Big Aspirations: Constructing a Collaborative Evaluation Culture to Support the Impact Goals of an Intermediary Organization

Citation

Permanent link
https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37373442

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story
The Harvard community has made this article openly available. Please share how this access benefits you. Submit a story.

Accessibility
Big Aspirations:
Constructing a Collaborative Evaluation Culture to Support
the Impact Goals of an Intermediary Organization

Doctor of Education Leadership (Ed.L.D.) Capstone

Submitted by:

Kofi A. Taha

To the Harvard Graduate School of Education
in partial fulfillment of the requirements for the degree of
Doctor of Education Leadership

April 2022
Dedication

To all my teachers, especially my parents.
Acknowledgements

Thank you...

To my advisor, professor and friend, Elizabeth City, for the respect of challenging questions, for the reminder that both love and strategy matter, and for always having my back.

To my advisor and committee chair, Martin West, for the thoughtful pushes and confidence, and for being willing to support me across this finish line.

To Karen Mapp, my faculty capstone committee member, for keeping me connected to my values and community, and for making completion possible.

To Christine Ortiz Guzman, my residency supervisor, capstone committee member and friend, for being true to your word, for sharing your bright vision with me and the world, and for inviting me to the party.

To Aislinn Betancourt, my sister in measurement, for the foundation upon which anything of value in this capstone rests, for your trust, and for keeping me laughing throughout this entire ordeal. You are up next.

To Aleiya Evison, Shakirra Meghjee, Samuel Allen, Issachar Curbeon, my friends and colleagues at EMD, for being so gifted at what you do, for encouraging me, for keeping me ever on my toes, for your examples of integrity in action.

To Deborah Jewell-Sherman, Irvin Scott and Paul Reville, my professors, for asking me to do better and for clearing the path.

To Maritza Hernandez, Drew Echelson, Nonie Lesaux, Matt Rose, Margarita Ruiz, Adelina Garcia, Lucretia Murphy, Heather Johnson, Lili Allen, Taylor Maag, Elizabeth Kidd McWhorter, David Fischer, Daphne Layton, Megan Fox, Cyn Rosario, Adam Parrot-Sheffer, Bonnie Lo, for your support, for your compassion, for stepping in when I asked for support. To Amy Loyd and Kyle Hartung for being builders and reminding me that I am a stone.

To Cohort 8, for your Cgr8ness. I feel so blessed to have made lifelong friends so late in life. #25in25out. To C7, C9, and C10, for all the ways you inspire.

To Rebecca, Keri, Stef, Jim, Courtney, Michael, Lance, and Mariel, for your friendship, gifts of insight and laughter, and for making me better.

To all the students and educators on my journey for all the things we learned together.

To Shanti, for endless tea, bountiful cariño, unwavering support, joy & peace.
To Isa and Satya, for reminding me of the three questions and to be a firekeeper.
To my big, beautiful family by blood, by choice, by belief for your patience and love.
Abstract

Intermediary organizations seek to make change in the world by strengthening the practice or building the capacity of direct service providers. In the education sector, these types of organizations often play important roles as technical assistance providers, capacity builders, additional administrators, investors, or funders. At the core of this capstone is a central question: how might intermediary organizations working at the institution-level define, articulate, and evaluate their own societal impact when most of the population-level changes they hope to see are so heavily dependent on actions taken by others?

My research suggests that answering this question is very difficult, but I argue that building an organizational culture that tries to do so is both an ethical obligation and a pathway to cycles of continuous improvement, greater strategic focus, and staff satisfaction. I explored these ideas through my residency at Equity Meets Design (EMD), a for-profit capacity-building company whose clients cumulatively work directly or indirectly with approximately fifty percent of all public-school students in PK-12 and higher education institutions in the United States.

This capstone chronicles my efforts to support EMD’s desire to use data more effectively as it works towards its mission to “increase the ability of organizations to (re)design interactions, interventions, and institutions towards increased equity in process and outcome.” While I do not arrive at a ready-made formula that offers quick and replicable answers, my analysis offers a roadmap for how intermediaries can 1) systematically embed evaluation practices into routine work, 2) practice using an equity lens to shift the definition, collection, and use of data, and 3) approach evaluation as an essential part of being in relationship with stakeholders.
Introduction

This capstone is fundamentally about accountability—how intermediary organizations can hold themselves more accountable for delivering on their mission, and how I can hold myself accountable for achieving my leadership goals. I attempt to distill a vast evaluation literature into some actionable steps for one organization, while also keeping track of what leadership lessons can be learned when trying to support an organization through uncertain times. The challenge of the Ed.L.D. residency experience is that we are tasked with bringing theory and practice together to lead a project in a real-world, real-time context that has its own vicissitudes, and to write a document that captures our progress within the confines of a fixed academic timeline. As I write this, I cannot say I completed all planned stages of my project within this tight schedule. My hope is that taking stock at this juncture in the process will lay the foundation for appropriate course corrections, inform the next phases of the process I am still leading, and offer some meaningful sensemaking about two ubiquitous challenges: 1) how to use an equity lens to better define, measure and articulate the impact of intermediary organizations ambitious enough to tackle complex social issues, and 2) how to lead change processes in dynamic, fast-paced organizations with finite time and resources.

I have had the great privilege of being placed as a resident fellow within Equity Meets Design (EMD), a “think-and-do tank” working to “increase[s] the ability of organizations to (re)design interactions, interventions and institutions towards increased equity in process and outcomes” (Ortiz Guzman, equityXdesign: Leveraging Identity Development in the Creation of an Anti-Racist Equitable Design Thinking Process., 2017). Its work is grounded in the core belief that “racism and inequity have been designed and so they can be redesigned” (2017, p. 30). Through workshops, coaching, online courses and publications, the five-year-old privately held
A conservative mapping suggests that EMD’s clients work directly or indirectly with approximately fifty percent of all public-school students at PK-12 and higher education institutions in the United States. Although not exclusively focused on the education sector, the genesis of EMD’s work was the capstone work of an Ed.L.D. alum who was specifically thinking about school and education system design and re-design. Most clients continue to be in the education space, though an increasing number are in other sectors.

Part of EMD’s success reflects its emphasis on design thinking as a tool that can enable clients to chart concrete next steps when their organization wants to “move from equity talk to equity action. Design thinking can be defined as a disaggregated problem-solving process with ordered, but non-linear, steps that include phases for problem framing and gathering information, idea-generating and imagining solutions, and building, testing and evaluating prototypes. While there are hundreds of representations meant to capture the design thinking process, Figure 1 is what EMD uses to highlight opportunities in design to focus on equity, i.e., in both the problem-definition and solution-finding phases.

**Figure 1: equityXdesign Depiction of the Design Process**

Source: equityXdesign Crash Course, 2020
EMD describes its work as “moving institutions and systems, not individuals,” meaning that while the methodology is individual capacity-building through EMD’s facilitation of design thinking workshops and tailored coaching support, the objective is for those individuals to transform their organizations and networks (Ortiz Guzman, 2017). Thus, if the end goal of a client is, for example, to have young people experience equitable learning environments, EMD does not directly train teachers or teach students; instead, EMD aims to build the capacity of its clients that design those learning environments (or, in some cases, that work to support those who directly design those environments) to use equitable design practices. In this way, EMD is a classic intermediary organization. Intermediary organizations are, “independently operating entities that work between multiple actors to facilitate communication and collaboration; build capacity and knowledge; and over time bring about change in the actors, their activities, and the results they achieve” (Center for Public Research and Leadership, 2017).

The clients EMD serves use a variety of strategies to pursue their diverse missions, including working at the system-level to bring evidence-based models and innovative solutions into practice; building the capacity of national networks of leaders to scale solutions, tools, resources, and strategies; shaping policies that strengthen federal, state, and local labor markets vis-à-vis education, workforce, and industry needs; and investing capital or re-granting resources to incubate and accelerate market-based solutions to pressing social problems. EMD pursues its own mission by strengthening how direct service providers, institutional leaders, and policymakers design and implement their work. This structural distance inherent to being an intermediary makes it challenging for EMD to accurately capture and communicate the impact it has on the organizations, communities, and individuals it ultimately seeks to benefit, even though this
potential impact is what is most motivating for everyone on the EMD staff (A. Evison, personal communication, June 12, 2021).

During my first few months on staff, I learned that this disconnect is a pain point. The seven members of the EMD team all care deeply about using their time and talents in pursuit of a mission that aligns with their values and that leads to tangible results. Because each of the individuals on the team has the skills and experience to be running their own consultancies or to choose to be employed elsewhere, the company’s founder recognizes that gaining clarity on impact is an important retention strategy. Recent efforts to tackle this challenge have led to the creation of important building blocks such as client surveys, engagement reports, and periodic data reviews but, to date, these tools have not been used consistently by staff or yielded the desired depth of understanding. The chief knowledge officer responsible for developing and deploying these tools has been appropriately thoughtful and skeptical about the validity, feasibility, and ethics of attributing changes in organizations and communities directly to EMD’s capacity-building activities given all the forces at work in complex socioeconomic environments, including the ways evaluation methodologies themselves can reinforce or mask inequity. This capstone describes my efforts to support her in revisiting unresolved questions around the design and implementation of EMD’s measurement and evaluation system.

The prevailing view at EMD is that addressing its own evaluation aspirations offers an opportunity to use an equity lens in both designing specific tools and defining “impact” itself. I place impact in quotation marks to emphasize, as I discuss later, that a uniform definition of what EMD as an organization, or the field more broadly, means by impact does not exist. Additionally, the hope is that improved evaluation tools and practices will contribute to improved knowledge-management capacity, more robust partner feedback, and data-informed improvement cycles. In
discussions with the team, there is no disagreement about the centrality of impact evaluation to the company’s strategic plan, but there is a palpable mixture of excitement and anxiety among the staff because of existing constraints on time and brain space. Zooming out further, the entire team is aware of the high degree of visibility, interest, and demand surrounding equity work following the murder of George Floyd, the global protests inspired by that crime, and the predictable white backlash that now animates national politics. This reality adds to the sense that careful, evidence-based claims about what EMD’s work is, what it isn’t, and what it contributes to are necessary.

My experience of this broader context, combined with consultations with the founder and chief knowledge officer, led me to propose a strategic project that I believed would simultaneously bring clarity to EMD’s “impact pain point” and yield useful strategies for the field. I suggested (and still expect) that this strategic project, if well-executed, would allow the entire organization to develop the feedback loops required to continuously learn and improve, prioritize, or deprioritize a given strategy, communicate effectively with the public, support new business development, and strengthen recruitment and retention of top talent. This interests me because internal organizational transition and core questions about efficacy are endemic to organizations seeking to bring big and important changes into existence. Thus, I set out to guide an inclusive, multi-stakeholder process that would be informed by cross-sector research, the lived experience of EMD facilitators conducting capacity-building workshops, and the experiences of clients. The end goal was to produce a useful framework or model for thinking about organizational effectiveness, to generate a set of tools that would make implementation within a busy organization like EMD feasible, and to establish the cultural conditions that would promote the use of these tools well beyond the completion of my residency. The shorthand we use for this endeavor is “making EMD a more data-informed organization.”
Though this effort is still in progress, my experience at EMD thus far suggests a roadmap for how intermediary organizations can become far more specific about defining, evaluating, and articulating their work. I also believe there are relevant implications for all types of intermediaries operating in the education sector, including those that view equity outcomes as part of their mission. This capstone documents my experience so far, beginning with a discussion of relevant evaluation literature basics and frameworks that helped me visualize the space that intermediaries occupy. This content research, combined with two leadership models, informed my theory of action for leading the strategic project. I synthesize some of the models and methods that I believe are most useful to consider given EMD’s context and then describe what I set out to accomplish, what has (and has not) happened up to this point, and why I think events have unfolded as they have. I close with a reflection on what I have learned from leading this process, focusing on implications for my own leadership, for EMD, and for the broader education sector.

**Review of Knowledge for Action**

My strategic project focuses on the central role that defining, measuring, and articulating impact can play in strengthening multiple dimensions of an intermediary organization, especially when that organization is moving through an internal change process. This section discusses the literature I relied on to guide my thinking in three areas: impact measurement, organization theory related to intermediary organizations, and leadership theory focused on change management. The literature on each of these topics is vast, so I limit myself here to summaries of two key concepts in each category that influenced my approach to designing and leading the project. From impact measurement, I use the *Theory of Change* approach to create a working definition of impact and the *Spheres of Influence* model to grapple with the limits of attribution. From organization theory, I employ Urie Bronfenbrenner’s *Ecological Systems* model to locate EMD’s position as an
intermediary and Mark Moore’s Strategic Triangle to frame the centrality of impact evaluation to strategic action. From leadership theory, I use John Kotter’s *Eight Step Change Model* and the Holding Environment concept from Ronald Heifetz and Donald Laurie’s *Adaptive Leadership* model to inform my own work plan. I conclude with my own theory of action for the strategic project; this theory is used later as a benchmark in the analysis section of the capstone.

**Impact Evaluation Method #1: Theory of Change**

Perhaps the most important thing to establish from the outset is that there is no single, universally accepted definition of “impact.” Often the concept of change or the resulting effect of some type of intervention is embedded in what is meant when we say “impact.” Sometimes we mean to imply that this change is positive, or we specify a desire to understand whether what has happened is considered a positive or negative change. Other definitions suggest that a dimension of time must be factored in and that, to qualify as impact, a change should be long-lasting and sustainable. There is little consistency in the definitions within the international development field, the sector that pioneered impact measurement. CARE International defines impact as “equitable and durable improvements in human wellbeing and social justice” (CARE International, 2019). For CARE, impact must be positive and sustainable over time. The Organization for Economic Cooperation & Development (OECD), on the other hand, has a much broader definition: “Positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended” (OECD/DAC Network on Development Evaluation, 2019). For OECD, impact includes all these effects, but they must be long-term.

“Impact” is often used to describe the scope or scale of a program, e.g., describing the impact of an education program by saying it has reached one million children; or it is used to talk about the degree or depth of change, e.g., describing what the educational achievements of those
children led to later in life in terms of earnings or other quality of life measures. In certain corners of the evaluation world, the word “impact” is only used in the context of a very specific type of evaluation: a randomized experiment with a treatment and control group. In these circles, “impact” refers to the difference in outcomes between the treatment group that received the intervention and the control group that did not (Duflo, Glennerster, & Kremer, 2007). These practitioners aspire toward scientific testing of the causal relationship between a program and the changes experienced by participants, introduce a counterfactual (what would have happened with no intervention at all), and seek to estimate the precise proportion of change that can be attributed to the project. The bottom line is that people use the same word but mean different things, which can be confusing, and context is essential. What became clear to me through the literature is that when communicating about impact, being explicit about how it is being defined is a critical first step.

My thinking about how to wade through this confusing terrain has been shaped by my colleague Laura Budzyna at MIT D-Lab; she leads D-Lab’s monitoring, evaluation and learning efforts as well as the Innovation Practice unit. As a university-based program focused on engaging communities in collaborative solution-finding in contexts where people are earning less than $3 a day, D-Lab takes the position that because resources are so scarce and the risk of harm to stakeholders is so high, understanding what is and is not working is a matter of integrity. Driven by this sense of responsibility, Budzyna has managed to effectively embed cycles of data collection, analysis, and sense-making in almost every project across the organization's three pillars (education, research and innovation practice). At the core of this effort is one of the simplest and most useful tools in the social impact field: the theory of change framework (see Figure 2). Essentially, a theory of change is an explanation of how a program or product intends to work. It reads like a logical statement: If we do X (input + output), then Y (outcome + impact) will happen.
A theory of change can be broken down into four key parts: inputs, outputs, outcomes, and impacts. The input is the “thing.” It is the product, service, or program—a solution—developed to address the problem at hand. The output is what happens when the “thing” is delivered to someone else, e.g., a product is sold to a customer, or a workshop is taught to a student. That is the output—the delivery of the solution. At this point, the “thing” is, quite literally, out of the hands of the implementer. Now that recipients or potential beneficiaries have the “thing”—the product or the workshop—what actions will they take? The actions taken are the outcome. Finally, if many people take those actions, how will the community or society change? In this very simplified breakdown, that is the impact—the change that happens—which could be an improvement, an unintended negative consequence, or no change at all (Budzyna, 2021).

A theory of change does three important things: first, it helps define what is within the control of the organization delivering the solution (inputs and outputs), and what actions, behavior changes and results are out of the control of the organization (outcomes and impacts). Second, a theory of change helps to test the logic, or the underlying assumptions embedded within each step of the theory, making it possible to isolate strengths or weaknesses in the theory. Third, a clear set of metrics or key performance indicators can be created for each stage of the theory (L. Budzyna,
A slightly more developed version of this simple model captures these additional strengths of the theory of change framework (See Figure 3).

### Figure 3: A Visualization of Theory of Change with Assumptions and Metrics

As I thought about EMD and its many programs, each with many activities, this broad framework set the groundwork for my efforts to discover the right definition of impact for the organization, the kinds of assumptions that would be important to test and track, and the metrics that are most relevant to the work. While I was hoping to use a more sophisticated framework with a magical algorithm that would help EMD quickly define its impact, no such incantation emerged from the literature. I chose to use this powerful but elementary framework as my guide primarily because EMD already uses the theory of change approach, the rationale being that this familiarity might facilitate adoption of new ideas developed during my residency.
Impact Evaluation Method #2: Spheres of Influence

A major tension within the Theory of Change framework, especially for intermediary organizations like EMD, is the last stage of linking outputs to outcomes and then outcomes to impacts in a causal chain. In complex social environments filled with multiple interventions, economic factors, and political dynamics, it is widely debated whether it is even possible to measure the effect of a single intervention on a given observed outcome (Evaluation Support Scotland, 2017). The consensus, especially among evaluation scholars focused on influence and advocacy initiatives, is that understanding *contribution* is more realistic than trying to assess *attribution* (Rogers P. , 2014). A claim of contribution infers that a given intervention (or cause) is one of many factors contributing to a change (or effect), whereas a claim of attribution infers that the change is solely due to the intervention. It is reasonable to argue that this is a distinction without a difference because both approaches require using a counterfactual (an estimate of what would have happened without the intervention) and other sound evaluation design practices that attempt to isolate the intervention’s effects from other influences, but the shift does have important methodological implications. Quantitative tools and randomized experimental design are the most compelling methods when claiming attribution, while qualitative tools are better suited to claims of contribution (p. 4).

In my early-stage conversations with long-time EMD staff members, this concern about reasonable claims to attribution for an organization that primarily works through partnerships was raised repeatedly. Unclear about how to approach this undeniable reality, I found the Spheres of Influence model (Figure 4) useful for conceptualizing EMD’s interventions along a continuum “with relatively declining influence of an intervention over time and space, within spheres of control, influence, and interest, where outcomes are defined as behavior change influenced by
changes in knowledge, attitudes, skills, and relationships of key actors in a system” (Evaluation Support Scotland, 2017).

**Figure 4: Spheres of Control, Influence & Interest**

![Spheres of Control, Influence & Interest Diagram](image)

Source: Adapted from Belcher & Halliwell, 2021.

A helpful dimension of this framework, especially given the sense of busyness that permeates EMD, is the prioritization that can be facilitated when different activities and actors are mapped and categorized in terms of what is in our control (inputs and outputs), what is in our zones of influence (outcomes) and interest (impacts), and what this looks like over the dimension of time (see Figure 5). Understanding these zones of influence reinforced the idea that there is a natural progression of activity that cannot be rushed, that different information is needed from different actors and stakeholders, and that tools developed through this strategic project would need to grapple with the relevance, feasibility, and plausibility of any given set of metrics.
Taken together, the Theory of Action and Spheres of Influence approaches provide clues about the kinds of clarifying language that can be helpful in building a shared understanding of the goals of an impact framework, but organizational theory is needed to develop a deeper understanding of how these general approaches to evaluation can specifically apply to an intermediary organization like EMD.

Organizational Theory #1: Ecological Systems

The more I have learned about EMD’s work, the clearer it has become that the wide range of clients means it touches different sets of stakeholders. Some work at the system-leader level, some provide direct services to students and families, while others work to affect state, regional or national policy. Recognizing that any impact framework developed as part of this strategic project would need to account for this diversity, and with the Spheres of Influence approach as a reference,
I looked to Uriel Bronfenbrenner’s Ecological Systems model to help me understand EMD’s location within the broader field (See Figure 6).

**Figure 6: Locating EMD in Bronfenbrenner’s Ecological Systems Model**

![Ecological Systems Model Diagram](image)

Source: Adapted from Bronfenbrenner, 1977.

Bronfenbrenner’s model is intended to illustrate the multiple layers of influence on child development, though I am deploying it here primarily to visualize the different spaces where EMD operates and, accordingly, what kinds of impacts (with associated metrics) might be reasonable to aspire towards evaluating. For these purposes, I am intentionally excluding the chronosystem, the dimension of time and history, that Bronfenbrenner later added to the model. The basic model moves from the *microsystem* of actors most proximate to a child (e.g., teachers, family members, direct service providers) and the network of those actors, called the *mesosystem*, to the different forces that indirectly affect that child without interacting with the child. This indirect but very influential layer is called the *exosystem* and includes policymakers, employers, system-level leaders, the media, social service agency leadership, and other actors. EMD would primarily be
considered an actor located in the exosystem, as indicated by the red arrow in Figure 6, because it does not primarily work directly with individuals at the center of the ecosystem, i.e., the youth and adults trying to thrive within the education system. Rather, it works laterally with, for example, networks of superintendents and provosts, technical assistance providers, policy advocates and foundations. What is interesting to me, however, is that although EMD itself is an exosystem actor working to build the capacity of other exosystem actors, it also works with mesosystem actors (e.g., fellowships and networks) that do work directly with individuals in the microsystem such as principals and teachers. Additionally, EMD works with actors that regularly produce reports and research intended to influence the outermost layer, called the *macrosystem*, which represents the attitudes, beliefs, and cultural norms that drive the behavior of all the actors in the ecosystem (Bronfenbrenner, 1977). The two light blue arrows in Figure 6 capture this multilayer sphere of influence that characterizes EMD’s work and, combined with all the other bi-directional arrows showing interrelatedness between layers, provide a map of the different spheres of influence any impact framework needs to consider. Each layer implies different types of metrics and performance indicators, reminding me that it would be a mistake to focus entirely on the changes that occur at the individual layer of the ecosystem. A measure of how, for example, young people advance through the education system to achieve family-sustaining careers is an incredibly important type of impact to keep track of, but if much of what EMD does focuses on other layers of the ecosystem, there are other important outcomes to keep track of to facilitate continuous improvement or strategy re-evaluation.

Organizational Theory #2: Strategic Triangle

While the ecological systems model helped me place EMD within the broader capacity-building space, none of the frameworks that I explored helped me to fully articulate the *importance*
of impact measurement. Certainly, many people within the organization articulated how wonderful it would be to have clearer impact information to share with clients or with the field, but for all the interest and positive affirmations, the fact remains that staff are extremely busy and have limited bandwidth to take on yet one more system or practice. How might the case be made that the extra effort of defining impact, collecting, and analyzing data, and using impact data for strategic reviews are all central to EMD’s mission? Although initially focused on government institutions, Mark Moore’s Strategic Triangle helped me think more critically about the public value that an intermediary organization like EMD offers (See (Figure 7)).

**Figure 7: Moore’s Strategic Triangle**

![Moore's Strategic Triangle](image)

Source: Adapted from Moore, 2000.

This model drove home the simple idea that EMD should be able to clearly articulate how what it does matters to society (public value proposition) and that ensuring the right skills and resources (operating capacity) are in place to properly document that value would be just as important as any framework that emerges from this strategic project process. There are, however,
two additional dimensions of this model that are worth highlighting. First, there is a direct relationship between effectively delivering on the public value proposition—whereby an organization like EMD demonstrates that it is “substantively valuable”—and increasing leverage within the authorizing environment. In other words, the better the proven results, the more responsive decision-makers and others with power are likely to be as an organization advances new methodologies, policy reforms, or transformative ideas. Second, being able to consistently document substantive value has internal benefits as well. People within the organization, especially organizations with big social impact goals like EMD, gain confidence that they are making a difference in the world. In these ways, Moore’s Strategic Triangle illustrates how a robust impact framework, supported by an organization-wide cultural commitment to using it, has the potential to strengthen institutional performance and sustainability in all areas. But how to get there?

Leadership Theory #1: Eight Step Change Model

While the preceding frameworks were reviewed to contextualize and organize my thinking around the content of my strategic project, i.e., theories related to impact evaluation for intermediary organizations, I also needed to think through the day-to-day work. Although my previous experience leads me to question its linearity, John Kotter’s Eight Step Change Model offered straightforward, unambiguous guidance for managing complexity within a fast-paced environment like EMD where the ultimate goal involves establishing a practice that requires a shift in organizational culture. There are three phases of work: Phase 1 and Phase 2 with three steps each, and Phase 3 with two steps that speak directly to the challenge of the durability of any change. Figure 8 illustrates the basic steps of the model:
**Figure 8: Kotter’s Eight Step Change Mode**

<table>
<thead>
<tr>
<th>Phase 1: Create the climate for change</th>
<th>1. Create a sense of urgency</th>
<th>2. Build a guiding coalition</th>
<th>3. Create a vision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 3: Implement &amp; sustain change</td>
<td>7. Hold the gains and build on the change</td>
<td>8. Anchor changes in the culture</td>
<td></td>
</tr>
</tbody>
</table>

Source: Adapted from Kotter, 2009.

This model informed my work plan and, with my background in collaborative design processes, encouraged me to think about how to include multiple voices without stalling the process.

**Leadership Theory #2: Holding Environment**

One of the most challenging parts of residency for me was selecting the right project for the present moment within EMD; there were many interesting potential choices, but I wanted to choose something that would be as much service to the organization as it would be to my own learning. During this assessment stage, it was impossible to ignore the context of the twin pandemics of COVID-19 and racism that permeated everything EMD experienced daily. On one hand, the new national focus on structural inequity was exciting and very welcome. On the other, we all were in mourning and grappling with deep trauma. In thinking about Ronald Heifetz and Donald Laurie’s concept of creating a *holding environment* during periods of transition when there is risk of falling outside of the “productive zone of disequilibrium” (see Figure 9), I wondered if
my strategic project could contribute to the holding environment that EMD leadership was working hard to provide. The holding environment is described by Heifetz and Laurie as, “[t]he cohesive properties of a relationship or social system that serve to keep people engaged with one another in spite of the divisive forces generated by adaptive work” (Heifetz & Laurie, 1999). I became curious about the way implementing a narrow project intending to produce a technical solution might support the organization’s leadership as it tackled omnipresent adaptive challenges.

Theory of Action for Strategic Project

Informed by my analysis of the context and the discussed research on impact measurement, organization theory, and leadership, my theory of action is that in this fast-paced environment, where staff can feel over-extended and overwhelmed by both the volume of their workload and the state of the world, the technical work of creating an impact framework will enable the adaptive work that is necessary for new ways of working to take hold. More specifically:
If I

- understand how EMD works as an organization and the process it currently uses to capture and communicate its impact
- research best impact evaluation practices used by other relevant intermediary organizations, including in adjacent sectors like philanthropy or health care
- build a collaborative process that staff experience as respectful and inclusive
- maintain close communication and coordination with, and secure buy-in from, senior leadership

Then

- we will create an equity-focused impact framework extending to EMD’s sphere of interest
- staff will adopt and incorporate the framework into their routine work
- all stakeholder voices within EMD that wished to participate will feel that they contributed their thinking to the development of the framework
- EMD’s senior leadership will be satisfied with how the organization can describe its impact to clients and the field

So that

- EMD will continuously improve its work because staff are better able to understand what is working and what is not
- EMD’s ability to improve will lead to better outcomes for its stakeholders and greater fidelity to both its organizational mission and theory of change
- there will be high levels of job satisfaction and talent retention within EMD because staff see clear alignment between their work and their values
- partner organizations and other intermediaries will have increased impact evaluation capacity because of EMD sharing its experience and expertise.

Description, Evidence, and Analysis of the Strategic Project

Description

A key concern for the entire EMD team is uncertainty about whether it is equipping social change actors to “diagnose problems, listen, and understand better, develop equitable relationships and communities, and create dramatically more effective solutions to equity problems (Guzman Ortiz, 2017)” While client satisfaction after workshops is consistently high, obtaining reliable data about whether clients go on to “create dramatically more effective solutions” is a major pain point. How does EMD know which of its many tools, offerings, and engagement strategies are most effective or if some should be abandoned in order to focus resources? As the intermediary field
becomes increasingly crowded and “equity” becomes the focus of political campaigns, how does EMD distinguish itself as best prepared to lead in this space? How does EMD continue to improve as an organization, recruit and retain talent, and advance work that not only impacts its own clients but demonstrably improves lives in the frontline communities that many EMD clients work with directly or indirectly? This project was designed to prototype and test some concrete answers to these questions, especially questions about how to conduct robust evaluation in the space between EMD’s sphere of influence and sphere of interest (recalling Figure 5 above).

I spent the first three months of my residency immersing myself in EMD’s client work and learning as much about my colleagues as possible. I participated in weekly all-staff meetings, semester close and strategy review sessions, and the semi-annual company retreat. I also held one-on-one conversations with all team members. While I participated actively in coaching and facilitation, I was initially afforded the luxury of being in a “second chair” position, i.e., not a project lead, while I learned the existing curriculum. This listening tour of sorts helped me gain a sense of both the core content of EMD’s work and the challenges the team experiences. I met regularly with my residency supervisor, who is the founder and chief executive of the company, and with the chief knowledge officer. I was searching for a project that would be a priority for leadership, had the potential to make a long-term contribution to EMD, and would meet some of my own learning goals. These personal learning goals included trying to understand the role of intermediary organizations like EMD as drivers of change in the education sector and deepening my understanding of impact evaluation as a potential lever for more just and equitable societies.

In conversations with every member of the team, I heard that the lack of crisp descriptions of EMD’s impact was a pain point, albeit for different reasons. For those focused on sales and growing the business, there was a strong desire to be able to more forcefully communicate about
documented changes in how clients work post-engagement; for those working directly with clients, there was a strong desire to know that our coaching and facilitation were leading to better outcomes for the frontline communities served by our clients or their stakeholders; and, almost universally, there was a strong desire to know that EMD was learning, growing, and improving with each client engagement. It became clear to me that a project focused on reviewing and upgrading EMD’s existing measurement and evaluation tools and practices would receive support from leadership, would meet the needs of all other staff, and would create an opportunity to build measurement-related curriculum for inclusion in EMD’s capacity-building offerings to clients. The widespread interest in a more satisfactory understanding of EMD’s impact was exciting, but this also made me cautious about assuming the project would lead to a clean framework or an easy set of answers. Indeed, there are an awful lot of smart people inside EMD and beyond who have thought about this extensively. Still, I entered the project believing that if my conclusion ended up being that EMD could not be more data-informed or could not do a better job of tracking progress towards its mission and theory of change, then these would still be useful findings.

With the umbrella objective being to co-create a feasible impact framework that supports EMD in becoming a more data-informed organization, I used Kotter’s three phases to organize the work: 1) create the climate for change; 2) engage and enable the whole organization; and 3) implement and sustain change (See Figure 10, highlighted in orange). Initially, I focused on building a guiding coalition and creating a common vision. Having identified and received authorization for my project in September, I established an aggressive timeline that aimed to align with key annual decision points for EMD, acknowledge the limited time people would have to contribute, and generate enough progress to fit within the tight capstone writing schedule.
Figure 10: Strategic Project Timeline with Kotter’s Phases

<table>
<thead>
<tr>
<th>Month Range</th>
<th>Activity Description</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>September-October</td>
<td>Establish guiding coalition; Set Vision; Gather Information (Creating the climate for change)</td>
<td></td>
</tr>
<tr>
<td>November-December</td>
<td>Research and external benchmarking; Share project at December closing meeting</td>
<td>(Engaging and enabling the whole organization)</td>
</tr>
<tr>
<td>December-January</td>
<td>Develop recommendation and tool prototypes (Engaging and enabling the whole organization)</td>
<td></td>
</tr>
<tr>
<td>January-February</td>
<td>Test prototypes; refine recommendations &amp; next steps; Share progress at January strategy meeting</td>
<td>(Engaging and enabling the whole organization)</td>
</tr>
<tr>
<td>March</td>
<td>Feedback loop; Develop internally facing implementation strategy; Develop client-facing content (Implementing and sustaining change)</td>
<td></td>
</tr>
<tr>
<td>April</td>
<td>Feedback loop; Review, refine &amp; test prototypes; Implementation plan presentation; Residency ends April 30&lt;sup&gt;th&lt;/sup&gt; (Implementing and sustaining change)</td>
<td></td>
</tr>
</tbody>
</table>

Sources: Adapted from Kotter, 2009.

Creating the Climate for Change

My approach to “creating the climate for change” at EMD was primarily relational: I shared my work informally during team meetings, had numerous one-on-ones intended to establish trust and mutual respect, and embedded myself in the day-to-day work of the organization. One of the first steps I took was to create what Kotter calls a guiding coalition comprised of “a team that has the power, information, and credibility needed to lead their organization in a significant transformation effort” (Kotter, 2009). At a large organization, this might mean including representation from all functions and programs, and engaging people with diverse backgrounds, levels of experience and positional power. EMD is a small, eight-person shop, so I put together a guiding coalition composed of the chief executive officer, the chief knowledge officer and me. For Kotter, starting with 3-5 people is fine; the common misstep is putting together a guiding coalition that “is not powerful enough” (p. 44). Having the two most influential people at the company interested in defining EMD’s impact bode well for buy-in and implementation in the future.
Due to everyone’s schedules, I made a conscious decision to minimize meetings and eliminate any pre-work requirements. After a kickoff meeting in which we aligned on a common vision of an EMD capable of using data to both inform decisions (e.g., changes to content, new offerings, client selection) and to understand if our theory of change is working, I convened the guiding coalition for one 60-minute meeting per month. These meetings were opportunities for me to share preliminary findings and musings, check for scope creep, and maintain momentum. They also provided accountability for me, which I found helpful. In between meetings, my primary task was to advance the group’s thinking by synthesizing external research, benchmarking against similar intermediaries as well as those in adjacent fields (e.g., health care, social impact investing, philanthropy, management consulting, data justice), and prototyping new tools. I also established weekly check-ins with the chief knowledge officer, who manages all existing efforts around data collection and use. This was not only a useful way for me to show respect for and learn about existing tools, but it also assured alignment with the person that would be directly responsible for carrying forward any measurement and evaluation system developed through my work.

Based on early-stage conversations, I felt confident that there was sufficient buy-in from senior leadership and urgency among the staff, so I did not specifically plan for a series of actions intended to generate momentum or more demand for the project. In other organizational change efforts, presentations, brown bag conversations, guest speakers, webinars or other activities intended to generate interest, awareness, and momentum are often necessary. At EMD, where every person on the staff is very mission-driven, there was only eagerness and enthusiasm expressed whenever I discussed this evaluation work, so the climate for change felt hospitable.

In the spirit of “creating the climate for change” internally, I also committed to a series of practices that I felt would sharpen my strategic focus and support my well-being during residency.
I scheduled an hour for weekly reflection, blocking out my calendar to ensure it did not get pushed aside. This was an important space for me to keep track of where I was in the process, reflect on insights or decisions, and document my emotional experience. I wanted to ensure that, in these times of trauma and change, I did not default to putting my head down and pushing through as is my habit; that approach did not always serve my health, well-being, and productivity during my time at HGSE. The other practices I committed to and tracked focused on my health: daily meditation and exercise, restful sleep, home-cooked meals, and quality time with friends and family. I used my reflection time to intentionally check in with myself on my consistency with these practices because if I learned anything in the Ed.L.D. program it is that my ability to sustain myself internally directly impacts the quality of my external leadership.

*Engaging and Enabling the Whole Organization: Understanding Existing Practices*

I used the time in between meetings to continue one-on-one conversations with all staff as a means of pressure-testing my assumptions about EMD’s existing data tools and soliciting preliminary feedback about the approaches I was learning about in my external research. As part of broader organizational outreach, I presented brief updates about the project in staff meetings, surveyed team members about their experience collecting and using data with EMD’s existing processes and led an hour-long conversation about emerging data practices at EMD’s semester closing meeting in December 2021. These actions allowed me to introduce the project and create a space for any interested staff members to engage with me further. In this phase, I more actively explored staff aspirations for a new impact framework and sought to understand their perceptions of the existing framework, its strengths and weaknesses, and why its adoption has been limited. Throughout, it was important to me to explicitly acknowledge that I was not working with a blank canvas; to the contrary, the chief knowledge officer had already developed a full suite of high-
quality surveys intended to accelerate both internal growth and the improvement of client-facing work. These internal systems include a well-structured set of reflection tools that are used on a regular basis (some weekly, others quarterly) to keep track of organizational culture and to encourage active re-design of ways of working that are not in line with EMD’s organizational values. For my strategic project, however, I narrowed my focus to client-facing tools and practices as these most closely align with understanding the impact of EMD’s work out in the world.

I learned during my one-on-one conversations that staff interest in being able to better describe the impact of their work was universally very enthusiastic. These discussions were generally casual, over Zoom but frequently with coffee or tea, and framed as an opportunity to explain my project while learning more about my colleague’s specific work. Using a semi-structured approach, I worked in four standard questions:

1. Do you use EMD’s current suite of evaluation tools?
2. If yes, what is working for you?
3. If no, what is not working for you?
4. How would you strengthen it?

The existing framework seeks to understand two basic buckets of information: how individual client participants experience their organizational work culture over time, and how individual client participants experience EMD’s content over time. The tools used are surveys and qualitative interviews. Facilitators are encouraged to build survey time into workshop agendas. Reports are generated by the chief knowledge officer and shared at times intended to encourage their use. For example, post-workshop survey analysis is provided to facilitators on the same day so that adjustments to the content can be made while client-satisfaction survey analysis is provided at the close of a contract (see Appendix 1).
The qualitative feedback suggested that the current impact framework could be more user-friendly, it is not implemented with fidelity in either the collection or use of data, and it is not yielding answers to questions around impact that are satisfying to the team. Some expressed confusion about the differences between some of the differences between various tools and when they should be deployed. Summary reports, by contrast, were unanimously described as “helpful” or “useful” but there was concern that meetings to review these reports were frequently postponed or deprioritized in order to respond to client needs. These data and conversations informed my decision (discussed in more detail below) to refocus my strategic project on actions that would strengthen EMD’s efforts to be a data-informed organization rather than rushing towards a more ambitious overhaul of its evaluation practice.

*Engaging and Enabling the Whole Organization: Exploring the Evaluation Landscape*

While conducting these interviews I also worked on a multi-sector analysis of impact measurement practice to serve as a resource and point of comparison for the team. Within Kotter’s framework, this was part of “enabling the organization” and “removing obstacles” by efficiently accelerating the learning curve within the context of a group of very busy people juggling many competing responsibilities. I used the following guiding questions:

1. Are there standard metrics that can be adopted by EMD?
2. Do other organizations/fields use methods that resolve the challenges of attribution and contribution that vex intermediary organizations?
3. Is there an existing evaluation methodology that is particularly well-suited to the kinds of projects EMD implements and its equity values?
4. Are there a core set of guiding principles that can be used as design specifications for a refreshed EMD impact framework?
I explored the measurement literature, examined peer organization websites and annual reports, and conducted interviews with evaluation practitioners. Specifically, I examined approaches taken in philanthropy (Gates Foundation, MacArthur Foundation, Ford Foundation), impact investing (Acumen, Bridges Ventures, Root Capital) and non-profit management consulting (Bridgespan, FSG, Strada). The Gates Foundation’s Actionable Measurement Framework was particularly useful (see Appendix 3). I reviewed the developmental evaluation approach in international development (USAID, Mercy Corps, Better Evaluation), and the social determinants of health model and co-produced research in public health (World Health Organization). The social determinants of health pushed my thinking beyond just intermediaries and raised helpful questions about what our evaluation approach reveals about our underlying mindset (see Appendix 4). I also explored the emergent field of data justice (Urban Institute, We All Count, Equitable Evaluation Initiative), which focuses on using an equity lens to structure how data is collected, analyzed, interpreted, and communicated (see Appendix 5).

The sector scan quickly led me to answer the first three questions the same way: there is no off-the-shelf method EMD can simply acquire and apply (see Appendix 6). However, when combined with what I learned about EMD’s context, the scan did yield a set of guiding principles or design specifications for a more robust impact framework. I presented a summary of this scan to the entire EMD team during our semester close meeting in December 2021. These are the ideas that I brought forward into Kotter’s implementing phase (though I consider this more of a prototyping phase):

- **Theory of Change:** This approach is the bedrock of virtually every other approach used across sectors and can be very helpful in making important linguistic distinctions related to impact (USAID, Better Evaluation).
● **Method Agnosticism:** A focus on what information is needed and for what purpose may be more useful than any given method (Gates).

● **Evaluation as Relationship:** Valuing partners includes focusing on limiting burden and maximizing mutual utility (Root Capital, We All Count). Embracing partner stories as a vehicle for communicating our impact even if it is “just” anecdotal (Bridgespan, Equitable Evaluation Initiative, Co-produced Research).

● **Evaluation as Culture:** The most effective, pain-free way to conduct detailed results tracking is to embed it in standard work processes (including reflection) and use it in decision-making (Mercy Corps, Gates).

● **Prioritization of Learning:** An approach to evaluation that focuses on improvement and future decision-making in the context of dynamic change is most effective (Gates, Developmental Evaluation, Mercy Corps).

● **Evaluation for Systems:** Though most evaluation tools were designed for project-level assessment, the complex interaction of different forces within an ecosystem calls for layered qualitative and quantitative methods to elucidate systemic problems and potential solutions (Urban Institute, Developmental Evaluation, Social Determinants of Health, Co-produced Research).

*Engaging and Enabling the Whole Organization: Prototype Development*

My review of evaluation practice, combined with my day-to-day experiences working with staff to support clients, prompted me to pause and reevaluate the objectives of my strategic project. It felt important to focus on what EMD most needed and not to just plow ahead with a pet project that matched my interests. I was increasingly feeling like “enabling the whole organization” in this situation might mean ensuring there would be fertile ground for planting a more ambitious set of impact evaluation tools and practices in a future beyond my residency timeline. In discussing with the guiding coalition in December 2021, there was agreement that there are some foundational components that do not yet exist or need strengthening that would be important to EMD becoming a more data-informed organization. As a result, in the prototype development phase I pivoted away
from trying to design a new evaluation framework aimed at understanding EMD’s impact in its spheres of influence and control. Instead, I turned my attention to addressing three core needs: 1) clarifying EMD’s theory of change, 2) ensuring that EMD content and existing survey tools align with the theory of change, and 3) building the collection and use of data into routine work and organizational culture. Though this did not change my entire theory of action for the strategic project, it did eliminate the top-level objective of creating “a refreshed, equity-focused impact framework tailored to EMD’s needs.”

EMD’s theory of change, in its simplest form, is that if it increases the capacity of people in organizations to use the equityXdesign framework to redesign their interactions, interventions, and institutional practices, then those organizations will produce outcomes for their stakeholders that reflect equity and liberation. These are my words, not an official theory of change that is known and understood by all staff. Given the importance of cohesion and alignment in organizational change efforts, I launched a review process of the existing theory of change that involved one-on-one and whole team conversations over three weeks. The goal was to design and test a prototype of a new EMD theory of change that is easily understood and translatable to day-to-day work. An aspect of this prototyping involved interrogating the “capacity-building” portion of the theory of change and specifying which capacities clients would gain through our content that would prepare them to design for equity and liberation in their work. I developed five competencies to capture what participants would work to develop and practice while working with EMD: 1) Equity Analysis, 2) Culture & Power Analysis, 3) Self-Awareness, 4) Direct Communication, and 5) Design for Equity. These competencies define what EMD means by capacity-building and, as a result, make it possible to track what skill development clients attribute to EMD more effectively. The intent is to provide EMD with a structured way to learn whether
those skills change how clients do their day-to-day work and whether there is a relationship between those shifts and changes that their stakeholders experience.

I prototyped competency-focused modifications to three surveys but will only go into depth about one here: the Power Score. This is an intriguing tool developed by Aislinn Betancourt that builds on *The Big 10+1 Ideas that Fuel Oppression* (see Appendix 7). The power score is a composite of baseline and endpoint survey answers to questions that ask respondents to categorize how they feel they are *perceived by others* along several dimensions (skin color, race, class, education, physical ability, etc.). It is not meant to be a proxy for individual identity but, rather, for the *intersectional experience of otherness*, meaning it reflects EMD’s theory that power is often conferred on individuals based on how they are *perceived and engaged by others* (and less often based on how they perceive or identify themselves). This tool is an important component in assessing some of the organization-level shifts that EMD’s theory of change hints at, i.e., the claim that building the capacity of individuals within an organization will shift how and to what extent equity-values show up both internally and externally. The competencies that directly relate to dynamics the power score aims to illuminate are direct communication, power & culture analysis, and self-awareness.

The second area of work in my revised strategic project is mapping these competencies onto EMD’s existing content and building new content to fill any gaps. Although it is inevitable that the content will grow and adapt as the needs of clients change, it is critical that there is always a clear and consistent linkage between the content and the competencies—this is the linchpin of any future impact evaluation efforts that aim to understand changes in the spheres of influence and interest. Accordingly, I initiated a curriculum review and revision process that, in addition to incorporating updates and improvements gleaned from client feedback, serves the purpose of
aligning every EMD activity and session with the competencies. This review revealed that although EMD uses data equity best practices to design and implement its own surveys, there is no content that offers clients an opportunity to learn and develop facility with data equity best practices. Accordingly, building a prototype of an entirely new piece of content to strengthen the foundational capacity of clients to do equity-minded evaluation of their own work became an additional component of my strategic project.

The third area of work in my refined strategic project focuses on EMD’s existing suite of evaluation tools and the accompanying practices that support or do not support the consistent collection, analysis, and use of data. As illustrated in Appendix 1, there are several existing survey tools and several practices designed to reinforce their use. What is on paper and what happens, however, are different things. A simple definition of organizational culture is the beliefs and behaviors that determine how an organization's people interact (Merriam-Webster, 2022). While this third area of work has some technical elements (revising specific surveys for user-friendliness), it is focused on bolstering EMD’s culture around data. As part of my reimagined strategic project, I prototyped ways to clarify for all staff what evaluation tool or practice is used when, and to move from an evaluation culture of compliance to a culture of enthusiastic learning. My guiding questions focused on reducing barriers to use:

1. What will make it easy and fun for my colleagues to collect data?
2. What will make it easy and fun to use data?
3. What would make my colleagues want to use and collect data?

Throughout the course of our work with clients I deployed several small prototypes aimed at promoting increased usage. Examples include taking advantage of the curriculum revision process to build survey times directly into workshop agendas; creating pre-made music playlists
for survey administration based upon glowing feedback from clients about our use of music; sending calendar invites for 15-minute debrief meetings after every facilitation; and hosting data parties at the close of each client engagement. At these remote data parties, each member of the team received a party hat, snacks, and a packet of synthesized survey data. We then reviewed the data together before celebrating the close of the contract. Most of these prototype adjustments to tools and practices were technical (or even silly) in nature, but the intended cumulative effect was to build a shared understanding across the organization of the value and benefit of collecting and using client data consistently over time.

In summary, my strategic project did not proceed along the timeline that I initially laid out, I pivoted to different goals three-quarters of the way through my residency, and significant work remains to achieve the long-term vision for a collaboratively generated framework and toolset that will allow this (or any) intermediary organization to evaluate impact in its spheres of influence or interest. Though this disappoints me, I have been offered a full-time role with EMD following my residency and intend to continue this work. With greater alignment between EMD’s theory of change, workshop content, and survey tools, and the beginnings of a pro-data cultural foundation, I am confident that a prototype impact evaluation framework tailored to the needs of intermediaries and rooted in the principles of both data justice and co-produced research will be ready for testing in September 2022.

Evidence

An important objective for me throughout my residency was to lead a collaborative process and to maintain good communication with the guiding coalition members. Collaborative design is second nature to me after years of fieldwork at MIT D-Lab and Olin College of Engineering, but one of the leadership tendencies that I have been working through since beginning the Ed.L.D.
program is holding challenges close to my vest and trying to work through them on my own. I can go silent, keep digging, and make myself a deeper hole. This is just bad practice, is usually ineffective, and would be disastrous in a small company like EMD. I intentionally kept track of how my communication with senior leadership progressed and surveyed them about their experience working with me. Starting in September 2021, our coalition committed to monthly meetings for updates, idea generation and strategy adjustments. We have met 6 out of a possible 7 months; the January meeting needed to be rescheduled due to a commitment the founder had. I view the consistency of our meetings as evidence of growth in my own leadership.

I also surveyed the two coalition members in mid-March 2022 and asked the following four questions:

1. On a scale of 1-5, with 1 being not at all and 5 being highly, how included was your perspective in developing the strategic project? (N=2, Average 5)
2. On a scale of 1-5, with 1 being not at all and 5 being highly, how informed about the strategic project did you feel throughout the process? (N=2, Average 5)
3. On a scale of 1-5, with 1 being not at all and 5 being highly, how collaborative was your experience working on the strategic project? (N=2, Average 5)
4. On a scale of 1-5, with 1 being not at all and 5 being highly, how enjoyable was your experience working on the strategic project? (N=2, Average 5)

Because these are two women who have absolutely no difficulty being direct and honest, I accept these data as evidence of my investment in relational work. This connects directly to two elements of my theory of action: 1) build a collaborative process that staff experience as respectful and inclusive; and 2) maintain close communication and coordination with, and secure buy-in from, senior leadership.
As described previously, in the prototype development phase I consulted with the guiding coalition and pivoted away from trying to design a new evaluation framework aimed at understanding changes in EMD’s spheres of influence and interest. I shifted my attention to addressing three core areas instead: 1) clarifying EMD’s theory of change, 2) ensuring that EMD content and existing survey tools align with the theory of change, and 3) building the collection and use of data into routine work and organizational culture. Below I summarize the outputs that resulted from my work in these three areas and some qualitative and quantitative data that suggests how they have been received by clients and colleagues.

Clarifying EMD’s Theory of Change

The one-on-one conversations that I had with colleagues early in my residency made me curious about the resonance of EMD’s current theory of change. Each person had a different take on what they hoped to accomplish, why they thought it mattered, and how they thought their own understanding matched the organization’s theory of change. My curiosity led to a brief survey in which I asked three basic questions:

1. On a scale of 1-5, with 5 being the highest, how confident do you feel in your understanding of EMD’s theory of change? (N=7, Average 2.5)

2. On a scale of 1-5, with 5 being the highest, how connected to the theory of change does your day-to-day work feel? (N=7, Average 3)

3. On a scale of 1-5, with 5 being the highest, how confident are you that the theory of change is a feasible way for us to pursue social change? (N=7, Average 2.85)

With this baseline established, I worked with the guiding coalition to refresh and simplify the theory of change. I also developed the competencies to make the focus in the theory of change on capacity building more tangible. On December 13, 2021, I facilitated a discussion session with
the entire staff on the new theory of change and on the competencies. The post-workshop survey asked the same questions and all averages increased significantly:

1. Confidence in understanding the theory of change increased (N=7, Average 4.7)
2. Connection to day-to-day work increased (N=7, Average 4)
3. Confidence in feasibility (N=7, Average 4.2)
4. I added 2 questions that were not on the pre-survey regarding the competencies using a Likert scale of 1-5 (“not at all” to “extremely”). I asked to what degree the competencies were understandable (N=7, Average 5) and were helpful in clarifying our objectives for clients (N=7, Average 5).

While these results are encouraging, they are very limited. I plan to survey this same group again at the end of the current client work cycle (June 2022) to explore how durable these changes are and to explore what it might take to get to 5’s across the board. Regarding the competencies, the prototype baseline survey questions that I designed were launched with 2 clients in Spring 2022 (see Appendix 2). As of this writing, we have not deployed the endpoint surveys yet so I do not have evidence to support my hypothesis that the competencies will help clarify EMD’s theory of change both internally and externally.

Theory of Change, Content and Survey Alignment

Capacity-building is the through-line between EMD’s theory of change, competencies, workshop content and evaluation tools. Evidence of alignment of these elements include 60 hours of collaborative work over 6 weeks to re-design the first three EMD design workshops; the outputs as of this writing are two revised design workshops. Every session within the workshops can now be mapped onto the competencies and many of the adjustments that were made were in direct response to client feedback captured in synthesized post-workshop or client satisfaction surveys.
This revised curriculum and surveys incorporating the competencies are currently being tested with two clients in Spring 2022; we do not have the endpoint survey results yet, but we do have the post-workshop surveys for the first 2 days of curriculum and the baseline surveys. Between the two clients, we received 68 responses:

- 66 of 68 respondents rated their workshop experience as either a 4 (very good) or 5 (excellent).
- On the baseline survey, the aggregated average response across all five competencies (ability to use an equity analysis, use a power analysis, use direct communication, practice self-awareness, and use design tools) was 2.75 on a 1-5 Likert scale (not at all confident to very confident). We expect to see significantly higher numbers in all five areas and, if that happens, will interpret this as evidence for successful alignment.

The other output in this area was a prototype of a new workshop focused on building the capacity of our clients to use data equity best practices in their own work. I co-developed a 4-hour workshop built around an assessment of current data practices and deployed this prototype with 33 participants in Solutions Worth Scaling, an online course offered by the Billions Institute and EMD (see Appendix 8). In a post-workshop survey (N=31), 95% of respondents rated their experience as either very good or excellent, and 90% of respondents indicated that they would recommend the workshop to a colleague. I did not design the post-workshop survey used, so additional data needs to be collected. The next iteration of this prototype will be run during the last week of my residency with representatives from six intermediary organizations that work with “nearly three-quarters of the nation’s public two- and four-year institutions” (Intermediaries for Scale, 2022). This output and very preliminary outcome data is evidence of progress towards
greater alignment and the bigger vision of an evaluation framework for intermediaries infused with data equity principles.

**Routine Work & Organizational Culture**

My main goal was to increase the use of existing tools for the collection and analysis of data. EMD’s client work calendar runs for 10 months and is divided into a 5-month fall semester and a 5-month spring semester. During my residency, we ran design cycle workshops with 14 clients total, 8 from the time my residency began in June 2021 through June 2021 semester, and 6 that are running in the current Spring 2022 semester. An analysis of data practices related to Fall 2021 revealed the following:

- Out of 80 surveys that should have been administered based on existing protocols, 52 were administered
- Of the 8 possible summary reports that could have been produced from these surveys, 5 were infeasible because of a lack of data
- Of the 8 possible after-action-review meetings to synthesize learnings from the summary reports, 4 meetings occurred
- When I asked my colleagues to rate their understanding of which surveys are used when, 2 respondents said *Excellent*, 3 respondents said *Fair*, and 1 respondent said *Poor*, and 1 respondent said *What surveys?* (N=7)

The prototypes I deployed for Spring 2022 workshops were small technical adjustments designed to encourage more consistent use. As noted in the description section, when my colleague and I worked to revise the curriculum, we were sure to build survey times directly into workshop agendas; I created music playlists on Spotify for use during survey administration; I established fifteen minute debrief meetings after every facilitation using calendar invites to easily and
efficiently request the time from relevant colleagues; and I transformed the after-action-review meetings into “data parties” complete with invitations, snacks and party hats. I also facilitated an all-team session on January 17, 2022, to review and discuss each of the survey tools, their timing and analysis, and how they work together to strengthen our work. After the session, I asked them to rate their understanding again. As of this writing, the data from all these actions looks like this:

- Out of 56 surveys that should have been administered based on existing protocols, 56 were administered (4 surveys remain for contracts that have not closed yet)
- Of the 4 possible summary reports that could have been produced from these surveys, 4 were produced (2 remain for contracts that have not closed yet)
- Of the 4 possible after-action-review meetings to synthesize learnings from the summary reports, 4 meetings occurred, 2 in the style of a data party (2 remain for contracts that have not closed yet)
- The “EMD Survey Playlist” on Spotify has 107 followers including all EMD facilitators; it was used in the administration of all 56 surveys
- When I asked my colleagues to rate their understanding of which surveys are used when, 4 respondents rated their understanding Excellent and 3 respondents rated their understanding Very Good (N=7)

I also conducted a short survey after each of the data parties. Of the 6 respondents, 100% rated the experience a 5 (excellent) on a 1-5 Likert Scale (poor to excellent), 100% said the experience was a good use of their time, and 100% said they would keep the data party format over the previously used after-action-review format. The following is excerpted from one of the open-ended responses: “This really felt like a collaborative experience! Digging into the data was fun and we identified some useful insights. I liked the time for solo reflection and then having
open conversation, and I thought the Jamboard was clutch. I low-key used to avoid these debriefs but I would absolutely do this again!” While there is still some outstanding work to be done this semester, these results are very encouraging evidence for small steps toward the strengthened organizational data culture that my strategic project is working towards.

The preceding evidence focuses entirely on the outputs and short-term outcomes related to my strategic project, but throughout my residency I have also been tracking what has and has not happened in relationship to my leadership commitments to myself. As mentioned above, these commitments included an hour for weekly reflection, daily meditation and exercise, restful sleep, home cooked meals, and quality time with friends and family. As of this writing, my personal data over the 37 work weeks during my residency are as follows:

- 37 weekly reflection entries in my journal
- 295 days of meditating tracked in the Calm app
- 588 total workouts tracked in my exercise app (on average 2 per day)
- 7.5 average hours of sleep tracked in the Sleepscore app

Beyond being pleased to see that I kept my commitments to myself regarding my well-being, my ability to be present and mindful as a leader was undeniable. In stark contrast from the extended period of cluster headaches that I experienced two years ago, I did not experience any headaches and felt energized throughout my entire residency. Though it felt unfamiliar and a little embarrassing to keep track of (and write about) these personal habits, I viewed it as a form of self-accountability. Besides, I take Audre Lorde seriously:

I had to examine, in my dreams as well as in my immune-function tests, the devastating effects of overextension. Overextending myself is not stretching myself. I had to accept how difficult it is to monitor the difference. Necessary for me as cutting down on sugar. Crucial. Physically. Psychically. Caring for myself is not self-indulgence, it is self-preservation, and that is an act of political warfare. (Lorde, 1988)
Analysis

The residency year has always been billed as a two-part experience: leading a strategic project and being self-aware about the leadership moves being made while leading that project. In my case, my strategic project itself had two parts: a period searching for an evaluation framework for intermediaries generally and then a post-pivot period focused on the evaluation infrastructure needs of EMD specifically. In this analysis of why I think my strategic project played out as it did, I look at the post-period strategic project work using Everett Rogers’ Five Characteristics of Innovations that Spread: relative advantage, compatibility, simplicity, trialability, observability (Rogers E., 2003). For the purposes of this analysis, the innovation I am going to discuss is the aggregated suite of prototypes that I developed and tested with EMD staff and for which there is evidence of uptake: the clarified theory of action, the competencies, the revised agendas and content, the data parties, the playlist. To look more closely at my leadership and how it affected how the entire strategic project played out, I will bring back Heifetz & Laurie (holding environment) and Kotter (eight-step change model), and I will explore Henry Mintzberg’s Emergent Strategy Model. The section closes with an imagined second opportunity to do this all over again and how I might revise my theory of action.

Five Characteristics of Innovations that Spread

A key takeaway from my years of work in the inclusive innovation space at MIT D-Lab is that added value and adoption are the defining features of what is or is not innovative. It does not matter that the gizmo is new or how enthralled the designer or media is with the gizmo. What matters is that the intended user believes that the gizmo is substantially better than the current way of doing whatever the gizmo does (added value), and that this experience-based belief leads them to replace the old way of doing things with the gizmo (adoption). In Rogers’ model, relative
advantage speaks to this question of whether an innovation is perceived to be better than the alternatives. A plausible explanation of why the EMD team embraced everything I put forward is that the possibility of a more data-informed organization represented by my prototypes was better than the inconsistent and ineffective evaluation practice that characterized the status quo.

Compatibility describes how well an innovation works within the existing context, especially in terms of needs, experiences, or values. In my view, the reason the prototypes I introduced were taken up so well is directly related to the way I framed evaluation as a form of integrity, addressed a pain point the staff was already experiencing, and reinforced the values that drive the company. On a practical level, everything I prototyped fit within existing structures and used the same computer programs. This connects to the third dimension of Rogers’ model which is simplicity, or the degree to which an innovation is understandable, easy to use and accessible. Because I was designing for busy people, I focused on prototypes that could become part of routine work and then zoomed in on ways to make collecting and analyzing data as painless as possible. Playlists may seem trivial or superfluous, but the evidence suggests that being able to pull up a pre-sourced list of quality songs and just press play, as opposed to having to think about song selection or sitting with 35 people in silence on Zoom. Building evaluation time into every workshop agenda, giving presentations about what to use when or about competencies, making data analysis a party—these were all aimed at making things as simple and understandable as possible. I believe my attention to these details contributed directly to the uptake of the full suite of prototypes.

Trialability is the idea that spread occurs when potential adopters have a low- or no-cost opportunity to gain first-hand experience with the innovation. Think about a 7-day free
subscription to a service. I intentionally called these “prototypes” because it communicates that these are in-progress experiments that welcome probing and feedback. Of course, this is the language of design that is already completely baked into everything EMD does, so beyond any leadership move that I made, I think the staff’s willingness to try on these prototypes is attributable to pre-existing organizational culture. In this context, exercising leadership meant reading this organizational DNA correctly and the uptake, as evidenced by both the increased roll out of surveys to clients and the high levels of staff participation in all my experiments and surveys, suggests that I did. Within Rogers’ model, the opportunities that I gave EMD staff to try different prototypes out help to explain why this portion of my project yielded positive results.

Observability is fifth and final element of Rogers’ model and it describes the persuasive power of seeing things in action, whether that is experiencing the innovation firsthand or seeing peers use and benefit from it. Once I pivoted the strategic project, I embraced the bias to action that is inherent in design work, not just for the purposes of getting feedback but to raise the visibility of the collection of prototypes for everyone in the organization to see. If I was co-facilitating a workshop and used the competencies as part of the framing, it was important to share that with the other four facilitators who were not present. Similarly, having people experience a data party for themselves and then also hear from others about how much they enjoyed it was key to gradually building momentum for the adoption of each new practice.

Although Rogers does not claim that one of these characteristics is more important than the other or offer a way to determine which is having the most effect when all are present, I do believe the model has explanatory validity given the presence of all five characteristics in the prototypes I rolled it out. There were other important contributing factors that had nothing to do with decisions I made, such as the EMD team’s predisposition to prototyping and the broader company
culture of experimentation. These provided fertile ground for the prototype seeds I planted and am harvesting in these last weeks of my residency. This part of implementing a project, honestly, is very much within my comfort zone, and I would be a bit embarrassed as a designer if I didn’t hit on all these characteristics. The challenge and growth for my leadership is best seen through analysis of both the pre-and post-pivot strategic project.

*Eight Step Change Model Revisited*

Reflecting on Kotter, one of the common mistakes he warns against is rushing through any of the stages of his change model (Kotter, 2009). The irony is that his model does not really include a phase dedicated to truly understanding the problem prior to creating a sense of urgency, building a guiding coalition, and creating a vision. What if you create a sense of urgency or a vision around the wrong thing? As an experienced designer, I know better, yet in framing my project within Kotter's model, I fell into the trap of jumping ahead to pursue a solution without understanding the context sufficiently. In the project description I discuss using the first three months to learn about how EMD works, but I really learned primarily about the way EMD works with clients. This was useful, but I did not pay enough attention to the existing tools, practices, habits, and beliefs specifically related to evaluation. Looking at my theory of change, the first statement is “if I understand how EMD works as an organization and the process it currently uses to capture and communicate its impact…” In practice, I did a good job of understanding the core functions of how EMD works but an inadequate job of understanding (or paying attention to and factoring in) its evaluation processes. The evidence for this is my realization more than halfway into my residency that EMD is not quite ready for a new evaluation framework to be laid on top of its existing foundations.
In my second pass at analyzing the context, I did a better job of seeing the strengths and gaps specifically in the evaluation tools and practices that already exist at EMD. When I came back down from my flight of fancy through the numerous approaches to evaluation used in other sectors, I was able to look at and see what was directly in front of me the entire time: a set of strong tools and systems that were only being sporadically implemented. I realized that I had become the proverbial hammer looking for a nail, where my hammer was the “refreshed, equity-focused impact framework tailored to EMD’s needs” referenced in my theory of change. When I stopped and looked for ways to strengthen and augment existing systems, the project started to gain traction and the behavior changes I described above began to be evident.

A related lesson from Kotter is the importance of orchestrating short-term wins (Kotter, 2009). I use the word “orchestrate” intentionally because Kotter’s point is that waiting for short-term wins to naturally emerge can sometimes be fatal and instead “rigging the game” is necessary to maintain momentum, especially in change processes that happen over years. A feature of the first half of my residency was that I did not ask much of my colleagues other than for one-on-one conversations. Conscious of being a bothersome addition to already busy schedules, I proceeded as if the strategic project was all mine for the first several months (a flavor of under communicating). The polite but relative lack of engagement during the first portion of my residency can be contrasted with the traction and participation that followed my pivot. In this second phase, I was much more active in asking people to use surveys, to provide feedback to me by taking surveys and to suggest pathways to improvement.

Overall, I do not think that using Kotter’s three phase, eight step change model was a misstep; it is too linear, it is very top-down, and it treats change almost like a one-off event, but it still helped me to think about what a successful pathway for this project might look like. The basic
steps of a) securing buy-in from leadership while simultaneously building broader support through formal and informal co-creative engagements with other staff, b) leading a cross-functional team that could facilitate piloting and implementation, and c) grounding the work in a review of both existing practices within EMD and external practices from multiple sectors were all productive and, in fact, these steps laid the foundation for the work that will be completed before and after my residency ends.

*Emergent Strategy*

One of the reasons the strategic project unfolded the way that it did was predictable and, in fact, articulated from the outset. Impact evaluation is *hard*. Impact evaluation for intermediaries is *harder* and there are no easy answers. Although I was very conscious from the beginning that many smart people have tried to tackle these questions before, my initial theory of change still reflects an over-ambition that did not sufficiently consider the time it takes to make progress in a fundamentally complex setting. Initially, I brushed aside any concerns about my expanding portfolio of clients and kept reaching for a grand unifying evaluation theory for intermediaries. I was slow to recognize that even if I knew exactly how to do robust impact evaluation for intermediaries, as a team we were not ready to onboard an entirely new body of work. I grappled with this realization—that I was trying to work on a big problem with important implications but did not think I had chosen the right thing to work on for EMD specifically—largely by myself for too long. I expand on this unhelpful habit of self-isolation in the section about implications for my own leadership. Would changing strategies be giving up, a kind of failure? These thoughts had me feeling a bit stuck for a couple of months, but because I had been holding myself accountable for clearer communication precisely for moments like these, I reached out to the guiding coalition for help. That is when we re-scoped the strategic project and that decision
unlocked all the prototyping, testing and feedback that made this a successful, albeit different, strategic project.

The work of Henry Mintzberg on emergent strategy predicts that these kinds of adjustments in strategy are commonplace (Mintzberg, 2007). Fundamentally, Mintzberg argues that all successful organizations, in the face of uncertainty and dynamic change, move through a progression from intended strategy to deliberate strategy, where unrealized strategies drop off and emergent strategies are added, ultimately leading to the realized strategy (see Figure 1). In this respect, one of the reasons why I had to shift strategies during implementation of this change project is that it is inevitable that strategy must adjust to conditions. This does not negate the fact that there were errors that I could have avoided with better analysis, and that more nimble leadership could have mitigated some of the challenges I faced, but I do think there is some truth to the fundamental need for re-evaluation and flexibility when leading change projects. In the words of Michael Quinn Patton, one of the founders of what became developmental evaluation, “We have to keep changing because the world keeps changing” (Patton, 2012).

**Figure 11: Mintzberg’s Emergent Strategy Model**

My initial theory of change was not wrong as much as it was embedded with untested assumptions about EMD’s context and my own ability to seamlessly carry a full client portfolio while leading the strategic project. The project has been neither completely successful nor a complete failure; after a reboot, I adjusted my strategy and leadership moves and made good progress towards a commonly shared vision. Additional contributing factors include a fantastic organizational culture that centers reflection, growth, and change. The production and rapid testing of specific outputs with clients exemplifies this supportive, experimental culture at EMD. Precisely for this reason, it is hard not to feel disappointed that more was not accomplished since many of the obstacles to change that are normally present in organizations were not at work here. Nonetheless, Mintzberg’s Emergent Strategy model has helped me embrace the idea that adjustments to strategy are always going to be a feature of leading change efforts.

**Holding Environment Revisited**

Bringing back Heifetz and Laurie’s concept of the holding environment and its critical role in stabilizing the productive zone of disequilibrium, clear and steady communication, or the lack thereof, had a major impact on why the project unfolded as it did. It is important to highlight here what has been happening all around us for the duration of this residency—successive waves of a deadly, global pandemic that has impacted every facet of life, ongoing extrajudicial killings by police, a massive white backlash from both extremist groups and elected officials against this country’s purported “racial reckoning.” Disequilibrium might be an understatement. In this context, I did a good job of building relational trust with my colleagues and being an active participant in organizational culture, but I did not do as well in specific regard to communicating with colleagues about my strategic project. I discuss this tendency to under-communicate in more detail in the implications section below, but in terms of trying to understand why things happened
as they did, I think under-communicating about the project added anxiety or confusion for me and others in an already stressful time. In my theory of change, I refer to countering this as “maintain[ing] close communication and coordination” and “build[ing] a collaborative process that staff experience as respectful and inclusive.” In practice, I did not always do a good job of this in the first phase of the project and the evidence of that is that my colleagues basically stopped asking me about what I was working on in the 3 months between when we first did one-on-ones and when I did an update during the December closing company meeting. I allowed the evaluation project to be crowded out by everything else going on in work and the world and, more importantly, I did not lead in a way that gave others confidence in a time of uncertainty. By contrast, over the last three months, there has been much better communication on my part about the tools and practices that have been updated, the support for using these systems, and how doing so fits in with the larger vision of both being more data-informed and understanding EMD’s impact. I have received inquiries via email and feedback on prototypes, which is a level of engagement that, frankly, I did not make sufficient room for in the initial phase of the project.

Although the holding environment is most frequently discussed in terms of a leader keeping their team within the productive zone of disequilibrium, there is also the dimension of the leader maintaining themself. As Heifetz writes: “The work of leadership demands that you manage not only the critical adaptive responses within and surrounding your business but also your own thinking and emotions” (Heifetz, Grashow, A., & Linsky, M., 2009). Three of his five recommendations for how to do this resonate: finding sanctuaries, reaching out to confidants, bringing more of your emotional self to work (p. 153). I believe firmly that I was able to recover from early missteps and successfully pivot because I maintained the practices that I identified as sustaining at the start of my residency. More than the number of workouts or reflections, the
strongest piece of evidence that these self-sustaining practices were essential to how the strategic project unfolded is the pivot midstream. At another time, I might have just tried to stubbornly push through. With sufficient internal reserves, I was able to see more clearly and make better choices, including the choice to work through alternative pathways with the guiding coalition.

Revised Theory of Action

One of the most powerful tools in the equityXdesign curriculum is called The Problems with Problems: A Guide for Using Problem Definition as a Tool for Equity (Ortiz Guzman, 2021). The core idea is that “we’re so caught up making sure we’re doing things that we aren’t making sure we’re doing the right things the right way. Too often entrepreneurs, innovators and philanthropists get into the work they do to bring an idea they have to life, ideas that are generated outside the context of where they will be used” (p.11). I coach on this material 4-5 times a week. I spent 15 years at MIT D-Lab focused on collaborative design and the democratization of innovation, all intended to end the all-too-common practice of “experts” parachuting in solutions for “beneficiaries.” Yet, in the kind of “do as I say, not as I do” moment I’ve only experienced as a parent, the fatal flaw of my original theory of action is that I embedded the solution in the definition of the problem. That’s Problem #1 in the 7 Problems with Problems rubric (p. 18).

While it is true that I spent my first 3 months at EMD learning about the context and seeking to identify a good strategic project, it is also true that the reason I was interested in doing a residency at an intermediary is that this question of how intermediaries can do robust impact evaluation had been with me for years. I designed a project and then created a theory of action that was locked into my particular vision for what solution EMD and other intermediaries need. Given an opportunity to start again the first part of my theory of change would focus on more intense problem identification and exploration phase that engaged all members of the organization. I
would also include an organizational readiness assessment step, either creating something or using a tool like the Organizational Readiness to Change Assessment instrument (Helfrich, Li, YF., Sharp, N.D., & et al., 2009). I would still include phases for outside research and reviewing internal processes, but I would put this on a much tighter timeline (2-4 weeks) to design for my tendencies to endlessly pull on threads that are interesting but non-essential to making decisions about the path forward. I think the rest of the theory of change is mostly solid and the majority of the “then” side came to fruition.

**Implications**

**Implications for Self**

A definition of leadership that resonated with me during the Ed.L.D program comes from Marshall Ganz: “Accepting responsibility to enable others to achieve shared purpose under conditions of uncertainty” (Ganz & Lin, 2011). Throughout this capstone I have tried to take responsibility for what has worked well so far and what has not worked as I initially planned, but what I have yet to really explore is why my timeline got so severely derailed. I was clear from the outset about the tight time constraints inherent in the residency year—10 months really, with writing deliverables shaving off at least 4 months from actual implementation of the strategic project. Still, I managed to immerse myself in all kinds of other EMD work in such a way that, regardless how productive I was as a facilitator and coach, delayed my strategic project from really getting going until the fourth month. While acknowledgement alone will not allow me to complete the initial vision for the project, I hope honestly laying bare my own contributions to what has transpired thus far will allow me to rest, refocus, and redouble my efforts in the final phase of my residency and beyond. Additionally, by digging deeper into why I did not more forcefully step into
my own leadership during the first 4 months of my residency, and why it took me so long to recognize the need to pivot, I hope to harvest some forward-looking lessons. Three come to mind: 1) inaction is an action; 2) complex problems are complex enough without any help, and 3) a sense of responsibility can be worn as weights or as wings.

One lesson is that while observing and taking time to learn about an organization is often the wise and appropriate step, it is possible to slip into an ambiguous zone where “orienting” and “learning” can become indecision and inaction. If Ganz is correct, and leadership is about accepting responsibility, then I must bring forward lessons about the consequences of taking too long to grasp that my original vision for this strategic project was not the right sequencing for EMD’s impact measurement journey. While I am glad that I eventually pivoted and adjusted to conditions, I can’t ignore that I could have accomplished more if I had not stalled by holding on so tightly to my initial idea. Fundamentally, as a leader, I must act and trust that, if I am working respectfully and collaboratively with people, we can create positive forward motion together. Stagnation and inactivity, however, has a corrosive, insidious energy to it that, if not course corrected, can seriously undermine important work that needs doing.

I have fallen into that kind of stuck-ness a few times over the course of the Ed.L.D. program and I can see that sometimes it is a web of my own making. An example from this project is how I took the already challenging issue of impact measurement, layered on questions around equitable data that could have been a project on its own, and then tried to think beyond EMD to all intermediary organizations. Things opened for me significantly on this project when I downsized the scope and re-focused my efforts on attending to my relationships with my colleagues, i.e., getting them the tangible things, they were telling me they needed if they were
ever going to start using data differently. It is important for me to remain aware of how overreach and overcomplication can be tendencies of mine that negatively impact my leadership.

A secondary dimension of this complexity issue is that I really want to tackle hard problems with quality work, and when I don't feel like my work is of high enough quality, I keep it to myself until I think the work is “ready.” It is a classic example of “letting the perfect be the enemy of the good.” Despite the many encouragements of the team at EMD, I sometimes kept this project a little too close at the expense of many low-stakes opportunities to share, benefit from feedback, and strengthen the data-positive culture I was trying to foster. I have been reminded through this residency that accepting the responsibility of leadership also means embracing the importance of clear communication. To meet my leadership goals, I must remain diligent about keeping track of my tendencies to over-commit and under-communicate; the combination is poison to any project or organization.

If I dig deeper, the root cause of these tendencies bumps squarely up against a version of survivor’s guilt that many people experience, albeit in different ways. My version goes something like this: I have been so privileged to live the life I have lived—to have had two loving parents that were able to navigate inhospitable, racist and classist systems to secure opportunities for me that then opened doors to academic institutions and employment opportunities that are inaccessible to most; to be in a position to have had jobs that enable me to both work on issues that align with my values and support my family; to have had the luxury of stepping away from the world of work to immerse myself in learning through this extraordinary doctoral program—and, therefore, in my mind, everything I do needs to justify this good fortune, this luck, this responsibility. I know whose shoulders I stand on—the cotton pickers, day laborers, launderers, porters, cooks, hairdressers, seamstresses, teachers, social workers in my family—what debt is owed, and what the other side
of the so-called birth lottery looks like. This awareness transforms into my own big aspirations. This capstone is part of one decades-long struggle to wrap my head around how to put all my passion, power and, yes, privilege, into service in the face of great personal, intergenerational, and societal trauma.

With that in mind, I see that I was drawn to this project about understanding impact because EMD wants to figure it out and many people on staff need the organization to figure it out, but on a personal level, I am perpetually seeking to understand how I will know if I am having a meaningful impact in the world. Or enough impact. On many occasions, instead of being fully present in the opportunity to build my capacity to be a more effective leader, I have pushed myself to the limits trying to make sure that doing this program has not just been a “selfish” endeavor. There have been family obligations and hardships that have called, teaching opportunities, service opportunities, venture opportunities, connecting with frontline community opportunities, and I have tried to answer all of them out of an internal compulsion to never let others down while pursuing a “fancy” degree. To never forget who I am. And as with a lot of self-sabotaging behavior, regardless of the intent or justification, the potential is there to get myself into a cycle that is hard to break out of.

The implication for my leadership has to do with keeping track of the emotional dimensions of leadership and the resourcing that is required to be a healthy contributor to any organization or sector. In trauma therapy, resourcing is defined as “identifying and instilling coping skills to help you deal with difficult reactions that you may experience.” At times, I made myself small in this project, not because I lacked the right analysis or the technical skills to adapt midstream; I made myself small because I let the sense of responsibility, I carry weigh me down instead of inspiring me to push through whatever adversity needed to be taken on. As embarrassing as it may be to
admit that publicly to all those who invested in me financially and otherwise, it is a truth that I 
need to take responsibility for as I step towards future leadership roles. If I want to lead, I have 
learned that I need to change my relationship to my privilege, or to my interpretation of what I 
need to do with it to be of service to others, that I need to attend to my physical, emotional and 
spiritual well-being, and that I need to allow myself to be however big the moment calls for.

Implications for Site

There are immediate steps discussed above that will advance the project beyond its current 
status. Together with leadership, I intend to identify a suitable client to test prototypes of a new 
evaluation system built upon the following ideas:

1. **Leverage Clients for EMD’s Evaluation System**

   One of EMD’s strengths is its relationships with clients throughout the education 
ecosystem. Used intentionally, this asset can serve as a foundation for an efficient and 
cost-effective source of data, insight, and feedback. Presently, EMD’s interaction with 
clients is primarily one-directional: EMD provides content and coaching while clients 
return little more than a few surveys focused on their own experience. Shifting to a 
data-informed culture could involve relying on a “partner-centric” approach to data 
collection that transforms what is normally an extractive act of compliance (surveys) 
into a mutually beneficial exchange around the experiences of stakeholders. To do this 
most effectively, I suggest that the subsequent phases of work on EMD’s impact 
framework engage clients in the design process, and eventually engage client 
stakeholders. A collaborative effort would increase the probability that what ends up 
being implemented is “rigorous, respectful, relevant, and right-sized” from everyone's 
perspective and ensures that clients (and their stakeholders) feel that it is in their best
interest to participate in EMD’s evaluation system. The evaluation field is increasingly moving in this direction. For example, a relatively young developmental evaluation approach called Collaborative Outcomes Reporting begins with the premise that stakeholders, not evaluators, should be the ones to define what the success, impact, or value of an intervention is, which results in a process centered around “most significant change” narratives generated by stakeholders (Dart & Roberts, M., 2022). These same principles animate the Data Equity Framework (We All Count) and the research co-production model in the UK and Canadian health care literature. EMD is well positioned to do this kind of collaborative work with clients given the duration (6-12 months) and intensity of a typical client contract. As relationship-based data development and use becomes a part of the “standard” work in a learning culture, a more pervasive, client-centered evaluation system has the potential to become a standard part of project design and contracting, thereby incrementally strengthening EMD’s evaluation capacity in a manner that pays for itself.

2. **Build out Design Cycle 4 as Evaluation Capacity Building**

Currently EMD has three primary design cycles (DC) that last anywhere from 3-6 months, depending on scheduling constraints. Clients work their way through one or two iterations of a prototype or a suite of prototypes that address pressing organizational equity problems. If the client wants to continue, we then custom design a DC4 and DC5 based upon whatever direction the client feels is most pressing. I believe that there is enough general demand for exposure to equitable data practices and enough alignment with our own evaluation aspirations to justify transforming DC4 to focus entirely on client’s designing or redesigning their evaluation tools and
practices. I have already begun building this material out—we have tested one assessment tool and received excellent feedback, will test a 4-hour session at the end of April, and we already have the internal capabilities to build out a full DC. Not only would this extend client capacity for their own work, but it would also serve as an important foundation for extending EMD’s evaluation reach into its spheres of influence and interest as recommended above. Equitable data practices also fit well within the Equity Analysis competency.

3. **Explore Social Network Analysis to Understand Sphere of Interest**

When thinking about measuring impact, it is advisable, as per the Gates Foundation, to remain method-agnostic and to carefully match approach to intended purpose. I agree with this way of operating and think that there is a range of low-cost methods available to EMD that meet most of its evaluation needs. That said, social network analysis tools may be particularly useful as we think about impacts on our clients’ stakeholders in our sphere of interest (Social Network Analysis, 2022). These tools are well-suited to tracking the spread of the kind of capacity-building EMD does and can help characterize the effectiveness of that work with network-specific metrics, e.g., health, centrality, connectivity, durability. What is powerful about this specific tool, and why it carries so much potential within a broader evaluation strategy for EMD, is that it can allow for focused attention and learning around movement building and other social change strategies. Social network analysis could also benefit from alignment with the principles of research co-production and data justice. Building specific capacity around network evaluation and equitable research co-production will position EMD to continue to learn how to improve one of its strongest assets (its
network) while also leading the field toward a clearer understanding of the strategic imperative of an ecosystem approach to social change in the 21st century.

4. **Continue Shifting Evaluation Mindset towards Learning**

EMD obviously already believes in the importance of learning—it is a capacity-building organization!—but when it comes to evaluation, everything is currently summative and backward looking. The received framing for this project was originally about communicating with clients and potential clients about what working with EMD can accomplish; this is too narrow a view of what evaluation can do for EMD and, frankly, a missed opportunity. EMD’s current culture is primarily focused on delivering quality support and information to clients and others in the design for equity field, and while it would be unfair to say this is entirely unidirectional, from my perspective, there are many more opportunities for EMD staff use data for decision-making purposes. There is a strong culture of reflection at EMD, but this is largely focused on personal growth, individual professional growth and interpersonal dynamics; these are all essential but distinct from using information regularly for organizational learning and decision making. A strong learning culture is about ongoing communication and informed decision-making within project teams, across programs, and throughout leadership. This can be framed as part of EMD culture—as a commitment to greater integration, learning, and “actionable measurement” in the language of the Gates Foundation—and it can be an explicit shift named and cultivated by leadership (see Appendix 3). There are so many wonderful existing learning features of EMD’s culture, from “back from break” kickoffs to each semester to mid-semester retreats to end-of-semester strategy reviews; these offer immediate
opportunities to meaningfully shift evaluation in this direction. Additionally, this strategic project necessarily focused narrowly on client-facing tools and practices, but I think it would be fascinating in future work to explore the relationship between EMD’s internally facing reflection tools/processes and its externally facing evaluation tools/processes, with specific focus on both value-alignment, how these tools interact with each other, and how they impact the way EMD carries out its external work.

**Implications for Sector**

While this capstone describes my efforts to lead a discrete process within a specific organization transitioning through a particular moment in its evolution, there are also clear implications for intermediary organizations trying to better understand the value they add to their field. We frequently talk about the scale of the complex problems within public PK-12 education—that we are trying to transform learning experiences for 50 million children across 14,000+ districts with endemic social, political and economic challenges so that a child’s zip code or ticket in the birth lottery does not predetermine their future opportunities. Given this daunting task, we cannot afford to ignore the potential of intermediary organizations that sit outside of school districts but seek to strengthen and contribute to their advancement. If EMD alone works with clients that indirectly and directly touch half of all public-school students in PK-12 and higher education institutions in the United States, it would be safe to estimate that a handful or two of intermediaries have the potential to influence the experiences of every student in the country. Given vast influence and resources, I believe it is reasonable to ask these intermediaries to be accountable for measuring their impact in their spheres of influence and interest. What intrigues me most of all, despite lessons from the serious downsides of the “accountability era” in education, is how well-considered
evaluation may meaningfully contribute to our efforts to advance system-wide equity and excellence.

Working on this project has activated my thinking about the potential role for evaluation using the Six Conditions of System Change framework and its two core assertions that 1) advancing equity requires “shifting the conditions that hold a problem in place,” and 2) that “shifts in system conditions are more likely to be sustained when working at three different levels of change: explicit, semi-explicit, and implicit” (Kania, Kramer, M., & Senge, P., 2022). Figure 12 illustrates these conditions and levels. If an intermediary like EMD can develop the capacity to do better impact evaluation that is learning-focused and partner-centric, others can as well, and if we seize the opportunity to develop shared metrics that are far more informative about the many

**Figure 12: Six Conditions of Systems Change**

![Six Conditions of Systems Change](image)

Source: Kania, Kramer and Senge, 2018
factors that contribute to the learning environment for our children, then perhaps we can set positive system-level changes in motion at all three levels of this model. Here I am reflecting on the social determinants of health framing that I believe is substantially enabled by the quality and quantity of data that is available to the public health community (see Appendix 5). NAEP and PISA scores are not going to cut it on their own.

What strikes me most about the social determinants of health model is the explicit acknowledgement that the domain of medical providers—protective interventions, clinical interventions, and preventative care—is insufficient and can only truly be effective through connections to other systems (housing, transportation, education, the environment) and institutions (governments, think tanks, media). Is it not also true that domain of educators on its own cannot address the needs of all students? Is it not also true that the domain of law enforcement on its own cannot address the safety needs of all members of society? The social determinants of health model is not usually drawn as concentric circles but is reminiscent of Bronfenbrenner’s Ecological Ecosystem model. I believe the data and change measurement system that has been in place since the 1990s in public health plays a major role in enabling systems thinking and an almost intuitive understanding of the network effect, which is desperately needed in social impact spaces like education.

Comprehensive evaluation developed in collaboration with students, parents, teachers, school leaders, social service providers, community-based organizations, health care providers, public institutions and the private sector might create deeper insights into the kinds of supports each child needs to successfully navigate not just school but the world around her/him/them. To my knowledge, this type of comprehensive evaluation does not yet exist, but if, as Pedro Noguera says, equality is everybody having shoes and equity is everybody having shoes that fit, then greater
visibility into the needs of all kinds of learners from all kinds of backgrounds is essentially a prerequisite to equitable education systems (Routman, 2018).

The way I see this connecting to the Systems Change framework is that actionable measurement, to again borrow the term used by the Gates Foundation, is a way of shifting practices at the structural/explicit level of the model; a vision of measurement and evaluation that highlights the ecological system every learner is embedded in and affected by would highlight the relationships and connections at the relational/semi-explicit level of the model; and without question, a holistic view of the experience of learners and how our values about housing, health care, transportation, nutrition, the environment, the arts, and work shape the efficacy of our classrooms requires a new mental model at the transformative change/implicit level.

As a thought experiment, take the Just City Index advanced by Toni Griffin through the Just City Lab at Harvard Graduate School of Design. The Lab has been collaboratively “developing and testing a set of core principles, values and metrics to assess and evaluate design’s role in achieving urban justice” (Veldacademie, 2022) Community meetings were used to generate principles and values-based indicators for what living in a just and equitable city means to everyday citizens. The metrics center around thirteen core principles: Acceptance, Aspiration, Choice, Democracy, Engagement, Fairness, Identity, Mobility, Power, Resilience, Rights, and Welfare. Within each of these principles, there are between two and seven values-based indicators and definitions for those indicators, including Belonging, Diversity, Cooperation, Accessibility, Accountability, Freedom, and (my personal favorite) Delight.

Now think about the standardized tests our children are subject to for 9-12 weeks out of the academic year in some districts, not to mention how that shapes the weeks of instruction before and afterwards. What if we had ways of measuring learning that reflected the principles and values
that we hold most dear? For me, these are the possibilities and implications that my strategic project holds out to the sector.

**Conclusion**

This residency and strategic project were important learning experiences for me personally and professionally, and though I have not yet completed my project to my satisfaction, I am committed to ensuring that EMD is well positioned to systematically deploy a set of tools that build its understanding of the contributions its capacity-building work by the conclusion of my residency. There are many more leadership lessons for me to harvest, particularly related to re-imagining my relationship to my own privilege and the responsibilities that comes with it. While difficult, the residency has helped me find a pain point in myself that is similar to the challenge EMD and the sector have with evaluating their impact: we have big aspirations, and it is only through right relationship to self and others that we have a chance of reaching them.
Bibliography


https://www.youtube.com/watch?v=b7n64JEjUUk&list=UUUi_6Ij8IgUAzI6JczJUVPA&index=3


Veldacademie. (2022, March 17). What is the Just City Index? Retrieved from Issuu:
https://issuu.com/veldacademie/docs/masterclass_design_for_a_just_city_web/s/10230780
Appendix

Appendix 1: EMD’s Existing Evaluation Tools and Practices for Client-Facing Work

<table>
<thead>
<tr>
<th>Survey tool, what it aims to evaluate and how it is used</th>
<th>When tool is deployed</th>
<th>Practices to support data collection + use</th>
</tr>
</thead>
</table>
| **Post-Workshop Survey (10 mins)**  
Participant feedback on what specifically went well/what could have gone better during the session, what questions remain.  
This information is used to help the EMD Project Team to make custom adjustments to content and facilitation. | After every facilitated session | Survey time built into workshop agenda. Facilitator Debrief (usually during weekly client review meeting) |
| **Client Satisfaction Survey (15 mins)**  
Participant feedback on what, overall, went well and what could have gone better  
This information is used to help EMD identify areas of improvement in content, facilitation, and project management. | At the end of contract (before the full project debrief with client project team) | Survey time built into workshop agenda.  
Debrief Day at the end of each semester. This is a high-level review of all client work with the goal of determining what should become standard practice moving forward. |
| **Progress Survey (10 Mins)**  
Captures how prepared individual participants feel to use eXd frameworks and tools  
This information is used to assess gaps in participant learning so that content and facilitation can be improved. It is intended to understand if we are building capacity. | During the last day of Design Cycle 3, 4 and/or 5 | Client-facing report generated before the Final Project Debrief with Client Project Team; reviewed internally first.  
“Rice & Beans” is the name for scheduled content improvement sessions. 5-8 meetings scheduled each semester.  
Best Practices meeting at the start of each semester. Entire EMD team reviews best practices, decides which to implement, what needs to be built in order to implement, and adjusts project management/onboarding. |
| **Client Baseline Assessment (15 mins) & Client Endpoint Assessment (25 mins)**  
Measures changes in organizational culture.  
This information is used to understand whether EMD is helping organizations to shift important elements of their culture. | Baseline during first workshop of 3-part Design Cycle 3  
Endpoint last day of Design Cycle 3 (3-6 months later) | Initial client-facing report generated before the start of DC3.2  
Final client-facing report generated by official close of contract (pre/post analysis still in beta and not offered at this time) |

**Source:** Adapted from EMD Surveys and When to Use Them, August 2021
Appendix 2: Competencies & Corresponding Survey Questions

- **Equity Analysis**: the ability to see and articulate the ways beliefs, practices, structures and systems produce disproportionate outcomes for historically excluded groups and to integrate that analysis into work projects that address unjust impacts of policies and practices.

- **Culture & Power Analysis**: the ability to see and articulate the ways our work culture reinforces oppressive ideas and practices, and to integrate that analysis into work projects that address the role of power in shaping decision making, design and implementation.

- **Self-Awareness**: the ability to utilize reflective practices that increase awareness of internal activation and emotional charge and bring empathy to interpersonal interactions with multiple group identities.

- **Direct Communication**: an ability to communicate clearly and directly with colleagues that is grounded in a commitment to work to resolve conflicts, to balance curiosity, respect and consent, to attend to both intent and impact in interactions, and to work constructively across differences.

- **Design for Equity**: the ability to apply and integrate the equityXdesign framework and its accompanying tools into work projects.

Pre- and Post-survey Questions. The surveys use a 1-5 Likert scale and ask respondents to self-report on their confidence in each skill or practice. Below are some examples (the post- version for questions 2-5 follow the same format as question 1B).

1. **Pre**: How confident do you feel in your ability to apply an equity lens to understanding a problem or the experience of people at the margins of a problem?  
   **1B**: Post: Compared to your level of confidence before taking this course, how confident do you now feel in your ability to apply an equity lens to understanding a problem or the experience of people at the margins of a problem?
2. **Pre**: How confident do you feel in your ability to identify and examine the ways cultural practices within your organization (including the ways you work) affect the outcomes of your work?  
3. **Pre**: How confident do you feel in your ability to maintain awareness of your own tendencies, habits, biases, and emotions as you engage in equity work?  
4. **Pre**: How confident do you feel in your ability to communicate directly with people in your organization or on your team about your thoughts, feelings, and aspirations related to engaging in equity work?  
5. **Pre**: How confident do you feel in your ability to use design tools and processes to define problems and create equitable solutions?
Appendix 3: Actionable Measurement at the Gates Foundation

![Diagram](image)


Appendix 4: Equitable Evaluation Principles

![Equitable Evaluation Principles Diagram](image)

Appendix 5: Five Social Determinants of Health & Relevant Interventions


Appendix 6: Map of Measurement Methodologies to Measurement Objectives

Source: So and Capanyola, 2016
## Appendix 7: 10+1 Ideas That Fuel Oppression

### The Big 10 (+1) Ideas that Fuel Oppression

<table>
<thead>
<tr>
<th>Idea One:</th>
<th>Idea Two:</th>
<th>Idea Three:</th>
<th>Idea Four:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lighter skinned people deserve more love, power, affection, wealth, grace, and dignity than darker skinned people.</td>
<td>Males are smarter, more trustworthy, better leaders, more responsible, stronger, and more honest than females.</td>
<td>Richer people are smarter, more trustworthy, more responsible, and deserve aspiration and grace more than poorer people.</td>
<td>Christians are more trustworthy, righteous, and more justified in their violence than non-christians.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Idea Five:</th>
<th>Idea Six:</th>
<th>Idea Seven:</th>
<th>Idea Eight:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterosexuals are more natural and deserving of love, dignity, humanity, and companionship than LGBTQIA+ folks.</td>
<td>Gender conforming people are more natural, deserving of love, companionship, dignity, and humanity than gender nonconforming people.</td>
<td>English speakers with dominant culture accents are more intelligent than non-English speakers or those with different accents.</td>
<td>People who are disabled are less intelligent than those whose abilities conform to the majority.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Idea Nine:</th>
<th>Idea Ten:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young adults are seen as smarter, more creative, more energetic, and more employable than older adults, teens, and children.</td>
<td>Adults with college degrees are smarter than adults without college degrees.</td>
</tr>
</tbody>
</table>

### Stability Clause:
These ideas cannot change and will continue to govern the relationships of future generations.

---

*Written by Caroline Hill*
Appendix 8: Prototype of 8 Problems with Evaluation Self-Assessment Tool

8 Problems with Evaluation

We think it is really important to ask the right questions at the right time as you work to develop your solution. What you want to avoid is starting to ask yourself if your solution really works when you have already rolled out manufacturing at scale or have fully deployed a project in multiple locations! Unfortunately, in our experience, these are not hypotheticals. The right time to ask this question, in our view, is in the prototyping phase, prior to piloting and definitely before scaling. So how do you assess or evaluate whether your solution works? You are all at different stages in your solution development journey, but consider this an opportunity to interrogate your thinking at any stage. What follows is a list of common problems we see when teams try to evaluate their solutions followed by a series of questions for you to work through as a team. Alternatively, consider looking through Continuous Improvement for Equity—Testing Interventions for a different set of self-assessment questions.

1. We decide what it means for our solutions to “work” in an echo chamber

Frequently, evaluating what works is done without any input or direction from those experiencing the problem. Not only is who sits at the table important, who sets the menu, who decides what is eaten, and who prepares the meal are all critical in determining what counts as validation of our solution.

- What does it mean for your solution to “work”?
- Who was involved in determining that?
- Who or what was excluded?
- How does your solution “work” in relation to existing systems that hold the problem in place?
  Does it challenge or disrupt these systems at all?

2. We only collect data to prove things, not to learn what can improve.

When we only focus on making sure that every piece of a solution is perfectly implemented “with fidelity,” we miss what we can learn from adaptations.

- Who uses the data you currently have and how do they use it?
- Does the data you collect help people at every level of your organization (from leadership to frontline staff) make decisions about what to keep doing, stop doing, or change/adapt?
- If not, why not?
- If not, what kinds of information would help you and your team make those
• decisions, how might you get that information, and who might you get it from?
• What might be an effective way to share that information, so that you and your
  team can access it when you need it?

3. We over-rely on quantitative methods and data.

While surveys can help us understand which outcomes have (or have not) occurred (and to what degree), qualitative methods can help us identify why these outcomes have (or have not) occurred. We like to imagine that quantitative data is somehow more objective and less subjective to interpretation; both assumptions are false.
  • What quantitative data do you currently have?
  • What assumptions are you making about what it means?
  • What questions still linger as you look at this data?
  • Where might qualitative data help you to check your assumptions and/or provide
    further insight into what the quantitative data cannot say?

4. We don’t include equity or justice when asking whether our solution “works”

This often means we haven’t understood and accounted for how our intervention affects or perpetuates existing drivers of inequity. The Equitable Evaluation Framework explicitly challenges us all to think about how “to approach [our] work in ways that contribute to equity.”

This includes asking:
  • Does our evaluation of our solution include an equity lens?
  • Does our evaluation ask whether our solution has been developed commensurate with the values underlying equity work?
  • Does our evaluation answer critical questions about:
    o the ways in which historical and structural decisions have contributed to the condition to be addressed?
    o effect on strategy of the underlying systemic drivers of inequity?
    o ways in which cultural context is tangled up in both the structural conditions and the change initiative itself?

5. We don’t explicitly account for culture

When we don’t name culture, we risk interpreting data through our own cultural biases or the dominant cultural perspective. This risks distorting our understanding of whether our solution works and, specifically, the spreadability of our solution. Read through the chart below and then consider:
  • What is your current understanding of the different cultures at play in your particular problem and solution spaces?
  • Do you have a good understanding of the ways that culture and power are at play?
• What steps can you take to eliminate bias as you try to understand if your solution works?
• What other methods might be useful for you to attend to cultural differences?

<table>
<thead>
<tr>
<th>AEA Statement on Cultural Competence Essential Practices</th>
<th>Definition of Equity-Focused Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledge the complexity of cultural identity.</td>
<td>A judgment made of the relevance, effectiveness, efficiency, impact, and sustainability – and, in humanitarian settings, coverage, connectedness, and coherence – of policies, programs, and projects concerned with achieving equitable development results. It involves a rigorous, systematic, and objective process in the design, analysis, and interpretation of information in order to answer specific questions, including those of concern to worst-off groups. It provides assessments of what works and what does not work to reduce inequity, and it highlights intended and unintended results for worst-off groups as well as the gap between best-off and worst-off groups. It provides strategic lessons to guide decision-makers and to inform stakeholders. Equity-focused evaluations provide evidence-based information that is credible, reliable, and useful, enabling the timely incorporation of findings, recommendations, and lessons into the decision-making process.</td>
</tr>
<tr>
<td>Recognize the dynamics of power.</td>
<td></td>
</tr>
<tr>
<td>Culture is not neutral. Cultural groupings are ascribed differential status and power, with some holding privilege that they may not be aware of and some being relegated to the status of “other.”</td>
<td></td>
</tr>
<tr>
<td>Recognize and eliminate bias in language.</td>
<td></td>
</tr>
<tr>
<td>Language is powerful. It is often used as the code for prescribed treatment of groups.</td>
<td></td>
</tr>
<tr>
<td>Employ culturally appropriate methods.</td>
<td></td>
</tr>
<tr>
<td>The methods and tools used for collection, analysis, interpretation, and dissemination of data are not culture free. … Culturally competent evaluators seek to understand how the constructs are defined by cultures.</td>
<td></td>
</tr>
</tbody>
</table>

6. We hide disparities - intentionally and unintentionally - instead of highlighting them.

We cannot address what we cannot see. We cannot see what we don’t disaggregate. We try to create simplistic representations or stories from data, over or underestimate the meaning of trends, and overuse things like averages.

• What do you know about the impact of your work across different demographics and subgroups?
• Do you collect and use demographic or subgroup data?
  o If not, what kinds of demographic or subgroup data might help you understand your problem, your solution or your stakeholders?
• When you analyze data, do you look for differences in experience/outcome based on demographics or subgroups?
  o If not, what differences in experience/outcome would be most useful to learn about?
• What kinds of demographic or subgroup data might help you understand how

From: Raising the Bar - Integrating Cultural Competence and Equity: Equitable Evaluation, p. 82
your solution is or is not working for your stakeholders?

- When you notice differences or disparities, do you do further research to understand what is causing them?
  - If not, how might you find out what’s causing these differences or disparities?
- Who might you talk to? What might you ask them?

### 7. We don't qualify our outputs, and worse, we confuse them with outcomes.

Outputs are things we can easily count (e.g. number of trainings we did, the number of attendees at each training, etc.) and generally are things within our control. Outcomes are changes that the people experiencing the problem want to see and feel in their lives (e.g. what happens for each attendee when they apply what they learned from the training, what differences emerged in their work/life/etc.) and are generally out of our control. Outcomes basically tell us what has changed. Oftentimes, for example, people think that the more trainings they do, the more outcomes they will have - but what if the training is ineffective? What if the people taking the training weren't the target audience?

It's important to know what outputs tell us and what they don't. Otherwise, we are at risk of manipulating data for simplicity or clarity of story. Check out these basic definitions and then ask yourself:

- What is the theory of change of our solution?
- What are the inputs, outputs, outcomes and impacts?
- What are your desired outcomes? Put differently, what are the changes that people experiencing the problem want to see and feel in their lives?
- How do you currently gauge whether or not those outcomes (changes) are happening?

### 8. We don’t account for power.

This often means we haven’t understood and accounted for power dynamics between those researching and those being researched. We don’t consider the needs or desires of those being subjected to our evaluation work. We create systems that are so burdensome that we turn the evaluation of the work into the work. We design systems to serve our needs and timelines. We don’t make data immediately available or useful to the frontline or community. All of these shortcomings can serve to distort our picture of whether our solution actually works and can instead simply confirm what we want to hear.

- How are you accounting for the power disparities that exist between those who are asking if your solution works and those who are being asked?
Additional Resources (Optional)

Continuous Improvement for Equity-- Testing Interventions

The critical inquiry questions, also known as equity considerations, aim to provide an entry point to examine processes with a list of equity-centered prompts for teams to consider. When using this resource, keep in mind: intentionally infusing, prioritizing, and sustaining an equity lens to an improvement project often requires a mindset change for those engaged in this work. This list reflects important considerations that teams must keep in mind during both planning and implementation of improvement processes, including:

• Starting with the self: Examining the layers of identity & perspective • Seeing the system: Examining context & history
• Examining the lens & locus of influence: Elevating authentic voice
• Shifting the paradigm: Valuing process & outcomes

The Data Equity Framework

When it comes to equity in data science projects, trying to find all the ways that bias, assumptions, unfairness and prejudice can sneak in may feel overwhelming. Trying to look at a whole project and see all the equity weaknesses and issues is almost impossible. That’s why We All Count has developed: The Data Equity Framework. You don’t have to recreate the way your team already works with data. You simply need to learn how to apply the lens of the Data Equity Framework, follow simple steps and checklists and apply practical tools. The Data Equity Framework is a systematic way of looking at data projects. It organizes every project into 7 stages.

10 Lessons from Health Care on Quality Improvement

Pitfalls and promising practices drawn from experimentation with quality-improvement methods and performance management in health care.

If You Don’t Know Who You’re Impacting, How Do You Know You’re Making an Impact?

Decades of social science research have underscored the relevance of aspects of identity on how individuals are likely to engage their environment, their community, and their world. These factors are not deterministic, nor static—they often interact in complex ways with many other factors—but neither are they trivial or irrelevant. So the analysis of the effect of any intervention on any community cannot fully be assessed if we do not collect and assess data along these parameters and use it to inform our efforts.

How to Write Better Demographic Survey Questions

“Demographic data collection can be somewhat intimidating. The desire to be sensitive to a diverse population can often conflict with a researcher’s need to segment their data. And there’s nothing more disheartening than completing a survey project and realizing that you omitted key demographic questions that would have given you a deeper understanding of your data. To avoid experiencing this sinking feeling, consider including at least a few demographic questions regardless of your survey’s subject matter.”
Respectful Collection of Demographic Data

Demographic data may be critical to your mission as a community center, legally required diversity disclosure of a corporation, or an idle curiosity of a blogger to understand their followers. Whatever your reason, this article establishes some guidelines for respectful design of your form and language for collecting demographic data:

1. Ask affected communities for their input.
2. Identify whether you truly need all of the information you ask for.
3. Explain your purpose and your privacy policy.
4. Offer multi-select checkboxes, not single-select radio buttons.
5. Allow users to self-describe.
6. Do not require a response.
7. Consider your defaults.
8. Consider the presentation and influence of your survey.

Developmental Evaluation

"Developmental Evaluation supports innovation development to guide adaptation to emergent and dynamic realities in complex environments. Innovations can take the form of new projects, programs, products, organizational changes, policy reforms, and system interventions. A complex system is characterized by a large number of interacting and interdependent elements in which there is no central control. Patterns of change emerge from rapid, real time interactions that generate learning, evolution, and development – if one is paying attention and knows how to observe and capture the important and emergent patterns. Complex environments for social interventions and innovations are those in which what to do to solve problems is uncertain and key stakeholders are in conflict about how to proceed."

VIDEO: Planning and Evaluating for Social Change

“Dr. Michael Quinn Patton, an internationally-acclaimed evaluation and social change innovation expert, explores the expanding landscape of evaluation within complex systems and environments. [His talk is] for educators, planners, researchers, evaluation practitioners, and anyone interested in how a developmental evaluation framework can act as a vehicle or language to connect people, enabling them to be more effective and strategic in their social change work.”

Why Am I Always Being Researched?

“If we do not address the power dynamic in the creation of research, at best, we are generating partial truths to inform decision-making in the social impact space. At worst, we are generating inaccurate information that ultimately does more harm than good in our communities...This guide was designed to help us level the playing field and reckon with unintended bias when it comes to the research. It is about shifting the way community organizations, researchers, and funders ask for, produce, and use knowledge. It is about restoring communities as experts. It calls on us all to stop, recognize, and question bias. It is an equity-based approach to research.”