It's common for participants to have minimal contact with their legal representation and be deprived of basic information regarding their own case, even including the dates of their next court hearings. All this is to say that the context of service provision was rife with problems bigger than what practitioners could feasibly handle in their day-to-day practice, and they grew tired of it. So, advocacy in these spaces was a way to remain proactive while also opening the space to potentially address important institutional barriers to implementation of the HaRP intervention. In what follows, I provide two extended examples of when practitioners and implementation scientists identified points of advocacy during the clinical trial, and, later, how these efforts were hindered by the commitments of scientific neutrality.

Example 1: Criminalization of Naloxone

During implementation science meetings, despite whatever was prescribed on the loosely knit agenda, our conversations would center on the challenges practitioners faced in their delivery of the HaRP intervention, and subsequently, for the sake of the science of implementation research, how we as researchers planned to not just document these obstacles but also find ways to address them. Practitioners are also asked to keep researchers abreast about what is happening 'on-the-ground' of implementation. Implementation scientists must intimately get to know the practitioners delivering the intervention, as well as their concerns, worries and frustrations. They must at least try to be deeply among practitioners, and make efforts to address contextual issues that arise, which may make the burdens of practitioners easier to carry. While some of the challenges implementation actors faced were simple enough to work around—deciding to conduct intervention visits over zoom versus in person, for example—others were more structural, more obstinate, and difficult to transform within the confines of the clinical trial.

In the midst of discussing adaptations and the tedious details of how we would go about documenting each change—a central preoccupation of many implementation science meetings—Amy, the clinical supervisor, interrupted the flow of the meeting in an exasperated tone, calling for advocacy on the part of the study team:

> Heather [a peer] told me about some stories from Sherwood County yesterday, and it's really appalling. So, the first 911 call to this house, it was these parents' daughter, who let somebody in their house overnight, who, you know, just used [an opioid] in their house, and she overdosed. They were sleeping, the daughter came yelling, asking for Naloxone. The parents called 911, and the daughter saved the young woman's life with Naloxone. And when the cops arrived, they said things like, "Why do you have Naloxone in the house? Are you encouraging people to use here?" yada, yada, yada, and [the police] also made a nuisance threat about their landlord, as if one call qualifies as a nuisance? And what they didn't know is that this, these parents lost their child, their other child to overdose. That's why they have Naloxone in the house.

The weight of the problem and that Naloxone might be stigmatized and penalized by police in Sherwood warranted, for her, straying our attention away from the seemingly endless demands of documentation demanded by the trial. While many police around the country are required to carry Naloxone in case they encounter an overdose event, it is common for police to harbor negative views toward Naloxone, understanding it as enabling drug use (Murphy & Russell, 2020). The nuisance threat could therefore very well deter a Sherwood resident from calling an ambulance in the future, putting them at risk of a future overdose. In fact, fear of losing housing, arrest or involvement in the child welfare system are the most frequently cited reasons that people do not call 911 during an overdose-related crisis (Wagner et al., 2019; van der Meulen et al., 2021). Police acceptance, therefore, is essential for promoting Naloxone's use and preventing overdose.

While Amy's more explicit critique was aimed at the practices of Sherwood police, what was also implicit was her eagerness to do more than what she felt the trial was achieving at that

moment– a seemingly endless loop of documentation, maintaining a paper trail of intervention activities, and then brainstorming best practices for maintaining and sharing said documentation. Raising these types of concerns was not out of the ordinary for Amy. In interviews, she lamented to me the glacial speed of academic production and how its forms of capital exchange tend to be in service of academic conferences and journal publications rather than the immediate struggles she encounters in her daily work.[4] During our meetings, most in the room tended to share this discomfort. The team asked Helen, the project manager, what type of appetite the leadership might have for taking this on as an advocacy issue. Helen, as part of her role, was charged with interpreting the presumed desires of PIs to the other research groups, as well as providing weekly updates to the PIs about any issues that arose throughout the study. Helen replied in a more dispirited tone than she usually carried: "I'm not sure." We would have to discuss this in an "all-team" research meeting, where most important decisions were finalized.

A couple of weeks after the implementation meeting described above, Amy and the rest of the implementation team met for our biweekly all-team meeting with the P.I.s, research staff and project managers to update the P.I. 's on our progress. Amy began to detail how the Sherwood police officers were punishing people who carried Naloxone (e.g., via nuisance fines), as she told the implementation team a week prior, and presented a list of concerns she had for how the carceral system in Sherwood had the capacity to cause *harm* to study participants: "Some examples include how they're charging people, and how people are ending up in jail for long periods of time for really tiny amounts of drugs, they're still being given [a charge of] delivery when they have like half a gram. It's preventing people from getting out, having their freedom, and accessing MOUD [medication for opioid use disorder]." Interestingly, Amy tethers

[4] During meetings, for example, Amy and others had the sense that productivity was measured in publications rather than the practical concerns of implementation.

these carceral practices to their potential negative effects on the outcomes of the clinical trial. My sense was that Amy was trying to enroll the P.I.'s into intervening by framing her concern in a way that aligned with what she perceived to be *their* interests. That is, she framed the problem as one that could very well affect one of the primary outcomes of the clinical trial, connecting to MOUD. Indeed, if people aren't released from jail, they won't be able to access treatment apart from what is provided in the jail, unless they are "lucky enough" to get into treatment court which tends to have strict eligibility criteria often favoring those deemed most likely to be successful. Amy's tone, however, one of escalating frustration, betrayed a sense that perhaps her concerns were less on meeting HaRP study metrics and more about the injustices she witnesses everyday while working alongside the HaRP practitioners in case consultations and clinical supervision.

Example 2: Chirping

During an all-team research meeting, just a month after Amy voiced her concerns to the implementation team about Sherwood's rejection of Naloxone, Helen was relaying updates on the clinical trial's progress. Her primary audience during the all-team meetings are the principal investigators, but clinical managers, research analysts, project managers and implementation scientists are also in attendance. While the PIs possess most of the decision-making power, they were relatively removed from the day-to-day operations of the study since most clinical trial jail sites are hundreds of miles away from their urban research institutions where they work. [5]

Helen began the all-team meeting with reports on how many participants had been recruited into the study during the past two weeks. There were cheers and excitement from

[5] This distance is one side effect of pragmatic clinical trials, where studies are meant to be implemented in putatively "normal" or "real world" settings as well as implemented by people with "typical" experience, as I will discuss more in chapter two (Ford & Norrie, 2016).

meeting attendees at the hint of increased enrollment since meetings prior. She relayed bureaucratic updates on hiring, turnover, and progress on establishing partner relationships, then turned to reports from the field. She described how some jails that served as sites of recruitment lacked enough people with opioid use disorder, a primary eligibility criterion. Participants also remained in jail longer than researchers had expected, which meant people weren't leaving jail even as they were enrolled into the study. Because participants often remained incarcerated after initial recruitment into the study, practitioners needed to maintain contact with participants to continue building a relationship. Practitioners also needed to gather information on their release date to efficiently connect them to services upon their release. But while initial Zoom enrollment in the study was free for participants, texting and emails were managed through a third-party, for-profit company (Combined Public Communications) that charged inmates for every message (called "Chirps") they sent and received. This was one of the more punitive ways jails have dealt with budget shortfalls, charging inmates for communication and other daily necessities (Wang, 2018). So, although enrollments into the study had increased, practitioners were not able to keep in contact with them in the way originally outlined by the study.

Chirping through the Combined Public Communications could become so expensive for inmates that the PIs felt it could interfere with participant engagement. As such, they were willing to allocate grant dollars toward funds for study participants to Chirp. While the study team wondered aloud whether paying for Chirping was something that would be feasible in "the real world" of service provision due to its high cost, Robert was willing to allocate study dollars with the promise of engaging people who needed treatment while maintaining the study recruitment number. As he stated, "and if we are somewhat ripped off by some stupid Chirping, it is worth it in terms of you know, if over the life of our study, we spent, we wasted $2,000 and

we got five more people that we successfully engaged that we wouldn't have engaged without the chirping, I'd be like: 'I'm happy I spent that money.'" Meanwhile, Helen the project manager, encouraged Robert to write an op-ed and advocate against the practice of charging people who are incarcerated to contact the outside world. Yet, as I will describe later in this chapter, Robert was not too eager to denounce the very institutions he relied upon to field the HaRP trial.

The Science of Advocacy: Implementation Research

The concerns of practitioners like Amy and even research managers like Helen are usually brought to the attention of and mediated by implementation scientists before they are reported to the larger research team. Indeed, contextual concerns like those described earlier, must be brought to those responsible for describing and intervening upon context in a clinical trial—the implementation scientists. Implementation researchers become the sort of "spokesperson" for context in a clinical trial like this one, particularly as context tends to be analytically separated from, and conceptualized as outside of, the intervention at hand. During meetings, the implementation science team reflected on how we might respond when confronted with the intransigent "structural barriers" that seemed endemic to the carceral system, including over-policing as well as financial extraction in the form of fines and fees. Again, working within the carceral system felt antithetical to the desires of many of the practitioners in the study. Rather than a place in which to invest, the jails we worked with seemed to, in many cases, perpetuate harm.

Carl, an implementation scientist on the HaRP trial, acknowledged Amy's concerns about the Sherwood police, and replied to the team that this was a reminder that, during early round stakeholder interviews, we needed to probe for complicating issues like this one—issues otherwise called "barriers to implementation" (e.g., police acceptance of Naloxone). Toggling

between treating Sherwood police's response to Naloxone as an issue for data collection and a social problem to be resolved, he also urged us to think carefully about the type of messaging researchers might use with police officers if the research team were to advocate or promote Naloxone's use among Sherwood constituents and carceral stakeholders. He asked the group rhetorically, "This is a social justice framework as well, right?" He then continued, "You can help the individual cope within this broken system, but you must also try to address the broken system." In this sense, for Carl, police and the carceral system were driving people to overdose by rejecting lifesaving medications like Naloxone, and he aspired to pursue more systemic change. Even still, it was not clear under what timeline, or through what mechanisms we would, or could, intervene given the constraints of the clinical trial. This, I argue, is a result of two factors that I will develop in the following sections: 1) The implementation science team is embedded in a clinical trial whereby the legitimacy of its evidence hinges on its "objectivity," and 2) implementation science is also an academic and scientific endeavor, which needs to lift data from the particularities of context to create generalizable findings.

Carl was calling for a greater recognition of and attention to the social, political, and economic issues in which implementation issues emerge, though he would shift to focusing on academic outputs like publications and the adaptation of frameworks. Indeed, there was constant pressure to produce academic outputs while doing the labor of implementation. Consider the following statement:

> And a lot of that [implementation science for health equity] is just different frameworks which refer to inequities or structures, *but we've actually just identified a really clear one* [charging people who are incarcerated for calls and texts]. And so, I think, capturing those in papers we're writing and then really arguing for being able to better tailor those models and frameworks, especially for our legal system work, will be important for scaling these kinds of projects up and out."

On the one hand, Carl is invested in health equity and aspires to make lasting change, but here his attention shifts from the immediate problem at hand to tailoring frameworks and writing academic publications. This is understandable. Implementation researchers face a variety of conflicting commitments in navigating how to respond to concerns like those of Helen and Amy. Indeed, the priorities of researchers, even those ostensibly invested in context, do not always squarely align with those of practitioners or activists (Hess, 2016). Implementation science researchers, in some sense, face a set of conflicting commitments perhaps even more complicated. Within a field of research that emerged from the highly localized concerns of clinical practitioners, implementation scientists must make sure to have one ear "to the ground" of implementation. They must intimately get to know the practitioners and service users as well as the organizational and institutional constraints that they face. Then, they must reframe practitioner concerns, however formidable, through the technical analytic of "barriers and facilitators." Political concerns become neatly packaged as external to the intervention or larger research project. This reporting provides the PIs with a sense of what might be "impacting" an intervention's success or failure. It also externalizes blame outside of the research apparatus, the intervention design and in some sense reifies the very idea that interventions are bounded entities that can exist outside of their political context.

Lastly, implementation researchers must focus on how local phenomena can be lifted from the particularity of its context to make generalizable claims both in order to innovate in the field of implementation science as well as to generate novel theories. The imperatives of science means transforming local concerns into issues that could be written about, generalized, and theorized. These conflicting commitments meant that during the HaRP trial, it was unclear when "contextual" problems of evidence-based program implementation become something that

researchers try to resolve versus simply data meant for documentation and theorizing. Carl would often lament struggling to separate the two in our efforts to conduct implementation science. When context is treated both as scientific fact and a political problem to be dealt with, how to proceed can be confusing. These decisions become more challenging still when the "barriers" identified are endemic to the very institutions researchers relied on.

Implementation science has started to align itself with broader health services research movements to reduce health disparities by addressing social determinants like structural racism as well as wealth and income equality (Bailey et al., 2017; Bor et al., 2017). At the same time, the field has become more actively invested in not just studying context but actively transforming it through implementation strategies. One way to mobilize the "advocacy" the implementation science team desired to pursue was through the mechanism of an implementation strategy, defined as "techniques used to enhance the adoption, implementation, and sustainability of a clinical program or practice" (Proctor et al., 2013, p. 2). Implementation strategies were also a way of framing a political act (advocacy) in the neutral terminology of science. The idea being that improving police officer acceptance of Naloxone could also improve the acceptability of the broader HaRP intervention. While the theory of change was a bit underdeveloped, the implementation team would eventually introduce this idea to the PIs for their approval.

Advocacy Deferred

Robert, along with Trevor, the PIs of the trial, had ultimate authority on how we would respond to issues like the police stigmatization and rejection of Naloxone or making public statements about the practice of charging people who are incarcerated to make phone calls. Ostensibly, both affect implementation and intervention outcomes. Robert replied to Amy's account of Sherwood police's rejection of Naloxone: "I go back and forth about how I would

want to respond to that story. On the one hand that's so emblematic of the ridiculous law enforcement response," he conceded, "and I would certainly want to educate our law enforcement partners that it's a good thing that the Naloxone was there; this person still would have used, and they would just now be . . . dead." At the same time, he was concerned that any type of advocacy might damage the research team's relationship with the jails and other carceral stakeholders who he relied upon for trial recruitment: "It's a little bit it's unclear," he said, "how we can be culturally competent with all of our different partners in pursuing those issues." By culturally competent he refers to the content and form by which researchers engage police stakeholders who, as he mentioned earlier, already have their suspicions of academic researchers. Researchers rely on a criminal-legal system, whose buy-in was integral to a clinical trial like HaRP. As researchers considered how they might enact change while also doing so in a way that is most palatable, the conversation began to focus on strategies of health communication. Health communication was a topic on top of mind of both P.I.s and other study leadership on the Zoom call that day who all happened to be fielding other studies on COVID-19 vaccine hesitancy, a major funding source as the public and federal medical research agencies became acutely concerned with rising anti-science rhetoric and anti-vaccine movements. Who would be a trustworthy messenger of these ideas, and how could we enroll police officers by aligning our interests with those of police officers? Advocacy became reinterpreted into the scientific, technical question of how best to communicate these ideas. Interestingly, while Robert felt that reform was considered more pragmatic and perhaps immediate than socio-structural change, in reality it was the technocratic mechanisms of the trial that subdued attempts at pursuing more immediate change. Moreover, framing this tension as an issue of cultural competence was one way to approach barriers to implementation in a way that was seemingly depoliticized, and

would mean respecting the values, and beliefs of carceral stakeholders, even if they came into conflict with the imperatives of the trial and those working hard to implement thee HaRP intervention. That is, while researchers might work hard to ensure their interventions fit with local culture, this still means deciding whose culture matters and whose is elided.

While Carl described to the P.I.s the implementation strategies, the bounds of the experiment started to expand and unravel. “I feel like we do a lot on the periphery,” said Trevor. “And I just see this never-ending scope; it’s too much.” That is, the scale at which we were trying to intervene began to expand and stray from the original purpose of the intervention. Implementation science desired to pursue strategies to address the very contextual problems they felt might help to prevent overdose and improve buy-in of the intervention. The field has made efforts to broaden the scope of its intervention. This makes sense given implementation scientists meant to be deeply attuned to the contextual issues facing practitioners, many of which are structural in nature. Yet, Trevor encouraged the implementation team to stick to “the most relevant context or maybe just one degree away from the intervention, and then stop there.” He was attempting to circumscribe the scale at which the implementation science team could plan to intervene. To institutionalize this labor through an implementation strategy with the goal of supporting intervention acceptability would exceed the confines of the clinical trial.

Conclusion

How do we draw lines around the scope of our intervention? Within a community-based clinical trial like HaRP, there is the sheer range of what constitutes “context,” and in this case, perhaps more importantly, what does not. This chapter shows how context is not preexisting before implementation, but actively contested and produced. One of my primary findings of this chapter is that trial actors operated within different temporal and scalar logics. That is, they had different

ideas for how broad, and under what timeline, social change should occur, which were professionally and disciplinarily motivated. In particular, the scale at which the diverse set of trial actors wanted to intervene at, and the timescale they anticipated for social change, diverged. Indeed, prior research that examines the relationship between researchers and social movement actors have shown that they tend to maintain a relationship of ambivalence, often because of their conflicting goals. Less have shown the interactive processes of these exchanges and their concomitant social effects. In the context of a clinical trial, I show how activism is deferred for the sake of academic pursuits, as well as maintaining institutional partnerships and experimental integrity.

As a trial that relied upon the expertise of peer workers and drug users, practitioners who delivered HaRP brought with them a social movement approach to their work. They did so to extend rights to participants in the trial, and improve access to the lifesaving overdose reversal drug, Naloxone. Activism tends to be sparked by those focused on a "specific and immediate problem for which they seek to gather information to be used in an adversarial political process" (Moore, 2006 p. 305). Practitioners knew, as most trial actors knew, that Naloxone worked. If placed in the right hands, including those of police, it could prevent overdose among people enrolled in the trial. They also knew that a strong network of support would be vital for people as they are released from jail, meaning they wanted to end Chirping and jail-based policies that prevented participants' communication with the outside.

As we saw in this chapter, the horizon of potential avenues for social change was circumscribed by the institutional commitments of academic research, the perceived culture of jail stakeholders and the desires of imagined policy makers ostensibly hoping for politically neutral evidence. For that reason, Robert tended to focus on outcomes that appealed to

policymakers and carceral stakeholders: reducing cost, and reducing crime, even if, as he states, these weren't the outcomes that motivated him and others on the study team.

As we heard from Robert, clinical trials tend to rely on a specific vision for social change: a causal story that occurs at the level of individual behavior and is therefore seemingly less immune to broader political variables. Indeed, while researchers may have political opinions, they must attempt to remain politically neutral in order to sustain their "neutral advice-giving capacity." (Hess, 2016, p. 122). Critical sociologists of the RCT have long questioned the ideal of "objectivity," less have shown the real-time social consequences when researchers pursue such ideals. This includes how objectivity is wielded and differentially pursued in ways that might quell activism and also, conveniently, sustain otherwise potentially contentious partnerships. Importantly, for PIs, practicing objectivity meant not questioning the institutions that were the very source of practitioners' political critique.

The emerging imperative of implementation science to address structural barriers placed implementation researchers, and the practitioners whose concerns they promoted, at odds with efforts toward scientific neutrality proffered by the clinical trial. The field of implementation has started to expand the scope of its own relationship to context, by considering the structural nature of inequities. Even with this recognition, as a field that often finds itself squarely embedded within the experimental paradigm of the clinical trial, efforts toward creating structural change were circumscribed by the demands of scientific neutrality and standardization. As part of the special call that would fund HaRP, jails and prisons were deemed reluctant adopters of medication for opioid use disorder (MOUD) and other evidence-based knowledge, rendering them ripe for implementation analysis. They were thus meant to evaluate and ostensibly improve adoption into criminal-legal settings. This meant that a larger critique about the role of the

criminal-legal system in providing care in the first place, was largely off the table. And, when context becomes something measurable, other important issues, like the scientific enterprise itself, can be relegated to the backstage, or outside the analyst's field of vision, as it did here.

Issues of advocacy brought forth with a sense of urgency became subject to the long-term, iterative nature of academic research. Larger problems were made into smaller, technical ones. Ambitions to end Sherwood police's rejection of Naloxone inspired conversations about academic movements, publishing articles, and the most scientifically backed ways to pursue conversations with carceral stakeholders. Within an RCT, the intervention must be bounded, and the contexts controlled. Any adjustments in the model or context, whether it be the dosage, the precise "ingredients" of the intervention, or the population under study must generally be subjected to further study. Compared to other forms of social change, timelines are also therefore longer and more expensive, even if the ensuing recommendations are considered more short-term and pragmatic. Perhaps compared to other forms of applied social science, implementation science is particularly ripe for this type of iterative practice. As a field that deals with "context," itself mutable, more research is required to adapt and tailor in order for an intervention to be deemed effective in more places. Indeed, it is common knowledge, and a refrain within implementation science itself, that just because something works in one place doesn't mean it will work in another.

Researchers on the clinical trial were motivated by ideals of social justice and human rights and implementation researchers desired to carry out the concerns of practitioners who intimately witnessed the harm of the criminal-legal system. However, it was the institutional obligations of the clinical trial that tended to distract trial actors from the larger vision of social change. More critical attention is merited to focus on how scientific processes themselves

become implicated in producing inequity. Here we observed how, despite implementation science's interest in pursuing advocacy, the field is constrained by its integration into the RCT and evidence-based policy more generally. Pragmatically, we might consider the incorporation of more democratic processes for designing interventions in pragmatic trials to ensure the concerns of trial actors, particularly those at the "frontlines" are more equally represented.

Chapter 2: Surviving the Trial: Negotiating "Real World" & Experimental Contexts

The Challenges of Generalizability

The proliferation of implementation science and its integration into standard randomized clinical trials (RCTs) emerged with the aim of making evidence more relevant to the contexts where it is delivered, commonly referred to as "the real world." A phrase used by implementation researchers, "the real world," refers to the vagaries of social and health service delivery that make it qualitatively different from the presumed artificial context of experiments. Indeed, RCTs tend to conjure the aseptic space of a laboratory, or a pharmaceutical trial where contexts are easy to control and therapeutic mechanisms appear relatively simple and isolatable. However, even in the context of pharmaceutical trials, their efficacy relies not just on the therapeutic compounds but also the people delivering the medication, the body of the person taking it, and other contextual factors and cultural expectations (Hardon & Sanabria, 2017). In the case of socio-behavioral interventions like HaRP, we encounter an even wider range of complicating factors. Indeed, the social and health services are often confounded by the task of making evidence-based programs "work" in their setting or local context. The following chapter pursues the question: How do researchers navigate the imperative of feasibility and scale with the scientific commitments of standardization and stability while fielding the HaRP intervention?

There are several reasons why a complex intervention like HaRP might not survive beyond the context of the originating trial. Prior literature can help us better understand this tendency, which tends to coalesce around three primary issues: (1) clinical trials often outperform their control because of the resources they bring to the intervention, (2) complex interventions are particularly difficult to scale because of their rather inextricable connection to the context of their deployment, and (3) the cause-effect relationship relies on "shields" or

"support factors" that may not be recognized or replicable beyond the confines of the experimental context. This is *not* to say that HaRP will not survive beyond the trial, but that prior literature, as well as this chapter, will show some of the complexities of creating evidence-based programs that seem viable elsewhere, that is, beyond the clinical trial.

To expand on the first point, randomized control trials (RCTs) like HaRP are highly resourced and run by highly trained professionals. Researchers introduce novel and expensive technologies, as well as complex procedures, to ensure that staff reliably adhere to intervention protocols. As described in the introduction to this dissertation, this is all part of what Montgomery (2017) calls their "material standardization." This "material standardization," though methodologically necessary, may render new innovations relatively unsustainable or not "feasible" in the "real world" of underfunded social service delivery. Feasibility is "the extent to which a new treatment, or an innovation, can be successfully used or carried out within a given agency or setting" (Proctor et al., 2011, p. 69). Underlining the effects of a trial protocol on practitioners and clinical sites, Montgomery argues that processes of standardization may frustrate sustainability and produce findings less "relevant" to everyday clinical research. Furthermore, by omitting the extensive labor required to produce standards, they may set up potential adopters for failure.

Secondly, HaRP is not just any socio-behavioral intervention, but a complex one taking place in jails across a midwestern state. According to the implementation science literature, complex interventions are difficult to define, combine multiple components, require coordination across a diverse set of actors, organizations, and types of service delivery, and *are highly dependent on the specific context of their implementation* (Bonell et al., 2012; Craig et al., 2008).

Critically, complex interventions are nested within multiple support structures, each structure adding complications to the act of scaling the intervention.

Thirdly, scholars grounded in science and technology studies (STS) focus on the "interactive" nature of efficacy, thus framing the problems of generalizability a bit differently. Rather than treating context as a backdrop for the cause-effect relationship to emerge, context interacts with a given intervention to bring about said changes. The former analytic tendency is rooted in the experimental model, which aims to identify and control for "extraneous" variables to distill causal mechanisms and measure outcomes (May et al., 2016). Variables like culture and climate are quantified and abstracted into categories and framed as either facilitative or as a barrier to intervention efficacy. Failing to appreciate the interactive relationship, it has been argued, oversimplifies how the cause-effect relationship manifests within a particular context. That is, by approaching context in this way, interventions are thought of as "separate from, yet shaped by, their implementations; as pre-existing but traveling objects made prior to their situated implementations" (Rhodes & Lancaster, 2019 p. 3). Instead, STS considers the situated practices that help make interventions efficacious, rendering any given intervention difficult *to analytically separate from its context* (Hardon & Sanabria 2017; Brives, 2016). For these researchers, it is not so simple as reproducing a given factor (e.g., training) because it is the localized, interactive effect that brings about desired outcomes.

Reidpath et al. (2022) have described this interactive process as "shielding." Drawn from philosopher Nancy Cartwright's concept of the "nomological machine," "shielding" is putting in place the right support factors so that a particular cause-effect relationship manifests. They argue that while an RCT may be useful in telling us whether a particular treatment caused an outcome,

this insight does not generalize beyond the trial without proper "shields." Shields might include, for example:

> the training of the people delivering the drug; the governance, policies, and infrastructure of the clinical setting; the wider health system; the levels of education of the patients; cultural views about drugs and disease; the incidence of disease; the socio-political, physical, and ecological environment that mediates the disease; and so forth (Reidpath et al., 2022, p. 3).

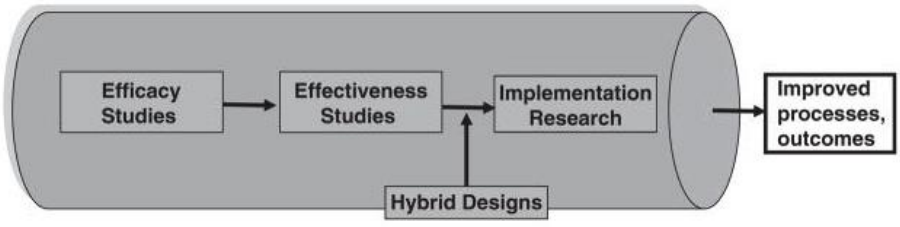
Importantly, these shields cannot be understood as a simple list of factors, but they are better understood as "ingredients" that work together (Cartwright & Hardie, 2012 p. 65). That is, a given intervention plays its causal role by *working with* the other ingredients, its support factors, to produce a contribution to the effect (Cartwright & Hardie, 2012). The cause-effect relationship is therefore "contextually bound," interactional, and dependent on a range of factors. In arguing this, this line of research emphasizes the need for greater attention to context and implementation processes. Similarly, it views generalizability not simply as a scalable cause-effect relationship (e.g., an evidence-based program), but rather as the reproduction of essential contextual ingredients (e.g., shields) that facilitate the desired outcome of an intervention.

Prior literature helps to diagnose the problems of RCT external validity and provides rigorous methodological and conceptual grounding to understand why RCTs of complex interventions might suffer such fate. This research, however, does not show us the types of tensions that emerge nor how researchers actively manage the imperatives of experimental success and developing "real world" findings to develop a scalable model, extractable from its context. The following chapter shows how researchers manage these tensions as HaRP is developed and implemented in criminal-legal contexts. This is particularly important to attend to given that these tensions are *actively built into* contemporary approaches to improve RCT

generalizability, that is, through hybrid and pragmatic clinical trials. Such trials actively combine the "here" of the trial with the "there" of the "real world."

Hybrid effectiveness implementation clinical trials combine questions of efficacy and pragmatic questions about an intervention's fit in the "real world" of service delivery (See Figure 2). Hybrid trials are therefore marked by their dual character. Communities where they take place must be rendered "lab-like" to make interventions comparable across sites. This requires the introduction of standards and other mechanisms to stabilize service delivery practices. Yet, jails and arguably all social service organizations are not "blank slates" but already filled with existing practices that may not line up to the logics of the social service technologies being introduced. To make the intervention as "real world" applicable as possible, researchers must attempt to blend research with service as usual. This makes it difficult for trial actors to distinguish the "real world" from the experimental one, and by extension precisely what the intervention *is*.

Figure 2: Research Pipeline



Source: Curran, G. M., Bauer, M., Mittman, B., Pyne, J. M., & Stetler, C. (2012). Effectiveness-implementation Hybrid Designs. *Medical Care*, *50*(3), 217–226. https://doi.org/10.1097/MLR.0b013e3182408812

This chapter delves into the tensions that emerge as researchers and implementation actors attempt to distinguish an intervention from its context to develop a stabilized, scalable intervention. In doing so, it reveals the moral and political justifications that researchers leverage

to expand the model, and the scientific justifications they used to restrict it.[1] While we know from research on standards that frontline professionals tend to alter them to better align with their own moral and ethical beliefs, we know less about how this happens, and with what consequences, during the implementation of clinical trials and within intervention science more broadly (Maynard-Moody, 2003). I argue that, ultimately, decisions made to help accomplish the experiment (namely facilitate recruitment and engagement) rendered the HaRP intervention less feasible in the "real world," and perhaps ironically, the model less stable. Contextual elements, political processes as well as material resources were obscured, sometimes to the dismay of practitioners and implementation scientists. Importantly, however, HaRP researchers were working with highly vulnerable populations, with pressing needs and many material "barriers" to recovery. Complicating blanket critiques of RCTs' lack of external validity, I show how researchers are pulled in multiple competing directions: to satisfy the moral and interpersonal desire to prevent immediate overdose by whatever means, the "real world" relevance of social service provision, and lastly, the scientific commitments of stability.

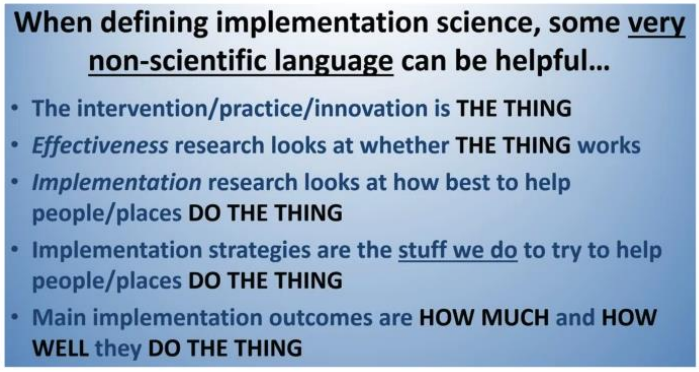
What is the "thing"?

The primary objective of an RCT, within textbook definitions, is to find *what works* and scale it elsewhere. But, in this familiar phrase—one made commonplace with the proliferation of evidence-based policy—the subject of the sentence is often obscured: what is the *it* that is working? In implementation science, at least at its most basic, researchers are often taught to think of interventions as the "*thing*." Meanwhile, implementation research "*looks at how best to help people and places do the thing*" (See Figure 3). Figure 3, oriented to the implementation

[1] While the term "model" is often used to describe a mathematical formula, the term was generally used by my interlocutors when they discussed the HaRP intervention and how it might be standardized (or not) across clinical trial sites to stabilize the intervention.

science novice, outlines the distinction between the intervention "the thing" and the context "the people and places" delivering it. Put this way, the intervention or innovation seems tangible and distinguishable from its context. However, as I will show in the following section, the "it" of HaRP was not always so easily stabilized, and often blended with its context.

Figure 3: "The thing"



Source: Curran, G. M. (2020). Implementation science made too simple: a teaching tool. *Implementation Science Communications*, *1*(1), 27. https://doi.org/10.1186/s43058-020-00001-z

To say HaRP *works* would mean we have already delineated what HaRP *is*. As described in the introduction of this dissertation and elsewhere, HaRP is a complex case management intervention that leverages elements of peer recovery and harm reduction to connect people who use opioids to medications for opioid use disorder (MOUD). But, once the intervention began to roll out in the summer of 2021, and as anticipated by prior literature, it was not so *easy* to stabilize HaRP. That is, researchers and implementation actors alike had to negotiate, circumscribe, and expand the HaRP model in response to the various dilemmas posed by the "real world." This is because of two interrelated conditions inherent to complex interventions: (1) complex clinical trials are notoriously dependent on their context–that is the network of organizations, types of collaboration and other elements involved in linking across systems of care, and (2) devolved clinical trials are difficult to monitor and control because they take place

in the community, far away from the research hubs heading the clinical trial. These two conditions make complex interventions like HaRP difficult to define, and the very process of definition requires researchers to rely on both scientific and extra-scientific (e.g., moral and political) factors that are critical to apprehend, yet usually elided in implementation science accounts. The following section will describe the ways in which the HaRP model was expanded or "let loose," and the extra-scientific as well as scientific decisions that motivated researchers' decisions.

Loosening the HaRP Model

There is no "clean slate" or "clean site" for implementation with socio-behavioral clinical trials. This meant that HaRP would look substantially different across each clinical trial recruitment site. Clinical sites, where interventions are meant to take place, are already filled with organizational stakes, interests, and routines. Jail recruitment sites, where HaRP performs its recruitment and engagement, were already populated with their own medical practices, other large-scale clinical trials, as well as substance use treatment and re-entry organizations looking to research and serve the jail population, respectively. For that reason alone, context can be challenging to manage because it inevitably overlaps with the efforts of other organizations, stakeholder interests, as well as established policies and practices. To gain buy-in and begin implementation, researchers must therefore negotiate these interests with the needs of the trial. In the case of HaRP, researchers had to adjust the protocol along the way not just to be collegial with neighboring organizations and study PIs, but to ensure buy-in and successful recruitment. The study team would also have to plan for "territorial conversations" with area reentry organizations in order to receive the green light to recruit from their "population." The tension being that local reentry organizations did not want to be understood as the "control" or treatment

as usual, as compared to HaRP, which brought an additional set of resources and would likely show better outcomes than the control. While area organizations claimed to be screening and serving all eligible individuals, the consensus of the study team and the stakeholders they spoke with, was that local reentry organizations were not serving everyone who needed it. HaRP PIs therefore had to collaborate extensively, and tread lightly with other clinical trial PIs to ensure the same individuals were not being recruited into multiple studies. To avoid recruiting the same individuals, HaRP researchers had to change their eligibility criteria, focusing on individuals who would return to specific area codes, distinct from those of the other clinical trials. That said, researchers were constantly needing to adapt to the "real world" context.

Each RCT site was qualitatively different, which, while part of the very experimental design, brought to the fore how we might understand, as we will see, precisely what HaRP was. While certainly not an exhaustive list, counties were marked by different treatment landscapes, different forms of collaborations and levels of coordination between the jail and treatment providers, as well as divergent types of policing practices. Both PIs had a sense that HaRP might look different and be shaped by different contexts. Robert, during a research team meeting after the immediate launch of HaRP, in Summer of 2021, described it this way:

> I think one of the tensions we are going to have is that this intervention will be different across sites, in an adapted way. There is just no way that x county and y county will be identical because *we must adapt to local circumstances*.

Trevor, the other PI, agreed the intervention would look quite different dependent on context. He stated that in his vision for HaRP, varying aspects of the intervention might be adapted in specific ways and replicated later—though such changes would largely be up to SafeLink, the study's translation center, and the implementation science team to keep track of and document.

He stated to me during an interview, after a little less than a year of implementation in the March of 2022:

> The ability to characterize differences in implementation likely due to setting will be really important for scalability, because I do think that, you know, people know their setting, and they can, they can interpret findings of a trial. And they can do it better if there's more information on the context of the setting. But perhaps it has broader generalizability, if there are different aspects of the intervention that are shaped by the setting, that people could then understand and then adapt most appropriately to their own setting.

Trevor's ideas about the future of HaRP echo Reidpath's (2022) notion that context matters and that, the more we can document a given context, the surer we can be at knowing how to produce the stated effect in other contexts. Trevor's sense was that stakeholders, down the line, might take up the task of interpreting and adapting the findings of HaRP to their specific context, knowing the ways in which it differs or not from the context of HaRP. Of course, this would necessitate SafeLink and the implementation team to thoroughly document implementation processes and contexts. This documentation would assist future stakeholders in determining the degree of divergence or similarity between their contexts and, consequently, the relevance of HaRP's findings. It would also require a precise understanding of what HaRP *was*.

While both Trevor and Robert would agree that HaRP might look different, requiring some flexibility and adaptation to context, this process was not so straightforward and was informed by multiple competing interests and commitments. As described in the introduction to this chapter, complex interventions are notoriously difficult to standardize and are often deeply enmeshed with and reliant upon their contexts. What I plan to show in the rest of this section is *how* complex interventions become reliant on their context, and how researchers negotiate the scientific commitments of standardization and stability (to develop an intervention) with the "real world" contexts where they take place. In what follows, I demonstrate how people and

institutions fielding HaRP would become part of the HaRP intervention, oftentimes with little capacity for any of the researchers or implementation actors to control it. Context would bleed into the intervention and, vice versa— the intervention would bleed into its context. The following examples highlight the difficulty of defining and stabilizing a complex intervention like HaRP in the sense of extricating it from context. I demonstrate the rationales that researchers would rely upon—both moral and methodological—to ground their decisions to expand or "loosen" the model. Ultimately, I show how researchers were invested in accomplishing the experiment and maintaining buy-in with practitioners, even at the cost of threatening its very scientific stability.

Example 1: Other Institutions Become Part of HaRP. By late Summer of 2021, we realized that, once enrolled into the study and released, many HaRP participants were being swiftly "caught" by, or transferred to, other institutions. HaRP was built on the assumption that participants, once released, would return to the community—that people would lose tolerance to opioids in jail and be made more susceptible to overdose after release was, again, the premise of the "problem" the intervention was trying to address. So, while the goal of the intervention was to build a relationship with participants and ultimately connect them to MOUD, many study participants were swiftly placed into residential treatment, sent to prison, or arrested again and placed back into jail. Indeed, a general symptom of the growing carceral state has been the swift "redistribution" of people from one institution to another (Lara-Millan, 2021).

During a research team meeting in September of 2021, Helen, the project manager who also helped recruit study participants, described this conundrum to the research team:

> I have a recent person who, and this is pretty common, I've heard this from more than one participant, they may be released to go to treatment with the understanding that if they successfully complete treatment, they will come back to

> court, and then likely get probation. But, if they decide to hightail it from the treatment center, after three days, the court will get a warrant for their arrest.

Participants who were thought to be imminently released to the community after jail would instead be mandated to residential treatment. That is, they would not be entering the "community" where the HaRP intervention was originally set to take place. While the study eligibility criteria included those meant to be soon released from jail, administrators and other stakeholders knowledgeable of the participants' case could rarely predict the fate of a given participant. Others, once recruited into the trial, and after their court hearing, would be sent directly to prison instead of getting released. One participant asked Heather, a peer worker, if she could have her number for her bible because "I can take my bible to prison with me." That is, the trial participant expected to continue receiving support from Heather, even once sent to prison. Meanwhile, these examples show the relative unpredictability of the carceral system– there was no way of predicting *when* participants would be released and, most importantly, *where* to. Such situations prompted researchers, practitioners, and project managers like Helen to ask what this would mean for the HaRP model as an intervention under study, if those ostensibly in the intervention were newly institutionalized elsewhere. Helen continued:

> And so that led us to kind of ask: What does that look like in terms of follow up with those folks who might get transferred particularly to the prison setting, without ever getting out, or ever working with us on anything outside of the correctional facility? and kind of … how we want to document that and what that looks like, because obviously, the intervention is going to look very different than for somebody who's being released in the community.

In this same meeting, Helen finally raised the question: should the participants who become (re)institutionalized get dropped from the study? And, if not, how can we say they are "receiving" the intervention if they do not interface with HaRP practitioners? To be sure, the clinical trial was an "intent to treat," that is, all participants who were randomized were included

in the statistical analysis, even if they were dropped from the study or didn't receive the treatment (McCoy, 2017). That said, a large number of participants were being re-routed to other institutions. Robert replied that he would like to reduce a sense of "moral injury" potentially felt by practitioners like Heather who would have to "drop" participants they'd worked to build a relationship with. Even if practitioners were not able to maintain contact with study participants, the act of dropping participants might cause a sense of guilt. Indeed, having fielded many clinical trials, Robert was deeply attuned to the ethical implications that arise when the scientific logics of the RCT (e.g., randomization) fail to match up to the logic of care that tends to drive practitioners who perform clinical trial interventions. Ultimately, the research team deferred to SafeLink and its practitioners, guiding them to do what seemed most "supportive" and "appropriate." In doing so, they were also gently suggesting to the SafeLink directors, Rhea and Francisco, that they keep these participants in the study. A similar example would occur a few weeks later, with a consistent line of reasoning for relaxing the model.

HaRP researchers were anxious to begin recruitment at Milton County, which would serve as the primary urban site in the clinical trial. Here, the projected number of people exiting jail with opioid use disorder could prove essential for completing the trial. Milton County, Trevor said, was the "workhorse" of recruitment and their involvement was necessary to reach statistical significance. But during meetings early on to obtain buy-in with the jail site, what the study team found was that Milton County was already contracting with a non-profit that delivered a very similar service—peer support and case management to people exiting jail. The one difference was that this organization was providing services for a shorter, three-month period, whereas HaRP was a 12-month intervention. Deciding to recruit and deliver HaRP at Milton County jail would trouble the model; HaRP would look substantially different in Milton

County, where participants are likely to also receive similar services, at least for the first three months. From those within the study, the question was: How could we account for the fact that some participants might already be receiving very similar services? Once again, the model was "under threat" and yet, the HaRP trial needed these recruitment numbers to meet statistical significance.

In cases where we were tasked with choosing between the model and meeting recruitment numbers, Robert would tend to say that he wanted to "maximize" the chance to engage those who could benefit from the intervention. In his words, "what really is important is that we *maximize the chance to help people*, as opposed to imposing a completely uniform model across the sites." Meeting statistical significance and maintaining the buy-in of practitioners overrode the scientific commitments of model stability. This meant that he wanted the research team to do what they could to recruit and engage participants into the study, even if it meant deviating from an ideal experimental model. It is decisions like this where researchers face a double bind: the more the model is controlled and circumscribed, the less "responsive" it is to the expressed needs of trial participants and the people who care for them. Importantly, invoking the model could also mean failing to reach the adequate recruitment numbers for statistical significance. However, releasing it, as the research team did here, could jeopardize the reproducibility of the model, prompting researchers and practitioners alike to question what exactly was being examined experimentally.

Example 2: HaRP "Stakeholders." People, integral to complex interventions and arguably all socio-behavioral ones, are difficult to control in the context of a clinical trial. In our case, the buy-in of jail administrators was an important contextual element to the successful implementation of HaRP. Invested jail administrators are more likely to marshal resources

toward an intervention and facilitate the "movement" of people into the study (or a given intervention). Buy-in is therefore an important contextual element to any implementation study. As outlined in the beginning of this chapter, Reidpath et al. (2022) argues that any given cause-effect relationship manifests because of specific "shields." That is, certain contextual elements are said to interact with a given intervention to bring about said change. The following section shows, however, how interventions can be difficult to disentangle from the "factors" meant to bring about the cause-effect relationship.

In the case of the jail superintendent of Sherwood County, Ryan, helps to elucidate this point. Ryan was a former jail guard but had worked his way up to jail superintendent. While originally skeptical of MOUD, understanding it as replacing one drug for another, he eventually came to see its value. Ryan was even eager to bring more services into the jail. He often felt frustrated that Sherwood County didn't seem to get the same attention and state resources as other more urban counties in the state. He also had a close family member tragically pass away from an overdose a few years prior, which animated his investment in people who use drugs and are incarcerated at Sherwood. And, though he had no formal training, he would become intimately involved in the lives of the incarcerated individuals at Sherwood County Jail.

In an interview with me a few months after recruitment began at Sherwood, he told me about the informal, supportive stance he usually takes with people who are incarcerated in the jail, through an extended example of how he might interact with a given person currently using drugs:

> I generally give them my personal cell phone and say, "Listen, you may relapse, through all this. Don't get scared. If you do, don't say, 'Oh, shit, now I'm done.' You're not done. Call me. Tell me, 'hey, I relapsed,' and you need help. We'll get you the help. Okay, if you're on probation, you know, we're not going to hide that from the state's attorney. We'll go together, we'll walk up there, we'll talk to the

> state's attorney about your case, you're going to tell them: 'I relapsed. I'm gonna go into treatment.'"

I was surprised by the level of involvement the jail superintendent had in the personal lives of people incarcerated at Sherwood. I also wondered about Ryan's own ideas about what might help a given individual into recovery. For Ryan, "owning up" to problems was critical to alleviate the putative burden of relapse. Confessing, however, would also mean, involving the state's attorney and other criminal-legal processes, and would likely result in sanctions and potentially reincarceration.

Around this same time in the trial implementation, a research manager said that Ryan was becoming very involved in selecting and engaging participants for the trial. To the dismay of some of the researchers, he was even "hand selecting" trial participants based on who he thought might be a "good fit" for the HaRP intervention. But, as we might expect from Ryan's quote above, Ryan did more than help bring in participants. In fact, in most conversations during case consultations at Sherwood County, Ryan would be named as a main actor in their service provision. Since Ryan had putatively built a relationship with specific incarcerated people, he would try to confer some of this trust he had built and guide them into the HaRP trial. He would also sometimes assume the role of case manager. Natalie, a case manager at Sherwood, described Ryan's "pilot program" to help one young woman get into treatment while incarcerated at the jail on charges of theft of a family vehicle:

> [The participant] has a court date coming up again in April. What Jason wants to do is he actually wants to find her [an] inpatient treatment facility, she will actually be taken to that facility in custody, and she will then go through treatment, and then she will still be in custody, when she comes out of treatment, she will have to report back to the south county jail, and then she'll go to court. Once she gets back from treatment, and then from there that they will take that into consideration, figure out her, you know, release conditions, and essentially, you know, she would be released after treatment. So that's what *Ryan is going for*.

Given the constraints the clinical trial faced, the struggles with recruitment, and the "real world" problem of overdose, it was difficult for any given researcher to ask Ryan to step back from this newfound, seemingly integral role in the HaRP intervention. In the case of Ryan, it was also difficult for any given practitioner to ask him to step down. It seemed to me that practitioners, at least sometimes, appreciated his support. For researchers, securing Ryan's buy-in would also mean little would be put in place to "control" him for scientific stability.

Importantly, Ryan's actions turned him from a contextual facilitator to a part of the intervention, but in ways that became difficult to account for. That is, it's worth considering how Ryan's investment, the way he explained the trial and intervention, and became intimately involved in participant treatment, would become part of the intervention *itself*. It is also worth considering his involvement and his theories of how people can change (e.g., being honest and open with the criminal-legal system to relieve burden) might have conflicted with the harm reduction model. That is, the goal of harm reduction, that researchers and clinical supervisors emphasized to practitioners, was that practitioners were to help participants *avoid* further criminal-legal involvement by not reporting to probation or prosecutors, a tension we will get into more in Chapter Three.

The lack of control in complex interventions is not usually the fault of any one researcher, or research plan, but likely endemic to complex and devolved clinical trials. Implementing at sites hundreds of miles away and managing so many actors confounds attempts to ensure delivery remains standardized. Complex interventions also trouble our ability to conceptualize interventions as distinct from their context, as in the case of Ryan. Lastly, researchers must manage difficult, oftentimes competing commitments—the "real world" needs of those imminently at risk of overdose, and the scientific need for stability and feasibility.

Circumscribing the Model

In the above examples, meeting the recruitment numbers, maintaining buy-in and accomplishing the trial took precedence over standardization and replicability of the HaRP model. Researchers were particularly attuned to interpersonal and moral factors, including the potential "moral injury" to practitioners faced with having to drop participants from the study. In the following section, I share an example of where researchers decided to circumscribe the model. Here, feasibility and replicability took precedence over the "real world" needs of study participants.

In this example, Francisco, a director at SafeLink, and I were asked to deliver the feedback we received from the community advisory board meetings (CABs) to the rest of the research team in March 2021, just a few months before implementation.[2] The meeting had taken place days earlier with a select group of service recipients at SafeLink. While Francisco facilitated the meetings, I attended to help take detailed notes of the CAB's responses. The CAB was set up in order to understand the needs of people as they are released from jail, their experiences of this transition, potential barriers and facilitators to jail recruitment and service engagement, and characteristics of service providers that foster a trusting relationship. While CABs are ideally set up to guide the entire process of a research project, Francisco's questions reflected a desire to understand the needs of people exiting jail and whether HaRP was designed to address such needs.

[2] Community advisory boards can be defined as "a group of people that formalizes the institution–community partnership and guides the research by providing a mechanism for community members to have representation in research activities. As a principle of good research practice, all clinical trials are expected to have a CAB to work as a link between the clinical trial's team and the community where the study is being implemented." (Mugenyi et al., 2021, p. 1).

What we found, and what Francisco reported to the research team, was that certain basic needs needed to be met before participants could consider getting involved in substance use treatment like MOUD, the primary outcome of the trial. Francisco described how people needed a safe, stable place to eat, bathe, and get a change of clothes. Supportive housing and other public services had reduced their beds to reduce COVID outbreaks, making securing housing even more difficult at the moment of HaRP's implementation. Voicing the needs expressed by the CAB, Francisco emphasized the importance of housing for people recently released, and how essential such stability would be in preventing overdose. In the following exchange, Francisco raised his concerns:

Francisco:

> And, you know, and an ID, you know, when we, in the work we do, getting an ID isn't that easy. And, you know, the other big thing that came out of that [CAB meeting] is like housing, some kind of, you know, not that a program should be able to do it. But that's a real key need that people have.

Robert:

> I mean, the one complication from a study design perspective is if we're providing a lot of benefit to people through things that are really not getting them into MOUD, we just have to be careful when we do the sort of, you know, year five analysis, that if we find that there are beneficial effects, it will be both coming from increased treatment uptake, and from other things that we're doing. So basically, the intent to treat estimates will be totally kosher. You know, anytime you're doing anything for people other than getting people into treatment, you have to be a little bit modest about making big causal claims like that 'we accomplished this by getting people into treatment.' But I mean, my biggest fear is not that we will have a huge benefit and not know how much of it was due to treatment.

Robert's reply to Francisco is that although housing may be necessary, its inclusion might disrupt the very model researchers were attempting to formulate and standardize. This would also make it challenging to demonstrate exactly how the study achieved its treatment outcomes. Adding too many services to HaRP, and introducing heterogeneity across sites, would make it difficult to make future, pragmatic recommendations about the actual causal effects of HaRP. Rhea, the

executive director of SafeLink, agreed with Robert, and had additional concerns about feasibility and replicability, if we were to include housing case management. Rhea concluded her statement, prompting the research team to consider what they hoped the HaRP model *to be*.

> Maybe part of the way we can think about it, because that's exactly what my concern is, is to kind of think about the factors along the sort of path to treatment that we can facilitate. And be very, you know, sort of thoughtful about what is in that box of things we provide. Yeah, and I was just thinking, from a community perspective, you know, down the line, if we're a recommended program for a community like Sherwood County, where would they get these resources from? We're not, you know, we're trying to develop a sort of a package of services that can be used. But also, *what is the hope for this model?*

For Rhea, adding housing case management would be yet another resource that seemed hard to justify given the usual, under-resourced nature of service provision. Rhea suggested the research team be "thoughtful" about the "box of things" included in HaRP. At the end of her response to Francisco and David, Rhea posed the question: "What is the hope for this model?" and, in thinking from the perspective of the community, she wondered what HaRP would look like, outside of the resource-rich context of the clinical trial. This was a question that troubled others on the study team, as we will see in the following section. Finally, Robert agreed with Rhea's concern about feasibility, stating that providing housing would likely not "survive the trial." In other words, Robert felt that providing housing case management would not be feasible beyond the trial.

The exchange between Francisco and study leadership is important for a couple of reasons. First, it reflects some of the challenges of negotiating the competing demands of flexibility *and* standardization and feasibility. Including housing case management as a primary intervention strategy could mean the model would be meaningfully different across sites, as Robert argues. At the end of the day, this would make it hard to say precisely what HaRP is and how the trial got to its outcomes. Lastly, as Rhea relays, providing this service may also not be

feasible, since this would take time and effort on the part of practitioners, and most community-based services don't have the resources to provide *everything* a given service user might need. In some ways, in complex interventions like HaRP, the line between what HaRP can provide and what would be "outsourced" to neighboring community-based organizations became blurry though nevertheless essential to negotiate and identify.

Alternatively, not including housing case management, as Francisco relayed, could also mean not meeting the material needs of participants. Amy, the clinical supervisor, would echo this point in an early interview in February of 2021, adding that while it is a difficult process, not providing what people need could even damage the therapeutic relationship, arguably the most essential component of any socio-behavioral intervention:

> So, let's say the person says, housing, which most everyone will, that's really hard to do. It's a lot of work, it takes a long time, are staff gonna invest in doing that? and how, how long are they going to invest in doing that? What if the person's not interested in working on their opioid use disorder at all? What does that mean for our staff time and investment and capacity... It's a sticky wicket. From a service provision standpoint, from my role, *I want to focus on what people want to work on*. I believe that's a way to establish rapport to get to the other things that we want to work on with people.

Here, Amy emphasizes the amount of labor securing housing would take, and how the flexibility required to meet client needs may be compromised by the pressure to meet specific goals delineated by the trial. Indeed, studies on re-entry from jail substantiate Francisco's findings from the CAB, and Amy's sense that securing housing would take a lot of work. People involved in the criminal-legal system as well as people with substance use issues often face exclusion from public housing authorities and discrimination in the housing market. This frequently results in homelessness, especially after peoples' social bonds with friends and family have been strained by incarceration and substance use (Augustine & Kushel, 2022; Miller, 2021). Rural areas, in particular, lack the resources and infrastructure to provide public housing (Buck-McFadyen,

2022). That said, the more the model was restricted, the less it could accommodate the needs of study participants.

Here, feasibility and the stability of the model was invoked by the research team while, as we saw earlier in this chapter, flexibility and participant needs took precedence. How might we account for the relative discrepancy in how the model was treated in this example versus earlier ones? For one, when we decided to "circumscribe" the model, this conversation took place before implementation began. This suggests that it might be easier to theoretically standardize than do so in real time, when facing the immediate needs of trial participants. This discrepancy might also be a symptom of a relative lack of control of a devolved clinical trial, where all decisions about who gets what simply couldn't be brought to a lengthy discussion. Instead, such decisions needed to be made ad hoc. That is, and as argued in section one, complex interventions are difficult to control. Lastly, it also seemed to me that when recruitment and engagement were at stake, the PIs and other decision-makers were more likely to forgo feasibility in favor of accomplishing the experiment, that is, "getting the recruitment numbers" and keeping the people who did not fall out of the sample engaged in the trial. That is, accomplishing the trial, at even scientific costs, seemed to take precedence. Yet, as we will see in the following section, letting the model loose made HaRP difficult to contain and less "real-world" feasible, which had consequences for those closest to its development and implementation.

Negotiating Experimental & Real World Contexts

The fact that the experiment and intervention started to blur was not just a simple methodological problem that researchers could later sort out using a sophisticated statistical model. It had immediate social consequences. Indeed, those closest to the intervention, tasked

either with delivering it or documenting its implementation processes, grew concerned, troubled even, that the intervention they were carefully developing did not mirror the "real world" reality of clinical practice. That is, real sites of service delivery had started to become qualitatively "unreal" While the replication crisis describes how scientific studies are difficult or impossible to reproduce, largely as result of an unprecedented increase in scientific productivity, what I call a crisis of feasibility, on the other hand, is the concern among implementation and research actors that the results of the trial may not be feasible beyond the study context. In other words, it is a concept that invites implementation actors to consider HaRP's ability to succeed in the "real world." In the case of HaRP, various implementation actors feared that HaRP would not "survive" outside its experimental context.

Importantly, however, actors were not unwavering in their investment in feasibility. What was *not* feasible prompted sociological discussions of scarcity in social services, or at other times, the critical needs of participants that overwhelmed the scientific logic of standardization. In fact, meeting the needs of participants and keeping them engaged in the trial often took precedence over ensuring its feasibility. In a way, the trial started to become an intervention in itself. That is, the point of the trial was no longer just about producing evidence for the future (as we saw in Chapter One), but instead, its resources were allocated toward meeting the critical and immediate needs of study participants. And, while there are too many examples to list the elements of HaRP that seemed artificial both to my interlocutors and myself, the following sections will detail some of the experimental resources and practices that vexed implementation actors and researchers trying to make HaRP "real world" relevant, while also meeting the demands of the experiment.

As I described in the previous section, for implementation researchers including myself and those closer to the "on the ground" delivery of the intervention, it became difficult to distinguish between what constituted "the thing" (HaRP) *versus* the experimental context. However, this is one of the longstanding inherited tasks of implementation science. So, the implementation science team would make efforts to imagine the HaRP intervention beyond the research study, extracted from the clinical trial apparatus. Yet, the implementation team knew very little about what HaRP would be beyond the trial. After the trial, would HaRP be implemented by jails or community-based organizations? How would service users be recruited into the intervention, without the bureaucratic infrastructure of the trial, and the research assistants? Would SafeLink continue providing clinical supervision and administrative support, since they were so deeply involved in HaRP's development? If not, who would be there to help HaRP live on through technical assistance once the trial ended? For many, it was difficult to extract the HaRP intervention from the experimental apparatus and imagine it living on its own beyond the clinical trial.

These questions meant that intervention actors—project managers, supervisors, and practitioners—would painstakingly document implementation practices and processes in order to have a sense of what the "it" was that was being delivered, while also attempting to neatly erase the influence of the experiment. And, this was not an easy process. The research team needed to constantly remind ourselves what was the "real world" of social service provision versus simply the experimental one.

What's in a context?

Implementation processes are usually laid out in specific sequential timeframes and scalar logics. Bounding the implementation process, and then categorizing it, seems to make it

more feasible to study. It also helps to establish clear boundaries between context that is pertinent to the intervention, or not. Yet, like Reidpath's concept of "shields," so many aspects of the HaRP trial seemed important for HaRP's success, yet difficult, if not impossible, to account for as an isolated intervention. Indeed, so many elements that came before implementation seemed important for HaRP's initial acceptability: university relationships, resources, legitimacy, and political work are just a few examples I will briefly describe in this section.

As described in Chapter One, jail stakeholders are known to be reluctant hosts to researchers. They may not always agree with the treatment ideals of social workers, or the political ideals they perceive to be represented by the majority of social scientists. Yet, the very idea that a large-scale research project would come to rural jails, with trained staff and resources ready to deliver an intervention that could reduce recidivism, made some jail administrators eager to participate in the study. In fact, in contexts of relative scarcity, jail administrators were less likely to see HaRP as an experiment, but instead as an additional source of support both material and symbolic: personnel, trainings, and even legitimacy. All of these could also, hopefully, help to curtail the rate of opioid overdose. "People are only implementing [HaRP] because of our research study and because we are funding these positions," stated Amy. "It won't tell us much about the viability of implementing it or sustaining it post-research because once we withdraw funding, there won't be take-up or continuation." While there is no way to tell whether Amy's prediction is correct at this point in time, she was overcome with a sense that the resources, both symbolic and material, rendered the process of buy-in and implementation a bit easier for jails administrators and other stakeholders to stomach. Indeed, stakeholders shared with me in interviews that they were more likely to partner with HaRP due to the legitimacy it conferred to them and their organization by working alongside academic research institutions.

Partnerships. Trevor, one of the principal investigators on the trial, already had a number of research projects underway with HaRP jail sites in Coleman and Sherwood Counties, helping to expand medical screening and treatment for HIV. This meant he had forged relationships with jail administrators and area treatment providers and helped to establish connections across this network of services. So, long before HaRP would begin, the jail signaled their investment in providing medical treatment, learned to coordinate new providers, and had established relationships with harm reduction organizations in the area. These existing partnerships, as well as Coleman and Sherwood's existing connections and investment, likely made HaRP more acceptable and feasible, making these connections a "shield" for successful implementation. However, the effects of research partnerships and existing projects are difficult to replicate and often left outside the scope of implementation analysis. How these extensive networks might be reproduced in the expansion or scale of an intervention is worth considering. Of course, researchers are increasingly encouraged to develop strong relationships with the organizations and institutions they study and from which they recruit. Indeed, scientists are increasingly asked to justify their research and show how it is meaningfully accountable to the society it's meant to serve. More attention is therefore merited to investigate and understand the dimensions of implementation science networks, including the community engagement apparatus that has grown alongside research institutions and universities.

Resources. Exposure to the clinical trial's abundant resources had many positive effects. Practitioners gained a range of new skills, and organizations accrued new technologies, personnel and even credibility. Community-based partners spoke highly of the benefits of participating in trainings, and organizations reported reputational gains by virtue of being associated with well-known academic institutions. Yet, for implementation scientists,

practitioners and PIs alike, these resources and connections also had to be negotiated with the "real world" relevance of the HaRP intervention, namely its feasibility. The following section will detail these negotiations.

"The problem with research," Amy said to me in an interview early on in HaRP's implementation, "is you get funding to build this beautiful program for three years and then it all dissolves." For Amy, researchers often focus on accomplishing the clinical trial at the expense of "real world" social service provision.[3] She continued: "Researchers sometimes have these blinders [...] 'my responsibility is the study,' and then it becomes not real world applicable, implementable or fundable." As suggested in the literature detailed in the introduction to this chapter, and underscored by Amy, the experimental context often does not match the "real world" of social service provision. Researchers make decisions that support accomplishing the study (e.g., improving recruitment and retention), which often render the interventions they produce only feasible with the support of trial resources. For, as Amy also argues, the apparatus and resources of the RCT tend to "dissolve" once the trial ends, and, along with them, the very intervention she helped build:

> There's no money to continue this work after the study, like the community partners that we're working with, they're not being given any money or grants to pay people to do this. And regardless of, you know, our findings about the importance of training and supervision, that doesn't change the culture of organizations or the resources of organizations to be able to shift their priorities and make that happen.

[3] To understand the sense of frustration and pessimism that ensued, I find it important to mention that much of the work of the implementation science team, including Amy, was dedicated to designing and implementing an implementation study that would help the research team to figure out how HaRP could survive the trial. Consider when we spent almost three hour-long implementation science meetings trying to account for the pre-implementation cost (e.g., salaries, resources) of a community partner's time needed to launch the HaRP trial. This meant tracking down the salaries of jail and community stakeholders and documenting the time they spent forging relationships with trial researchers and filling out paperwork like memorandums of agreement. This was done to estimate the pre-implementation cost of delivering HaRP for those wishing to evaluate it in the future. In essence, the implementation team was investing the majority of time considering and promoting HaRP's future.

What Amy is suggesting here is that the funding and resources provided by HaRP didn't seem to match the types of resources she found in "the real world." Even if the team could show how important a given "shield" like training might be, it didn't seem realistic that it would be provided beyond the trial context. The following example helps to elucidate Amy's point.

Once released from jail, participants, as like many of those "re-entering" into the community, needed critical resources: housing, money, and transportation. One participant, for example, needed to get from Sherwood County to a Milton County recovery home, a trip which typically took ninety minutes. It was common for participants in the more rural areas of the study to need transportation, both around the city as well as to gain access to treatment centers which tend to reside in the more highly resourced suburban and urban counties in the state. The problem was, there was no public transportation and this participant did not have access to a car. Amy continued: "And so Helen [the project manager] tends to say 'yes,' right? So, Helen paid for a $100 Uber ride to their recovery home, but he didn't have any money, so Helen paid for their rent. But that's not something that's replicable. Right? Like most social service agencies can't do that." On the other hand, she said, "that's what people need." A tension Amy had been dealing with over the past few months was how to negotiate the critical needs of participants with her desire to also see HaRP survive the trial.

It wasn't just practitioners like Amy who had to make difficult decisions about managing the tension between feasibility and meeting the critical needs of study participants. Recall that in Chapter One we found out that Sherwood County jail was charging participants to make phone calls through what is called "Chirping." A Chirper is a device used to allow incarcerated people to communicate with the outside at a high cost. When the research team was apprised of the issue, one way to handle it was to simply pay for participant's use of the service. This could

ultimately keep people engaged in the study, and seemed like the ethical thing to do, at least for Robert:

> From a study perspective and a money perspective, I am delighted to pay for it. It seems to me that the really important thing is to make sure we have good human connection with people [...] So you know, I'm not thrilled about it, but I think it's money that should be, should be spent because I just think we have to connect with people. That's just, there's no other way.

We see here that replicability and feasibility were not always top of mind for researchers; the needs of study participants sometimes overrode how the intervention might function in the "real world"—contexts in which resources were likely to be more limited. In the context of a well-funded academic institution, a highly resourced clinical trial, at a moment with low recruitment and high attrition, it can be hard *not* to provide the resources necessary for participants to succeed. In making the decision to pay for Chirping, Robert himself referred to the ample resources of the clinical trial. Practitioners and PIs alike were also sensitive to the needs of participants and the relative incongruence between participants' needs and the resources the research team could provide. In that sense, the moral imperative of providing resources seemed to outweigh the imperatives of the scientific model and its feasibility.

Infrastructure. For many practitioners, to survive "the real world," the HaRP intervention would require new organizational structures, personnel, and a drastic change in both local and state funding priorities. This disparity was something many implementing actors were sensitive to. Consider Jessie, who started working as a case manager at SafeLink but had moved up the ranks and was charged with managing practitioners. After Amy and the HaRP practitioners, Jessie was the most familiar with what HaRP looked like "on the ground." This is because Jessie was tasked with figuring out how to document practitioners' work. She listed for me some of the elements that had to be recorded for study purposes: "time spent on calls, time

spent on meetings, time spent on coordination." And yet, she too was plagued with a sense that the infrastructural complexity and scale of HaRP would make it difficult to translate to the types of community organizations she's worked at for years: "it's very intimidating to see this very big academic project and go like, 'Oh, yeah, that worked. And, and to say, like, we're going to do that here in our small organization.'" Arguably, part of the role of implementation scientists and other implementation actors like Jessie in a trial like HaRP is precisely to document these realities: to show the labor, infrastructures and training required to implement quality evidence-based practices. Yet, it is nevertheless worth considering how practitioners who are implementing and documenting this work might perceive the unsustainability and discrepancy between the services provided in a clinical trial versus those provided under usual care. Jessie was particularly concerned about the paperwork, the extensive and expensive data management systems required to document clinical practices, and other forms of administrative burden that seemed necessary to implement HaRP but were likely out-of-reach for most implementing organizations. Again, as this had been a primary task of many implementation actors on the trial, failing to imagine HaRP surviving the trial was understandably met with disappointment and frustration.

Anti-Feasibility

Some elements of HaRP were even *decidedly* unfeasible. That is, some researchers were willing and eager to meet the needs of participants at the expense of feasibility. Training, group consultation, and one-on-one clinical supervision represented an *unprecedented* investment in practitioners, at least from the perspective of those more accustomed to the "real world" of service delivery. There were many exogenous factors that would come into play that would affect the implementation of HaRP and its perceived effectiveness. To name just a few: changing

jail protocols and policies, COVID-19 restrictions and lockdowns, changes in drug markets, and the local availability of treatment providers and other essential services. Often, implementation actors, including myself, felt helpless at times about the factors that would frustrate implementation. What implementation researchers *could* attempt to control was how practitioners were trained to enter the field and deliver HaRP. That said, the training and support offered to practitioners through HaRP, as Amy once put it once, was "luxurious."

Training covered each of the evidence-based practices promised in the grant application, as well training on both substance-use and co-occurring disorders, overdose prevention training, as well as MOUD. Combined, it took approximately two months for practitioners to complete training. Case managers and peers received bi-weekly case consultations that provided the opportunity for practitioners to offer and share resources, experiences, and support for one another. Practitioners were also asked to bring case scenarios for discussion to the group in order to explore and problem solve challenges. As will be described in Chapter Three, this was also the place for the study team to evaluate fidelity. Lastly, each practitioner received weekly, individual supervision with Amy to discuss cases as well as interpersonal and systems issues.

Amy was adamant that practitioners receive the support they needed *throughout* the implementation process, not just during the initial months of training. Carl, too, saw the extensive training and practitioner support as a careful effort to "make a case" that this *should be* an organizational element to everyday service delivery. When I asked about the feasibility of the training and support provided to the practitioners, Carl responded that peers and case managers "deserve kind of like support, right? Not just top-down supervision, but support, right, like investment in their skill set and their well-being." Emphasizing feasibility also meant upholding the status quo that included overworked, and under supported social service professionals. In

witnessing the extensive resources allocated to HaRP practitioners and thinking it was first and foremost "unfeasible," I also naturalized ongoing systematic disparities in social service funding by asking the question above. Indeed, feasibility is a goal many of us blindly pursue, and stress even, without considering the potentially pernicious social effects.

In the ongoing process of task-shifting across domains of social service and medical practice, specific types of labor are delegated, usually from highly credentialed professionals to those with "shorter training or fewer qualifications, including laypeople" (Orkin et al., 2021, p. 1). Peers are often called upon to engage and build rapport with service users, because it is assumed by providers that they have more of a common ground than a professional or highly credentialed expert. Traditional hierarchies, however, tend to prioritize such credentialed experts over those with experiential or lay expertise, which is undervalued in the service delivery labor market. Peers and case managers are often not provided the same investment in their training and are often employed by community-based organizations that lack critical funding. That said, Amy and others wanted to change this customary practice and finally invest in peer workers and case managers, and the HaRP trial was one vehicle to do so. Importantly, even if the implementation team understood that this level of management, training and consultation did not mirror the "real world," or wasn't feasible, they wanted to show that something like this *should* exist for peers and case managers. Its lack of feasibility took the form of a political critique about under-resourced social service organizations and under-supported lay personnel.

What is the future of HaRP?

For the implementation scientists, ensuring HaRP lived on beyond its trial was not just a methodological imperative. Implementation science, after all, emerged to ensure the findings of science are made applicable to everyday social service provision out of a normative and

pragmatic imperative that research impact “the real world” in a meaningful way. Since its inception, the field has focused on closing the iconic seventeen-year gap between the determination of evidence and its influence on practice. Concern about HaRP’s future prompted the implementation team to ask the PIs, Trevor and Robert, about their goals for the end of the experiment. The following exchange during a research team meeting took place a few months before recruitment began:

Amy:

> How would [HaRP] actually be replicated? So, there's like certification of sites and that kind of thing. And so we're wanting to know from the PIs, like what do you envision after this study is completed? [Do you envision] being able to, or wanting to provide training and technical assistance to people who want to replicate this intervention? What's your vision, post-study?

Robert:

> Wow, that's a hard one. [...] I guess I didn't envision that we would personally be sort of doing the sort of activities that you just described, I envision more that we would implement this in a very transparent and manualized way, so that others could draw from it.

Amy:

> Yeah, I mean, how will people know that they're replicating what we've created if they don't have access to experts to consult with?

Trevor:

> If we take, if we step back and evaluate the Hub and Spoke component, then I do think then it would probably be, you know, the SafeLink leadership team that might answer those questions.

This conversation is interesting for a few reasons. For one, it suggests that there were few plans in place for thinking about the afterlife of the trial. This makes sense. The PIs had a myriad of other issues to consider while attempting to field this very complex intervention across the state, in partnership with the carceral system, during a pandemic. In other words, their duty was to accomplish the experiment. Importantly, it was the work of the implementation science team, on the other hand, to decide HaRP’s fate after the trial ended. This is what Trevor suggests when he says the future of HaRP might best be mapped out by those at SafeLink, considering they hire the practitioners, clinical supervisor and arguably closer to HaRP’s implementation. Despite the

plethora of decisions made that would make HaRP less feasible, and the model less scalable, SafeLink and the implementation science team would be charged with dealing with HaRP's afterlife. With the emergence of hybrid trials, research teams tend to become compartmentalized with those invested in the experiment on the one end and those invested in "context" on the other. Through this delegation of tasks, they each have their own "professional vision," which are also, at times, at odds with one another.

Second, Amy suggests that a training manual would not be enough to implement HaRP, and that positions like hers (i.e., the clinical supervisor), and other built-in elements like training, case-consults and supervision would need to be replicated beyond the trial. Her work, and the ongoing support and training for practitioners, was critical to its success and not a dispensable "extra" that could be shed after the experiment. Arguably, these might be important implementation strategies that the implementation science team could develop and evaluate. Yet, given their work in developing the intervention, the implementation team desired more involvement in HaRP's life beyond the trial. A list of manual or list of strategies, for Amy, wouldn't be enough. The implementation team met a few days after this conversation to discuss the results of this meeting:

Amy:
> I like, I get it, but what is the point of developing an intervention for replication if people can't get the information to replicate it?

Carl: Especially given that we're doing *all* of this implementation work.

Amy:
> Yeah. I mean, training and technical assistance work was all I did for the last eleven years prior to this and it's definitely super labor intensive. I mean, we're going to make an implementation guide and manual. But that's *not enough* for people.

Emily:
> It makes me wonder how many intervention RCTs like to go this route versus a different route?

Amy:

> Yeah, it's a great question. Well, I think part of this is, it shows how problematic the model of doing this kind of expensive research is and not really, it doesn't get replicated. It takes 20 to 30 years for an evidence-based practice to get adopted. And this is why.

Carl:

> I feel like we're making a strong *ethical argument* to push or to gently support thinking about if this is effective, that there should be technical assistance for this to live in the real world. (...) a lot of my work in the past has been doing intervention studies, publishing it and moving to another thing. *We want this stuff to live in the real world, and we need to build an infrastructure around it*."

This conversation reveals a growing disenchantment with academic social science and its pressure to take on more studies and publish more articles without thinking about the "real world" significance or impact. As we saw in chapter one, imminent problems to be resolved were often deferred to the timescales of academic ones. We were also disappointed that more thought and attention had not been placed on HaRP's future. As put by Amy, the focus of PIs seemed to be accomplishing the experiment at all costs. That said, ensuring that HaRP lived on, and "building an infrastructure around it" became an ethical case to pursue for the implementation science team.

Conclusion

In this chapter I argue that attempts to accomplish the trial tended to override "real world" concerns of feasibility, or even scientific ones like ensuring the model. That is, everyday decisions about the model, whether to let it 'loose' or circumscribe it, tended to focus on accomplishing the trial, rather than considering the intervention's future beyond it. Importantly, HaRP researchers faced a double bind: the more the model was controlled and circumscribed, the less "responsive" it was to the expressed needs of trial participants. Yet, letting the model "loose" threatened the RCT model and its perceived feasibility. Critically, these decisions were

often motivated by extra-scientific factors, including researchers and practitioners' own moral and political views, as we might have anticipated from literature on discretion (Lipsky, 1980).

Trial actors had to cope with the pressing needs of trial participants who could face imminent overdose without needed resources. Frustrated by the pervasive disinvestment in social services, researchers and practitioners began to understand the experiment and trial resources as an intervention *in itself*. By intervention I mean (1) the trial resources were not delivered solely to create a stabilized intervention but to also meet the urgent needs of trial participants (e.g., through transportation or chirping) and (2) a political intervention about the underfunded nature of peer work in particular and social service delivery in general. Importantly, it showed that this clinical trial was not just a site for HaRP's implementation, but the very grounds for its development and re-making.

The training, practitioner development, and support offered by the HaRP trial might not have aligned with the practical realities of everyday practice. However, the research team felt an ethical obligation to invest in the practitioners delivering HaRP and advocated for this investment to become a standard expectation of support. I showed how implementation actors had to wrestle with the exigencies of feasibility, while also aspiring to see what they've developed to continue to live out in the "real world." They wanted to promote the types of systems and supports (e.g., case consultation, supervision, and robust training) that they felt *should* be in place for practitioners to thrive. In doing so, I show the extra-scientific factors that may trouble both scientific stability as well as the feasibility of a given intervention.

I started this chapter by showing the challenges of controlling the model in a complex intervention, and how HaRP was difficult to extricate from the context of its deployment. Through hybrid trials, this tension is in some ways built in. To improve generalizability, the

experimental context is made to be more "real world" like. In the face of managing such complexity, implementation scientists have suggested a variety of technical strategies: Providing flexibility, extensive documentation, and developing a theory of change in order to apprehend the effectiveness of complex interventions in ways that do not require strict standardization. Documentation is the key to recording "permissible" adaptations, variability in intervention delivery, as well as a good faith effort at recording the intervention as it is rolled out to enable its replication (Lambert et al, 2019). Moreover, to better acknowledge the variability in delivery, implementation scientists have suggested that we refrain from considering case management interventions like HaRP a *single* intervention but instead a set of "interconnected actions" taking place within a complex system of organizations (Lambert et al., 2019). Ostensibly, this makes complex interventions easier to scale without the rigidness of strict standardization and more flexible to accommodate diverse settings. This recognition has also prompted efforts to focus on theory and causal mechanisms over the strict delivery of activities, a more flexible and responsive format compared to requiring rigid intervention "forms" (Bonell et al., 2012). For example, to enhance practitioner's knowledge of evidence-based practices, skill-building would be considered a function, rather than requiring the format of a workshop or training. This means finding the causal theory behind the practices, rather than using the practices themselves as a reference point or target to evaluate a given interventions' stability and therefore putative validity.

Yet, this chapter shows limitations of these strategies, at least on their own, to contend with this complexity. First, in the process of standardization, practitioners and researchers alike are faced with complex, political, and moral issues that are difficult to anticipate beforehand and may require additional attention and support beyond a particular model or framework.

Interventions blurred with their context and varied across sites in ways that exceeded even the common factors meant to ground them as stabilized, replicable interventions. Even the concept of shielding, for example, suggests it is possible to easily separate a given intervention from its context. That is, the concept of shielding does not account for the political, scientific, and moral negotiations required to produce and maintain a given intervention model. In the case of HaRP, I argue that this contextual dependence made it difficult for researchers to distinguish and define essential elements prior to the study's roll out. This required active negotiation throughout implementation, or sometimes eliding altogether what might have made HaRP work (e.g., Ryan's labor). This chapter adds to the literature by examining the tensions that arise as well as how researchers actively navigate the imperatives of experimental success, making an intervention "real world" compliant, while producing a putatively scalable model. This chapter showed how we might better anticipate the challenges posed by complex interventions in our attempts to extricate them from their context and scale them to new sites and populations. Lastly, it also brings to the fore whether complex interventions can be successfully scaled, and what types of processes, histories and contexts might be lost in attempts to universalize them.

Chapter 3: On Being Faithful: The Politics of Intervention Fidelity

Certain refrains are common with the U.S. helping professions, and the therapeutic landscape in particular—one of which is to "meet people where they're at." Taught to professionals in positions of relative authority—the teacher, the child welfare worker or in this case, case managers and peer workers—imperative is that professionals relinquish their own expectations and that of the organization they represent to instead take the time to intentionally listen and understand the values, desires, and needs of the service user. However, the directive to "meet clients where they are at" and be "client centered" is often compromised by the fact that, within contemporary social service contexts, practitioners are also required to direct clients toward "normative, professionally supported behavioral change goals" (Carr, 2023 p. 32). Indeed, the ideological foundation of the strategies and practices that practitioners are asked to leverage might not always match up to the social contexts in which they operate: in the context at hand in this dissertation, a clinical trial that took place in jails.

Harm reduction espouses "non-judgmental and non-coercive provision of services" ("Harm Reduction Principles," n.d.). Practitioners of harm reduction are often asked to adhere to its values before deploying attendant strategies. This might mean taking more time to listen to a person who uses drugs, allowing them time to speak about complexities in their lives, and affirming their beliefs instead of hastily following study protocols or deploying evidence-based techniques to meet service outputs (Hassan, et al., 2022). In fact, harm reduction tends to reject normative clinical outputs, expected outcomes and the accountability structures that impose them. Instead, it tends to acknowledge "any positive change," which is determined by the individual rather than an external party. Peer work too is meant to draw on one's own life experiences—the embodied, vast stock of specific, contextualized knowledge, both tacit and

otherwise, which gives peer practitioners the intuition and ability to support others through similar life experiences. Their status as a peer is meant to help neutralize the perceived authority of social service professionals who might otherwise deter the engagement of participants by their very social and demographic differences. Yet, the newly institutionalized status of peery recovery and harm reduction and their integration into public health and criminal-legal systems remains ideologically complicated if not outright fraught. This was the case with the HaRP clinical trial.

The harm reductionists and peers were hired from around the state to deliver HaRP. Some, particularly those from grassroots harm reduction organizations, began to feel that their approach for building relationships with people who use drugs was compromised by the demands–and carceral context–of the clinical trial. The imperative to "meet clients where they are at" and honor their self-determination was complicated by the fact that peers were not only directing participants toward therapeutic goals but also specific study targets (e.g., outcomes, recruitment) and physically meeting them in a place to which they were legally confined (e.g., the jail cell). Perpetually behind on recruitment, the trial imposed a sense of pressure on the part of HaRP practitioners to engage participants and maintain them in the study. Yet, as much as practitioners felt their practices being compromised by the study protocol and its mechanisms of stability and accountability, researchers were enthusiastically attuned to ensuring the protocol remained flexible around the vagaries of context and practitioner experiences. So, how do we account for this discrepancy?

Treatment Fidelity and Standards

With theological roots, the term fidelity generally connotes loyalty to a spouse or partner. It is also used to describe the exactitude by which something, like a text, is reproduced. Like

ensuring the right compounds in a pharmaceutical drug, fidelity is usually described as the degree to which therapeutic actions, processes and essential ingredients match up to what is specified by evidence-based program developers. In evidence-based policy, the concept of treatment fidelity is vital. Within the model of evidence-based practice and its imperatives of "scale," researchers are expected to establish a set of practices that can be exported from the local particularities of their research context into new ones. In other words, fidelity is the degree to which the therapeutic actions, processes, and "essential elements" specified by the developer and laid out in published reports are replicated in practice (Carroll et al., 2007, p. 3). Although methodological strategies to promote adaptation are now commonplace, it is still expected that specific core elements of an intervention will be enacted faithfully across different contexts.

In a randomized clinical trial (RCT), fidelity is the mechanism by which researchers can stabilize a set of clinical practices to be sure (1) experimental outcomes or "what works" can be attributed to the intervention as prescribed, and (2) so that the intervention can be exported from the trial context and translated into practice elsewhere. Therefore, one could argue that fidelity, as a methodological construct, is principally a mechanism of stabilization. Failing to stabilize HaRP practitioners' work and allowing too much variation across contexts could cast doubt upon the entire clinical trial.

In truth, the fidelity protocol's reach extended beyond practitioners, and its effects were not solely methodological. It was also a system of accountability, a managerial tool for supervisors, and for those being trained to cultivate their practices to fidelity, an avenue for professionalization. Yet, despite these different functions, fidelity was a thorn in the side of almost every actor in the study—from the PI's concerned about meeting it for internal validity, to the implementation researchers charged with creating protocols to measure it, to program

managers and supervisors tasked with encouraging it, and, perhaps most importantly, the intervention actors asked to adhere to it.

The practice of ensuring fidelity during a clinical trial necessitates behavioral change. Practitioners are asked to acquire and implement new skills and try to break old habits that may not align with the evidence-based practices (EBPs) they've been tasked with implementing. We know from science and technology studies (STS) that in order to study something, it has to be in some ways altered, whether it is made to be more uniform, standardized or more explicit (Latour & Woolgar, 1983). As we learned in chapter two, to meet research requirements, HaRP retrained practitioners, adopted novel reporting and data management technologies, and established new types of relationships both with experts and service users. Training to fidelity is also one such way that practices are altered. Ethnographic studies of clinical trials have shown how experimental conditions can transform care as usual (Petty & Heimer, 2011), while others have shown the extensive compliance and surveillance work acquired to achieve these types of transformations (Winther & Hillersdale, 2020).

In other analyses, standardization in the form of evidence-based practice, reduces clinical encounters to rigid procedures and, as a result, overlooks social and cultural aspects of care and replaces the practice expertise (Lambert, 2006; Gone, 2013). Evidence-based practices, for others still, might miss essential ingredients that are more tacit and elusive to measurement, like the practices involved in building what is generally considered the most essential element to effective mental health treatment: the therapeutic relationship (Kirmayer, 2012). Alternatively, ethnographic studies of evidence-based protocol introduction in medicine have cogently demonstrated how the introduction of standards is a situated practice that requires collaboration and flexibility (Timmermans & Berg, 2010; Mol, 2002) and often requires "tinkering, subverting

and circumventing" on the part of standard users (Lampland & Star, 2009, p. 4). This insight is also widely recognized within implementation science—that context matters and must be studied and managed accordingly for any standard–or intervention–to thrive. An entire sub-field of implementation scholarship has focused on adaptation, usually the development of frameworks that help to, perhaps somewhat ironically, standardize adaptation.

Yet, despite fidelity's proliferation as a mechanism both in clinical trials and everyday social service practice, little research has examined how researchers, practitioners, and other clinical staff experience this rather novel form of standardization and technology of accountability. What's more, few critical examinations on the practice effects of research standards, and the social effects of integrating evidence-based practices more specifically, have considered emerging trends in health services research toward "pragmatic" clinical trials, where flexibility is not only welcomed, but actively desired. Considering this fundamental shift in contemporary health services research, further investigation is merited to understand the negotiations and tensions that emerge as practices are made to be both flexible and contextually relevant on the one hand and stabilized and generalizable on the other.

In what follows I argue that despite the aspiration from all levels of the research team toward flexibility and ensuring the protocol remained contextually-relevant, the practitioners' sense of how best to build a therapeutic relationship was compromised by the exigencies of the RCT, including the pressure to meet study goals, and recruitment targets as well as the pressure to deploy certain evidence-based practices that failed to line up with practitioners' own sense of their role, including their 'lived experience.' This emphasis on flexibility, I add, might very well obscure fundamental tensions in social service provision, particularly as it relates to self-determination and coercion, made ever more acute in the context of a clinical trial which took

place within carceral systems. While in Chapter Two we saw how the model was 'let loose' to meet study targets and accomplish the trial, here, practitioners' labor was the one element of the research process that could be more readily and feasibly controlled. Given the vast differences in local policies toward drug use across sites, diverse jail protocols and policing practices, and other myriad external contingencies across study sites, the one aspect that researchers in this case might at least attempt to replicate was practitioner behavior. Finally, I show how the fidelity protocol, and the evidence-based practices it promoted, did more than (attempt to) stabilize practitioner's work, it also helped to ideologically obscure how relational practices are fundamentally reconfigured—and at times compromised—by the clinical trial protocol.

The following chapter will explore the concept of fidelity: how it was thought about, operationalized, implemented and lastly, its social effects on practitioners. I begin by describing how PIs and implementation researchers aspired to provide flexibility and actively defied some of the logics inscribed in the concept. Due to methodological and managerial imperatives, however, implementation actors needed to stabilize HaRP. Next, I will show the complexity of developing fidelity to a complex intervention that involves diffuse, political, and conceptually "full" concepts like harm reduction and peer recovery. Lastly, I show how practitioners experienced the fidelity protocol and the evidence-based practices it promoted, including how practitioners, particularly peers and harm reductionists, felt their role compromised by an amalgamation of the clinical trial, the fidelity apparatus and the carceral context.

Flexibility

Fidelity was designated as the primary implementation outcome with clinicaltrials.gov, a federally managed website where investigators are mandated to report clinical trial outcomes in advance of the trial's rollout. This reporting is done for purposes of transparency and to ensure

the timely public access of study results. Considering fidelity was so essential to the integrity of the trial design, it was surprising that the lead principal investigators of the trial were ambivalent and even skeptical not only about the fidelity system developed by the implementation science team, but of the very idea of fidelity. And, while the PIs would voice their concerns, the implementation science team would require fidelity for both epistemological and managerial reasons.

We can see this clearly in a meeting I attended in 2020. Asked to provide updates on study progress, Francisco stated solemnly that he had to call an ambulance earlier that day after a few people showed up to a mobile unit with severe Covid-19 symptoms. Francisco was a co-director of SafeLink, which managed many mobile units across Milton County. They provided drop-in services, including syringe exchange, rapid HIV and STI testing as well as referrals to substance use treatment. During another Covid-19 wave, Robert expressed his concern and asked how staff might adjust their practices to ensure everyone's safety. Robert also wondered aloud how these types of issues might affect the HaRP trial once recruitment began. Indeed, a majority of 2020 had been spent scrambling to surmise how the research team would adapt to the new reality of the Covid-19 disease and adjust the study protocol in line with university as well as WHO and CDC guidelines. While the study team discussed remote options, those closer to the field were concerned about people's access to technology, particularly in areas with less infrastructure. They desired in-person meetings, despite what was prescribed by the IRB. In other words, researchers had to balance the liability inherent to human subjects research, while also remaining true to the "hard to reach" populations that the HaRP trial was ultimately meant to support.

After an extensive conversation about the IRB, we addressed the fidelity evaluation the implementation science team was developing. This too showed the struggle to balance the clinical trial imperatives with the flexibility required by a local context. Carl was in charge of describing the fidelity plan, which he went over in exacting detail. After listening intently, Robert said: "I sort of don't like the word fidelity, because that implies that there's a theological standard, you know, at some subconscious level, that there's this theological standard out there that we're all trying to be faithful to." A second PI, Trevor, echoed this concern, and said that he too was uncomfortable with the concept of fidelity, at least as it had been defined by implementation science.

Trevor wanted us to think about fidelity as a form of quality assurance centered on "improving practice" rather than "strictly adhering" to any one model. Unlike a pharmaceutical trial, he argued, studying fidelity in a "socio-behavioral" trial involves studying real people and professionals rather than the therapeutic mechanism of a drug: "But it's like that A to B [...] And what's the thing to be doing right, like, how do you get there [the outcome]? That's not the evidence base, the evidence base is that ACE inhibitors work for people with diabetes, right?" For Trevor, evidence is "in" the drug. Here, the drug and its evidence-based are seemingly separated from the labor and practices required to make it work in a given human, the administering, prescribing or, in the case of HaRP, the case management. Getting participants onto MOUD, on the other hand, is largely up to the practitioners who should leverage the suite of tools available to them, adhering to a few core elements:

> You know, I think if the goal is to get them into MOUD, right, then it's going to take a person-centered approach, and a peer is doing the needs assessment, identifying those needs and addressing them. And so, it may be that the intervention component is a really comprehensive assessment, and a toolkit of resources that could be used to address the particular barriers identified in the assessment, as opposed to, like, "at least four phone calls a month."

On the one hand, Trevor's idea could be read as wanting to afford practitioners flexibility and autonomy. In place of scripting how practitioners work or counting their visits, researchers, he argues, should largely leave it up to them on how to get participants onto MOUD. On the other hand, his thoughts on fidelity could also be read as sidelining the labor of HaRP practitioners and their expertise, indeed evidence is not just "in the drug" and a whole body of literature in evidence-based practices grounded in psychology and social work research has emerged to professionalize and bring evidence to social service labor. That is, if efficacy is simply "in the drug," then this would also discount the training, professional credentials and lived experience of the practitioners implementing them. Importantly, both Robert and Trevor were in effect urging the implementation scientists to consider allowing practitioners' some discretion and autonomy. Of course, this defies much of the implementation science literature as well as fidelity literature on socio-behavioral trials, but it also points to a level of flexibility that was aspired for, and yet, ultimately impossible due to the methodological constraints of the trial.

Both PIs were also alluding to common debates about evidence-based practices and the standardization of social service labor. They wanted to ensure that the intervention remained flexible and attuned to the preferences and autonomy of practitioners carrying out the intervention. Doing would also help them to avoid scrutiny. That is, establishing a high standard of fidelity would also mean placing the trials' perceived internal validity at risk if practitioners could not meet such standards. In fact, Robert had professed several times his negative experiences fielding past clinical trials in partnership with nonprofits that disseminate evidence-based programs. Trainers had admonished his research team for not meeting the standards of fidelity prescribed by the intervention, even though the research team felt they had simply been adjusting the program to fit their context.

The implementation team was also cognizant of how practioners might perceive their fidelity monitoring. Carl in particular was invested in the professional autonomy of practitioners who are otherwise often treated as "automatons" who "can't be trusted" in processes of ensuring fidelity. Here, she adhered to some of the sociological concerns about the fidelity tool. Similarly, Amy, in advocating for some flexibility, described the intervention less "like a cookie cutter intervention, that everybody's getting the same thing," but more like "a suite of things that we think are important that then [the practitioners] use based on what people need." It was up to practitioners to mold the intervention to the service needs of participants.

In this sense, the research team wanted to allow case managers and peers some autonomy in how they mobilized the suite of practices in which they were trained and that made up HaRP. But, if there was no routine or standardized implementation of the intervention, then how could the research team either establish its effectiveness or replicate it beyond the trial? The imperatives of evidence-based practice and trialing social interventions means that interventions must, to some degree, still be regimented across sites (Winther & Hillersdal, 2020). As stated by John, one of the implementation scientists, in an interview:

> But there does become a point where you have to say "these are the key principles that we're going to measure" and in terms of using that information, I think it needs to be used throughout the study as a form of audit and feedback. Basically fidelity is telling you what the black box is. And without fidelity, you know, if you come out of that, and you say "this did or didn't work," but who knows, if they were actually delivering the thing that you were thinking they were delivering?

Stabilizing an intervention is not just important for generalizability, but it is necessary for methodological integrity. Here, John is referring to type III error, which is defined as "failure to implement an intervention as planned, leading to an erroneous conclusion that null results are due to attributes of the intervention itself, rather than to its mal-implementation" (Akiba et al.,

2022, p. 335). Avoiding type III error is a commonly cited motivation for studying fidelity in clinical trials. Yet, the implementation science team's investment in fidelity wasn't just methodological.

Despite both PIs' investment in HaRP's flexibility, the implementation team also generally believed that those delivering "evidence-based" interventions do not always do so correctly. In fact, a common statistic motivating fidelity evaluations is that fewer than 50% of providers deliver evidence-based practices with fidelity after training (Wiltsey Stirman et al., 2012). This was a statistic, or at least some version of it, that supervisors would refer to when motivating the need for both fidelity as well as ongoing technical assistance. As such, the implementation science team felt that once HaRP was set free into the "real world" of social service provision, practitioners would require some mix of monitoring, encouragement, and accountability to ensure the programs' effectiveness. With both epistemological and managerial tasks, to stabilize and hold practitioners accountable, the implementation team would need to develop a fidelity protocol. But, as a complex intervention that relied on a variety of conceptually dense and amorphous practice values and principles, deciding what the intervention was, would also need to be established. However, such a definition would require difficult conversations and some political finesse.

Developing Fidelity to Loaded Concepts

HaRP had been laid out in a grant proposal, extensively presented over PowerPoints, and described in trainings for its practitioners. However, the intervention's specific contours would only be brought into relief by those most involved in carrying it out. It wasn't until the implementation team had started to develop the fidelity protocol that the intervention took "concrete" form. I spoke to Amy, the clinical supervisor, during the fidelity development process

in the winter of 2020, and her voice dropped to a whisper, as though she were delivering highly confidential information when she said, “HaRP is a shell of an intervention created for a grant” and “nobody seems to know what it is, or how we gonna assess it, or how we are gonna develop it.” Amy often worried about the institutional hierarchy of the study—decisions made “at the top” by PIs were left to be dealt with by those “on the ground,” providing services. The implementation researchers and clinical staff therefore had to contend with the bureaucratic inertia of, at least what they presumed to be, prior decisions made about the study’s design. Therefore, the team worked to establish a system for ensuring and measuring fidelity to an intervention they were only just beginning to conceptualize, a common feature of hybrid trials where researchers are meant to both stabilize and localize research practices simultaneously. This process was made even more difficult by the concepts the team was meant to operationalize.

HaRP was loaded with capacious concepts, including, but not limited to peer recovery and harm reduction, each of which tended to mean different things to different actors in different contexts (See introduction). Creating a fidelity protocol to concepts like harm reduction and peer recovery would mean concretizing amorphous, rather political concepts that might otherwise benefit from their imprecision and malleability. Indeed, without fidelity, people could attach their own meanings to the concept, but abide by different versions of it in their own discrete practices. This is, after all, how standards tend to function and thrive (Star & Griesemer, 1989). Permitting this relaxation in standards, however, could place the study’s integrity at risk.

Peer recovery, for one, is notoriously difficult to define. How the implementation team might measure fidelity to peer recovery therefore became a frequent topic of concern. Deciding on who constituted a peer was particularly challenging: Is a peer someone with experience of incarceration, substance use, or a family member of someone who has experienced these? Are

they actively using drugs or have they "recovered?" Other conversations centered around how practitioners might navigate the tensions inherent to their role as both professional researchers and people with lived experience. Fidelity was one way to–at least attempt to– manage this tension. That is, it is through the fidelity mechanism that such decisions are made and operationalized.

The implementation team, which generally deferred to Amy for her clinical expertise, was particularly concerned with ensuring that peers did not impose their own experiences–particularly their "recovery journey"–onto study participants. Amy often emphasized that practitioners should not "mistake similar for the same." To be sure, people often tend to rely on their own experiences of resilience to try and develop the trait in others who may not have the same material, social or emotional advantages to enact it. It was somewhat contradictory, then, that we came to define the primary role of the peer as their ability to "leverage their recovery journey." This placed peers in a complicated position: they were asked to use their recovery story, but not in such a way that might impose a particular narrative of recovery onto study participants. It was these types of contradictions, perhaps inherent to the role itself, that seemed difficult to manage yet seemingly important to grasp for the purposes of intervention stability.

As shown in the first chapter, harm reduction was also open to multiple, competing interpretations, for it is what Campbell (2020) describes as a "hybrid concept." In one meeting, Carl asked Amy to define harm reduction, and how we might consider measuring practitioner's adherence to it. She responded:

> Well, it's generally like a stance, right? A philosophical stance. There are some things we can operationalize such as, you know, "What are you offering when somebody is struggling with substance use? Are you saying, 'Oh, well, your options are to go to treatment or MOUD [medication for opioid use disorder], regardless of what people are interested in?' Or, are you offering a whole buffet of options?" There's no existing tool. But there are some things that I would want

> to look for, for example, "Are you offering options? Are you engaging in substance use management, planning and overdose, prevention, safety planning," that kind of stuff.

How would the team boil down a "philosophical stance" into a bulleted list of strategies that could be confirmed as having stayed true and faithful to the protocol? Harm reduction, for Amy, was about offering up a choice to study participants on how they might reduce their harm. This also was part of a concerted effort to refrain from imposing a particular recovery journey onto study participants, even if there were normative goals regarding their participation and movement through the study. Importantly, Amy did not offer up what goal this buffet of options was in service towards, and this was intentional. She had been insistent that while medication for opioid use disorder (MOUD) was a primary study outcome, it was not necessarily a clinical goal of HaRP. This was, as she had put it in a variety of ways over our many conversations, a way of honoring the self-determination of people who use drugs. It was also part of a common harm reduction stance that means to acknowledge "any positive change," versus the prescriptive goals of a given service provider, or in this case clinical trial. For a variety of reasons, and as we will see throughout the rest of this chapter, such "honoring" was tricky to do in the context of a clinical trial, particularly one that took place in the criminal-legal context. In a later implementation science meeting, Amy offered up a working definition of harm reduction for the fidelity evaluation, which included a range of possibilities and strategies aimed at reducing harm:

> Options for participants who use [opioids] include (1) not using by yourself; (2) not sharing needles with others; (3) having access to naloxone and knowing how to use it, (4) getting tested for HIV and HepC; (4) using needle exchange sites; and (6) incorporating constructive personal rituals.

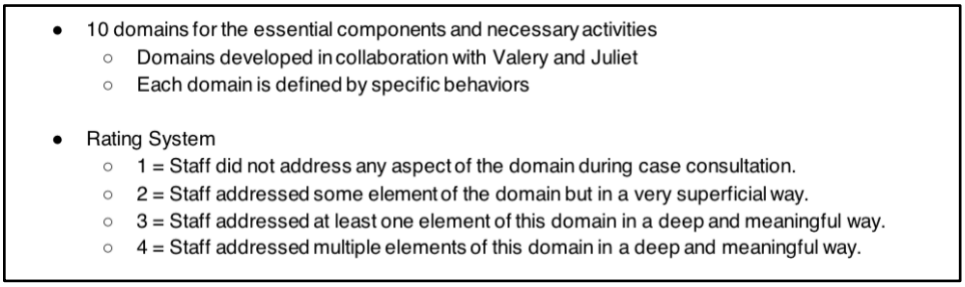
Carl immediately noticed that MOUD was missing from that list of strategies and reminded us that MOUD was a primary outcome of the study. I too had noticed the absence of MOUD and agreed with Carl that it should be included. Amy replied by admitting that "it's definitely our

goal," but she also insisted that at the same time "we're not pushing anybody into that." For Amy, meeting the study targets through MOUD ran against ensuring that the practices she supervised matched up to the values of harm reduction—not "pushing" anybody into a particular recovery trajectory. In response, Carl provided a detailed example showing how it might be possible to work with participants in a way that both offers up suggestions and validates a participant's decision not to pursue MOUD (or any recovery trajectory). It might mean, he stated, helping to prepare the participant for a future crisis and "stay supported" during the stressful time that was the COVID-19 pandemic. Amy replied by expressing concern at how "nuanced" practitioners would have to be in Carl's example, and how difficult it would be to not only set up a fidelity standard to harm reduction, but also then evaluate what Carl had proposed. This discussion brought to the fore the difficulty of both allowing autonomy while also reinforcing specific study goals.

Then of course came the methodological and technical complexity of measuring fidelity to HaRP, which would take place across the state, and in various types of locations as described more in depth in Chapter Two. This made capturing the intervention in real-time a difficult endeavor. The eventual solution was to record case consultations with case managers and peers led by Amy, the clinical supervisor. The implementation team decided that fidelity would be evaluated via case consultations as an "analogue" for how practitioners work with study participants. That is, how practitioners describe and talk about working with study participants might be a good indicator for what they do in real life practice. In clinical consultations the primary objective is to describe—or as my interlocutors put it, "unpack"—relevant aspects of a participants' background, concentrating on patterns of substance use, and then identifying goals and concrete, evidence-based strategies to attain them. Analyzing how practitioners

conceptualized participant problems, as well as staff's capacity to describe—mobilizing evidence-based practice terminology—how they would usher the client toward "positive change" became how the implementation science team would evaluate fidelity.

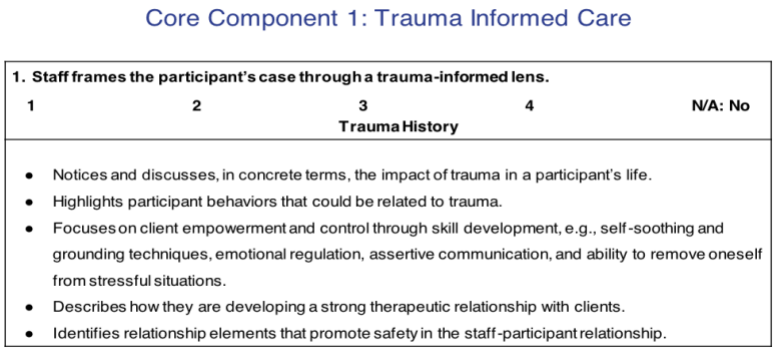
Figure 4: Fidelity Rating System



Boiling down capacious and politically loaded concepts like harm reduction and peer recovery into measurable activities or strategies, for many of us on the study, felt reductive. In fact, it was difficult–if not impossible–to maintain their essence once they were diluted to a list of behaviors or strategies. Figure 4 demonstrates how frontline staff's actions would be assessed for fidelity during case consultations with the example domain of trauma informed care. Practitioners would be rated on a scale from 1–4, with a 1 signifying that they did not describe or enact any of the necessary elements and a four signifying they described multiple elements in a "deep and meaningful way." Figure 5 demonstrates the list of elements that would be expected if the practitioner was adhering to trauma informed care. A practitioner might receive a four if they connect a traumatic life incident with current behavior that indicates trauma, rather than framing the participants' behavior as difficult or callous. The research team would therefore try to assess how practitioners contextualized and understood participant behavior and whether they were framing the case through a trauma informed lens. As much as fidelity adherence was a way of encouraging practitioners to perform particular practices, it was also a way of getting

practitioners to see and apprehend participants in a particular way— in this case, detecting and acknowledging the role of past traumatic events on a participant's current behavior.

Figure 5: Fidelity Rating System: Trauma Informed Care



Core Component 1: Trauma Informed Care

1. Staff frames the participant's case through a trauma-informed lens.

1	2	3	4	N/A: No
		Trauma History		

- Notices and discusses, in concrete terms, the impact of trauma in a participant's life.
- Highlights participant behaviors that could be related to trauma.
- Focuses on client empowerment and control through skill development, e.g., self-soothing and grounding techniques, emotional regulation, assertive communication, and ability to remove oneself from stressful situations.
- Describes how they are developing a strong therapeutic relationship with clients.
- Identifies relationship elements that promote safety in the staff-participant relationship.

To meet fidelity, staff needed to reach 80% of the total score, a number that had been suggested to the implementation team by a consultant on the grant. Among the developers, including myself, discussions about how to enumerate this score were often followed by giggling and wry smiles, as though we were all in on the same joke that we rarely confronted until we had to: the comic and absurd challenge of quantifying real-life practices and interactions in this way. Some fidelity standards reflected a high level of rigidity, and some were so vague and subjective they required the trained eye of a seasoned clinician to evaluate. But more importantly, our conversations reflected aspirations for a level of flexibility that was not necessarily represented by the final outputs of the assessment.

Practitioner Reception of the Fidelity Protocol

The training on fidelity took place in May of 2021, just a few months before recruitment would begin. The training oriented practitioners toward how their work would be evaluated and stabilized during the HaRP clinical trial. As I will show in the following section, as practitioners were introduced to the fidelity system, they felt that they were being asked to do too much in the

critical moment of building a relationship with a given study participant. Stabilizing the work of peers in this way seemed antithetical to their role and professed function. In response, implementation scientists would refer to the "gestalt" and "spirit" of HaRP as a way of providing flexibility. This led to questions about how tacit nature of Harp could be measured.

Consider Marvin, who felt the fidelity system seemed divorced from the reality of his daily work, particularly given the critical moment in a person's life at which he was meant to forge a working relationship: a person's release from jail during the midst of a pandemic. After learning about the various evidence-based practices peers would be asked to deploy in meetings with study participants, he felt overwhelmed:

> So, I'm trying to figure out what it's going to be like, because I have no clue what kind of trauma people are coming out of the correctional system with, and what that looks like in combination with the pandemic, Black Lives Matter, exceptional isolation, because there have been no visitors. And there's all this preset data and preset of fidelity. And I'm trying to work out in my head how it all fits together. So much has been generated without that in mind. So, when it comes to actual implementation of what it's going to look like, it's… it is overwhelming. Yeah. And I'm a pretty smart guy, and a hard worker. But in my brain trying to put all this together, it's . . . something.

For Marvin, the practices encouraged through the fidelity protocol seemed divorced from what he perceived to be the needs of study participants exiting jail, particularly during a moment of political struggle, civil rights activism, and a pandemic. Marvin, as he is expected to do through his role as a peer, points to the context of jail, the isolation, suffering and anger, and wonders aloud to the study team why the fidelity system seemed to not account for this. How could practitioners be expected to integrate these techniques during such a critical time in someone's life, at such a crucial moment in the working relationship?

Important here is how Amy replied to Marvin's exasperated "something." "I hope that what we've designed," she said, "is an intervention that has these elements that can hold

whatever is thrown at it." This meant, she went on, that "having a trauma-informed framework should allow you to hold whatever kind of trauma is thrown our way." In other words, it was an arsenal of tools that practitioners were expected to pull from as they engaged a given participant. The trauma-informed lens is one such tool that could help attend to structural context. As was often the case when the research team was faced with the structural conditions of participant predicaments, they would refer to the trauma-informed lens as a way of gesturing towards or dealing with structural problems that seemed altogether too complex and all-consuming for behavioral intervention alone. Using a model that itself diagnoses participant problems as environmentally produced, a trauma-informed lens became one way of attending to practitioners' concerns. It was also frequently the case that when there appeared to be a disconnect between what was developed for practitioners, and what practitioners felt was feasible, researchers would count on those practitioners who were often glossed as "on the ground" to "bring that context in"—that is, to explain how the research or the intervention might be improved with "place" in mind. As Carl would ask, amid this line of questioning by Marvin and others: "So, when you think about operating in a not business as usual, America, our question for you is, what would you like to see added?"

In an interview conducted shortly after Carl's training on fidelity, I followed up with Marvin about his concerns with the fidelity system. They didn't seem resolved by the discussion we had a week prior about leveraging trauma informed care: "The world has changed since the initial conception of this project," Marvin said. The fidelity system, he argued to me, would place undue pressure on him, and the practitioners he worked with, to do something, expect something, and collect information for the purposes of meeting fidelity targets, rather than what he thought was warranted at the moment of participant engagement. He wanted to "allow the

client to kind of vent a little bit," particularly considering that "as they leave the penal institution system, this may be their one connection to support or a formidable form of support." He concluded, "so I think that because that wasn't built in, it would be really good to build that in." All this is to say, there was an overarching sense that the fidelity system was overprescribed. Practitioners felt they needed to follow the guidelines of the trial, asking and prompting participants to reveal information, over allowing time for active listening. Tensions like this one were only amplified by the hierarchical nature of a federally funded and academic-led experiment, where practitioners, in this case peers and case managers, were less able to marshal the discretionary power assumed to be awarded to practitioners facing the standardization of their labor. Therefore, they felt a sense of anxiety to meet the outcomes, outputs and fidelity targets mandated by the clinical trial.

The Spirit of HaRP

Practitioners being trained on the fidelity system were also anxious about the consequences of having their language, or how they talk about a case, evaluated during case consultations. In one meeting, Carl, after going through how we would evaluate the case consultations, opened the floor for questions. A long pause ensued, verging on uncomfortable silence, until Carl broke it, to say he took their silence as a "no," that there were no further questions. Again, it was Marvin who jumped in to say he was processing this information and that it was "definitely not a no." He went on to ask the research team to clarify precisely how practitioners' language would be evaluated for fidelity. Marvin, as well as other practitioners who I spoke with, were concerned about enacting specific evidence-based practices the "right way" by using the correct, or fidelity-consistent language. During one interview, directly after

the fidelity training, a case manager named Natalie revealed the anxiety it induced as she began to strategize how she might manage the pressures of fidelity:

> Sometimes we don't mention everything, right? And Amy is able to identify which things those are. So, I had to really think about that when I have those conversations and they're being recorded, to shift gears or do some code switching to make sure that I need at least . . . I think it's . . . 80% of the criteria you have to meet. So, I was thinking about how to do that. And, as I had conversations with others, we were having some of the same thoughts. Okay, how do we do this, and make sure that we say that's what we're doing?"

It was often in moments like these, times when the implementation researchers had to confront the reality of standardizing practitioners' work as prescribed by the fidelity system, that Amy and the team would invoke more intangible, less easily graspable concepts like "the spirit," or the "gestalt" of HaRP. Carl replied to Marvin during the fidelity training:

> I think we're less concerned about language and more concerned about the spirit with which folks are talking because there's variability in how everybody talks and describes their relationship with their clients. They're really looking for the spirit of, you know, what is harm reduction.

In this case, it wasn't so much the language used by practitioners but the spirit with which they carry out their work with study participants. It is common for evidence-based practices to embody both spiritual and scientific elements. Carr, in her ethnographic study of Motivational Interviewing (MI), an evidence-based practice that was used in HaRP, showed how, rather than a set of techniques, MI tends to be thought of as both a "mind-set and heart-set" (Miller & Rollnick, 2013, p. 14 as cited in Carr, 2023, p. 99). Performing MI, and doing so with MI spirit, is much more than using the right techniques or phrases. Instead, the spirit "manifests in interaction, subject to what others can see, hear and feel" (p. 99). While Amy and others were not directly referring to MI when calling upon the spirit of HaRP, they were equally invested in the underlying "heart" of the model, the intangible aspects that Amy could perhaps see and feel

but not directly measure. Not to mention, MI was an active evidence-based component to the HaRP model.

As a seasoned clinician, Amy was likely aware of just how integral the therapeutic relationship is to successful mental health treatment, however defined, she said as much during the training. One could also argue that invoking "the spirit" was also an attempt at providing flexibility and honoring practitioner autonomy amid concerns about professional overreach and practitioners' loss of discretion, a problem Amy herself experienced as she frequently lamented having little control over the design of the intervention. Providing flexibility was one way of maintaining buy-in with staff, even if it meant suspending, at least momentarily, a dedication to uniformity and the methodological demands of fidelity. Of course, it was never totally clear, at least to me, what "the spirit" was referring to, which left me wondering how this would, or even could, be measured.

In a later Zoom call in the Summer of 2022, I followed up with Amy, asking her how she understood "the spirit" of HaRP. She replied:

> I mean, I think it's, it's total radical, compassionate acceptance of people where they're at and you know, that they have value and deserve respect, no matter what choices they're making. I think that goes for both harm reduction and motivational interviewing and a belief in people's ability to survive, make changes and take care of themselves and in one another. Like all of that is sort of the stance, the spirit in the stance, regardless of what we might be working on with somebody at that time.

So, the spirit of the intervention, what Amy and Carl cared about first and foremost, wasn't necessarily whether practitioners adhered precisely to each bullet point in the fidelity protocol. Despite the extensive strategies outlined in the fidelity form, there was, at its core, a right way to do this intervention, one that relied primarily on the therapeutic relationship, having "radical compassion" and approaching individuals, no matter their story, with a lack of judgment. But

how could this be measured, particularly using the "analogue" of case consultations? This was a question that practitioners kept returning to when discussing the standardization of their work.

In conversations with practitioners, there was a sense that the fidelity we created would miss important mechanisms of the therapeutic process. In my conversations with practitioners, they described how their work with participants relied on nonverbal signals, sensing, and other complex but imperceptible interactive processes that would seemingly be impossible to measure yet essential to the very mechanisms of their work. How would the study capture the relatively elusive nature of relationship building, something you might see in action but fail to define? Marvin, in an interview with me, put it like this:

> It's kinda like, you almost need to be a fly on the wall to see that happen? Yeah. Yeah, that's what I say about that. Because I think human beings are allowed. Like insects, we all have some vibration, some energy, just like ants, they walk past and they can buzzer do whatever they do to share energy. And it's, it's simply accepted.

Like all tacit and embodied knowledge, the work of engagement is difficult to capture with language or written standards like fidelity. Instead, it's something we see "in its action," as Marvin describes, using the example of a "fly on the wall" (Polanyi, 2002, p. 49). The implementation team was thus torn between ensuring that HaRP was socially relevant and that we stick to what matters: building a therapeutic relationship. But the pressure to stabilize and ensure that this is done in a way that is somehow standardized, would mean efforts needed to be put in place to evaluate how these relationships are developed and maintained, a critical mechanism of HaRP and, in particular, the role of the peer. Yet, the essence of peer work was hard to capture and the very process of measuring it, appeared to change it. As I will show in the next section, the clinical trial context often compromised practitioners' careful efforts to promote participants' self-determination.

Fidelity in Practice: Compromised by the Protocol

As we saw with Marvin, practitioners were troubled by the discrepancy between their lived experience and "on the ground" expertise with what was demanded of them by the study protocol. Heather, a peer on the intervention team, grew up in Sherwood County and had extensive experience dealing with the criminal-legal system having had family in and out of jail and prison for drug-related offenses since she was young. After she lost her job during the initial wave of COVID in 2020, she volunteered for Meals on Wheels and was then hired on the study part-time through a state grant. Heather was recruited to be part of the study because she worked with a local harm reduction organization in Sherwood County called Love, Inc. (LHR). Her decision to take on the role came, however, with some hesitation. She was reluctant to turn activities she had long done as a matter of course, throughout her life, into a profession. Heather had helped her own family and friends navigate social service and carceral bureaucracies for decades.

Though relatively new to the role of a peer worker, before being trained through the study, she was experienced in the delicate art of "hanging out and listening", as she put it, because "people just need to talk" and "talk to themselves through me." She confessed: "You know, I'm not a big question asker." Heather's description of her support process sounded a lot like Marvin's, allowing people the space to vent rather than directing them toward normative therapeutic or social service goals. And, while Heather might explain her work in this way, what was also clear in talking with her was that she had very specific ideas, based on her practice experience, about how to most effectively build relationships with service users at LHR. And, more specifically, people who use drugs who had been caught in the criminal-legal system. Relationships, and the people "walking in the door," came first. This mode of engagement,

however, didn't seem plausible once Heather started working for HaRP. She said as much during an interview when she stated:

> As a peer worker, like, their interaction with me, I guess, it's hard to define. If we're talking about LHR I don't chase people. You know what I mean? Like, it's totally voluntary. […] Someone comes to me because they are like, "Hey, can you help me with this." And then if they ghost me, they come back in two months. But like, with the study part of it, it's a little different because we're required to do this follow up and required to try to get a hold of them so it does kind of change the dynamic of like the recovery coach. Um, but yeah, sometimes chasing people down just does the opposite. [..] I don't hound them.

The study was asking her to do her work differently than she would otherwise at LHR. Given her experience, pursuing participants, hounding them, was counterintuitive. That is, she was accustomed to allowing people to come to her at LHR. She also felt that those entering the study don't "know what they've signed up for," compared to at LHR, where people walk in the door asking for support. Heather, and others too, were uncomfortable with carrying out certain organizational imperatives prescribed by the clinical trial. This tension became more acute in the context of recruiting in a jail, where participants had complex, material needs and faced structural hurdles in attaining them. Despite the philosophical imperative of meeting people where they are at, recruiting and engaging participants overwhelmingly became the primary focus of case consultations early in the trial. In Heather's experience, effective support meant not chasing. But many of the study's roleplays centered on finding clients who weren't answering calls, identifying strategies for following up with those who "ghosted," and being taught ways to democratically identify goals for participants who didn't appear to have any besides the obvious: getting out of jail.

While American jails are increasingly recognized as sites for treatment and service provision, particularly in rural contexts, it's important to emphasize, though obvious, that participants are there against their will. In Sherwood County, participants are shackled to a chair

in jail as they consented to the study. This was described rather cavalierly one day by Helen, project manager, as she reported back to the research team her initial thoughts about the study enrollment process in the jail. Some participants, she says, show interest in the study out of abject boredom. The two-hour interview of personal questions seemed more enjoyable than sitting in the cell with nothing to do. In-person visits, as opposed to Zoom visits, Heather told me, provide the opportunity for incarcerated participants to "walk further down the hallway" providing a small, albeit meaningful, sense of freedom. Practitioners must therefore manage the imperatives of recruitment, meeting service provision requirements made more commonplace with drives toward accountability, all while "meeting people where they are at." Meeting certain targets might be good for study outcomes, or meeting fidelity, though they might also place undue pressure on practitioners who see other, more important, immediate needs to take to task, including taking the time to listen and build rapport.

As indicated in the sections above, fidelity is often conceived of as that which stabilizes therapeutic practices—a mechanism to ensure that interventions can be tested and made transferable to new contexts. What I found, however, was that the fidelity process, and the practices it promoted, also attempts to stabilize practitioners, obscuring and smoothing over their concerns with the ways HaRP, and the clinical trial apparatus, fundamentally transformed the therapeutic relationship. Take, for example, motivational interviewing (MI), a main component of HaRP and a popular therapeutic modality found in the United States and beyond. MI is often deployed to establish goals in a way that, at least at the surface, appear client driven. In training, MI is touted as an intervention that will make things easier for the practitioner—the promise being that, with the use of MI strategies, they will no longer have to "fight" with their clients to instill change. It does so by trying to "equalize the interactional terms of once hierarchical

relations," between professional service providers and service users (Carr, 2023, p. 56). Amy, the clinical supervisor and the MI trainer, described it this way: "the beauty of motivational interviewing is that it's not confrontational, right? It's more curious and getting the client to confront themselves." The promise of MI is that it can help produce change seemingly without the explicit direction of its practitioners.

These MI strategies would be promoted through the HaRP trial and through the mechanism of fidelity in particular. During a clinical consultation, early in the study's implementation, Heather described how she struggled to engage a study participant still incarcerated in Sherwood County jail. She characterized this participant as "very guarded." Heather felt the jail context, where she was meant to forge a therapeutic relationship, was not a conducive environment for participants to share their story. After all, people with criminal-legal histories, she argues, are often encouraged to keep quiet and not "open up," particularly while they are incarcerated. Heather was drawing on her own "lived" experience to suggest that she needed to take her time with this participant. Heather felt stifled.

During the case consultation, Heather wanted guidance from Amy and her colleagues on how to gather the information needed for the study and to establish clinical goals. She wanted to do so, however, without pressuring the participant to disclose information, given the participant did not appear comfortable sharing what Heather surmised as painful moments of her past: "she slips facts at you. Without emotion. [...] and then will deflect and tell a funny joke, or bring it somewhere else immediately." To this, Amy replied:

> Which, Heather, is a really good indicator of trauma, right? People are telling you the most horrific things you've ever heard with totally flat affect, right? They're not, they don't have the appropriate affect that you would expect. That's a coping mechanism. It isn't bad. It's what keeps people alive and functioning, right? And the use of humor to deflect, Be–AU-tiful.

Heather:

> So like, so do I . . . do I just like, what's the best approach? Do I just let that play out the way it does? Or do I try to guide it? Me being me. I'm just like, I hear what you tell me. And whatever else you want to tell me. I'll be here to listen, but I'm not big on asking questions about "how did your mom die?" You know what I mean? I'm more like, "I understand she passed away." And then we went, we went somewhere else. Yeah, rather than, like, "I'm interested . . . I'm curious," but I don't want to ever ask questions that are going to shut somebody down as well.

Heather's worries are striking for a number of reasons. First, she's doubtful that she can convincingly and effectively deploy MI lines of questioning, as she readily lists them: "I'm interested . . . I'm curious." She suggests that certain modes of information gathering inhibit her capacity to listen. And, perhaps most importantly, she is concerned that asking questions or "digging," as she later describes the act of probing, an MI strategy, might actually shut the participant down, thereby harming the therapeutic relationship—the very goal of her assumed role as peer.

Yet, the clinical trial engenders a sense of pressure to gather information on participants, be it through formalized modes of data collection or thoroughly delivering the details of a clients' case in service toward identifying and meeting study goals. As such, peers and case managers are urged to ask more and know more. MI provides a way to do so, ostensibly, without confrontation. In response to Heather's apprehension, Amy suggests that Heather ask permission (another MI strategy) from the study participant to probe further about the participant's mother: "Would it be okay, if I asked you to tell me a little bit more about your mom? And you know, when did she die? How did she die? What impact did that have on you?" Heather still wasn't convinced:

> I think the world would be very different if I knew [the participant] didn't have to go back to that pod. Yeah, like if I knew she was gonna go home to a safe place somewhere where she can go home and cry. You know, I think that uh, I know, sometimes I just feel like, I don't ask enough questions. But I have not felt comfortable digging somewhere where it feels like it hurts.

Again, Heather suspects asking more might actively harm the participant, particularly in the participants' current context of having to share a jail cell (or pod) with several women with whom she doesn't get along, as she elsewhere described during case consultations. Having visited the jail extensively both for friends, family, and her work, she was intimately aware of the lack of privacy in "the pod" and the potential harm of unsettling the study participant emotionally before she returned. In other words, she continues to feel uncomfortable about the line of engagement prescribed by the HaRP intervention to achieve fidelity targets.

Amy offered another set of MI skills that she felt could bring Heather closer to gaining access to the participant's story, while also smoothing over Heather's discomfort with "prying." Amy encourages Heather to use a reflection based on the participant's story. A reflection is often used as a "check-in" with clients, to see if practitioners have accurately grasped the client's situation. The check-in also projects the image of a professional wholly dependent on the client's verbalizations, what they say about their own experience (Carr, 2023). This helps to cast the client as the expert, rather than the practitioner. Amy suggests that, instead of an open-ended question, that Heather use a reflection:

> Yeah, and questions do feel like digging and so I'm going to remind you of motivational interviewing skills and like while open ended questions are the best kinds of questions, but reflections can be just as powerful and feel less like you're digging in, right? So like, reflecting like, wow, you've, you've seen a lot of horrifying things in your life and for her to be like, yeah, life has been a horror, right? Like, just that reflection. That's, that's less like poke, poke, dig, dig. And it's more like I'm feeling you, I am here feeling with you.

Here, MI strategies are introduced to assuage practitioner concerns with prying, particularly in the context of jail recruitment. While Amy suggests a reflection might also be used as validation, the idea is that the client–or in this case the participant– are the experts of their own experience. As described earlier, MI is an evidence-based practice that tries to even the playing field between client and provider. Yet, if Heather's role is to understand the participant's context, having lived

it herself, the idea of projecting inexpertise and naivete, might compromise her role as someone who has lived through similar experiences, the very essential element that would help her to ostensibly forge a working alliance. In some ways, this way of engaging encourages Heather to elide the very context peers are meant to "bring in" to the working relationship. Heather felt as though she knew where the participant was at, and she was not meeting her there.

Conclusion

A key difficulty facing the HaRP study and the development and implementation of the fidelity measure were multiple—and sometimes contradicting—therapeutic modalities and professional commitments, which inevitably had to be sorted out 'on the ground.' Such 'hybridity' is not just particular to HaRP, but is instead conditioned by funding structures and bureaucratic interests, which call for the amalgamation of evidence-based programs often in the name of innovation. Sociological critiques of clinical trials and the evidence-based programs that they produce often center on their potential to undermine the autonomy of social service organizations and the clinical authority of professionals who inhabit them (Mosley, Marwell & Ybarra, 2019). Meanwhile, scholars of STS have shown the ways in which the very act of researching something, tends to transform it in meaningful ways. Yet, they also require flexibility to properly function. Here, I show how the trial, from the perspective of practitioners themselves, fundamentally remade and transformed clinical care. This brought to the fore the question of what the HaRP intervention was after being transformed by the trial apparatus and its fielding in the carceral context, despite all efforts at flexibility.

Given what has been written about evidence-based practices and their assumed effects on social service professionals, I was surprised by the level of flexibility that was promoted by most researchers on the HaRP trial. Yet, in talking with practitioners, and observing case

consultations, it became clear that practitioners still felt constrained. After all, one could argue that every practice or policy intervention is infused with a moral code or "script," to use Akrich's (1996) terminology. Akrich (1996) argues that technological innovations bring with them a vision or a script for how they will play out in the real world. She describes how, in creating new technologies, designers also imagine a particular public and the contexts that will take them up. Designers imagine "specific tastes, competences, motives, aspirations, political prejudices [...] and a large part of the work of innovators is that of 'inscribing' this vision or prediction in the technical content of the new object" (p. 2018). And, while the therapeutic practices suggested or at times mandated by fidelity protocols may be developed with users in mind through imperatives of flexibility and 'responsiveness', in practice they did not always align with the political inclinations, motives, or morals of its intended users. This tension brings to light the potential limitations of implementation science concepts like "the adaptable periphery," and how, they may fall short given certain political elements are infused into a given intervention and the context where it is delivered.

Practitioners on the HaRP trial described an ideal form of engagement that did not seem possible through the HaRP trial. They described a form of care that emphasized attending to people, and participant context, before hastily following protocols. Such accounts match up to anthropological accounts of care, which emphasize attention (Mol, 2008). Qualitative studies of those engaged in connective labor have found that this type of engagement, "mirroring another person's emotional perspective" as one of the more meaningful aspects to their work, allowing them to "witness others [e.g., clients or students] not just to teach or heal but in order to promote understanding, purpose, and dignity" (Pugh, 2023, p. 1780). What HaRP practitioners found

most meaningful about their work, and arguably what is potentially most "efficacious" about their role, seemed compromised by the study at hand.

Lastly, I argue that what practitioners did in fidelity-consistent language not only brought them closer to fidelity, but it also had the effect of diffusing–or at the very least obscuring–perceived tensions that surfaced during implementation. This suggests that fidelity did not just serve the instrumental goal of stabilizing peer labor, but also helped to stabilize the very people performing evidence-based practices. By this I mean, the fidelity protocol encouraged practices that helped to resolve or obscure ideological conflict. More generally, it suggests that certain evidence-based practices, and their stabilization through the mechanism of fidelity, may have extra-scientific effects beyond simply establishing a uniform set of practices.

As federal funders and service providers increasingly prioritize building trust and credibility with service users, peer work and other forms of care, such as harm reduction, are gaining prominence for their potential to facilitate this process. It is worth considering how their integration into such systems might change the very essence of their role and to what extent their work might defy the types of standards and outputs necessary to render it amenable to an intervention RCT like HaRP. Indeed, scholars of performance management have documented how performance mandates that solicit nonprofit professionalization suppress community-oriented perspectives resulting in goal incongruity among staff and service recipients, ultimately affecting nonprofits community legitimacy (Smith & Lipsky, 1993). At the same time, such formal integration, as described in Chapter Two, might bring about the material benefits of professionalization. At the very least, more attention is merited to investigate how peer workers and harm reductionists entering these types of roles might cope with these transformations, or

how we might better apprehend the “value” of their work (Barrenger, Stanhope & Miller, 2019).

Conclusion

This dissertation analyzed the social processes of fielding a complex, multi-sited clinical trial, and more broadly the social processes of scientific reform. It focused on the ethical, political and epistemological challenges of negotiating the imperatives of scientific inquiry with social and contextual relevance. Throughout, I examined how researchers and providers negotiated moral and political tensions where the imperatives of research, clinical practice, and activism converged. The organizing question of this dissertation was: How, and with what social effects, do researchers negotiate the principles of scientific inquiry with the exigencies of the "real world," and the diverse set of organizational commitments that make it up?

I answered this question through analyzing the case of the Harm Reduction Peer Recovery (HaRP) clinical trial, an ongoing hybrid effectiveness-implementation clinical trial that brought together public health, criminal-legal, and harm reduction actors working to stem the tide of overdose in their communities. Throughout, this dissertation highlighted how local context and scale, objectivity and activism, and standardization and tacit, embodied forms of care were negotiated with experimental logics like standardization, fidelity and objectivity. Ultimately, I argued that the imperatives of the clinical trial, coupled with the institutional obligations to which it was tethered, often ran against the ethical and political commitments of practitioners and researchers themselves regarding the most effective strategies for preventing overdose.

Social movement actors and activists have long maintained an ambivalent relationship to researchers and have even taken on antagonistic stances towards those fielding RCTs. While HaRP practitioners were not activists per se, many did bring a social movement background and approach to their work, having been hired from local harm reduction organizations. Critically,

they aspired to effect change in ways that could trouble researchers' relationships with criminal-legal stakeholders and cast doubt upon the neutrality of the experimental findings. While sociological critiques of RCTs have argued that researchers' perceived objectivity might marginalize other ways of validating social policy on a macro scale, I show how this chilling effect occurred in real-time, when the experimental commitments overrode alternative approaches to social change.

In Chapter One, we observed that the potential pathways for social change were constrained by the entrenched institutional commitments within academic research, the perceived culture prevalent among stakeholders in the carceral system, and the preferences of hypothetical policymakers seeking politically impartial evidence. Indeed, while researchers hold political opinions, they must also attempt to remain politically neutral in order to sustain their "neutral advice-giving capacity." (Hess, 2016, p. 122). Consequently, Robert tended to prioritize outcomes that would resonate with policymakers and stakeholders in the carceral system, such as cost reduction and crime reduction, even if he and others on the study team were not inspired by these specific outcomes. Meanwhile, in Chapter Two, we saw how housing case management was not provided to maintain the model, when other times the model was "let loose" for purposes of accomplishing the trial and meeting the immediate needs of certain participants.

Throughout, I demonstrated how implementation science's growing investment in confronting structural inequity has set implementation researchers and the practitioners they advocate for in opposition to the scientific norms like impartiality often pursued in clinical trials, particularly those taking place in politically contentious contexts like the criminal-legal system. Implementation science was stymied in its endeavors to effect structural change by needing to appeal to both the commitments of scientific neutrality as well as to criminal-legal stakeholders.

Yet, there were also aspects to the field itself that lead to this "stagnation." Even as implementation researchers were mobilized to develop implementation strategies for Naloxone's acceptance among Sherwood police, issues of advocacy, presented with a sense of urgency, were subjected to the prolonged and iterative process of academic research. Structural problems were broken down into smaller, technical ones, often to the dismay of HaRP practitioners.

I also argued that efforts to accomplish the trial took precedence over "real-world" concerns such as feasibility, and sometimes even scientific commitments like ensuring the model's integrity. Everyday decisions regarding the model tended to prioritize completing the trial at the expense of HaRP's long-term future. In some ways, this made HaRP both more responsive, but also less likely to "survive" the trial itself, bringing into question the very point of the trial among implementation actors themselves. As described in Chapter Two, HaRP researchers encountered a dilemma: the more control and restriction imposed on the model, the less "responsive" it became to the needs expressed by trial participants. However, allowing the model to operate freely posed a threat to the integrity of the RCT model. Even so, I demonstrated how the imminent needs of study participants led trial actors to compromise model stability and feasibility. Critically, I argue that current theoretical models and frameworks do not account for the political and moral negotiations required to produce a scalable intervention. Implementation actors and researchers alike faced numerous competing obligations as they defined the contours of the intervention versus its context. Such negotiations are critical to apprehend as we attempt to develop and scale complex interventions like HaRP.

Lastly, I demonstrated how the trial, as perceived by practitioners themselves, fundamentally reshaped and transformed clinical care. This raised the question of what the HaRP intervention *was*, once altered by the trial apparatus and its implementation in the carceral

setting. Researchers promoted flexibility and responsiveness to the local trial contexts. However, in practice, such interventions, even with adaptations, may not consistently align with the political inclinations, motivations, or values of practitioners. This tension highlights the potential limitations of "essential ingredients," and underscores how ideological elements infused into an intervention may hinder adaptation or the "adaptable periphery" from fully realizing the locally relevant promise of adaptability.

Together, each chapter demonstrated the ways in which the "real world" is managed through science, and to what effects. The pursuit of objectivity, generalizability, and control were met with the vagaries of the real world: local activism, contextual idiosyncrasies and versions of care and social intervention that failed to match up to the logics inherent to the clinical trial and the carceral contexts where it took place. While the commitments of science were not uniformly abided by, as we saw in Chapter Two, accomplishing the clinical trial, tended to override "real world" relevance like feasibility, even if this was not a goal all trial actors aspired to.

What We Learn from a Given "Case"

The tensions and practices that are addressed in this dissertation, while one "case," are not particular to the HaRP trial. Certainly, HaRP researchers were presented with remarkably unique challenges. HaRP began recruitment in 2021—about one year into the pandemic. Arguably, this is precisely the one quality that all contexts have in common: they are mutable, diverse, and ever-changing. No matter what a particular context has in store, researchers, and particularly implementation scientists, are tasked with managing and accounting for it. Perhaps it is the unprecedented context I outlined above that brings into sharp relief the focus, and central tension of this dissertation: how researchers might deal with context while also attempting to

pursue science. That is, the pandemic, and other social processes, underscore the practices imbricated in managing context: the world is transformed overnight and yet researchers persevere to create something that is ostensibly scalable. It demonstrated the strength of the scientific impulse to endure despite complexity.

A second critique about the generalizability of my findings might be that there was something particular about the researchers I studied, and that perhaps they did not do science correctly or made errors that astute intervention or implementation scientists might question. As a social work researcher with experience in intervention science and deeply engaged in the implementation science literature, I was surprised by some of the choices made by researchers on the project. As a research assistant, I did try to deliver my opinion even if it was not always heard. In saying this, I want to emphasize that this dissertation is not a critique of how "good" these scientists are, but instead I hope that we look at it as an investigation 'under the hood' of some of the tensions of dealing with the "real world" scientifically, a perennial issue for applied social scientists. Second, by most academic metrics, the researchers I studied were wildly successful scientists and the funding they procured to deliver HaRP was competitive, and largely only given to accomplished researchers. Lastly, suggesting HaRP's researchers made mistakes, conceding variance, reinforces the idea that there is *no* standard science. That is, the invocation of "bad science" itself suggests that scientific practices tend to vary in meaningful ways. What might be important, then, is to investigate how and why scientific practices might vary. What might be generalizable are the types of practices and negotiations that are required in producing it, the central focus of my study.

Implications

Framing

As we know from policy literature, the strength of a given policy solution has a lot to do with the way it is framed in the first place. The way policies are framed also dictates who can (or wants to) be at the table. Schattschneider, for example, claims that "*nearly all theories about politics* [are about] *who can get into the fight and who is to be excluded*" (1975, p. 20).[1] This is done, for example, by constraining what is defined as socially sanctioned, legitimate problems (Guidry & Sawyer, 2003) or more simply a matter of representational bias (Gilens, 2005). As with the majority of socio-behavioral clinical trials, and as explained by Robert in Chapter One, the causal story of RCTs tend to be framed at the scale of individual behavior. As has been argued within the sociological literature, focusing on individual behavior disentangles the problem framing from more structural issues at hand (Adams 2016; Ramachandran, 2020). This helps researchers to develop ostensibly politically neutral advice and maintain credibility in the political field. Critics have also argued that it may limit acceptable forms of social change, as only specific types of social change tend to be validated through the mechanism of the RCT. My dissertation shows how trial actors, even within a single study team, may diverge on how they desire to pursue meaningful social change. Trial actors felt constrained by the commitments of the clinical trial apparatus even as they desired to intervene through mechanisms that might supersede it.

As described in the introduction to this dissertation, the larger federal initiative that funded HaRP understood the rising rate of opioid overdose among populations released from jail

[1] It is important to consider the manpower, the social capital, and the immense bureaucratic and scientific infrastructures required to be awarded and pursue projects with NIH funding. In the application process for NIH grants, researchers must prove their capacity, and connections, including their past procurement of federal grants, as well as mentorship by recognized scientists in the field. Seventy five percent of NIH fundings goes to only 13 schools of social work (Corvo et al., 2008). The review process is largely black boxed and grant reviewers are selected from the pool of funded researchers, which only serves to perpetuate the exclusion of those who have historically been under-funded as well as stifle innovation by re-funding research and supporting research that is already ideologically aligned.

or prison as due to a loss of tolerance during their confinement. The framing of the problem meant federal resources would be spent trying to improve availability of medication for opioid use disorder (MOUD) in and around American jails and prisons. Inevitably, this meant investing in jails and prisons. An alternative way of approaching the rise of overdose among people recently released would be to reconsider the criminalization of drug use and enforced abstinence in the first place. The way the HaRP trial was framed and its inevitable investment and partnership with jails was a fundamental source of tension that surfaced throughout the trial. For some harm reductionists and peer workers, it meant working in the very place that they understood as a fundamental source of harm (Chapter Three). Meanwhile, as described in Chapter One, harm reductionists and implementation researchers desired to pursue advocacy, however these efforts were deferred out of the need to maintain buy-in with jails and future policymakers. While hybrid designs and implementation science have emerged to better contextualize a given intervention, more attention should be placed on the original study design, and how seemingly simple or politically neutral questions might, once on the ground, be a source of political or ethical tension.

The original framing of the HaRP trial was circumscribed at the start by the intervention's design. However, my time with trial researchers revealed that developing HaRP and selecting outcomes, though shaped by the RFA, was also a process of negotiation with jail stakeholders. Researchers spent time anticipating the needs and desires of jail administrators in order to secure buy-in, so researchers needed to select outcomes that were of interest to jail stakeholders. This meant focusing on crime, cost-benefit, and recidivism, a commonly used statistic to measure if someone released into the community has been re-incarcerated (Chapter One). A concentration on recidivism, however, as has been argued elsewhere, is an outcome that

both “reinforces an attachment to law enforcement and surveillance” while also undermining the broader, more positive goals of “bolstering education and the welfare state, investing in economic development, and restoring citizenship” (Gottschalk, 2016 p. 101). That is, focusing on recidivism also means investing in the infrastructures and personnel that support its production (e.g., surveillance systems). As argued in Chapter One, it also deflects attention away from structural problems and causes, including the role of criminal-legal system exposure and the role of racialized disparities and over policing on a given individual’s ability to “make good,” or not recidivate. Meanwhile, the technocratic focus on evaluating policy though cost-benefit may detract from broader political goals as it is challenging to translate broader goals (e.g., social cohesion) into quantifiable indicators (Gottschalk, 2016). Yet, this dissertation shows how it is not so simple to alter how we design studies and select outcomes given the institutional pressures researchers faced in selecting them. Chapter three also demonstrated how providing flexibility and encouraging adaptation may still not be enough to help “localize” or make interventions “real world” relevant, when the very institutions they are asked to operate within are themselves a source of harm. That is, while we might make efforts to frame interventions and study designs differently, we might also more broadly reconsider the very institutions we partner with.

On Dealing with Context

This dissertation showed how sites of service provision are fundamentally transformed through the RCT and implementation science, as well as what new subjectivities and power asymmetries are produced in the process (Brives, 2016). Implementation science, understanding the integral role of context, must explicitly account for and transform beliefs, practices, relationships and environments or take part in ‘reassembling the social’ to facilitate intervention uptake and to solidify replication (Latour, 2007). In the case of HaRP, for example, care was refashioned (Chapter Three), activism was deferred (Chapter One), and the very concept of

feasibility was contested in order to promote more robust social services (Chapter Two). As we saw in Chapter One and Two, even the designation of what constituted "context" was actively produced, with social and scientific consequences. In both chapters, researchers and practitioners had to negotiate when the experimental context would end and the transition to the "the real world" would begin, choices that were often grounded in moral and political concerns.

Yet, the technical nature of implementation science, its grounding in positivist epistemologies, and lack of attention to real-time processes may obscure the very politics and social conditions required to better apprehend how a given intervention might work within a particular context. As described in the introduction to this dissertation, implementation science as a field is intimately connected to the larger evidence-based practice movement. This means that it is also invested in the experimental methods, largely understanding context in variable terms, and as a bounded "thing" that can ostensibly be accounted for or managed to help pave the way for intervention success (or failure). Therefore, within the field of implementation science, context tends to be treated as an add-on to the research process rather than a fundamental aspect to a given research project. In doing so, the field, at least within the context of an RCT, tends to figuratively "crop" context borders, ultimately missing important elements. As we saw in Chapter Two, context tends to be circumscribed in ways that facilitate the research project, while eliding important factors that may contribute to the success of a given intervention (e.g., institutional connections). Meanwhile, as we saw in Chapter One, it was not always clear when context should be treated as a barrier to examine through research, versus a problem to be resolved. Likewise, elements of the study design, including outcomes, partnerships, and the molding of the "model" were decided upon through research practices generally missed in

implementation science reports. These types of negotiations are important to apprehend in order to improve the development and implementation of social programs.

We might productively understand implementation processes like HaRP through the lens of street level scholarship, which tends to ethnographically and qualitatively attend to the 'on the ground' practices of implementation, looking beyond the four corners of the legislative process to aspects of policy delivery otherwise 'under the hood' (Brodkin & Marston, 2013). Rejecting normative assumptions about the policy hierarchy, street-level scholars understand implementation, though in the context of formal policy, as a major site of political contestation (Lipsky, 1980). HaRP's development, as I demonstrated in this dissertation, was not a linear process of simply taking up the evidence-based practices and programs outlined in the grant proposal, but was instead developed through a consistent process of negotiation and *in situ*, with diverse social effects. That is, HaRP was not developed prior to the study, instead the study sites served as the grounds for its negotiation and subsequent development. Pragmatically, more empirical, processual knowledge of the types of constraints trial actors face as they field complex interventions helps to not only improve their development, but also anticipate challenges and conduct better social science.

Policy and Practice Recommendations

There is a risk of critically analyzing novel approaches like implementation science and other models that have emerged to contend with context and the complexities of real world service delivery. As critiques are delivered, new approaches seem to emerge out of the woodwork; models and frameworks that appear to resolve the very tensions that serve as the focus of this dissertation. Indeed, implementation science continues to develop new frameworks in ways seemingly more attuned to the real world and the "translational" problems that occur

because of it.[2] I am also aware that the form of critique pursued in this dissertation puts me in the rather nihilistic position of having to defend against approaches that seem to resolve, or at least better deal with, the context of social service delivery. However, even the field of implementation science itself has recognized the overabundance of frameworks it has generated, and the ways in which doing so has led the field astray from its original mission (Beidas et al., 2022). Perhaps compared to other forms of applied social science, implementation science is particularly ripe for this type of iterative practice. The seemingly perennial issue of context, itself mutable and open to different interpretations, makes it an evergreen focus for ongoing research and primed for innovation. While I will not offer a framework, the following section will present implications based on the observations made throughout this dissertation.

Investment

Federal interest in harm reduction and peer recovery is growing, and such practices will be imminently scaled across the United States. Chapter One as well as Chapter Three showed some of the tensions of integrating harm reduction and peer work into the context of a clinical trial. It perhaps left us wondering: can the aspirations of grassroot activists, and practitioners be better integrated into clinical trials like HaRP, or are their underlying logics and values too incompatible? On the one hand, adding to a given "evidence-base" can lead to further professionalization with additional funding, training opportunities, and other material benefits (Chapter Two). What we *did hear* from practitioners and implementation actors alike is that we need to invest in peer recovery and harm reduction practitioners as though we are performing a rigorous and well-resourced clinical trial like HaRP. Researchers were dismayed by the

[2] For example, intervention science has advanced beyond the traditional RCT and the bounded concept of the intervention with frameworks like experimental therapeutics, which conceptualizes interventions not as a bounded "thing" but as a group of components each with their own mechanism of action (Raghavan et al., 2019). This is meant to make interventions more flexible and attuned to the real world.

widespread lack of investment in social services, and the relative disparity between what HaRP could provide compared to the "real world" of social service delivery. As we saw in Chapter Two, both researchers and practitioners began to view the experiment and trial resources as an intervention itself. The resolutely unfeasible nature of some of HaRP's elements represented a political intervention highlighting the underfunded nature of peer work, and other forms of community-based care. Amy and other trial actors knew that HaRP was not sustainable, and its resources did not represent the entrenched inequities in social service funding, but nevertheless understood this investment is what *should be done* in otherwise under-funded community-based and grassroots social service delivery. In other words, as voiced by practitioners themselves, we need further investment in community-based social service workers.

Collaboration

My research empirically demonstrates that researchers lack guidance on how to negotiate locally embedded social and professional values and growing activist commitments with the obligations of scientific inquiry during socio-behavioral clinical trials. As science becomes more integrated into society, the number of those evaluating "good science" also necessarily increases (Jasanoff, 2007). That is, scientists are increasingly asked to justify their research and show how it is meaningfully accountable to the society it's meant to serve. As such, researchers are increasingly encouraged to develop strong relationships with the organizations and institutions they study or from which they recruit. Even still, while collaboration and community participation is an expected part of most federally-funded research projects like HaRP, they have the tendency to be used as "end of the pipe legitimation devices" that do not allow meaningful reframing beyond the original research design (Wynne, 2006). While there are a plethora of exceptions of meaningful collaboration, a point I will elaborate on in the following pages,

HaRP's community advisory board (CAB) meetings tended to focus on seeking legitimation rather than more substantial feedback about the project. That is, the CAB was used to tweak the intervention design, rather than ask meaningful questions about the very premise of the HaRP project. Similar to the findings of Kerr et al. (2007) practitioners' concerns tended to be treated as "adjunct rather than an alternative" to the expertise of researchers (p. 406). Meanwhile, as we saw in Chapter Three, practitioners were asked to "bring that context in," when they noticed a disconnect between the intervention design and the "real world" of service delivery. At the same time, however, their own experiential expertise was relegated when asked to deploy evidence-based strategies in ways that, to them, seemed context-naive.

As has been argued elsewhere, researchers should try to ensure that research objectives and practices do not compromise the agenda of the organizations they serve, which can be done by working with organizations from the inception of the study framing process, rather than seeking feedback once the project and interventions are designed (Bodini et al., 2020). Likewise, researchers should seek to develop long-term relationships with community-based organizations, which may allow for more impactful research that can transcend a given short-term research project. Researchers should attempt to embed their research projects locally, and re-center leadership to include service users as well as practitioners engaged in implementing a given intervention. Real-time mechanisms of feedback should be put into place, not just for intervention implementation, but also the research process. This might help to bring different actors together and create meaningful space to discuss and navigate ethical and moral challenges, like the ones that surfaced throughout HaRP. Of course, this would mean more funding and built-in incentives need to be allocated to the organizational processes involved in collaboration, allowing researchers and implementation actors to improve communication and power-sharing.

The researchers I worked with and studied were particularly interested in issues of trust and mistrust in the organizations and service users they were meant to partner with. Research assistants and practitioners needed to build trust with participants, to enroll them into the study. Researchers needed to build trust with carceral stakeholders. Practitioners needed to build trust with study participants. Yet, practitioners and research assistants alike struggled to recruit, get participants to "open up" and the study altogether suffered from attrition. In Chapter Three, we saw that trauma and other individual participant issues were diagnosed as a cause of this mistrust. We also witnessed mistrust from practitioners and implementation actors themselves, who put forth valid sociological critiques of the HaRP trial, including its interest in accomplishing the experiment at the expense of its "real world" viability. It seems to me researchers are astute at diagnosing mistrust, including its conditions and the reasons that generate it, without reflexively examining the research apparatus and academic institution as itself, at least potentially, an originating source. Barriers to trust and the take up of science—issues like culture, miseducation or lack of understanding become yet another domain for further research. As put by Wynne (2006), "the public is imagined, constructed and projected, in a reflection of the unspoken needs of the institutionally powerful" (p. 218). The concerns of the public, the ones researchers are meant to try and understand, become reframed as another instrumental, scientific goal that obscures their very role in the process of generating such troubled relationships. That is, this framing, and the methods to comprehend and intervene upon mis-trust come at the cost of potentially understanding how the structural conditions of research might also engender mistrust and hostility. That said, to improve collaboration with practitioners, and forge trust with study participants, research institutions should attempt to reflexively

understand the moral and political implications of their research instead of seeing it as somehow "outside of" the research process.

Future Research

Future research could seek to understand how research itself, particularly intervention and implementation science research, is affected and transformed when working in collaboration with social movement actors, activists or in pursuing community-based research projects. We saw that implementation scientists on the HaRP project had desires to pursue advocacy in Chapter One, and understood the very lack of HaRP's feasibility as a political point to pursue. Further research should help us to understand how research questions change, how conceptions of social change are transformed, how methods are adjusted, as well as how ethical issues and issues of trust and mistrust are navigated when working in partnership with practitioners and social movement actors. Often, as argued in this dissertation, the research processes and practices tend to be "black boxed" yet are critical to apprehend to better understand and improve implementation practices. Pragmatically, this could help researchers to improve practices of collaboration by anticipating barriers as well as identifying best practices. In doing so, we might also better frame our original study designs to be more inclusive of the practitioners performing the work, study participants as well as even the researchers eager to conduct more transformative research.

References

Adams, V. (2016). *Metrics: What Counts in Global Health*. Duke University Press.

Akiba, C. F., Powell, B. J., Pence, B. W., Nguyen, M. X. B., Golin, C., & Go, V. (2021). The case for prioritizing implementation strategy fidelity measurement: benefits and challenges. *Translational Behavioral Medicine, 12*(2), 335–342. https://doi.org/10.1093/tbm/ibab138

Augustine, D., & Kushel, M. (2022). Community supervision, housing insecurity, and homelessness. *The ANNALS of the American Academy of Political and Social Science, 701*(1), 152–171. https://doi.org/10.1177/00027162221113983

Bailey, Z. D., Krieger, N., Agénor, M., Graves, J., Linos, N., & Bassett, M. T. (2017). Structural racism and health inequities in the USA: Evidence and interventions. *The Lancet,389*(10077), 1453–1463. https://doi.org/10.1016/S0140-6736(17)30569-X

Barrenger, S. L., Stanhope, V., & Miller, E. (2019). Capturing the value of peer support: Measuring recovery-oriented services. *Journal of Public Mental Health*, 18(3), 180–187. https://doi.org/10.1108/JPMH-02-2019-0022

Beidas, R. S., Dorsey, S., Lewis, C. C., Lyon, A. R., Powell, B. J., Purtle, J., Saldana, L., Shelton, R. C., Stirman, S. W., & Lane-Fall, M. B. (2022). Promises and pitfalls in implementation science from the perspective of US-based researchers: Learning from a pre-mortem. *Implementation Science*, *17*(1), 55. https://doi.org/10.1186/s13012-022-01226-3

Ben-Menachem, J. (2023, September 19). Our Evidence-Based Obsession. *Inquest* https://inquest.org/our-evidence-based-obsession/

Berk, R. A., Boruch, R. F., Chambers, D. L., Rossi, P. H., & Witte, A. D. (1985). Social Policy Experimentation: A Position Paper. Evaluation Review, 9(4), 387–429. https://doi.org/10.1177/0193841X8500900401

Berman, E. P. (2022). *Thinking like an economist: How efficiency replaced equality in U.S. public policy*. Princeton University Press.

Biehl, J., & Petryna, A. (2013). *When people come first: Critical studies in global health*. Princeton University Press.

Binswanger, I. A., Stern, M. F., Deyo, R. A., Heagerty, P. J., Cheadle, A., Elmore, J. G., & Koepsell, T. D. (2007). Release from prison — a high risk of death for former inmates. *New England Journal of Medicine*, *356*(2), 157–165. https://doi.org/10.1056/NEJMsa064115

Bodini, C., Baum, F., Labonté, R., Legge, D., Sanders, D., & Sengupta, A. (2020). Methodological challenges in researching activism in action: Civil society engagement

towards health for all. *Critical Public Health*, *30*(4), 386–397. https://doi.org/10.1080/09581596.2019.1650892

Bonell, C., Fletcher, A., Morton, M., Lorenc, T., & Moore, L. (2012). Realist randomised controlled trials: A new approach to evaluating complex public health interventions. *Social Science & Medicine (1982)*, *75*(12), 2299–2306. https://doi.org/10.1016/j.socscimed.2012.08.032

Bor, J., Cohen, G. H., & Galea, S. (2017). Population health in an era of rising income inequality: USA, 1980–2015. *The Lancet*, *389*(10077), 1475–1490. https://doi.org/10.1016/S0140-6736(17)30571-8

Boulton, R., Sandall, J., & Sevdalis, N. (2020). The cultural politics of "Implementation Science." *The Journal of Medical Humanities*, *41*(3), 379–394. https://doi.org/10.1007/s10912-020-09607-9

Brass, C. T. (2006). *Congress and program evaluation: An overview of randomized controlled trials (RCTs) and Related Issues*. Congressional Research Service, the Library of Congress.

Breslau, D. (1998). *In search of the unequivocal: The political economy of measurement in U.S. labor market policy*. Bloomsbury Academic.

Brives, C., Le Marcis, F., & Sanabria, E. (2016). What's in a context? Tenses and tensions in evidence-based medicine. *Medical Anthropology*, *35*(5), 369–376. https://doi.org/10.1080/01459740.2016.1160089

Brodkin, E. Z., & Marston, G. (2013). *Work and the welfare state: Street-level organizations and workfare politics*. Georgetown University Press.

Brown, M. (2019). Constructing accountability: The development and delegation of outcome evaluation in american social work. *Social Service Review*, 93(4), 712–763. https://doi.org/10.1086/706468

Brown, P., Zavestoski, S., McCormick, S., Mayer, B., Morello-Frosch, R., & Gasior Altman, R. (2004). Embodied health movements: New approaches to social movements in health. *Sociology of Health & Illness*, *26*(1), 50–80. https://doi.org/10.1111/j.1467-9566.2004.00378.x

Brown, T. M., & Fee, E. (2014). Social movements in health. *Annual Review of Public Health*, *35*(1), 385–398. https://doi.org/10.1146/annurev-publhealth-031912-114356

Brownson, R. C., Kumanyika, S. K., Kreuter, M. W., & Haire-Joshu, D. (2021). Implementation science should give higher priority to health equity. *Implementation Science*, *16*(1), 28. https://doi.org/10.1186/s13012-021-01097-0

Buck-McFadyen, E. (2022). Rural homelessness: How the structural and social context of small-town living influences the experience of homelessness. *Canadian Journal of Public Health, 113*(3), 407–416. https://doi.org/10.17269/s41997-022-00625-9

Camic, C., Gross, N., & Lamont, M. (2011). *Social knowledge in the making*. University of Chicago Press.

Campbell, N. D. (2020). *OD: Naloxone and the politics of overdose*. MIT Press.

Carpenter, D. (2014). *reputation and power: Organizational image and pharmaceutical regulation at the FDA*. Princeton University Press.

Carr, E. S. (2023). *Working the difference: Science, spirit, and the spread of motivational interviewing*. University of Chicago Press.

Carr, E. S., & Obertino-Norwood, H. (2022). Legitimizing evidence: The trans-institutional life of evidence-based practice. *Social Science & Medicine (1982), 310*, 115130. https://doi.org/10.1016/j.socscimed.2022.115130

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science, 2*(1), 40. https://doi.org/10.1186/1748-5908-2-40

Cartwright, N. (2007). Are RCTs the gold standard? *BioSocieties, 2*(1), 11–20. https://doi.org/10.1017/S1745855207005029

Cartwright, N. (2011). A philosopher's view of the long road from RCTs to effectiveness. *The Lancet*, 377(9775), 1400–1401. https://doi.org/10.1016/S0140-6736(11)60563-1

Cartwright, N., & Hardie, J. (2012). *Evidence-based policy: A practical guide to doing it better*. Oxford University Press.

Cartwright, N. (2013). Knowing what we are talking about: Why evidence doesn't always travel. *Evidence & Policy*, 9(1), 97–112. https://doi.org/10.1332/174426413X662581

Casper, M. J., & Clarke, A. E. (1998). Making the Pap Smear into the 'right tool' for the job: Cervical cancer screening in the USA, circa 1940-95. *Social Studies of Science, 28*(2), 255–290. https://doi.org/10.1177/030631298028002003

Chambers, D. A., & Emmons, K. M. (2024). Navigating the field of implementation science towards maturity: challenges and opportunities. *Implementation Science, 19*(1), 26. https://doi.org/10.1186/s13012-024-01352-0

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2008). Developing and evaluating complex interventions: The new Medical Research Council guidance. *The BMJ, 337*, a1655. https://doi.org/10.1136/bmj.a1655

Curran, G. M. (2020). Implementation science made too simple: A teaching tool. *Implementation Science Communications*, *1*(1), 27. https://doi.org/10.1186/s43058-020-00001-z

Curran, G. M., Bauer, M., Mittman, B., Pyne, J. M., & Stetler, C. (2012). Effectiveness -implementation hybrid designs. *Medical Care*, *50*(3), 217–226. https://doi.org/10.1097/MLR.0b013e3182408812

de Souza Leão, L., & Eyal, G. (2019). The rise of randomized controlled trials (RCTs) in international development in historical perspective. *Theory and Society*, 48(3), 383–418. https://doi.org/10.1007/s11186-019-09352-6

Des Jarlais, D. C. (2017). Harm reduction in the USA: The research perspective and an archive to David Purchase. *Harm Reduction Journal*, *14*(1), 51. https://doi.org/10.1186/s12954-017-0178-6

Donovan, K. P. (2018). The rise of the randomistas: On the experimental turn in international aid. *Economy and Society*, *47*(1), 27–58. https://doi.org/10.1080/03085147.2018.1432153

Drug Overdose Deaths | Drug Overdose | CDC Injury Center. (2023, August 22). https://www.cdc.gov/drugoverdose/deaths/index.html

Dryden-Palmer, K. D., Parshuram, C. S., & Berta, W. B. (2020). Context, complexity and process in the implementation of evidence-based innovation: a realist informed review. *BMC Health Services Research*, *20*, 81. https://doi.org/10.1186/s12913-020-4935-y

Epstein, S. (1995). The construction of lay expertise: AIDs activism and the forging of credibility in the reform of clinical trials. *Science, Technology, & Human Values*, *20*(4), 408–437. https://doi.org/10.1177/016224399502000402

Epstein, S. (1996). *Impure science: AIDS, activism, and the politics of knowledge*. University of California Press.

Eyal, G. (2019). *The crisis of expertise*. Polity Press.

Ford, I., & Norrie, J. (2016). Pragmatic trials. *New England Journal of Medicine*, *375*(5), 454 –463. https://doi.org/10.1056/NEJMra1510059

Fortney, J. C., Curran, G. M., Lyon, A. R., Check, D. K., & Flum, D. R. (2024). Similarities and differences between pragmatic trials and hybrid effectiveness-implementation trials. *Journal of General Internal Medicine*. https://doi.org/10.1007/s11606-024-08747-1

Frickel, S., Gibbon, S., Howard, J., Kempner, J., Ottinger, G., & Hess, D. J. (2010). Undone science: Charting social movement and civil society challenges to research agenda setting. *Science, Technology, & Human Values*, *35*(4), 444–473. https://doi.org/10.1177/0162243909345836

Frickel, S., & Moore, K. (2006). *The new political sociology of science: Institutions, networks, and power*. Univ of Wisconsin Press.

Garriott, W. C. (2011). *Policing methamphetamine: narcopolitics in rural America*. NYU Press.

Gauchat, G. (2012). Politicization of science in the public sphere: A study of public trust in the United States, 1974 to 2010. *American Sociological Review*, 77(2), 167–187. https://doi.org/10.1177/0003122412438225

Ghose, R., Forati, A. M., & Mantsch, J. R. (2022). Impact of the COVID-19 pandemic on opioid overdose deaths: A spatiotemporal analysis. *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, *99*(2), 316–327. https://doi.org/10.1007/s11524-022-00610-0

Gilens, M. (2005). Inequality and democratic responsiveness. *The Public Opinion Quarterly*, 69(5), 778–796. https://doi.org/10.1093/poq/nfi058

Glasgow, R. E., Vinson, C., Chambers, D., Khoury, M. J., Kaplan, R. M., & Hunter, C. (2012). National Institutes of Health approaches to dissemination and implementation science: Current and future directions. *American Journal of Public Health*, *102*(7), 1274–1281. https://doi.org/10.2105/AJPH.2012.300755

Gone, J. P. (2013). A community-based treatment for Native American historical trauma: Prospects for evidence-based practice. *Spirituality in Clinical Practice*, *1*(S), 78–94. https://doi.org/10.1037/2326-4500.1.S.78

Gottschalk, M. (2016). *Caught: The prison state and the lockdown of American politics*. Princeton University Press.

Greenhalgh, T., & Papoutsi, C. (2018). Studying complexity in health services research: Desperately seeking an overdue paradigm shift. *BMC Medicine*, *16*(1), 95. https://doi.org/10.1186/s12916-018-1089-4

Greenhalgh, T., & Wieringa, S. (2011). Is it time to drop the "knowledge translation" metaphor? A critical literature review. *Journal of the Royal Society of Medicine*, *104*(12), 501–509. https://doi.org/10.1258/jrsm.2011.110285

Gresham, G. K., Ehrhardt, S., Meinert, J. L., Appel, L. J., & Meinert, C. L. (2018). Characteristics and trends of clinical trials funded by the National Institutes of Health between 2005 and 2015. *Clinical Trials (London, England)*, *15*(1), 65–74. https://doi.org/10.1177/1740774517727742

Guidry, J. A., & Sawyer, M. Q. (2003). Contentious pluralism: The public sphere and democracy. *Perspectives on Politics*, 1(2), 273–289. https://doi.org/10.1017/S1537592703000203

Gustafson, P., Abdul Aziz, Y., Lambert, M., Bartholomew, K., Rankin, N., Fusheini, A., Brown, R., Carswell, P., Ratima, M., Priest, P., & Crengle, S. (2023). A scoping review of equity-focused implementation theories, models and frameworks in healthcare and their application in addressing ethnicity-related health inequities. *Implementation Science, 18*(1), 51. https://doi.org/10.1186/s13012-023-01304-0

Hailemariam, M., Bustos, T., Montgomery, B., Barajas, R., Evans, L. B., & Drahota, A. (2019). Evidence-based intervention sustainability strategies: A systematic review. *Implementation Science, 14*(1), 57. https://doi.org/10.1186/s13012-019-0910-6

Hansen, H., Jordan, A., Plough, A., Alegria, M., Cunningham, C., & Ostrovsky, A. (2022). Lessons for the opioid crisis-integrating social determinants of health into clinical care. *American Journal of Public Health*, *112*(S2), S109–S111. https://doi.org/10.2105/AJPH.2021.306651

Haraway, D. (1988). Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist Studies*, *14*(3), 575–599. https://doi.org/10.2307/3178066

Hardon, A., & Sanabria, E. (2017). Fluid drugs: Revisiting the anthropology of pharmaceuticals. *Annual Review of Anthropology*, 46(1), 117–132. https://doi.org/10.1146/annurev-anthro-102116-041539

Hariton, E., & Locascio, J. J. (2018). Randomised controlled trials—the gold standard for effectiveness research. *BJOG: An International Journal of Obstetrics and Gynecology, 125*(13), 1716. https://doi.org/10.1111/1471-0528.15199

Harm Reduction Principles. (n.d.). *National Harm Reduction Coalition*. Retrieved February 12, 2024, from https://harmreduction.org/about-us/principles-of-harm-reduction/

Harvey, G., Rycroft-Malone, J., Seers, K., Wilson, P., Cassidy, C., Embrett, M., Hu, J., Pearson, M., Semenic, S., Zhao, J., & Graham, I. D. (2023). Connecting the science and practice of implementation – applying the lens of context to inform study design in implementation research. *Frontiers in Health Services*, *3*. https://doi.org/10.3389/frhs.2023.1162762

Hassan, S., Tourmaline, & Brown, Adrienne Maree. (2022). *Saving our own lives: A liberatory practice of harm reduction*. Haymarket Books.

Hess, D. J. (2016). *Undone science: Social movements, mobilized publics, and industrial transitions*. MIT Press.

Indicators of Mental Health Problems Reported by Prisoners and Jail Inmates, 2011-2012 | Bureau of Justice Statistics. (n.d.). Retrieved March 28, 2024, from https://bjs.ojp.gov/library/publications/indicators-mental-health-problems-reported-prisoners-and-jail-inmates-2011

Jasanoff, S. (2003). Technologies of humility: Citizen participation in governing science. *Minerva*, *41*(3), 223–244. https://doi.org/10.1023/A:1025557512320

Jasanoff, S. (2006). *States of knowledge: The co-production of science and social order*. Routledge.

Jasanoff, S. (2007). Technologies of humility. Nature, 450(7166), 33–33. https://doi.org/10.1038/450033a

Jasanoff, Sheila. (2011). The practices of objectivity in regulatory science. *Social Knowledge in the Making.* Ed. Charles Camic, Neil Gross, and Michèle Lamont. University of Chicago Press, 2021, 307-338.

Kerr, A., Cunningham-Burley, S., & Tutton, R. (2007). Shifting subject positions: Experts and lay people in public dialogue. *Social Studies of Science*, 37(3), 385–411. https://doi.org/10.1177/0306312706068492

Kirk, S. A., & Reid, W. J. (2002). *Science and social work: A critical appraisal*. Columbia University Press. https://doi.org/10.7312/kirk11824

Kirmayer, L. J. (2012). Cultural competence and evidence-based practice in mental health: Epistemic communities and the politics of pluralism. *Social Science & Medicine (1982)*, *75*(2), 249–256. https://doi.org/10.1016/j.socscimed.2012.03.018

Kitson, A. L., Rycroft-Malone, J., Harvey, G., McCormack, B., Seers, K., & Titchen, A. (2008). Evaluating the successful implementation of evidence into practice using the PARiHS framework: Theoretical and practical challenges. *Implementation Science*, *3*(1), 1. https://doi.org/10.1186/1748-5908-3-1

Krauss, A. (2021). Assessing the overall validity of randomised controlled trials. *International Studies in the Philosophy of Science*, *34*(3), 159–182. https://doi.org/10.1080/02698595.2021.2002676

Lambert, A.-S., Legrand, C., Cès, S., Van Durme, T., & Macq, J. (2019). Evaluating case management as a complex intervention: Lessons for the future. *PLoS ONE*, *14*(10), 1-18. https://doi.org/10.1371/journal.pone.0224286

Lampland, M., & Star, S. L. (2009). *Standards and their stories: How quantifying, classifying, and formalizing practices shape everyday life*. Cornell University Press.

Lara-Millán, A. (2021). *Redistributing the poor: Jails, hospitals, and the crisis of law and fiscal austerity*. Oxford University Press.

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Harvard University Press.

Latour, B., & Woolgar, S. (2013). *Laboratory life: The construction of scientific facts*. Princeton University Press.

Levenson, J., Textor, L., Bluthenthal, R., Darby, A., Wahbi, R., & Clayton-Johnson, M.-A. (2023). Abolition and harm reduction in the struggle for "Care, Not Cages." *International Journal of Drug Policy*, *121*, 104163. https://doi.org/10.1016/j.drugpo.2023.104163

Lipsky, M. (1980). *Street level bureaucracy: Dilemmas of the individual in public services*. Russell Sage Foundation. https://www.jstor.org/stable/10.7758/9781610447713

Loudon, K., Treweek, S., Sullivan, F., Donnan, P., Thorpe, K. E., & Zwarenstein, M. (2015). The PRECIS-2 tool: Designing trials that are fit for purpose. *BMJ*, *350*, h2147. https://doi.org/10.1136/bmj.h2147

Maynard-Moody, S., & Musheno, M. (2003). Cops, teachers, counselors: Stories from the front lines of public service. University of Michigan Press. https://doi.org/10.3998/mpub.11924

May, C. R., Johnson, M., & Finch, T. (2016). Implementation, context and complexity. *Implementation Science*, *11*(1), 141. https://doi.org/10.1186/s13012-016-0506-3

McCoy, C. E. (2017). Understanding the intention-to-treat principle in randomized controlled trials. *Western Journal of Emergency Medicine*, *18*(6), 1075–1078. https://doi.org/10.5811/westjem.2017.8.35985

McKim, A. (2017). *Addicted to rehab: Race, gender, and drugs in the era of mass incarceration*. Rutgers University Press.

McLean, K. (2011). The biopolitics of needle exchange in the United States. *Critical Public Health*, *21*(1), 71–79. https://doi.org/10.1080/09581591003653124

Miller, R. J. (2021). *Halfway home: Race, punishment, and the afterlife of mass incarceration*. Little, Brown.

Mitchell, P., Curtin, K. M., & Magliocca, N. R. (2023). Race, rurality and geographic accessibility to medication for opioid use disorder in the U.S. *Journal of Maps*, *19*(1), 2270632. https://doi.org/10.1080/17445647.2023.2270632

Mol, A. (2002). *The body multiple: Ontology in medical practice*. Duke University Press.

Mol, A. (2008). *The logic of care: Health and the problem of patient choice*. Routledge.

Montgomery, C. M. (2017). Clinical trials and the drive to material standardisation: 'Extending the rails' or reinventing the wheel? *Science & Technology Studies*, *30*(4), 30–44. https://doi.org/10.23987/sts.59226

Moore, K. (2006). Powered by the people scientific authority in participatory science. In K. Moore & S. Frickel (Eds.), *The New Political Sociology of Science: Institutions, Networks, and Power*. Univ of Wisconsin Press.

Mosley, J. E., Marwell, N. P., & Ybarra, M. (2019). How the "what works" movement is failing human service organizations, and what social work can do to fix it. *Human Service Organizations: Management, Leadership & Governance*, *43*(4), 326–335. https://doi.org/10.1080/23303131.2019.1672598

Mugenyi, L., Mijumbi, A., Nanfuka, M., Agaba, C., Kaliba, F., Semakula, I. S., Nazziwa, W. B., & Ochieng, J. (2021). Capacity of community advisory boards for effective engagement in clinical research: a mixed methods study. *BMC Medical Ethics*, *22*(1), 165. https://doi.org/10.1186/s12910-021-00733-0

Murphy, J., & Russell, B. (2020). Police officers' views of naloxone and drug treatment: Does greater overdose response lead to more negativity? *Journal of Drug Issues*, *50*(4), 455–471. https://doi.org/10.1177/0022042620921363

NIH funds $945 million in research to tackle the national opioid crisis through NIH HEAL Initiative. (2019, September 25). National Institutes of Health (NIH). https://www.nih.gov/news-events/news-releases/nih-funds-945-million-research-tackle-national-opioid-crisis-through-nih-heal-initiative

Nowotny, H., Scott, P., & Gibbons, M. (2003). "Mode 2" revisited: The new production of knowledge. *Minerva*, 41(3), 179–194. 10.1023/A:1025505528250

Okpych, N. J., & Yu, J. L.-H. (2014). A historical analysis of evidence-based practice in social work: The unfinished journey toward an empirically grounded profession. *Social Service Review*, *88*(1), 3–58. https://doi.org/10.1086/674969

Orkin, A. M., Rao, S., Venugopal, J., Kithulegoda, N., Wegier, P., Ritchie, S. D., VanderBurgh, D., Martiniuk, A., Salamanca-Buentello, F., & Upshur, R. (2021). Conceptual framework for task shifting and task sharing: an international Delphi study. *Human Resources for Health*, *19*(1), 61. https://doi.org/10.1186/s12960-021-00605-z

Petty, J., & Heimer, C. A. (2011). Extending the rails: How research reshapes clinics. *Social Studies of Science*, *41*(3), 337–360. https://doi.org/10.1177/0306312710396402

Phelan, J. C., Link, B. G., & Tehranifar, P. (2010). Social conditions as fundamental causes of health inequalities: theory, evidence, and policy implications. *Journal of Health and Social Behavior, 51 Suppl*, S28-40. https://doi.org/10.1177/0022146510383498

Polanyi, M. (1998). *Personal knowledge: Towards a post-critical philosophy*. Psychology Press.

Porter, T. M. (1996). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton University Press.

Prins, S. J., & Reich, A. (2021). Criminogenic risk assessment: A meta-review and critical analysis. *Punishment & Society*, *23*(4), 578–604. https://doi.org/10.1177/14624745211025751

Proctor, E. K., Powell, B. J., & McMillen, J. C. (2013). Implementation strategies: Recommendations for specifying and reporting. *Implementation Science*, *8*(1), 139. https://doi.org/10.1186/1748-5908-8-139

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., & Hensley, M. (2011). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health*, *38*(2), 65–76. https://doi.org/10.1007/s10488-010-0319-7

Pugh, A. J. (2022). Emotions and the systematization of connective labor. *Theory, Culture & Society*, *39*(5), 23–42. https://doi.org/10.1177/02632764211049475

Pugh, A. J. (2023). Constructing what counts as human at work: Enigma, emotion, and error in connective labor. *American Behavioral Scientist*, *67*(14), 1771–1792. https://doi.org/10.1177/00027642221127240

Rabeharisoa, V., Moreira, T., & Akrich, M. (2014). Evidence-based activism: Patients', users' and activists' groups in knowledge society. *BioSocieties*, *9*(2), 111–128. https://doi.org/10.1057/biosoc.2014.2

Ramachandran, M. T. (2020). The logic of randomised controlled trials in the social sciences. *Social Scientist*, *48*(1/2 (560-561)), 41–52. https://www.jstor.org/stable/26899498

Rayzberg, M. S. (2019). Fairness in the field: The ethics of resource allocation in randomized controlled field experiments. *Science, Technology, & Human Values*, *44*(3), 371–398. https://doi.org/10.1177/0162243918798471

Reid, W. J. (1997). Research on task-centered practice. *Social Work Research*, *21*(3), 132–137. https://doi.org/10.1093/swr/21.3.132

Reidpath, D. D., Allotey, P., Barker, S. F., Clasen, T., French, M., Leder, K., Ramirez-Lovering, D., Rhule, E. L. M., & Siri, J. (2022). Implementing "from here to there": A case study of conceptual and practical challenges in implementation science. *Social Science & Medicine (1982), 301*, 114959. https://doi.org/10.1016/j.socscimed.2022.114959

Rhodes, T., & Lancaster, K. (2019). Evidence-making interventions in health: A conceptual framing. *Social Science & Medicine (1982), 238*, 112488. https://doi.org/10.1016/j.socscimed.2019.112488

Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences, 4*(2), 155–169. https://www.jstor.org/stable/4531523

Rubin, R. (2023). It takes an average of 17 years for evidence to change practice—the burgeoning field of implementation science seeks to speed things up. *JAMA, 329*(16), 1333–1336. https://doi.org/10.1001/jama.2023.4387

Sanabria, E. (2016). Circulating ignorance: Complexity and agnogenesis in the obesity "epidemic." *Cultural Anthropology*, *31*(1), 131–158. https://journal.culanth.org/index.php/ca/article/view/ca31.1.07

Schattschneider, E. E. (1975). *The Semisovereign People: A Realist's view of democracy inAmerica*. Dryden Press.

Shapin, S. (1994). *A social history of truth: Civility and science in seventeenth-century England*. University of Chicago Press.

Shelton, R. C., Adsul, P., & Oh, A. (2021). Recommendations for addressing structural racism in implementation science: A call to the field. *Ethnicity & Disease, 31*(Suppl 1), 357 –364. https://doi.org/10.18865/ed.31.S1.357

Small, M. L., & Calarco, J. M. (2022). *Qualitative literacy: A guide to evaluating ethnographic and interview research*. Univ of California Press.

Smith, S. R., & Lipsky, M. (1993). nonprofits for hire: The welfare state in the age of contracting. Harvard University Press.

Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, "translations" and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science, 19*(3), 387–420. https://www.jstor.org/stable/285080

Tanenbaum, S. J. (2005). Evidence-based practice as mental health policy: Three controversies and a caveat. Health Affairs, 24(1), 163–173. https://doi.org/10.1377/hlthaff.24.1.163

Steeger, C. M., Buckley, P. R., Pampel, F. C., Gust, C. J., & Hill, K. G. (2021). Common methodological problems in randomized controlled trials of preventive interventions.

Prevention Science: The Official Journal of the Society for Prevention Research, *22*(8), 1159–1172. https://doi.org/10.1007/s11121-021-01263-2

Steen, S., & Bandy, R. (2007). When the policy becomes the problem: Criminal justice in the new millennium. *Punishment & Society*, *9*(1), 5–26. https://doi.org/10.1177/1462474507070548

Stein, H., Cunningham, S., & Carmody, P. (2021). The rise of "behavioral man": Randomized controlled trials and the "new" development agenda. *Human Geography*, *14*(1), 62–75. https://doi.org/10.1177/1942778620987068

Strauss, A. L. (1987). *Qualitative analysis for social scientists*. Cambridge University Press. https://doi.org/10.1017/CBO9780511557842

Thorpe, K. E., Zwarenstein, M., Oxman, A. D., Treweek, S., Furberg, C. D., Altman, D. G., Tunis, S., Bergel, E., Harvey, I., Magid, D. J., & Chalkidou, K. (2009). A pragmatic -explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *Journal of Clinical Epidemiology*, *62*(5), 464–475. https://doi.org/10.1016/j.jclinepi.2008.12.011

Timmermans, S. (2005). From autonomy to accountability: The role of clinical practice guidelines in professional power. *Perspectives in Biology and Medicine*, *48*(4), 490 –501. https://doi.org/10.1353/pbm.2005.0096

Timmermans, S., & Berg, M. (2010). *The gold standard: The challenge of evidence-based medicine*. Temple University Press.

Timmermans, S., & Epstein, S. (2010). A world of standards but not a standard world: Toward a sociology of standards and standardization. *Annual Review of Sociology*, *36*, 69–89. https://doi.org/10.1146/annurev.soc.012809.102629

Title IV-E Prevention Program. (2023, June 23). https://www.acf.hhs.gov/cb/title-iv-e-prevention-program

van der Meulen, E., Chu, S. K. H., & Butler-McPhee, J. (2021). "That's why people don't call 911": Ending routine police attendance at drug overdoses. *International Journal of Drug Policy*, *88*, 103039. https://doi.org/10.1016/j.drugpo.2020.103039

Wagner, K. D., Harding, R. W., Kelley, R., Labus, B., Verdugo, S. R., Copulsky, E., Bowles, J. M., Mittal, M. L., & Davidson, P. J. (2019). Post-overdose interventions triggered by calling 911: Centering the perspectives of people who use drugs (PWUDs). *PLOS ONE, 14*(10), e0223823. https://doi.org/10.1371/journal.pone.0223823

Wahlberg, A., & McGoey, L. (2007). *An elusive evidence base: The construction and governance of randomized controlled trials*. https://ora.ox.ac.uk/objects/uuid:bccbbca4-6f7e-490a-9439-164950f851ca

Wang, J. (2018). *Carceral capitalism*. MIT Press.

Wiltsey Stirman, S., Kimberly, J., Cook, N., Calloway, A., Castro, F., & Charns, M. (2012). The sustainability of new programs and innovations: A review of the empirical literature and recommendations for future research. *Implementation Science, 7*(1), 17. https://doi.org/10.1186/1748-5908-7-17

Winther, J., & Hillersdal, L. (2020). Balancing methodological purity and social relevance: Monitoring participant compliance in a behavioural RCT. *BioSocieties*, *15*(4), 555–579. https://doi.org/10.1057/s41292-019-00163-7

Wynne, B. (2006). Public engagement as a means of restoring public trust in science– hitting the notes, but missing the music? *Community Genetics*, 9(3), 211–220. https://doi.org/10.1159/000092659

Yao, Q., Li, X., Luo, F., Yang, L., Liu, C., & Sun, J. (2019). The historical roots and seminal research on health equity: A referenced publication year spectroscopy (RPYS) analysis. *International Journal for Equity in Health*, *18*(1), 152. https://doi.org/10.1186/s12939-019-1058-3