USING BEHAVIORAL INSIGHTS TO STRENGTHEN LABOR PROGRAMS
LESSONS LEARNED

Jaclyn Lefkowitz*
Matthew Darling*
Gregory Chojnacki
Irma Perez-Johnson
Samia Amin
Mikia Manley

*ideas42

MATHEMATICA
Policy Research

ideas42

Introduction

The U.S. Department of Labor's (DOL) Behavioral Interventions (DOL-BI) project adds to a growing body of evidence indicating that relatively small changes in how programs operate can lead to striking improvements in their performance. In three trials that tested applications of behavioral science, the project team found substantial benefits for three DOL programs. We have published detailed technical reports on the design and findings of each trial. (See trial summaries below.) This brief focuses on the lessons learned by the team as it identified opportunities for behavioral trials and implemented each one.

The three trials differed from one another in terms of scope and program focus, demonstrating the broad applicability of behavioral interventions. And although the trials were implemented in labor programs, the findings have potential implications for the use of behavioral interventions in programs funded by other agencies as well.

> **About Behavioral Science**
>
> Behavioral science studies how people make decisions and act in a complex world. It draws on decades of research in the social sciences to provide a more realistic model of how we make decisions and act in real life. Other approaches commonly assume that we consider all available information, weigh the pros and cons of each option, optimize our choices, and then reliably act on them. In practice, however, people often decide and act with imperfect information or fail to act altogether, even when they may want to. Behavioral interventions test whether aligning policies, programs, and products to these human tendencies can result in improved outcomes.

Employee Benefits Security Administration (EBSA) and DOL Human Resources Division (DOL-HR). To prepare for retirement, DOL employees can contribute to the Thrift Savings Plan (TSP), which offers an employer match of up to 5 percent. Despite the TSP's efforts to educate employees about the plan's benefits through its "Take Five for Your Future" campaign, more than 25 percent of eligible DOL employees were not contributing enough as of fall 2015 to receive the full match. In a two-phase trial, we tested the effectiveness of emails designed to address the behavioral factors that might reduce employees' participation in the TSP. The first phase tested a single email, while the second phase tested the original email as well as a variant that framed its message in terms of social norms on savings and added reminder emails sent six weeks after the initial emails. Sending emails incorporating behavioral insights increased the number of trial participants contributing enough to receive the full employer match by 7.5 percentage points.

Occupational Safety and Health Administration (OSHA). Employers who are cited by OSHA for health and safety violations do not always respond as intended. Of 24,000 citations issued with penalties in fiscal year 2013, 22 percent were ultimately referred to the national office for debt collection; 75 percent of these cases were a result of employers failing to respond to the citation. A two-phase trial tested whether OSHA could improve employer responsiveness by (1) distributing a handout on its citation process during workplace inspections, (2) revising its citation cover letter, and (3) following up with reminder postcards and phone calls. After an initial phase that tested the impact of all three components, the trial's second phase tested the effect of the handout, revised cover letter, and postcard, excluding the phone calls, to determine whether an approach that required fewer resources could yield similar benefits. These changes boosted employer responsiveness in both phases; in the second phase, we saw a 5.4 percentage-point increase in the number of employers that responded to citations.

Employment and Training Administration (ETA) with Michigan Works! Southwest (MWS) and W.E. Upjohn Institute. To continue receiving Unemployment Insurance (UI) benefits, claimants must schedule and attend a series of up to three one-on-one Reemployment and Eligibility Assessment (REA) sessions at their local MWS office. Despite evidence of the benefits of the REA program, 43 percent of claimants selected to complete the REA program at Michigan Works! Southwest did not schedule their first appointment before the deadline, and 19 percent did not attend their first scheduled session. In this trial, we partnered with MWS and the W.E. Upjohn Institute to test the effect of a series of emails on participation in REA. The first email introduced claimants to REA services using a friendly, positive tone and provided clear instructions for scheduling and attending the first REA appointment, and follow-up emails provided timely reminders about additional sessions. The emails led to a 14 percentage-point increase in the share of claimants who completed the REA program.

Snapshot of Lessons Learned

Lessons Learned When Selecting Behavioral Projects

- Conduct broad outreach to identify candidate problems

- Develop selection criteria based on goals, resources, and priorities

- Apply these criteria to select promising candidates

Lessons Learned When Implementing Behavioral Projects

- Interventions can be implemented on both local and national scales

- Dedicating time to understanding the problem and associated behavioral barriers strengthens designs

- Exploring administrative data can shed light on the problem and confirm or challenge assumptions

- Ongoing attention to operational details is crucial when designing behavioral interventions

- Allowing time for iteration and multiple trial phases can improve outcomes

- Engaging program staff of differing positions and levels during design and implementation is critical

- Adopting the user's perspective facilitates each step of the process

- Engaging practitioners allows them to rapidly inform programs and apply findings

Lessons Learned When Selecting Behavioral Projects

A successful intervention requires an appropriate behavioral problem to solve as well as partners who are willing and able to collaborate with the intervention design team. This section describes how we worked with DOL's Chief Evaluation Office (CEO) to recruit partner agencies and identify problems appropriate for behavioral interventions.

Conduct broad outreach to identify candidate problems

As the DOL-BI project was getting underway, we spoke with representatives from numerous DOL agencies to explore promising opportunities to apply insights from behavioral science to a variety of program areas. We knew that a broad outreach effort would be necessary because many programs that appear to be promising in preliminary discussions might not advance to the implementation and testing stages. We reached out to over a dozen potential programs and pursued six that seemed promising. However, after gathering detailed information about each program, we ultimately identified three that were suitable for a behavioral trial.

Develop selection criteria based on goals, resources, and priorities

To systematically select problems that would be the most appropriate for a behavioral trial, we gathered information from program administrators on the five criteria below.

> **Problems Identified by DOL Agencies**
>
> Although the agencies we contacted to discuss trial opportunities differed in the programs and services they offered, they often wanted to address similar problems.
>
> Two in particular stand out: (1) the low use of programs by individuals despite evidence showing the benefits of the program, such as access to an employer match on retirement savings or job-search assistance to speed a return to employment, or (2) the failure of individuals to follow through on key next steps after signing up for a program, such as scheduling and attending appointments or filling out relevant paperwork.

High-priority problems
- What urgent, important program performance areas and outcomes are DOL agencies looking to improve?

Problems with behavioral dimensions
- Which problems show evidence of common behavioral challenges (e.g. a disconnect between what an individual intends to do, and what they actually do)?

Problems with scale and relevance
- Which problems affect large populations of interest to DOL or involve key programs?

Program processes that can be adjusted
- Are there existing program elements and procedures that could be modified relatively easily and quickly?

Feasibility of evaluation
- Are program performance and outcome data readily available and accessible to test the impacts of behavioral interventions?

The DOL-BI team used the first three criteria to focus on problems that are urgent, affect many people, and show signs that behavioral insights may point to novel solutions. These criteria are important because we encountered several problems in our initial outreach to agencies that were high on their lists of priorities but did not appear to involve behavioral factors. In contrast, we learned of problems that might be solved by applying insights from behavioral science but that affected fewer people than the number in the trials we ultimately conducted.

Before investing resources in diagnosing which behavioral barriers might be contributing to a problem, we wanted to ensure that we would be able to translate that diagnosis into program changes that could be implemented and tested quickly and inexpensively, allowing multiple rounds of testing where appropriate. The DOL-BI team used the fourth and fifth criteria to identify such opportunities.

We used these criteria to shape our initial discussions with prospective partners. Specifically, we asked agencies a variety of questions about their programs, their target populations, the accessibility of data, the problems they were experiencing, and their goals for improvement. Ultimately, the trials we proceeded with had problems that met all five of these criteria.

Some of these criteria are specific to the goals and constraints of the DOL-BI project and may not be applicable to additional work in which behavioral interventions are applied to DOL programs. For example, the fixed budget and timeline of the DOL-BI project made it necessary to work with agency partners that could change their process within a few weeks or months, whereas other projects may be able to work with longer time horizons.

Additionally, the DOL-BI project sought to rigorously test the effectiveness of our interventions through randomized controlled trials, which we conducted rapidly, to determine how well the intervention was working. Others applying behavioral insights may choose not to limit opportunities to those that will support such trials, however, for a variety of reasons. For example, interventions could be tested using less rigorous methods (such as measuring behavior before and after the intervention, without a control group). In other cases, where direct measurement is impossible, behavioral interventions that have been demonstrated to work across multiple contexts can be implemented with little risk. In general, the higher the cost of implementing a change or the more resource-intensive it is, the more reason there is to invest in a rigorous design to test its effect.

Apply these criteria to select promising candidates

Although all the problems we looked at were important to DOL agencies, there were some that we did not pursue for several key reasons, discussed further below. These reasons also apply to other partnerships that did not advance past an initial discussion.

Occasionally, we found that some problems had elements that could be considered behavioral, but they lacked existing procedures that could easily be modified to incorporate insights from behavioral science. For example, service take-up or program participation may be low because of behavioral barriers, but if there are no existing mechanisms for follow-up communication with participants, implementing a trial would require a larger degree of technical assistance that was beyond the scope of this project.

In several other cases, we found that programs were not appropriate for a randomized controlled trial because the population of individuals eligible to participate in a trial was too small to support a rigorous test of the behavioral solution's impact. Without a large enough group of individuals to test the impact of the intervention on, we wouldn't be able to tell with certainty whether it was working, how well, and for whom. In addition, several potentially promising opportunities did not advance beyond preliminary discussions because it was not clear if they had administrative data that could be used to assess the impact of a behavioral intervention. Although DOL agencies collect performance and financial data across a wide variety of programs, some agencies were not collecting data on the specific outcome that our intervention would be intended to change.

Although not every discussion we had with DOL agencies led to a behavioral trial, we sought to make the engagement process informative and useful for each agency. If we determined that a given problem was not appropriate for a behavioral trial, we provided the agency with a summary of any findings from our research on the problem and feasibility of a trial. If we had advanced to the initial intervention design, we also offered preliminary ideas for program modification that the agencies could further explore on their own.

When considering whether to run a behavioral trial with a program, it is important to review the five criteria with agency leaders. Having consensus on which problems are the highest priority establishes the foundation necessary to work successfully through project completion. Leaders and members of the project team will be more willing to invest time in diagnosing behavioral barriers and in designing potential solutions if everyone agrees that the problem is high enough on the agency's list of priorities.

Lessons Learned When Implementing Behavioral Projects

A great advantage of behavioral interventions is that they often improve program effectiveness at comparably little cost. However, it is still important to invest time and resources in identifying the behavioral bottlenecks affecting people's choices and actions, carefully designing the intervention, and supporting implementation. Below, we describe several overarching lessons we learned while implementing trials in the DOL-BI project.

Interventions can be implemented on both local and national scales

Our trials not only tackled a range of problems; they also varied in the number of people involved and in their geographic scope. We implemented the MWS trial on a local scale, involving UI claimants from four counties in southwest Michigan. In contrast, the OSHA trial was implemented in 8 of 10 agency regions nationwide. Because some OSHA regions vary in their policies and procedures, we coordinated with regional administrators to ensure that our intervention's design would be feasible for each region. Specifically, we spoke at length with the administrators to deepen our understanding of their operations, to identify key regional differences, and to adapt the intervention to regional conditions when necessary. With the right partners at the table, it still was possible to identify and test behaviorally informed changes within a relatively short timeframe.

Dedicating time to understanding the problem and associated behavioral barriers strengthens designs

The behavioral design process must include three key steps: (1) develop a thorough understanding of existing program operations, (2) determine where issues tend to occur, and (3) identify the behavioral factors that may contribute to the problem. Our detailed understanding of the context for each of our behavioral trials allowed us to tailor designs to the situation at hand.

For example, at the beginning of the OSHA trial, we held several discussions with national, regional, and local office managers. We also analyzed existing administrative data to gain a better understanding of the citation process, to see where employers dropped off in the process, and to learn more about these employers. Through this research, we found that many employers never responded to OSHA to begin with (as opposed to responding but not following through to resolve their citations). We also learned that the employers who tended not to respond were often small firms and were often run by Spanish speakers, according to OSHA regional and local staff who work regularly with employers to resolve citations. We leveraged these insights to (1) design an intervention that targeted employers in the early stages of the citation process, (2) present the response options clearly for busy citation recipients, and (3) provide Spanish translations of handouts, letters, and postcards used in the trial.

Exploring administrative data can shed additional light on the problem and confirm or challenge assumptions

DOL's internal programs were in a good position to benefit from behavioral trials given the abundant administrative data the DOL collects. The agency's data systems played an important role in all stages of the DOL-BI project. We used administrative data to:

- Learn more about our focal programs and their procedures and target populations

- Assess the scale of the problems we were trying to address (for example, how many program participants are affected by a problem)

- Uncover behavioral factors influencing the problem

- Challenge our own hypotheses and assumptions about potential behavioral factors

- Refine our diagnosis before brainstorming solutions

- Evaluate the effectiveness of our interventions using a variety of short- and long-term outcome measures that were already being collected in administrative data

Ongoing attention to operational details is crucial when designing behavioral interventions

When deciding how to address behavioral barriers, it is critical to be aware of contextual factors related to program operations that limit the solutions that are feasible and how they can be implemented. There may be limitations related to data or technology capabilities as well as staff capacity. For example, when designing solutions, we worked closely with our DOL agency partners to think through the details of the proposed changes to operations and to determine how they could be implemented. We discussed facets such as who would be sending emails and postcards and tracking responses, how they would do this, and whether the effort was reasonable given their usual responsibilities.

Even with a careful review of the details, it is difficult to fully anticipate how a solution will function in practice until it is implemented. For example, we conducted a short pilot of the changes to OSHA's citation process and shortened the script for reminder phone calls to employers based on feedback from the office in which the pilot was conducted. Even so, the frontline staff who made these calls during the full-scale trial still found them to be too time-consuming. In a second phase of the OSHA trial, we dropped the phone calls and examined the benefits of the remaining changes to the citation process. In the MWS trial, our initial recruitment numbers were lower than expected. This threatened our ability to rigorously assess our intervention's effectiveness. Our research partners, the W.E. Upjohn Institute and MWS, worked with us to include more counties in the trial in order to ensure that we reached enough people. This change also made it necessary to modify our initial email messages to include contact information for the REA program locations we added to the trial.

Allowing time for iteration and multiple trial phases can improve outcomes

For each trial, we used an iterative approach to design the intervention components. We shared prototypes of our proposed behavioral solutions with the agencies and our other partners for review. Staff commented on whether the designs corresponded to their strategic goals and how the trials would function within existing operations. We then revised the design several times, soliciting feedback at each stage. We also sought input from the program's target population whenever possible. Incorporating these diverse sources of input at key points during the development of the interventions helped us refine our interventions and improve their scalability.

Feedback is not only useful during the early stages of design; it can also inform refinements after implementation has begun. As noted, feedback from OSHA frontline staff about the time and effort needed to complete follow-up phone calls guided our decision to conduct a second phase of the intervention without that component.

Similarly, in the EBSA trial, we refined the email designs several times as we received feedback from EBSA, DOL-HR, and the Financial Literacy and Education Commission on the specific language that might be most effective for DOL employees. After the first phase of the trial, we ran a second phase to see if we could further improve outcomes by modifying the message and pairing it with a reminder email after six weeks.

Considering the context for Phase 1 of the EBSA trial—including a possible government shutdown, the Pope's visit to Washington, DC, and religious holidays—we also suspected that its timing may not have been conducive to helping employees focus on their retirement. This provided another rationale for the trial's second phase, in which we found even greater impacts than we had measured in the first. Testing the intervention at different times allowed us to see results that we would not have uncovered had we tested the emails only once.

Engaging program staff of differing positions and levels during design and implementation is critical

Our trials benefited from having champions in high-level positions who were enthusiastic about testing behaviorally informed changes, but we also involved staff from a wide array of departments and positions. We consulted with local and national departments as well as data specialists and frontline program staff. Incorporating a variety of perspectives led to concrete improvements in the solutions we tested. For example, staff in the national office may have the best understanding of current agency-wide priorities, whereas regional or local staff have a better understanding of which aspects of program operations can be changed and the level of effort required. Frontline program staff shared their qualitative understanding of potential behavioral bottlenecks, which could be confirmed (or refined) through data obtained from information technology (IT) staff. In the OSHA trial, the IT staff also provided invaluable guidance on the capabilities and constraints of OSHA information systems. This information was crucial when we were designing the revised citation cover letter for that trial, since the letter needed to be integrated into the system so that local staff could automatically generate the new letter as they did with the previous one. Maintaining regular touchpoints and feedback loops with a wide variety of staff during the full length of the trial provides several opportunities for input and iteration.

Adopting the user's perspective facilitates each step of the process

Whether the goal is to identify behavioral barriers or create solutions, adopting the target user's perspective is key. Whenever possible, we observed or worked through the steps in the existing program or process ourselves, as if we were the intended users. This firsthand experience strengthened our understanding of the context and behavioral factors that could be inhibiting desired outcomes. For example, in the EBSA trial, we worked with a DOL staff member to walk through the steps for changing one's contributions to the TSP online. In addition, we supplemented our observations of this process by interviewing target users and program staff. Before launching the trial, we sent our intervention email to a small sample of DOL employees and then asked them about their reactions when they received and read it. This helped us to confirm that the email was clear, easy to understand, and likely to be read.

It is also important to think through any intervention from the point of view of the staff who would be implementing it. Adopting this perspective allows the design team to communicate effectively with frontline staff about the trial's procedures and goals, which may make them more receptive to changing current operations. Communicating clearly with frontline staff members can also empower them to provide insight

into why the intervention is working or, more important, why it may not be working. For example, in the OSHA trial, local office staff informed the design team that the reminder postcard might pose a problem because it was intended to incorporate the citation delivery date, but some offices do not receive confirmation of citation delivery until weeks or even months later. This insight allowed us to provide uniform direction on how to use the postcard when local offices need to send it before they know the citation delivery date.

It might not always be possible to directly observe processes, to interview participants, or to test designs with users and implementing staff, but this approach can deepen our insight into a problem's context and help us to generate more effective solutions.

Engaging practitioners allows them to rapidly inform programs and apply findings

Once our partner, MWS, worked with us to design the intervention and then learned of our interim findings, they began applying key aspects of the behavioral design process, such as adopting a user-centered perspective, to other programs that they run. Management at MWS noted that the trial had motivated their staff to reassess their current ways of engaging with customers and identify opportunities for implementing small changes that might improve clients' experience of their programs. They have shared findings from the trial with colleagues at other workforce agencies and reported to the DOL-BI team that their experience with this behavioral trial has whetted their appetite for future research collaborations. These reactions have the potential to speed the adoption of research-based best practices and foster further opportunities to test novel applications of behavioral science.

Looking Ahead

The DOL-BI project provides clear evidence that modest modifications grounded in behavioral science can enhance the effectiveness of a wide variety of DOL programs and help agencies to better carry out their missions. Although designing and testing behavioral interventions may be new for many organizations, they can often take small steps to get started. Identifying persistent problems and building consensus among staff about the importance of addressing these problems is a critical starting point that can raise enthusiasm among employees about behavioral solutions. The problems need not be the most difficult ones. In fact, it may be best to begin with a problem that is less complex and to focus on modifying processes over which you have full control (versus needing to coordinate with other teams or third parties). Starting small and setting up a solid foundation can help build the momentum that drives the later development and implementation stages. Even if an organization does not have the capacity to complete the testing stage (for example, by designing a randomized controlled trial and collecting and analyzing the data), much can be learned early in the process. And there may be less costly ways to gather basic information on whether a change had an effect on important outcomes measured in existing administrative data (for example, by comparing participants' outcomes before and after a change is made).

Finally, although our trials focused on solving problems in existing programs, behavioral science can be applied in broader ways. Instead of using it to modify how existing programs function, the behavioral design approach can be used at the outset of a program or during policy development. By continuing to conduct rigorous evaluations and applying the findings at scale, we can create social programs and policies that are better aligned with human behavior and yield improved results for many people we serve.

To learn more

Our findings from implementing these trials contribute to a growing body of work on how insights from behavioral science can improve the performance and outcomes of labor and other social programs. For others who are interested in further exploring our approach to these trials, the DOL-BI team has developed the following resources to support your work:

1. **The** *Practitioner's Playbook*, an easy-to-use reference guide for program managers interested in applying the behavioral diagnosis, intervention design, and testing process.

2. **Two** *Promising Practices Handouts*, which provide concise tips on solutions informed by behavioral science that can help improve forms, informational materials, communications, and other aspects of program operations regardless of whether you implement a behavioral trial.

These resources are available for download at DOL CEO's behavioral insights web page:
https://www.dol.gov/asp/evaluation/BIStudy/

www.mathematica-mpr.com

Improving public well-being by conducting high quality, objective research and data collection

PRINCETON, NJ ■ **ANN ARBOR, MI** ■ **CAMBRIDGE, MA** ■ **CHICAGO, IL** ■ **OAKLAND, CA** ■ **WASHINGTON, DC**

ideas 42