Insights from the Health Profession Opportunity Grant Program’s Three-Armed, Multi-Site Experiment for Policy Learning and Evaluation Practice

by
Laura R. Peck
Hilary Forster
Nicole Constance

for
Special Issue of Evaluation Review on Multi-arm Randomized Control Trials in Evaluation and Policy Analysis

May 15, 2018

Abstract
Administered by the Office of Family Assistance (OFA) in the U.S. Department of Health and Human Services’ Administration for Children and Families (ACF), the Health Profession Opportunity Grants (HPOG) Program provides education and training to Temporary Assistance for Needy Families (TANF) recipients and other low-income individuals for occupations in the healthcare field. The impact evaluation for the first cohort of grantees (HPOG 1.0) leveraged the program’s implementation across many locations to add a second treatment arm in some places as a way to examine whether any of three selected enhancements contributed to the program’s overall impact. This article tells the story of the evaluation and draws lessons from that experience—identifying and discussing eight criteria—for future implementation of multi-armed experiments in a multi-site evaluation.

Acknowledgements
The authors gratefully acknowledge input on earlier versions of the manuscript from Mark Fucello and Naomi Goldstein of the ACF Office of Planning, Research, and Evaluation; from Kim Stupica-Dobbs and Susan Golonka of the ACF Office of Family Assistance (OFA); and from Alan Werner and Larry Buron of Abt Associates. We are also grateful to Stanley Koutstaal and current and former staff in OFA whose partnership was instrumental in implementing the HPOG enhancements at the center of the three-armed experiments. The paper was funded, in part, under the Career Pathways Intermediate Outcomes study (Contract No. HHSP23320095624WC/HHSP23337019T).
Insights from the Health Profession Opportunity Grant Program’s Three-Armed, Multi-Site Experiment for Policy Learning and Evaluation Practice

Multi-site evaluations provide opportunities to exploit both natural and planned variation to learn what about a program associates with its impacts. Prior analyses have considered this cross-site variation qualitatively (e.g., Riccio, Friedlander, & Freedman, 1994) and non-experimentally (e.g., Bloom, Hill, & Riccio, 2003; Greenberg, Meyer, & Wiseman, 1994). However, a recent, major national evaluation has done so using an innovative experimental evaluation design, which added a second treatment arm in several locations as a way to identify key aspects of the program’s impact.

Initially authorized in 2010 and administered by the Office of Family Assistance (OFA) in the Administration for Children and Families (ACF), the Health Profession Opportunity Grants (HPOG) Program provides education and training to Temporary Assistance for Needy Families (TANF) recipients and other low-income individuals for occupations in the healthcare field. It does so both to improve the employment prospects of these individuals and to meet local demand for workers in the healthcare field. The first round of HPOG grants was awarded in 2010 (HPOG 1.0), and a second round was awarded in 2015 (HPOG 2.0).

HPOG was authorized as a demonstration program with a mandated evaluation to assess the success of the initiative. ACF’s Office of Planning, Research, and Evaluation (OPRE) leads a multi-pronged research and evaluation portfolio to answer questions about the implementation, outcomes, and impact of HPOG. The portfolio encompasses research and evaluation projects that aim to identify what types of program approaches work well, for whom, and in what circumstances. The HPOG 1.0 Impact Study is a central component of this portfolio. The study includes 42 distinct local HPOG programs operated by 23 of the 32 HPOG 1.0 grantees. The
study enrolled a sample of 13,717 participants in treatment and control groups whose outcomes were compared to estimate the impact of the collection of programs.

The federal evaluation gave HPOG 1.0 programs the opportunity to receive additional funding to implement one of three selected program components—peer support, emergency assistance, or non-cash incentives—on the condition that they would ration access to the component in support of providing experimental evidence on the component’s contribution to overall program impacts. In total, 19 of the 42 programs ran one of these three-armed experiments in which one of the selected program components was added to the standard HPOG program. The study’s first impact report was released in 2018 (Peck et al., 2018).

From a design perspective, what is distinctive about the HPOG 1.0 Impact Study is that it randomized access to selected program components to be evaluated on top of a standard program. This is what has been labeled an “enhanced treatment” design and operates in contrast to what has been labeled a “competing treatments” design (Bell & Peck, 2016; Peck, 2015). An enhanced treatment design compares a standard version of a program to a scaled-up (or scaled-back) version of that same treatment to understand the contribution of the chosen enhancement (or reduction) to overall program impacts. A completing treatments design compares fully alternative treatment models as a whole (cross-reference other article/s in the special ER issue?). Both of these designs can function with or without a control group, and both are multi-armed designs that help in understanding “what works.” This paper provides insights on the implementation of an enhanced treatment design in practice.

This paper considers various practical and technical design and implementation issues associated with this kind of experimental test, one in which a subset of many sites, all running multi-faceted programs, take on a second treatment arm. We first provide additional context for
the HPOG 1.0 programs and describe how the three-armed experiment evolved and took place. We then generalize from that experience to identify and explain several criteria to use for identifying whether conditions are right for operating a multi-armed experiment in practice.

**Context: The HPOG 1.0 Programs**

As noted above, HPOG was authorized in 2010 to meet the dual goals of increasing the supply of healthcare workers while creating vocational opportunities for low-income, low-skilled adults. The HPOG Program uses the career pathways framework of postsecondary occupational training and education, which organizes instruction as a series of manageable and well-articulated steps to facilitate participants’ movement to successively higher levels of credential-bearing training and employment.\(^1\) Career pathways programs include four key strategies: comprehensive assessments of skills and needs; promising approaches to basic skills and occupational training; academic and non-academic support services; and approaches for connecting students to employment (Fein, 2012). Using this framework, HPOG grantees develop pathways that meet employer demand and other needs in their local communities. HPOG grantees vary in regards to the healthcare occupational training offered but the most common training programs include those for nurse aides, home health aides, licensed and vocational nurses, registered nurses, medical assistants, pharmacy technicians, and phlebotomists. Many TANF recipients and other low-income adults lack the basic academic skills needed to enroll in occupation-focused training. In response, HPOG programs are encouraged to provide basic skills training (either prior to or integrated with occupational training) to ensure participants have the

---

\(^1\) As noted above, HPOG 2.0 grants were awarded in 2015 and are currently operating, whereas the HPOG 1.0 grants have ended. For this reason, we tend to discuss the HPOG Program in the present tense and the first-round evaluation activities in the past tense.
foundational academic skills necessary to enter and complete the program and obtain long-term employment.

HPOG’s authorizing legislation requires that programs “provide eligible individuals with financial aid, child care, case management, and other supportive services.” Accordingly, HPOG programs provide comprehensive supportive services to assist participants’ program retention and completion, and attainment of employment in their field. Almost all HPOG grantees cover all or part of participants’ tuition costs. In addition, many grantees facilitate access to other sources of financial assistance for participants including Pell Grants and Workforce Investment Act Individual Training Accounts. Virtually all HPOG programs employ case managers who work intensively with program participants, providing counseling, connections to supportive services, financial advice, and other guidance. In addition, most programs provide a wide variety of employment assistance services including soft skills training, resume building, interview preparation, and career readiness classes. Lastly, most HPOG programs offer services to address participants’ personal and family material needs that could interfere with stable training participation and completion, such as transportation vouchers or childcare assistance.

Implementation: The HPOG 1.0 Story

In 2011, OPRE competitively selected Abt Associates as the independent evaluator for the HPOG 1.0 Impact Study. Collectively, OPRE and Abt decided that rather than choosing a

---

small number of large sites for site-specific impact evaluations, we would engage nearly all of
the grantees in a national impact study.³ The main rationale for this “all in” evaluation strategy
was that it would permit additional learning from the substantial cross-site variation that existed
across HPOG programs without the need for substantial additional funding. Fortunately, OPRE
and OFA included the following language in the Funding Opportunity Announcement (FOA) for
the 2010 HPOG grants regarding expectations for evaluation:

“ACF plans a multi-pronged strategy to assess the success of approved
demonstration projects. This strategy will include federally sponsored evaluations
of a selected subgroup of grantees and grantee-level documentation of activities,
outputs and outcomes based on uniform data collection guidelines. As a condition
of acceptance of a grant award, all funded grantees are required to agree to
participate fully in a federal evaluation, if selected, and to follow all evaluation
protocols established by ACF or its designee contractor. Fully participating in a
federal evaluation may include supporting and complying with special data
collection requirements, providing additional administrative data on program
participation or service receipt, facilitating on-site meetings and observations,
including interviews with program and partner managers and staff as well as
participants, among other activities. In addition, a federal evaluation may include
a rigorous evaluation design using random assignment of individuals, groups or
organizations to new demonstration services or existing services. Selected
grantees may be required to facilitate the random assignment process. Grantees
selected to be in a federal evaluation will receive technical assistance and other
support in meeting evaluation requirements.”⁴

This FOA language simplified the evaluation planning process by reducing the costs associated
with recruiting grantees to be part of the impact evaluation. Instead, grantees were expected to
participate in an ACF-supported evaluation as a condition of accepting their grant award. As a
result, the project did not have to incur costs associated with site recruitment into the evaluation,
which permitted the evaluation to take on a larger number of sites. Further, under the “all in” site

³ Of the 32 HPOG 1.0 grantees, 20 participated in the HPOG 1.0 Impact Study. The remaining grantees were not
required to participate because they were already engaged in some evaluation research and ACF wanted to avoid
undue burden on grantees and participants. A complete description of OPRE’s HPOG Evaluation Portfolio is
available at https://www.acf.hhs.gov/opre/research/project/evaluation-portfolio-for-the-health-profession-
opportunity-grants-hpog.
⁴ Funding Opportunity Announcement: Health Profession Opportunity Grants to Serve TANF Recipients and Other
Low-Income Individuals, HHS-2010-ACF-OFA-FX-0126.
strategy, a large portion of data for the study’s implementation analysis would come from comprehensive grantee, staff, and stakeholder surveys being conducted by a coordinating component of the HPOG research and evaluation portfolio, rather than by intensive implementation field work needed to understand site-specific impacts. Consequently, there was less of an added cost associated with the implementation evaluation of a greater number of sites, given a reduced need for multiple, costly site visits.

From the outset, both OPRE and OFA were committed to the evaluation, and that commitment evolved as the nature of the evaluation evolved. OFA and OPRE operate in close partnership to administer the HPOG Program and to conduct the evaluation efforts, respectively. Both offices are invested in maximizing the learning from HPOG. As such, both OPRE and OFA were keenly interested in the HPOG 1.0 Impact Study being able to report on specific aspects of program effectiveness, given the potential for such evidence to inform practice, policy, and the research field. The inclusion of a larger number of programs in the study yielded the opportunity to use a three-armed design to act on this interest. In response, the evaluation team decided to incentivize some of the programs to add one of three selected components to their program offerings and ration access to that addition in order to generate experimental evidence on the effectiveness of the selected program components. With this idea in place, the team needed to

---

5. The HPOG National Implementation Evaluation (NIE), which included all 27 non-tribal HPOG 1.0 grantees, conducted these surveys as part of its implementation, outcome, and systems change analyses. More information about the NIE is available at [https://www.acf.hhs.gov/opre/research/project/national-implementation-evaluation-of-the-health-profession-opportunity](https://www.acf.hhs.gov/opre/research/project/national-implementation-evaluation-of-the-health-profession-opportunity).

6. It is possible that an impact evaluation with only a small number of sites could have included a third arm. However, in HPOG’s case, including a larger number of sites made the possibility to include a third arm much more feasible for two reasons—power and variation: having many more sites involved meant the possibility of sufficient power to operate the three-armed experiments; and having many sites also meant much more programmatic variation on which to capitalize.
identify *which* program components would be good candidates for this test, and *which* grantees would be good candidates as test sites.

*What to Evaluate.* The team started by identifying a long list of possible program components. This list included 12 specific potential “impact drivers,” which had at least suggestive evidence to support their value to career pathways programs. With this information, the team prioritized the program component candidates in order to narrow the list to be part of the three-armed test. Identifying higher-priority candidates involved considering additional criteria including grantee demand to test the component, the feasibility of randomizing the component, whether it could be implemented consistently across programs, and whether it could be implemented in a relatively short time frame within an ongoing program. The feasibility criterion involved ensuring that the standard and enhanced treatment groups could be kept distinct to minimize the likelihood of contamination, either through formal cross-overs or through spillover effects. Thus, in the case of HPOG, identifying what to evaluate involved considering how adding the enhancements into the existing program would work.

Subsequently, as part of each grantee’s continuation application process in summer of 2012, the evaluation team invited input from grantees regarding interest in adding particular components from this narrowed list to their programs. Through the process of reviewing the

---

7 The complete list of possible program components included proactive case management, strong personalized support, staff/participant ratio; social, peer supports (“learning communities”); accelerated courses; well-articulated and short-term learning modules; flexible scheduling; contextualized developmental education; strong interagency partnerships; active employer involvement in multiple program aspects; employment/work experience opportunities during training (either paid or unpaid); instructional supports (e.g., tutoring); support services (all grantees provide these, but there is variation in in types and/or accessibility); and financial supports (again, all grantees provide these, but there is variation in the types, accessibility and/or scale).

8 The HPOG Impact Study began in the second year of the five-year HPOG 1.0 grant period, and the three-armed experiment did not launch until the third year of the grant period. Therefore, grantee demand and quick implementation were two criteria used in HPOG but that are not necessarily relevant to a broader application of this design in practice.

9 To receive funding for Years 2 through 5 of the five-year grant period, grantees were required to submit non-competitive continuation applications. Ongoing awards were based upon the satisfactory progress of the grantee and/or a determination that continued funding would be in the interest of the government.
research literature and practical lessons from programs, soliciting feedback from the grantees, and considering evaluation feasibility, the team arrived on a short list of program components that it would offer as randomized-to enhancements: emergency assistance, non-cash incentives, and facilitated peer support.

Four steps followed the final selection of program components: (1) articulating in detail what each component comprised, so that programs selected to participate in the three-armed experiment would have a clear sense of what to implement; (2) identifying the programs eligible to implement a given enhancement; (3) recruiting from among eligible programs; and (4) supporting programs in their implementation of their chosen enhancement.

Before the evaluation team committed to these four steps they undertook a series of power analyses to determine the minimum sample size at which the three-armed experiment would be viable. This analysis took place on a rolling basis as eligible programs were identified and recruited. The team considered the number of individuals that each program would add to the pooled sample for the three arms and identified whether the end result would provide value. Partway through the recruitment, when it became clear that anticipated effect sizes were plausible and policy-relevant, ACF and the evaluation team fully committed to moving forward with these multi-armed tests.

Where to Evaluate. Table 1 lists the HPOG 1.0 programs, highlighting which of them were eligible to add a given enhancement. A program was “eligible” if it did not already have the given program component in place. The cells in Table 1 that are shaded grey and labelled “offered” indicate the programs that were not eligible for testing that component because they already offered the given component as part of their standard program. The remaining programs were eligible. The evaluation team reached out to the eligible programs and invited applications
for participating in the three-armed experiment\textsuperscript{10}. The results of that application and selection process is reflected in the black-shaded cells labeled “tested”: those are the programs that added the given component to their programs as part of the three-armed experimental evaluation.

Using additional funding from ACF, the selected programs added the given program components, following guidelines that the evaluation team had prepared regarding what each component comprised and the programs’ plans as proposed in their applications. The team provided technical assistance and support for implementation to ensure that there was maximum implementation contrast and consistency. This technical assistance drew on the set of guidelines for each enhancement and helped programs implement the component in line with those guidelines. Although programs adhered to the general guidelines, factors outside of the guidelines still varied; therefore, implementation was in practice standardized but with slight variations across the programs. The evaluation team also added fields to the HPOG 1.0 management information system in order to ensure complete tracking of treatment group members’ participation in each of the program enhancements. The enhancements had been implemented for at least 15 months by the time the evaluation team analyzed the outcome data.

\textsuperscript{10} The “application” was somewhat of a formality; programs needed to demonstrate their commitment to operating the three-armed experiment and adherence to the guidelines for each component prepared by the evaluation team. However, the evaluation team worked with all applicants to ensure that they developed strong applications that ACF and the team would approve to move forward. This was, essentially, the beginning of a technical assistance and monitoring process that focused on implementing the components in the field.
Table 1.
Location of Program Components, by Program and Component Type

<table>
<thead>
<tr>
<th>State</th>
<th>Grantee – Program Operator</th>
<th>Emergency Assistance</th>
<th>Non-Cash Incentives</th>
<th>Facilitated Peer Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>AZ</td>
<td>Pima County Community College District</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>San Diego Workforce Partnership - MAAC South</td>
<td>OFFERED</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>San Diego Workforce Partnership - Metro CTS</td>
<td>OFFERED</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>San Diego Workforce Partnership - North County Lifeline</td>
<td>OFFERED</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td>The WorkPlace</td>
<td></td>
<td></td>
<td>TESTED</td>
</tr>
<tr>
<td>FL</td>
<td>Pensacola State College</td>
<td></td>
<td></td>
<td>OFFERED</td>
</tr>
<tr>
<td>IL</td>
<td>Will County WIB - Central States SER</td>
<td>OFFERED</td>
<td></td>
<td>OFFERED</td>
</tr>
<tr>
<td>IL</td>
<td>Will County WIB - College of Lake</td>
<td>OFFERED</td>
<td></td>
<td>OFFERED</td>
</tr>
<tr>
<td>IL</td>
<td>Will County WIB - Instituto del Progreso Latino</td>
<td>OFFERED</td>
<td></td>
<td>OFFERED</td>
</tr>
<tr>
<td>IL</td>
<td>Will County WIB - Jewish Vocational Services</td>
<td>OFFERED</td>
<td></td>
<td>OFFERED</td>
</tr>
<tr>
<td>IL</td>
<td>Will County WIB - Joliet Junior College</td>
<td>OFFERED</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KS</td>
<td>Dept of Commerce - Heartland Works, Inc.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KS</td>
<td>Dept of Commerce - Southeast KANSASWORKS, Inc.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KS</td>
<td>Dept of Commerce - Workforce Alliance of South Central Kansas</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KS</td>
<td>Dept of Commerce - Workforce Partnership</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KS</td>
<td>Dept of Commerce – WorkforceOne</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KY</td>
<td>Gateway Community and Technical College</td>
<td>OFFERED</td>
<td>TESTED</td>
<td></td>
</tr>
<tr>
<td>LA</td>
<td>WIB SDA-83 Inc.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MO</td>
<td>Full Employment Council</td>
<td>TESTED</td>
<td>OFFERED</td>
<td>OFFERED</td>
</tr>
<tr>
<td>NE</td>
<td>Central Community College</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NH</td>
<td>New Hampshire Office of Minority Health</td>
<td>OFFERED</td>
<td>OFFERED</td>
<td></td>
</tr>
<tr>
<td>NJ</td>
<td>Bergen Community College - Bergen Community College</td>
<td>TESTED</td>
<td></td>
<td>TESTED</td>
</tr>
<tr>
<td>NJ</td>
<td>Bergen Community College - Brookdale Community College</td>
<td>TESTED</td>
<td></td>
<td>TESTED</td>
</tr>
<tr>
<td>NJ</td>
<td>Bergen Community College - Community College of Morris</td>
<td>TESTED</td>
<td></td>
<td>TESTED</td>
</tr>
<tr>
<td>NJ</td>
<td>Bergen Community College - Essex County College</td>
<td>TESTED</td>
<td></td>
<td>TESTED</td>
</tr>
<tr>
<td>NJ</td>
<td>Bergen Community College - Hudson County Community College</td>
<td>TESTED</td>
<td></td>
<td>TESTED</td>
</tr>
<tr>
<td>NJ</td>
<td>Bergen Community College - Middlesex County College</td>
<td>TESTED</td>
<td></td>
<td>TESTED</td>
</tr>
<tr>
<td>NJ</td>
<td>Bergen Community College - Passaic County Community College</td>
<td>TESTED</td>
<td></td>
<td>TESTED</td>
</tr>
<tr>
<td>NJ</td>
<td>Bergen Community College - Sussex County Community College</td>
<td>TESTED</td>
<td></td>
<td>TESTED</td>
</tr>
<tr>
<td>NJ</td>
<td>Bergen Community College - Union County College</td>
<td>TESTED</td>
<td></td>
<td>TESTED</td>
</tr>
<tr>
<td>NJ</td>
<td>Bergen Community College - Warren County Community College</td>
<td>TESTED</td>
<td></td>
<td>TESTED</td>
</tr>
<tr>
<td>NY</td>
<td>Research Foundation of CUNY-Hostos Community College</td>
<td>TESTED</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NY</td>
<td>Buffalo and Erie County WDC</td>
<td>OFFERED</td>
<td>TESTED</td>
<td></td>
</tr>
<tr>
<td>NY</td>
<td>Schenectady County Community College</td>
<td>OFFERED</td>
<td>OFFERED</td>
<td></td>
</tr>
<tr>
<td>NY</td>
<td>Suffolk County Department of Labor</td>
<td>OFFERED</td>
<td>TESTED</td>
<td></td>
</tr>
<tr>
<td>OH</td>
<td>Eastern Gateway Community College</td>
<td>OFFERED</td>
<td>OFFERED</td>
<td>OFFERED</td>
</tr>
<tr>
<td>PA</td>
<td>Central Susquehanna Intermediate Unit</td>
<td>OFFERED</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SC</td>
<td>South Carolina Department of Social Services</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TX</td>
<td>Alamo Community College District and University Health System</td>
<td>TESTED</td>
<td></td>
<td>TESTED</td>
</tr>
<tr>
<td>WA</td>
<td>Edmonds Community College</td>
<td>OFFERED</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WA</td>
<td>Workforce Development Council of Seattle-King County</td>
<td>OFFERED</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WI</td>
<td>Milwaukee Area WIB</td>
<td></td>
<td></td>
<td>OFFERED</td>
</tr>
</tbody>
</table>

Notes: “OFFERED” refers to programs where the component existed as part of the program from the beginning of the grant period; whereas “TESTED” refers to programs where the component was added as part of the three-armed experimental evaluation, partway through the grant period.


It is important to note that in some cases, the programs struggled with implementation.

For example, one grantee decided that it would add an enhancement component in each of its ten...
programs, but the program staff themselves were less eager to do so. This required the evaluation team to engage more intensively with these programs first to “convince” them to take on the enhancement, and then to provide them with technical assistance to establish the component. Programs that implemented the peer support enhancement found that some participants randomized into the enhanced treatment group needed convincing to participate in the peer support group. This led some programs to change the “opportunity” to participate to a “requirement” to participate, in order to ensure great enough take-up for a strong test. There were also implementation success stories. Staff in one program were very eager to give as many enhanced treatment group members as possible the benefit of the added emergency assistance available; they ended up disbursing assistance to about 60 percent of them, under a broad definition of “emergency,” which included emergency prevention as well. These examples highlight that—although the evaluation aimed at relatively consistent and high-quality implementation of the selected components—there was variation in that implementation in practice.

**Criteria for Designing and Operating a Three-Armed Experiment**

From the HPOG experience, we identified eight practical and technical criteria for consideration in implementing a three-armed experiment to set up a rigorous, reliable, and relevant test of program components. Beyond these criteria, there may be other factors that are important to address when choosing a program component to test in a particular circumstance. For example, as noted above, an important factor in the HPOG case was that the program component could be implemented relatively quickly; this would not necessarily be an important factor in another experiment. Additionally, the HPOG-specific implementation of a three-armed
design had to account for the fact that the enhancements were added well after the program began. To a certain extent, this might be desirable in future applications of the design because it allows programs to mature in their standard implementation before considering adding on. In other situations, an evaluation might work with programs from the beginning to ascertain how to carve out the right pieces to test for optimal program learning.

The criteria discussed below are generally universal conditions to adhere to for a successful three-armed experiment, whether from the beginning of a program’s implementation or after the standard program is in place. Table 2 summarizes the criteria along with observations about their particular application in the HPOG 1.0 Impact Study. We then discuss each of the criterion in more detail following Table 2.

**Table 2. Criteria for Designing and Operating a Three-Armed Experiment as Gleaned from the HPOG Experience**

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Definition</th>
<th>HPOG’s Implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Funding Commitment</td>
<td>The funding agency and/or program administrator(s) has the commitment and resources to support the evaluation.</td>
<td>The Federal Government funded the HPOG Program and, per the authorizing legislation, mandated an evaluation. In addition to the legal obligation to evaluate, agency staff were eager to push the evaluation to learn as much as possible, and took the opportunity to invest in the three-armed experiments to test three distinct program components.</td>
</tr>
<tr>
<td>2. Relevance</td>
<td>The funding agency has interest in learning about the program component.</td>
<td>Policy relevance was a main criterion in deciding what to carve out for the multi-armed tests, where learning about the chosen components would be of value to program managers and policymakers alike.</td>
</tr>
<tr>
<td>3. Evidence of Potential Effectiveness</td>
<td>Prior/suggestive evidence of effectiveness justifies testing this program component.</td>
<td>Although no rigorous evidence existed on any of the components under consideration, at the very least some prior research on them flagging their potential effectiveness was needed. That is, we did not want to implement something with no hope of making a difference. In some cases, that evidence was “program staff think it’s really important.”</td>
</tr>
<tr>
<td>4. Randomization Fidelity</td>
<td>Randomization can be established and maintained (e.g., no spillover).</td>
<td>Because of the commitment to operate a rigorous, experimental evaluation, randomization needed to be feasible. This ruled out some kinds of components.</td>
</tr>
<tr>
<td>5. Implementation Consistency</td>
<td>The component can be implemented consistently across programs.</td>
<td>Beyond the fact that the given component could be implemented, the HPOG evaluation also aimed to ensure relative consistency of the enhancements’ implementation across programs.</td>
</tr>
<tr>
<td>Criterion</td>
<td>Definition</td>
<td>HPOG’s Implementation</td>
</tr>
<tr>
<td>---------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>6. Documented Implementation</td>
<td>The component has demonstrated successful implementation in practice.</td>
<td>This was desirable in the HPOG case for proof of concept reasons and for methodological reasons. HPOG grantees that had not yet implemented one of the selected components needed the proof of concept assurance that came from other grantees’ prior successful implementation. Methodologically, there was a benefit to having both experimental and non-experimental evidence on the same components.</td>
</tr>
<tr>
<td>7. Implementation Contrast</td>
<td>The program where the test would take place does not have the given component in place, and so adding it would create a program contrast worth evaluating. Alternatively, a program that already offers the component could change access rules in order to create a program contrast worth evaluating.</td>
<td>This was a criterion in the HPOG case because the evaluation added program components not already in operation at that program in order to create an implementation contrast.</td>
</tr>
<tr>
<td>8. Sufficient Power</td>
<td>There is sufficient power to test the component’s effectiveness.</td>
<td>The study undertook a series of power analyses to identify whether the expected sample sizes would be sufficient to detect impacts of the size that might be likely to arise from the test and would also be policy relevant to detect.</td>
</tr>
</tbody>
</table>

**Criterion #1: Funding Commitment**

One of the principal conditions for conducting a three-armed experiment (and any evaluation in general) is having the commitment and financial support of the funder—both of which the HPOG Program was fortunate to have secured. ACF significantly invests in rigorous evaluation of the HPOG Program. Its commitment to doing so is driven by HPOG’s authorizing legislation (discussed below) and the opportunity HPOG presents to inform policy, practice, and research (discussed under criterion #2).

As noted above, HPOG was authorized as a demonstration program and the legislation mandated an evaluation to assess the success of the initiative. This legislative mandate signaled that evaluation is a fundamental component of the HPOG initiative, conferred upon ACF the authority to make participation in evaluation activities a condition of grant award, and provided the requisite funding to support these activities. OPRE recognized that the confluence of these conditions presented an exciting opportunity for rigorous evaluation of wide scope and rich
depth. OPRE has a long history of studying ACF programs and the populations they serve through rigorous research and evaluation projects. Additionally, OPRE established the ACF Evaluation Policy,\textsuperscript{11} which guides ACF’s culture as a learning organization and confirms its commitment to using evidence to inform policy and practice. Moreover, OFA was supportive of rigorous evaluation and willing to serve as a key partner in coordinating the execution of the HPOG Program with the evaluation activities, to ensure the efforts learn from and inform each other. Thus, both OPRE and OFA were open and eagerly willing to heed the legislation’s call for rigorous evaluation of the HPOG Program.

HPOG’s funding level is sufficient to support both program administration and rigorous evaluation. Since its authorization in 2010, the Program has had an annual appropriation of about $85 million, with $10 million of this amount set aside for evaluation activities. This level of resources enabled OPRE to design and implement a comprehensive research and evaluation portfolio for HPOG. Having a three-armed experiment as a component of this portfolio promised to both maximize the learning from HPOG and advance this innovative methodology. Therefore, ACF invested both program and evaluation funds to support this study feature, provided that the components selected were germane to HPOG (and other career pathways programs both in the healthcare field and more broadly) and had the potential to generate evidence relevant for policy, practice, and research (as discussed under the next criterion).

\textit{Criterion #2: Relevance}

Another factor central to garnering support for a three-armed experiment and selecting which components to test is the relevance of anticipated findings. As noted above, ACF

recognized that HPOG in general—and an experimental test of program components specifically—presented a very appealing opportunity to inform policy and practice, and to advance the research field.

Both OPRE and OFA had a genuine interest in understanding not only whether HPOG was effective but also why—what about the model made it work? Such information could inform ACF programs and policy by helping to refine program expectations and requirements for future generations of HPOG (or similar) programs, and by offering guidance for service provision under other ACF programs (such as TANF). Such information would also be highly valuable to program practitioners in the field. Knowing what specific program components are effective and for whom could help guide program design and service delivery, use of operational funds, and continuous quality improvement efforts.

Additionally, an experimental test of specific program components would be an important contribution to the research field. HPOG offered an opportunity to “get inside the black box” to explore what about a program influences impact magnitude (e.g., Bloom, Hill, & Riccio2003; Greenberg, Meyer, & Wiseman, 1994). Many employment and training interventions are bundles of services; often studies cannot disentangle which specific piece of the bundle drove or influenced impacts. HPOG’s size (in terms of the number of grantees, programs, and participants) and structure (being comprised of a range of different services) allowed for the intentional selection and addition of discrete components for an experimental test to identify what particular strategies may be especially effective across outcomes, interventions, and populations.

OPRE recognized that a three-armed experiment under HPOG would also be relevant for the mounting career pathways research field. When HPOG was authorized, the career pathways
framework was gaining attention as a promising approach to promote education, training, and workforce advancement among low-income and low-skilled individuals. Yet the framework was relatively new and its effectiveness had not been rigorously evaluated. OPRE was leading the field in evaluating career pathways programs with its Pathways for Advancing Careers and Education (PACE) project, which had launched in 2007 prior to the authorization of the HPOG Program. HPOG provided OPRE with an opportunity to build on this portfolio of research and evaluation around the career pathways approach. Identifying specific, effective components of career pathways programs would inform practitioners hungry for concrete career pathways design and implementation guidance, the investments of several federal agencies that pledged support for the approach, and the research field eager to fill gaps in knowledge regarding the structure and effectiveness of this new approach.

**Criterion #3: Evidence of Potential Effectiveness**

Program component(s) selected to test in a three-armed experiment should have prior evidence that suggests the component is effective and is worthy of being examined more closely for its contribution to the program’s impacts. In the case of HPOG, in order to meet the needs of their target populations and the requirements of the FOA, HPOG programs assembled a bundle of services to help participants enroll in and complete training, and secure employment. These services included basic skills education, healthcare occupational training, financial assistance, employment assistance, and academic and personal support services (as described above). Prior

---

12 PACE is a multi-site, random assignment evaluation of nine programs following the career pathways framework.
13 In April 2012, the U.S. Departments of Labor, Education, and Health and Human Services formed a federal partnership and issued a letter of joint commitment to promote the use of career pathways and the coordination of activities in this area across agencies. In 2016, an updated letter was circulated indicating the federal partnership had grown to include the White House National Economic Council, the Office of Management and Budget, and ten additional federal agencies.
to adding the third treatment arm, the HPOG 1.0 Impact Study’s focus was to examine the
impacts of the entire bundle of services offered by HPOG programs.

Several HPOG programs already offered the program components that were selected as
program enhancements. We discuss below why that is an important consideration in its own
right. Nevertheless, that some programs built the components into their initial program model
suggested there was reason to believe the components might contribute to participant success.
Several HPOG grantees provided anecdotal support for the program components that were
ultimately selected when the evaluation team gauged interest in launching the three-armed
experiment. For example, one program that included a strong peer support component in its
original program model cited the support and accountability associated with a group experience
as one of the program’s most important elements. Beyond anecdotal evidence, prior studies had
found more favorable outcomes among students that have established meaningful social
relationships (Karp, 2011), and that better adjusted students are more committed to their
educational goals (Grant-Vallone, Reid, Umali, & Pohlert, 2004).

Similarly, program staff provided anecdotal evidence that unanticipated financial issues
were a barrier to program completion. Research echoes this belief, indicating that financial
constraints are a significant barrier to low-income students’ enrollment in and completion of
postsecondary training and that even a small amount of financial aid can increase student
retention and completion (Goldrick-Rab, Harris, Kelchen, & Benson, 2012). HPOG program
staff and the evaluation team thought that providing students with easy access to emergency
funds could prevent students with unanticipated financial issues from dropping out. Earlier
researchers reported that college administrators and students claimed that emergency financial
assistance had helped students stay enrolled in college (Geckeler, 2008). Finally, in selecting
non-cash incentives as a program component to test, the evaluation team heard from programs that they were interested in implementing a non-cash incentive structure in an effort to increase success with program outputs, such as retention and credential attainment. In this case, the component was lacking prior evidence, though there was evidence that cash incentives increase college credits earned, retention in college, and measures of student motivation (Richburg-Hayes et al., 2009).

In all three cases, the empirical evidence on the effectiveness of the components was limited but suggested that the components could be an important factor in the bundle of services offered by HPOG programs. This helped to justify adding them as enhancements to programs as a way to test the unique contribution of each to HPOG’s impacts.

**Criterion #4: Randomization Fidelity**

An important consideration of any experimental test is that randomization can be established and maintained without spillover between the experimental groups. The programs implementing this type of three-armed test have to ensure that control group members receive no program services to preserve the overall test of the program’s impacts. To preserve the test of the specific program component, they also have to ensure that, among those receiving program services, only those assigned to the enhanced treatment group have access to the enhanced program services. This creates an added challenge for programs in the evaluation.

One way to maintain randomization fidelity is to select program components that are discrete and “stand-alone,” and therefore easily withheld from participants in the standard treatment arm. The program components selected for HPOG’s three-armed test all share this quality. It is straightforward to offer emergency assistance and a non-cash incentive program to
some participants and not others, and it is straightforward to organize a facilitated peer support group among only a select group of students. On the other hand, testing a program component such as intensive case management is of great interest to the field, but poses some challenges to maintaining randomization fidelity in a three-armed test. Specifically, it would require a single program to implement two simultaneous approaches to case management while also ensuring no crossover between the standard and enhanced treatment groups. To do this, one option would be to ask individual case managers to provide some services to some clients and to provide other services to other clients. However, it is a tall order to ask a staff member to operate two models concurrently. Given the complexity with that approach, another option would be to randomize which case workers provide which kind of case management. In that case, service provision is perfectly correlated with its service provider, and—unless there is a relatively large sample size of providers—the resulting evidence would be limited. Although these are not impossible challenges to overcome, this kind of program component makes it more difficult for a program to implement and maintain separation between the two treatment groups.

**Criterion #5: Implementation Consistency**

Another important reason to choose program components that are discrete is that they need to be implemented consistently across programs selected for the three-armed test. For the larger experimental test of HPOG as a funding stream, it was less important that all 42 HPOG programs implemented their programs uniformly. Programs were expected to provide a bundle of services that met general criteria delineated in the grant FOA (e.g., offering healthcare training for occupations that were in-demand to the local labor market, providing academic and
nonacademic support services, etc.), but were given the flexibility to develop programs that were suitable to the local labor market, service delivery area, and target population.

However, the program components tested in the three-armed test needed to be operationalized and implemented consistently across programs to provide an adequate test of the specific program component. The program components selected as enhancements to the standard HPOG program could all be operationalized and implemented relatively consistently across programs. To ensure this, the evaluation team created guidelines for each of the three program components and used those guidelines in working with programs to establish the components. For example, emergency assistance was defined as providing direct financial aid to students facing an immediate crisis or need, such as utility shutoff, vehicle repair, etc., to help them remain enrolled in training. The implementation guidelines specified procedures and assistance limits that the programs agreed to follow, which also helped narrow the scope of the definition of “emergency assistance.” Although programs set up different procedures for students to access emergency assistance and had different definitions of what constituted an emergency, the general principle of “emergency assistance” was discrete enough that it could be implemented consistently across programs. That said, while all of the programs adhered to the general guidelines, there were still implementation challenges and successes (described above) that created variation as well as some slight variations in practice. For example, administrative structures meant that some locations had slower reimbursement processes than others. In addition, an “enthusiasm factor” meant that one location was simply much more eager than the others in getting the word out to enhanced treatment group members about the availability of assistance, and that location had a relatively higher take-up rate.
Criterion #6: Documented Implementation

The component(s) being tested must also have demonstrated successful implementation in practice. A three-armed experiment is not a kind of evaluation design that should be used for a new demonstration but instead to build evidence on something that already has some suggestive evidence behind it. Epstein and Klerman (2012, 2016), for example, suggest a threshold test or “tollgate” for identifying when a program is ready for evaluation, and having been implemented with fidelity to its model is part of that identification.

In the case of HPOG, this documented implementation came from the grantees who had chosen to integrate the given components as part of their program design from the beginning of the grant. As such, there was evidence from the field that the component could be implemented as part of the multi-faceted HPOG program. Existing implementation in the field also helped inform guidelines that the evaluation team developed to aid in implementation of the enhancements. Prior implementation success was desirable in the HPOG case (practically) for proof of concept reasons and (technically) for methodological reasons. Practically, it was helpful to be able to point to other grantees’ success in operating the components when recruiting grantees to test the effectiveness of the selected program components. Methodologically, eventually having both experimental and quasi-experimental evidence on the effectiveness of the same program components would help advance the evaluation field. Choosing a program component that already existed in the HPOG program world meant that there was both natural variation across sites in those components’ existence and there would be experimental evidence of their effectiveness as well. Those two kinds of evidence could be brought together to learn more, such as what non-experimental methods best replicate experimental results (Bell & Peck in the ER special issue; Peck et al., 2018).
Criterion #7: Implementation Contrast

It is also important that programs testing a given program component do not already have that component in place, so that adding it creates a program contrast worth evaluating. In HPOG’s case, a program was “eligible” for implementing and testing one of the program components only if it did not already have the given component in place (the “OFFERED” programs identified in Table 1). This eligibility requirement supported implementation contrast. This can be established from the beginning of a program or established after a program’s initial implementation.

Considering a similar three-armed experiment in another setting, the criterion of implementation contrast could be met if a program would be willing to ration access to a program component that is already has in place. This would involve discontinuing or changing access to a given program component for some treatment group members in order to create a sufficient contrast worth evaluating. The HPOG evaluation team briefly considered this possibility but concluded that grantees who included the component in their original program design would likely not be willing to withhold the component from some treatment group members in the interest of an evaluation; instead, they would likely only be willing to add a new component to test.

Nevertheless, this is an option for future evaluations to consider: changing eligibility rules and the way participants gain access to certain parts of a program—whether it is to gain access to a newly added enhancement or to change access rules for a given program component. For example, consider that internship opportunities are limited and, currently, people are selected to participate either based on certain characteristics or on a first-come-first-serve basis. If
programs were willing to discontinue the first-come-first-serve access, they could replace it with a lottery, giving equal chance to anyone who met eligibility rules, even if that meant denying the opportunity to some people who might have received it under prior rules. That is, a meaningful contrast could be established by changing program rules to impose a random assignment process that would allow some people access to an in-demand service or opportunity (such as internships) in order to learn more about the impact of that program component.

**Criterion #8: Sufficient Power**

The final criterion is that there is sufficient power to test the component’s effectiveness, as with experiments in general. The HPOG evaluation team undertook a series of power analyses as eligible programs were identified and recruited to ensure the anticipated effect sizes were plausible and policy relevant. The conclusion from these analyses was that the magnitude of effect that could be detected—given sample size and other assumptions—would be large, but not unreasonably so; and that if those effects would arise, then it would be relevant knowledge for policymakers and program administrators.

**Discussion: Implications for Policy Learning and Evaluation Practice**

This paper has focused on practical and technical criteria for consideration in implementing a three-armed experiment within the context of a larger, multi-site experimental evaluation, using the HPOG experience as an example. The major practical requirement for any rigorous evaluation is the commitment—political, financial, and programmatic—from a variety of stakeholders to ensure its success. Beyond that general commitment to evaluation, there are additional considerations when assessing the possibility of mounting a multi-arm study.
From the evaluation team’s perspective, additional considerations were mainly technical in nature and related to the integrity of the design. For example, randomization fidelity (criterion #4) was essential to the evaluation team: we wanted a strong evaluation where random assignment into the enhanced and standard groups could be established and maintained over time. Additionally, the evaluation team also wanted to ensure that the test had sufficient power to detect likely effects (criterion #8).

From the local programs’ perspective, it was important to test components that were relevant to the types of services they deliver and that had the potential to improve the experiences and outcomes of their participants. Grantees were eager to learn which aspects of programs are the essential ingredients to their success; such knowledge could inform decisions about program design and investment of resources in program operations. It could also support sustainability of their programs beyond the HPOG grant period. Evidence regarding not only the effectiveness of their overall program model but of specific components of that model would provide a strong and compelling case to elicit support from community stakeholders for continuing the program (or aspects of it).

From the government’s perspective, it was important to capitalize on the research and evaluation opportunities afforded by HPOG. As discussed above, ACF was in the fortunate position to have leadership and financial support for a comprehensive and rigorous evaluation of the HPOG Program; such conditions do not necessarily exist for other federally-funded programs. In response, ACF intentionally designed the evaluation portfolio to maximize learning from the HPOG Program, to inform the design and operation of the Program itself and to inform other relevant areas. For example, we can use what we learn in the HPOG case to generalize to other programs that serve similar populations (such as TANF recipients) or operate similar
program models (such as other career pathways programs). Further, evidence generated by HPOG can inform policy regarding postsecondary education and training approaches for low-income populations. In a world of fewer resources and increasing pressure for evidence-based policy, there is growing appreciation of the need to move beyond an evaluation framework that considers only whether a program works, and instead consider why a program does or does not work and for whom (Solmeyer & Constance, 2015). HPOG’s three-armed experiment advances this agenda. Its experience can serve as a model for other large-scale evaluations and its findings can inform where future research should focus.

The Results of HPOG’s Three-Armed Experiment

At this point in the discussion, our readers are probably wondering: what were the results of the HPOG three-armed experiments? The study’s analysis indicated there were not detectable impacts for two of the three enhancements, and there was a detectable but unfavorable impact on the third. Table 3 reports the impacts detected, the assumption-based minimum detectable effects (MDEs) estimated during the project’s design phase, and the data-based MDEs estimated after the fact using actual project data.

Table 3. Enhancements’ Impacts, Estimated MDEs, and Post Hoc MDEs

<table>
<thead>
<tr>
<th>Enhancement</th>
<th>Estimated Impacts on Q5 earnings (standard error)</th>
<th>Assumption-based Advance MDEs [planned sample]</th>
<th>Data-based Post Hoc MDEs [actual sample]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emergency Assistance</td>
<td>$154 (190)</td>
<td>$659 [532]</td>
<td>$472 [1,585]</td>
</tr>
<tr>
<td>Non-Cash Incentives</td>
<td>-104 (191)</td>
<td>540 [951]</td>
<td>474 [1,289]</td>
</tr>
<tr>
<td>Facilitated Peer Support</td>
<td>-432* (251)</td>
<td>395 [1,543]</td>
<td>625 [901]</td>
</tr>
</tbody>
</table>

Sources: Harvill et al. (2018), Appendix Exhibits F.2, F.4, F.6; November 21, 2013 internal project memo. Statistical significance levels for two-sided tests are indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.
Post hoc impact analyses revealed that the impacts on the first two enhancements would have needed to be three to four times larger than what we observed to be detected in the case of the emergency assistance and non-cash incentives enhancements. In both cases, the study’s actual sample size exceeded the planned sample size used in assessing whether the anticipated MDEs would be of the size to be detected and policy relevant. As a result, the study’s actual MDEs were smaller than anticipated, but the enhancements’ actual impacts did not materialize in the magnitude needed to be detected.

One reason for these null effects is that the programs implementing the enhancements experienced implementation challenges (as described above) that resulted in tests that were not as strong as might have been ideal (Peck et al., 2018). Considering the negative effect estimated for facilitated peer support, as described earlier the programs implementing this enhancement cited challenges with student attendance to the support group meetings. By making peer support group meetings mandatory to try to address the attendance challenge, it is possible the programs added yet another time commitment for participants already overburdened with school, work, and family responsibilities.

We caution about the generalizability of results from these enhancement tests to the potential effectiveness of these three program components more generally. Peck et al. (2018) and Walton, Harvill, and Peck (2017) analyze the study’s impacts among programs that offered the program components from the beginning of the grant period, added the component as enhancement tests halfway through the grant period, or neither initially offered the program component nor later enhanced their programs with it. These analyses flag that those three sets of programs are quite distinct from one another. As a result, the impacts of adding program enhancements should be expected to differ from the impacts of having the same component fully
integrated into the program from the beginning. This is a lesson relevant to future applications of this sort of three-armed experiment: although the experiments provided rigorous evidence, they did so for a selected subset of the programs. Thus, in the HPOG case, the experimental evidence should be generalized only to those programs that would be interested in adding the component later on as opposed to those programs that integrated the component from the beginning. Other project analyses consider, via non-experimental analyses, the more general case of the availability of various program components and implementation features across the full set of HPOG locations.¹⁴

Conclusion

Multi-armed experimental evaluations have the potential to provide important “black box opening” insights, with four key beneficiaries in mind: evaluation scholars, policymakers, program practitioners, and program participants. The value to scholars comes from design innovations and associated challenges: evaluators can learn from these experiments in practice in order to improve the quality of future applications. The technical challenges in any one test can be tackled and streamlined in future tests. For example, future evaluations could account for the administrative challenges in dispensing emergency assistance experienced in HPOG’s test prior to implementation of their own test. Future tests might also sort out challenges not encountered in HPOG’s case but likely to arise in other three-armed experiments, such as no-shows and cross-overs.

¹⁴ Chapter 7 of the study’s Implementation and Short-Term Impacts Report (Peck et al., 2018), the report’s Appendix G (Harvill et al., 2018), and a project special topics paper (Walton, Harvill & Peck, forthcoming) provide details of these additional analyses and their results.
The value to policymakers comes from the substantive learning about the best investment of resources. Knowing whether and why a program works and for whom can guide the investment of scarce resources to maximize the efficiency and effectiveness of program operations. The value to program administrators comes from the direct learning about which program components are essential ingredients for their programs’ success. Such information can guide programmatic decisions around service delivery and resource allocation, and help build a case for stakeholders’ continued or novel investment in the program.

Finally, a fourth set of beneficiaries cannot be overlooked—the participants in the programs that are subject to rigorous evaluation. The value that three-armed experiments offer to the public is the evidence they generate to inform smart program design and practice. This evidence helps ensure that the programs that receive funding operate in the best way possible to improve their participants’ lives.
Works Cited


Author Bios

LAURA R. PECK, Ph.D., is a Principal Scientist in the Social & Economic Policy Division at Abt Associates and also Global Lead for Abt’s Research, Monitoring, & Evaluation Capability Center. Dr. Peck specializes in innovative ways to estimate program impacts in experimental and quasi-experimental evaluations, and she applies this to many social safety net programs.

HILARY FORSTER, M.P.P, is a Senior Social Science Research Analyst in the Office of Planning, Research, and Evaluation at the U.S. Department of Health & Human Services’ Administration for Children & Families. Her work at OPRE focuses on career pathways programs, employment and training programs, coaching interventions, family self-sufficiency, and conducting research with tribal communities.

NICOLE CONSTANCE, Ph.D., is a Social Science Research Analyst in the Office of Planning, Research, and Evaluation at the U.S. Department of Health & Human Services’ Administration for Children & Families. Her work at OPRE focuses on career pathways programs, employment and training programs, family self-sufficiency, programs for parents who are justice-involved, and programs for noncustodial parents.