# Notes From a Fellow: Making Secondary Data a Priority

February 26, 2021



*At last count, the Office of Evalution Sciences estimates there to be more than 200 teams worldwide working on applying behavioral insights to improve government services. More and more psychologists are likely to join these teams in the coming months and years. While PhD students and postdocs don't have to look farther than their nearest advisor or instructor to find examples of an academic career, they may have a harder time getting a feel for what it would be like to work with a government team.*

The end of January 2021 finds me approaching the midpoint of my fellowship with the Office of Evaluation Sciences (OES). In many ways, the time has flown by; it has me realizing how short a year can be and hoping that some of the projects we're talking about now will make real progress before I rotate off the team. Many former fellows stay involved with OES as academic affiliates, working part-time on specific projects—and I hope I will, too. But still, I'd love to see at least one of these fledgling ideas become something concrete that I'll be free to talk about when I'm back to teaching and university life. (Most government work stays under wraps until it's complete.)

Heather Kappes (New York University)

One element of OES work that's somewhat unique is that the office only uses administrative data that's already being collected by government agencies. Other behavioral insights teams sometimes use this kind of data, but they also conduct surveys or online experiments to measure new outcomes. At OES, one of the building blocks to deciding if a project is a good fit is whether there are administrative data capturing the outcome of interest.

> *One thing I like about using administrative data is that you can be pretty sure that the outcome is something that policymakers care about, since they're bothering to measure it.*

The government collects a lot of data, so usually something relevant to a particular research question exists; it might be physicians' vaccination rates, the number of tax credits that filers claim, or energy use in public housing units. But just because the data exist doesn't mean they're easy to get ahold of, even for another government entity like OES. I've been learning about the world of data-sharing agreements. If you're curious about these agreements, you might enjoy the relevant chapter from a [new handbook](#) from the Abdul Latif Jameel Poverty Action Lab. As the author explains, negotiating the boundaries of who can receive what and under which restrictions they can use it can be a months-long process. That's quite a change if you're used to Mechanical Turk data that arrive an hour after you post your study.

One thing I like about using administrative data is that you can be pretty sure that the outcome is something that policymakers care about, since they're bothering to measure it. What's less convenient is that this kind of data usually doesn't give much insight into the psychological mechanisms that underpin an effect. We can see that someone clicks a box or submits a form, but we don't know what they're thinking or feeling while they do it. Just another reason why—from my perspective—lab and online experiments, and even qualitative data, are a necessary complement to big field experiments with administrative data.

[Learn more about OES in "Federal Agents of Change: Behavioral Insights Power Evidence-Based](#)

[Efforts to Improve Government."](#)

To come back to the "time flies" theme, OES recently concluded recruiting for the 2021–2022 set of fellows. Sharing the advert on psychology listservs and with psychology contacts made me think about things that psychologists could do to prepare for this type of role. One of those things is gaining comfort analyzing secondary data. That doesn't have to mean negotiating a data-sharing agreement. There are lots of large survey-response data sets (the World Values Survey, for instance) that are publicly available. Some people in our field specialize in working with this type of data; the challenges include finding proxy measures (it's rare that a survey question was asked exactly the way you would have asked it) and cleaning and managing large data sets. If I were advising my grad-school self, I'd tell her to get experience doing this. (Oh, and APS has a [useful resource](#) for ideas about that!)

*Feedback on this article? Email* [apsobserver@psychologicalscience.org](mailto:apsobserver@psychologicalscience.org) *or post a comment.*