

Board of Advisors

Robert Boruch
University of Pennsylvania

Jonathan Crane
Coalition for Evidence-Based
Policy

David Ellwood
Harvard University

Judith Gueron
MDRC

Ron Haskins
Brookings Institution

Blair Hull
Matlock Capital

Robert Hoyt
Jennison Associates

David Kessler
Former FDA Commissioner

Jerry Lee
Jerry Lee Foundation

Dan Levy
Harvard University

Jodi Nelson
Bill & Melinda Gates
Foundation

Diane Ravitch
New York University

Howard Rolston
Abt Associates

Isabel Sawhill
Brookings Institution

Martin Seligman
University of Pennsylvania

Robert Solow
Massachusetts Institute of
Technology

Nicholas Zill
Westat, Inc.

President

Jon Baron
jbaron@coalition4evidence.org
202-683-8049

Increasing the Success of Evaluation Studies in Building a Body of Effective, Evidence-Based Programs:

Recommendations of a Peer-Review Panel

Earlier this year, the Coalition for Evidence-Based Policy convened a peer review panel to assess the Department of Labor Employment and Training Administration's (ETA) evaluation efforts. As part of its deliberations, the panel developed two main recommendations, designed to increase the efficiency of ETA's evaluation efforts in building a body of workforce programs, practices, and strategies that have credible evidence of effectiveness. We believe these recommendations may be applicable beyond ETA, and have prepared this brief to share them with the broader community of government and philanthropic officials interested in advancing evidence-based social spending.

The recommendations are attached (two pages), along with two short appendices: (a) a summary of an ETA evaluation that we believe is an excellent example of the approach the panel is recommending for wider use; and (b) a list of the evaluation studies that the panel reviewed.

Context for the recommendations:

- The recommendations were developed by a peer review panel comprised of –
 - (i) Rebecca Maynard, University of Pennsylvania ([bio](#))
 - (ii) Larry Orr, Johns Hopkins University ([bio](#))
 - (iii) Jon Baron, Coalition for Evidence-Based Policy ([bio](#))
- The panel was convened by the Coalition under a contract with ETA's Office of Policy Development and Research, which provided the panel with a set of 30 evaluations to review (appendix B). These evaluations comprise most of the larger studies that ETA has sponsored since 2006. ETA officials did not participate in the panel's review activities nor in its development of the recommendations.
- The recommendations focus on efficient evaluation strategies, but not on the contract or grant mechanisms (e.g., under the Federal Acquisition Regulation) needed to implement these strategies. The example in appendix A illustrates, however, that ETA has successfully implemented these strategies in at least one evaluation initiative.
- The Coalition for Evidence-Based Policy is a nonprofit, nonpartisan organization that has no affiliation with any program or program models, and no financial interest in any of the policy proposals that it supports.

Recommendations for the DOL Employment and Training Administration's (ETA) Evaluation Program

Offered by the Peer Review Panel on ETA's Evaluation Program¹

Based on our review of ETA's evaluation program, we offer the following recommendations to increase the program's efficiency in building a body of workforce interventions (i.e., programs, practices, or strategies) that have credible evidence of effectiveness. Our main recommendations flow from a recognition that the clear pattern in past experimental evaluations in workforce development, as in other fields (e.g., medicine, education, crime prevention), is that most interventions produce weak or no positive impacts compared to services-as-usual. Taking account of this pattern can lead to substantial economies in ETA's evaluation program.

Recommendation 1

Where feasible, fund impact evaluations in “seamless” stages, using relatively low-cost methods in the initial stages. This has two main advantages:

- (i) **It would improve efficiency by allowing mid-course adjustments in the evaluation plans** (including, in some cases, reducing scope to omit portions of study plans that are not likely to yield credible and useful evidence); and
- (ii) **It would increase the number of interventions that can be rigorously tested within ETA's evaluation budget.**

For a given intervention, these stages might include the following:

- **First, an experimental evaluation to measure the intervention's impact on the primary outcome(s) of interest (e.g., earnings) – if possible, using low-cost administrative data.** This first stage is thus designed to answer the most policy-important question: does the intervention produce the main hoped-for effects?

We suggest that this initial evaluation also obtain basic information on the intervention being evaluated, to help inform its replication should it be found effective. This might include: (a) data on its operations – e.g., number of workers trained, how many hours of training they received, what skills were taught; and (b) data on the intervention's cost – obtained, for example, by aggregating the cost of labor, materials, and other resources used to deliver the intervention. In this initial stage, the study might also measure the intervention's impact on any secondary outcomes for which data can be obtained inexpensively (e.g., receipt of public assistance).

This recommendation is geared toward avoiding large investments in elaborate implementation evaluations and data collection to support exploratory analyses (e.g., to learn about the mechanisms through which impacts occur, the factors that may interfere with effectiveness, and implementation challenges) for programs, policies, and practices that will not, in the end, be sufficiently effective to warrant adoption.

- **Subsequent stages involving more comprehensive data collection and analyses would go forward only if the findings of impact in the first stage are sufficiently promising.** For interventions that have modest to no impacts relative to their cost, traditional implementation

¹ Rebecca Maynard, University of Pennsylvania; Larry Orr, Johns Hopkins University; Jon Baron, Coalition for Evidence-Based Policy

research would have limited value. If any further study of the intervention is warranted, it would likely be more efficient to focus it on examining possible sources of breakdown in the logic of the intervention, for example, by conducting studies of variations in design, targeting, and/or implementation strategy.

- **Examples of the types of strategies that may be warranted in these subsequent stages, conditional on the first-stage findings, include the following:**
 - **Implementation research, conducted retrospectively in the sites that produced the largest impacts and those that produced the smallest impacts, to identify factors that might account for the different size impacts.** Such research would help inform efforts to disseminate the intervention, and/or to develop revised or enhanced versions of the intervention to test in future evaluations. Revised models intended to improve effectiveness in some or all sites might warrant “retesting” in future low-cost evaluations.
 - **A second impact evaluation, to try to replicate the initial findings of effectiveness in other sites, and carry out more comprehensive data collection and analyses.** Such an evaluation might include, for example, collection and analysis of data on secondary outcomes and key subgroups, documentation of services received by both the intervention and control groups, longer-term impact estimates, and benefit-cost analyses. Such information and analyses can help: (a) confirm, through replication, that the intervention is indeed effective, and that its effects generalize to other sites; (b) inform policy decisions about whether to fund larger-scale implementation of the intervention, considering the likely impacts and costs; and (c) identify conditions and populations in which the intervention is most likely to be effective.

Recommendation 2

In selecting interventions to evaluate in a given area (e.g., assistance for disconnected youth), strategically search the existing evaluation literature to identify the strongest candidates – i.e., those most likely to produce sizable positive impacts. We recommend this approach because, as noted above, the clear pattern in past experimental evaluations in workforce development is that most interventions produce weak or no positive impacts compared to services-as-usual. Thus, to increase the likelihood of meaningful positive findings, we suggest clearly focusing evaluation efforts on interventions with compelling evidence or reason to believe they could be the exceptions that are effective.

Appendix A

Illustrative Example of the Panel's Recommended Approach: ETA's Evaluation of the Reemployment and Eligibility Assessment Program

The Coalition has prepared this brief overview of the Reemployment and Eligibility Assessment (REA) evaluation, which we believe is an excellent example of the evaluation strategy that the peer review panel recommends ETA adopt more widely.

Program Description:

REA is a program for Unemployment Insurance (UI) claimants that includes an in-person review of the claimant's eligibility for UI, and personalized job-search and other reemployment assistance.

Specifically, REA is a mandatory program for UI claimants deemed likely to exhaust their UI benefits, who comprise approximately one-third to one-half of the UI claimant population. The program is funded by the Department of Labor at about \$60 million each year, enough to reach only a small segment of the UI claimants who could potentially benefit. The program is delivered by a trained interviewer at a DOL one-stop center, and provides:

- An in-person UI eligibility review to verify that the claimant is eligible for benefits (e.g., has been actively seeking employment) and prevent overpayment;
- Labor market information (e.g., on job openings, wage trends); and
- Development of an individual reemployment plan, which includes referral to reemployment services (e.g., job-search and resume assistance) and an assessment for training.

The REA program costs approximately \$50-200 per participant.

Evidence of Effectiveness:

A. The program was evaluated in a well-conducted randomized control trial, carried out in four states – Florida, Idaho, Illinois, and Nevada – with a sample of 134,550 UI claimants.¹

Between July and December 2009, these claimants were randomly assigned to either (i) a group required to participate in REA, or (ii) a control group that was not offered the program. Our review found that this study meets widely-accepted criteria for a well-conducted randomized controlled trial, generating strong evidence about program effectiveness.²

B. The study found the following effects over a 12-18 month follow-up period, all of which were statistically significant:

- **\$180 in net government savings per claimant from reduced UI payments.³**
- **Especially large savings in Nevada – \$604 per claimant** – possibly due to distinctive features of Nevada's REA program that could be replicated elsewhere (as described below).
- **An increase in job earnings of \$2,600 (18%) per claimant in Nevada – the one site that obtained a reliable estimate of the effect on earnings.** The study also found a smaller (5%) increase in earnings in Florida over a 12-month period, but the study's analysis suggests this finding may not be reliable.⁴

Why We Regard This Evaluation As Exemplary of the Panel's Recommended Approach:

- A. REA was a promising candidate for a rigorous evaluation, based on earlier research findings about the value of job search assistance – a key element of REA.** Such assistance has been found effective in some earlier rigorous evaluations. For example, the UI Job Search Assistance Demonstration – an ETA-sponsored randomized trial of mandatory job search assistance for UI claimants – found a sizable increase in earnings for UI claimants, and net savings to the government, at one of the two study sites.⁵ (Based on the study report, the lack of effects at the other site may have been due to weak program implementation.) Job search assistance has also been found to be an effective reemployment strategy in some other policy contexts, such as welfare-to-work.
- B. The study used a staged approach, obtaining a reliable estimate of program impact at very modest cost, and investing in further research only after effectiveness had been established.** Specifically, the study's cost was quite modest because it measured all outcomes using administrative data on UI receipt and earnings that the states collect already for other purposes – as opposed to the more expensive approach of locating each sample member and administering an interview to obtain his or her outcome data. As a rough estimate, the study cost \$320,000 through the initial 12-18 month follow-up, based on informal communications with the principal investigator. This is a fraction of the usual multimillion-dollar cost of such large impact evaluations.

Based on the especially promising findings in Nevada, ETA commissioned a second stage of the evaluation – namely, longer-term follow-up in that state (20-26 months) – which found that the sizable effects in Nevada were sustained over the longer period. This stage of the study was also conducted at modest cost (approximately \$80,000), by using administrative data to measure the outcomes.

Although these first two stages of the study did not include a formal implementation research component, the researchers identified distinctive features of the Nevada program that they theorize may account for its larger impacts. Specifically, in Nevada, both the UI eligibility assessment and the reemployment services were mandatory, and were typically provided in the same interview session, whereas in other states the reemployment services were voluntary and usually provided by a separate agency or organization.

As a third stage in the evaluation, ETA recently awarded a \$1.9 million contract to Abt Associates to evaluate which aspects of the REA program are most effective, drawing in particular on the Nevada approach.⁶ The goal, presumably, is to see if the large Nevada impacts can be replicated elsewhere, and to build knowledge needed to disseminate the program in its most effective form.

In summary:

This evaluation effort has produced a credible finding of effectiveness that is of practical importance for policy decisions – and it did so at remarkably modest cost. This was achieved by (i) focusing on an intervention backed by promising prior evidence, and (ii) carrying out the evaluation in stages that used rigorous, low-cost methods to establish initial evidence of impact, and made further research investment only after such impact had been found. Thus, we believe this study illustrates the value of the peer review panel's recommended evaluation strategy.

References

¹ Eileen Poe Yamagata, Jacob Benus, Nicholas Bill, Hugh Carrington, Marios Michaelides, and Ted Shen. *Impact of the Reemployment and Eligibility Assessment Initiative*. Impaq International, June 2011. Marios Michaelides, Eileen Poe-Yamagata, Jacob Benus, and Dharmendra Tirumalasetti. *Impact of the Reemployment Eligibility Initiative In Nevada*. Impaq International, January 2012.

² The criteria we used to assess whether the trial was well-conducted are summarized in the [RCT Checklist](#), and include such items as: (i) the program and control groups were similar in their pre-program characteristics; (ii) the study had low sample attrition, and similar attrition rates for the program versus control group; (iii) the study measured outcomes for all individuals assigned to the program group, regardless of whether or how long they participated in the program; and (iv) study outcomes were assessed with valid measures.

³ This is an average of the net savings across the four states in the study, weighted by the sample size in each state. Net savings within each state were calculated by subtracting the per-person program cost from the per-person savings in UI payments. This estimate is meant to be conservative since it assumed a per-person program cost of \$201 in all states, which was the highest estimate of costs observed across the four states.

⁴ The study measured the program's effect on job earnings in the Florida site, and found that it produced a statistically-significant \$476 increase in earnings per claimant over a 12-month follow-up period, roughly offsetting claimants' loss in UI benefits. However, this was a regression-adjusted effect on earnings; the unadjusted effect was near zero and not statistically significant. Because the effect differed under these two different estimation approaches, we believe the positive findings for earnings in Florida are best viewed as tentative, and need corroboration in future studies before being accepted as valid.

⁵ Paul Decker, Robert Olsen, Lance Freeman, and Daniel Klepinger. *Assisting Unemployment Insurance Claimants: The Long-Term Impacts of the Job Search Assistance Demonstration*. Mathematica Policy Research, 2000.

⁶ Abt Associates' press release on the evaluation is [linked here](#).

Appendix B

ETA Evaluation Studies Reviewed by the Peer Review Panel

- Beneficiary Choice Program - Evaluation
- Career Advancement Account Evaluation
- Community-Based Job Training Initiative, Phase I, II, and III – Evaluation
- Enhanced Transitional Jobs Demonstration
- Green Jobs and Health Care Impact Evaluation
- Improving Measures of Efficiency for Employment and Training Programs
- Individual Training Account Experiment
- Job Corps Preliminary Study
- Jobs for Veterans Act, Priority of Services (POS) - Evaluation
- Latino Coalition’s Youth Offender Project - Evaluation
- Mentoring, Educational, and Employment Strategies to Improve Academic, Social and Career Pathway Outcomes in Persistently Dangerous Schools - Evaluation
- National Agricultural Worker Survey
- National Assessment of Adult Literacy Secondary Analysis
- National Fund for Workforce Solutions
- New American Centers Demonstration - Evaluation
- Post-summer 2009/2010 ARRA Summer Youth Initiative - Evaluation
- Prisoner Re-Entry Initiative – Evaluation
- Ready4Work Analysis
- Reemployment and Eligibility Assessment (REA) Program Evaluation
- Re-Integration for Ex-Offenders (RExO) Program - Evaluation
- School District-Based Strategies for Reducing Youth Involvement in Gangs and Violent Crime Through a Workforce Development Approach - Evaluation
- Self-Directed Services - Evaluation
- Technology-Based Learning (TBL) Demonstration - Evaluation
- Twin Cities Rise! – Evaluation
- Workforce Innovation in Regional Economic Development (WIRED) Demonstration - Evaluation
- Workforce Investment Act (WIA) Adult and Dislocated Worker Programs - (WIA Gold Standard Evaluation)
- Workforce Investment Act (WIA) Quick Impact Study
- Workforce Investment Act (WIA) Waiver Evaluation
- Young Parents Demonstration
- YouthBuild Program