An Interview with Chelsea Mauldin, Executive Director of the Pubic Policy Lab

M: How do you ensure that your work is "good" for the people you're setting out to help?

CM: We make a whole series of affirmative efforts to do no harm. We have an ethical research and design process that makes it clear what we're doing, why we're doing it, and what potential damage could arise from participating. This is necessary because part of our work involves going back to powerful people and telling participants' stories. Our research is not aggregated to the complete point of anonymity. It is possible that a person could be recognized, and they need to agree that they're okay with that possibility.

If you don't realize the potential for harm, you can't plan for how to mitigate it. Most of the time when we are working in public sector context, there are a lot of procedures to protect the agency and not the end user, but that's changing.

M: You talk about measuring an outcome "en route." How do you do that?

CM: As well as we can! There are projects where we've said to our partners, "we would like to measure the process of implementing and the ultimate impact of that work," and essentially people have just said "no," because there is often no time, budget, authority, or access to do M&E as a part of this work.

There is also the question of how to collect data on these projects. There are no control groups. We are almost always working inside a complex service delivery context, proposing changes to systems, policies, and procedures. So, to design a 'clean' RCT is not viable, given the scale of time and budget available.

So how do we tell if we are doing anything useful if we can't do a post-implementation evaluation? We can do a process evaluation to determine if people were able to use the thing we designed as we designed it. Or, we do a field test where we are making an evaluation of our design product.

M: Were you able to integrate implementing M&E at the same time?

CM: The NYC DOE project "Connecting Families to Public Benefits" was 16 weeks from start to finish, from preliminary scoping to delivering the report. We did research with frontline staff and experts, then went back out to schools with two rounds of prototypes,

getting to five different streams of potential deliverables, then we trained staff at three schools to use these different design outputs. We observed while they used them for a month and conducted surveys during the same time to get a sense of how many people our prototypes reached and what their experience was with it. We did more qualitative interviews with the school staff at the end of month and got input on how they were able to implement. Staff members were excited that they were able to track the process that well.

M: How did school staff respond to PPL reaching out and asking for feedback?

CM: They were happy to respond because they knowingly signed up to be part of the pilot program. We did lightweight surveying throughout the pilot, plus   qualitative exit interviews with key participants at the end of the field test that lasted 30 minutes, and people were happy to give us feedback.

At the end of the project we presented all the activities of the project to leadership from the DOE, funders, and project participants, such as the community school directors that participated in pilots. It was an interesting thing for people in an education environment to see this way of working.

M: Did you get push back on the sample size or the way the data collection was done?

CM: We did say 'proof of concept' over and over. We wanted to make sure people understood the small scale and preliminary scope of the engagement. There certainly were questions about the sample, but people understood once we explained the goals of this initial effort.

There's a tension between the desire for outcomes data and what it requires to collect it. A funder asked these two questions, in this order: "Do you think you could collapse field test and piloting phases to save budget? Do you think you'll have enough time and data to evaluate the project?" It is hard to demonstrate clear impact if you only have four weeks for field testing. The conversation around this is changing, but the funding is not changing at the same pace.

M: Was the measurement piece in this work different than what you've done before?

CM: We have gotten a lot smarter about telling people in advance that we're building a measurement process into the work. There is always pressure as to the amount of time

and budget. What we've come to see is that it is more effective to say, We will run research, design, and lightweight field-testing as a single process, which gives us some evaluation we can use.

One more reflection : When we started the organization, we thought we would spend most of our time designing services for implementation. What has happened over the years is an increasing demand for designers as strategic thought partners.   .We are able to deliver recommendations n D, but doing new things is really hard. New things involve change, and most people aren't comfortable with new things that they didn't come up with themselves. We now have a set of methods to ensure that our partners work with us to generate change.

1.   We embed partner staff in our design team, so we have members of the agency staff actively participating in field researchd, processing field notes, generating design concepts. This is extremely helpful. We get the advantage of agency staff members' deep organizational knowledge of what will and won't work. We offer government staff members a new way of working that they find emotionally engaging and satisfying. The people embedded will have ownership over the implementation of the design product once Public Policy Lab leaves. They have internalized a bunch of the approach and reasoning on why the product is what it is.

2.

3.   We do design pin-ups with agency leaders, as well – standing meetings, giving critical feedback, and putting up notes and information on a wall. The goal is for leaders to participate in a hands-on way. At each point we ask them for feedback on the viability of carrying the ideas forward to implementation. We say, "If the stuff we're doing leads to a product you can't use, tell us early and often."

4.   The point of measurement isn't to measure – the point is to find out how to improve people's lives. Even if we don't have the scope to do the kind of post-assessment we'd like, we can do a lot to build implementation success into the design process.