MATHEMATICA Policy Research

ideas42

Department of Labor Behavioral Interventions Podcast

with Greg Chojnacki (Narrator), Samia Amin (Project Director, Mathematica Policy Research), and Matthew Darling (Behavioral Science Lead, ideas42)

This project was funded by the Chief Evaluation Office of the U.S. Department of Labor (or DOL) under Contract # DOLQ129633249/DOL-OPS-16-U-00126. The views expressed are those of the authors and should not be attributed to the Federal Government or the Department of Labor.

Greg: Each one of us makes an enormous number of decisions every day. Even if you're not making a major decision about your personal life or job, you'll tackle many small decisions—what clothes to wear, what to eat for breakfast, or which task to take on first when you get in to work. Think about it, how many decisions have you already made today?

Many of these decisions are guided by habit, the information available at the time, or other constraints. And some of our choices, including those with the potential to have the greatest impact on our lives, are shaped by a predictable set of reactions to the world around us. Everything from the TV programs we watch to how much money we save every month could be influenced by these reactions.

Applied behavioral science is becoming increasingly influential in program and policy design. This is a field that draws on research in behavioral economics, psychology, and other social sciences to provide a more nuanced understanding of people's interactions with policies and programs and identify ways to improve them. Researchers in behavioral science study how a range of factors—such as limits on time, attention, and the ability to quickly process complex information—may systematically cause people to behave in ways that defy the predictions of our standard decision making models which assume that we are always acting perfectly rationally. These insights from behavioral science are being applied to cafeterias, health insurance plans, and even in governments in the U.S., Britain, and other countries.

My name is Greg Chojnacki, and today we'll be looking at a new study funded by the U.S. Department of Labor's Chief Evaluation Office to test the use of behavioral science in DOL programs. Here to dive into this evaluation with us is the study's director, Samia Amin, a senior researcher at Mathematica Policy Research, and Matthew Darling, vice president at ideas42 and the behavioral science lead on the project.

Samia, your team ran three trials, all of which addressed very different issues in the labor context. Can you give us a little background?

Samia: Sure Greg, I would love to give you an overview of this very exciting work we've been doing over the last three years! You know, when we began, the objective was to see whether insights from behavioral science could improve programs and outcomes for a range of different labor programs and we wanted to do this for programs that were operating in different contexts. And we've been really lucky to find terrific agency partners who've been eager to test innovations with us and have been instrumental in the tests we've run, which we often refer to as trials. So why don't I give you some background on the three trials we've run. Our first partnership was with the Occupational Safety and Health Administration, or OSHA. And their priority was to test pilot changes that could improve workplace safety,

specifically by prompting employers who were cited for health and safety violations to resolve those citations. Then, our next partnership was on an equally important topic: increasing employees' willingness to save for retirement. We partnered with the Employee Benefits Security Administration, or EBSA, and DOL's Human Resources division to test whether targeted emails would lead to DOL employees contributing more to the agency's retirement savings plan. And then for our third trial we collaborated with the Employment and Training Administration for a partnership with Michigan Works! Southwest, a local workforce agency, and the W.E. Upjohn Institute. And here our partners were passionate about understanding if simple ways to encourage take up of a reemployment program could result in positive outcomes. This program is mandatory, it's offered for free, it's been shown to work, but many unemployed workers just weren't showing up and our partners wanted to understand if we could change that. So as you can see, these are three very different efforts operating at different scales. And we're at that stage where we've just wrapped up all three trials and are really excited about sharing what we've learned.

Greg: Before we learn more about the trials themselves, let's dive a little deeper into what exactly behavioral science is, and how insights from it can be applied to the strategies agencies might already be using. Matt, why don't you talk to us a little bit about that?

Matt: Sure Greg! So we all know that people sometimes act in ways that are surprising or perplexing, or can just make mistakes. Researchers in behavioral science have identified a number of reasons for this, such as limited attention, undervaluing future benefits compared to those in the present, or our tendency, when we're faced with difficult decisions, to substitute a simpler problem and then solve that instead of the one we're actually confronting. They have also shown that the context in which people make decisions can have a surprisingly large impact on what decisions they make.

These insights can help program designers to find ways to address barriers—or behavioral bottlenecks—that go beyond what is typically included in traditional program design. Solutions to behavioral bottlenecks can be small, such as tweaking a letter or email you might be sending to include a checklist of action steps, or large, like changes to the way in which options are framed–for example, setting the default to be participating rather than not participating.

These behavioral strategies account for the fact that in the moment when people come across an agency's letter or email, for example, they may have competing demands for their attention and time. The letter may have been one of 10 that came in the mail that day, which they're trying to open while helping a small child eat dinner. To complicate matters, people also don't have all the information at their fingertips that they might want in order to make a decision and act on it. As a result, they often respond differently than how they would in an ideal world in which they have unlimited time, attention, and information. Addressing bottlenecks like these can help people take actions that are in their own self-interest AND that align with the goals of DOL agencies.

Greg: That's really interesting. Could you illustrate that a bit more with an example from one of your trials?

Matt: Sure, so okay, so we all know that saving for retirement is important—everyone is always encouraging us to save right? Yet it seems that every week, there's another story in the news about how we under-save. In our retirement savings trial we found that more than a quarter of DOL employees were either not saving or not saving at least 5 percent of their salary to get a full employer match. And we wanted to know why people were leaving money

on the table. Now traditional economic theory would assume that every employee balances his or her desire for a comfortable retirement against the downside of having less money in the short term, and then follows through on this calculation to choose an optimal level of savings. So this problem should not be occurring, but it is. Behavioral science and our findings suggest a number of behavioral bottlenecks can come together to prevent people from acting as this traditional economic approach would predict. They might include the absence of clear information on how to set savings levels, anxiety about making financial decisions, or just the sheer number of things on everyone's to-do lists. When there are multiple bottlenecks that appear to be preventing people from acting as they otherwise might, then there's a real potential for behavioral insights to make a difference. In situations like these, even small, low-cost changes can improve people's outcomes—in this case, for DOL-HR, these types of changes can support employees' ability to save for retirement.

So that was a central part of what we did in the retirement savings trial. We sent people two emails, a month apart. Those emails started by bringing the benefits of adequate savings—and of getting the full employer match—to employees' attention. The emails also showed employees that taking a few simple steps would suffice to change their contribution level and reassured them that they could change their contribution at any time. This simple pair of emails led to a significant increase in the number of people who began to save enough to get the employer's, that is, DOL's, full matching contribution, which translates into meaningful savings over the course of their careers.

Greg: Wow, and Samia, am I right in recalling that all your trials yielded similar positive results?

Samia: That's right. You know, we found that across the board when it came to many of the outcomes that these programs care about, and that our partner agencies cared about, some pretty modest changes in how the programs contact their target audiences led to measurable improvements. So Matt's already spoken to you about the findings from the EBSA and DOL-HR trial. So let me round that out with our findings from the remaining two trials. So, do you remember the trial I mentioned where we used handouts, letters, and reminder postcards to contact employers? We actually managed to reduce the number of employers who were referred to debt collection by OSHA because of unresolved workplace safety and health violations. And [in] the third intervention—where we sent a series of emails informed by behavioral science to people on unemployment—we managed to expand their participation in a reemployment program.

We find these results really exciting because they yield insights on how other programs facing similar problems might test solutions. All the innovations that we tested were scalable because they were easy to implement, they were low-cost, and they could really be adapted to the constraints and opportunities experienced by our implementing partners. You know what we found when we couldn't do a high-effort fix (for example redesigning a confusing website), we could do a low-cost hack such as just coming up with a one-page visual with screen shots to help people navigate the website easily. So that was pretty empowering for our team. And we think these will be of interest to a wide audience. For example, the findings from our retirement savings trial are likely to be useful to businesses, policymakers, federal/state/local benefits administrators, and to others thinking about retirement security. So what we are hoping is that others will learn from our efforts and move forward. We are sharing more information about the intervention designs and findings for each trial at the Department of Labor's Chief Evaluation Office website. We have a full technical report, a brief, a poster, and a video on each of these three trials.

Greg: So in hindsight, the value of this effort is clear because the programs you worked with improved their outcomes. But at the beginning, what was the motivation for DOL to initiate these trials? What did the Department hope to gain?

Samia: Well, Greg, as you mentioned at the beginning, insights from behavioral science are being applied more and more widely every day, across both the private and public sectors. In sponsoring the trials, DOL shared with us that their motivation was to rigorously test whether behavioral insights could improve the performance of the Department's programs. Ultimately, DOL wanted to generate evidence that [could] help labor agencies meet their mission of advancing opportunities for American workers, as well as inform the behavioral science field more broadly.

DOL's starting premise was that many of their programs already engage in communication and technical assistance with the general public, and they reach out to specific communities, populations, industries or firms, a lot of actors really, so there might be a lot of scope for making program improvements that are guided by insights from behavioral science. But first they wanted to rigorously test if that was actually true, and understand how behavioral insights could lead to these improvements, in what contexts, and for whom.

So we started reaching out to various DOL agencies and program administrators to see what problems they had and find out if behaviorally informed innovations could be implemented in a way that their results could be rigorously measured. And at this stage, DOL was keen on looking for options that could be rapidly implemented and could yield measurable, short-term outcomes. So why this emphasis on short term outcomes? This was driven by the goal to determine quickly whether innovations informed by behavioral science were working, so we might have the opportunity to further iterate and test alternatives within the scope of this study. And fortunately for two of our three trials, we ended up implementing a second phase, and we have richer and more nuanced findings as a result.

Greg: So many of these improvements that you both have described just sound like common sense solutions to these problems. I'm curious, how are the changes you tested in these trials different from what someone might do simply just by following common sense?

Matt: Thanks Greg, I can answer that. What's important to note is that our first step was to diagnose behavioral bottlenecks and to design solutions that are grounded in rigorous evidence. That's not always the case when it comes to common sense. The evidence we relied on gave us good ideas about what works in practice and what doesn't—and those findings weren't always those that "common sense might suggest." Let me give you an example. Common sense might suggest that sending a job seeker an email or letter with a long list of job openings available in their area and their sector might motivate them to apply. And this might work if they mistakenly believe that there are no jobs available. But if the behavioral barrier is different, and in fact they are not putting sufficient time into looking for work because they are overconfident in their prospects, then that email will actually be counterproductive. Or if they are overwhelmed by the prospect of having to look into so many opportunities and figure out the right fit, then that email again might make them panic and delay. So understanding the true drivers for behavior first, *before* opting for a common sense approach is really important.

For all of our trials, we based our diagnosis on real data and information about the individuals whose behavior we were hoping to change. Understanding these individuals' perspective helped us to find the barriers that programs didn't know they had and really hone in on where people's behavior was departing from what we might otherwise expect. I should also say that we didn't do this work of diagnosing behavioral bottlenecks alone. This project really relied on insights from DOL. Their program managers and documentation of their procedures were incredibly important, and were something we used along with administrative data and interviews with program participants to get a full picture of the environment in which people are engaging—or not engaging—with the program.

Greg: So what did you learn about applying behavioral insights specifically in the Labor context?

Matt: We found it interesting that there were some barriers that showed up in all three trials. In particular, we found reason to believe that the simple fact of people's limited time and attention may be preventing them from responding to all three programs as the agencies would like. For example, we learned from OSHA that employers who did not respond to a citation were more likely to be small businesses, and we know that in these types of businesses, the owners wear many hats, so the attention that they can devote to any one task may be limited. Once we have this potential barrier in mind, we have a starting point to address it by piloting changes to program operations or communications.

And then we learned a lot about devising behaviorally informed strategies that align with existing agency procedures and priorities. For example, with two of our partners, OSHA and the local workforce agency in Michigan, we needed to develop solutions that complemented the compliance or enforcement aspect of those agencies' mission. How the compliance or enforcement activities were conducted and the language used in existing communications needed to stay intact. But in working with agencies we found that this did not prevent us from adding messages that were informed by behavioral science (for example, personalized, friendly approaches to providing summary information) that could help recipients understand and act on the more formal standard communications that they were receiving.

Greg: So you just hinted at the behavioral design process a bit, Matt. Samia, can you talk to us more about any common elements of the intervention you designed and perhaps talk through the features of one of the interventions so we know what they look like in practice?

Samia: Well, all of the interventions we designed involved creating or refining communications like letters, emails, postcards, or even phone call scripts. That's one common way to translate information and attempt to address those barriers. But to answer your question more generally, once we knew the barriers that we suspected were getting in people's way, we identified specific components of existing operations—like the citation cover letter a company receives or communications sent to employees about their benefits—that offered opportunities to address those barriers. We figured out which of these lent themselves to modest changes that could be implemented quickly, taking care to minimize as much as possible the burden on program staff. And then we targeted those procedures in our interventions.

So why don't we take a close look at the workplace safety trial. One of the key barriers we identified there was complexity. For example, after being cited for a safety violation, employers receive citation packages in the mail that are designed to help them remedy the problems and are pretty comprehensive in the information they provide. But, as you might imagine, we found these packages were both long and complex, they required time and attention to fully understand and take action on. As Matt mentioned before, most of the

employers who weren't responding to citations run small firms, like construction contracting companies. They may not have HR personnel or other designated staff whose jobs are to handle these types of issues. So, they were often missing deadlines for resolving their citations and that's not surprising. Our strategy when designing the changes that OSHA piloted for this trial was to use a simple, straightforward cover letter for the packages, so firms could better understand what was in them and what they had to do. The cover letter had simple, streamlined text, and it clearly indicated payment amounts, due dates, and other crucial next steps in a checklist. We also designed some other communication tools to be used in conjunction with this revised cover letter, such as a handout to be shared at inspections and a reminder postcard to follow the citation. When we tested the effect of these changes, we found evidence that they worked: like in offices using the revised cover letter, handout, and reminder postcard, fewer employers were, in fact, referred to debt collection for not resolving their citations. So that was pretty exciting.

Greg: So as you just mentioned, Samia, one of the goals was to rigorously test the interventions you were implementing. So how did you do that and why? And were there any surprises that came up since you were working in the real world rather than [in] a controlled environment like a lab?

Samia: You know, Greg, one of the wonderful things about piloting innovations like these is that more often than not there are opportunities for conducting experiments that yield rigorous and exciting results. What I mean by that is if the population you want to serve is large enough in number and you have ways of implementing your innovation selectively (that is, giving it to some and not to others) then you can use random assignment, and random assignment is essentially a flip of a coin to create two comparable groups. And if you test out your intervention by giving it to one of these groups and not the other and you compare outcomes across the two groups, you can be confident that the differences you observe are caused by the intervention and not other factors. And that information is really valuable both for programs and for the wider field. And the beauty is that this can all be done fairly quickly and easily if you can access data on short-term outcomes and what you are doing is designed to affect those outcomes. And if you have enough time and data, you can readily measure long-term impacts too. So we were able to use this random assignment approach in all three of our trials and what was really wonderful was that for each of them we just leveraged the existing data the programs were collecting as part of their routine operations and so that part was pretty easy.

Matt: And Greg, let me answer your question about surprises. We did have hiccups even after a great deal of planning with staff at multiple levels. So let me give you one example. For the retirement savings trial we had worked very hard to find timing that would be optimal given our partner agency's workloads, HR events that might vie for employees' attention and so on. Even after these efforts, there were still elements of the unexpected. The first phase of the trial was launched during the week before a potential government shutdown in Fall 2015. Which was also the week of the Pope's visit to Washington, DC. And that ends up affecting the daily commuting and work routines of many federal employees in the DC area. That's a lot of distraction from dealing with emails about retirement savings. When we ran a second phase of the trial at a time when other events weren't competing so much for people's attention, the impact we measured was much larger. So one of the lessons of these trials was that given the possibility of unforeseen challenges, we really found that it was useful to build in time for feedback, iteration, and even multiple phases into the timeline of a trial.

Greg: Ok, so earlier you mentioned that you had great, willing partners in the frontline staff and program managers during these trials. Samia, how would you describe the value of this type of work to someone who is more skeptical?

Samia: You know, the beauty of doing rigorous trials is that you can go in with an enthusiastic or a skeptical mindset of what you are testing and you can allow the results to inform whether the interventions worked, and how well. The kinds of trials we ran allow you to learn about what works and what doesn't work, for whom, and in what contexts. I would also say that when they do work, these kinds of investments can have pretty large benefits. So, take the interventions we tested in the trial with OSHA, those results translate to roughly 1,000 fewer cited employers being sent to the national office for additional enforcement each year. That's significant and it means a fair amount of staff time and resources are being saved. Finally, sometimes the benefits of the trial spill over beyond the effort itself. We've talked to our partners, and they've had some very positive things to say about their experiences applying lessons from behavioral insights. Here's what we heard from one of our partners, an administrator at Michigan Works! Southwest, whom we partnered with on the employment program participation trial:

> **[Clip from Eric Stewart, Michigan Works!]** This trial helped us take a giant step back from our typical customer process. The new lens has encouraged us to rethink how we approach all of our work. Finding simple changes like this is becoming kind of a standard of what we are applying widely across the organization. These principles have infused every aspect of what it is we do. This little intervention created a large change and was really a catalyst for our involvement in additional pilot initiatives and has made us even more interested in and excited about partnering in evaluation.

Greg: So for someone who's listening to this podcast who wants to put these insights to use, Samia, what would you recommend as a concrete first step they should take?

Samia: What a great question, and a great chance to apply our research findings by providing clear next steps. You've been paying attention!

So the very first thing you should do is write down one or two very specific problems you care about that you've been trying to solve.

And after that go download the tip sheets we've put together—they're on the DOL Chief Evaluation Office's behavioral insights webpage. Those sheets might give you pointers on what issues may be relevant and on potential solutions that you might think of applying. We also wrote a playbook that explains in more detail some of the key steps in the process of defining a problem, diagnosing what types of behavioral bottlenecks might be in play, and designing and implementing a solution. And we have some recommendations on how to test the effectiveness of your solution.

We'd also recommend finding at least a couple of like-minded collaborators and then working with them to figure out which of the high-priority problems you've identified might be worth thinking more about. But, we advise people to choose a problem where the things you might want to change are in your control, they're not run by another office or third party, or they're not set in legislation or law. If your solution is successful, you can then build on that success. If not, it's always a good idea to talk to people who were directly involved to see what might have gone wrong.

And I would say, if you have the opportunity, I would say really consider seeing whether you can draw on the expertise of behavioral scientists and/or evaluation specialists. It is important to think carefully about how these interventions address the problems you're seeing in your unique context. If you use a strategy informed by behavioral insights that is targeting the wrong bottleneck, you probably won't see much improvement. That is why working with behavioral scientists or evaluators is helpful; those experts can help point you to aspects of design, testing interventions, or understanding further implications for program improvements that you might not have considered by yourself. And many researchers are looking for opportunities to partner with practitioners so that their research can have a more direct impact on people's lives.

Greg: Alright, well thank you so much, Samia and Matt. It's been great hearing more about this study. For those of you listening, head over to the DOL Chief Evaluation Office website to check out more information and resources from the project. The url is http://www.dol.gov/asp/evaluation/BIstudy/. Thanks for listening!