



Considerations for Assessing the Applicability of Evaluation Evidence

Adam Travis, September 2023



Findings from causal impact evaluations can provide valuable evidence about the effectiveness of reemployment programs, such as the Reemployment Services and Eligibility Assessments (RESEA) program. Becoming an informed consumer of such evidence creates opportunities to deliver more effective services to unemployment insurance (UI) claimants. One common challenge when using existing evaluation evidence is determining whether findings from an evaluation conducted elsewhere—even one that has been conducted using rigorous methods—are likely to be applicable to *your* program or program context. Often, you may find that the only available evaluation evidence on an intervention's effectiveness may come from a study in another state, where program or contextual features might differ.

External validity refers to the extent to which that evaluation's findings are applicable beyond the specific context (e.g., time, place, population) in which the study was conducted. That is, are findings on an intervention's effectiveness generally valid beyond the context of the evaluation, or do findings only hold under specific conditions? An evaluation with a high degree of external validity is one where there is confidence that findings are **generalizable**. That is, findings from an intervention studied in one context are likely to apply outside of that context.

When using evaluation evidence to inform your RESEA program, you will want to consider whether you are drawing on high-quality, rigorous studies with credible findings that are likely to be relevant for *your* state program. This tip sheet offers three tips for assessing whether an evaluation's findings are likely to be relevant for your RESEA program. **The tips:**

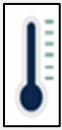
1. Look for consistent patterns of favorable findings from **multiple** rigorous evaluations of the same intervention.
2. When only one rigorous evaluation of an intervention is available, consider the specific ways that program features or contexts might be different from your own. If you decide to implement a program or program component for which limited evidence is available, **consider running your own evaluation**.
3. When evaluation reports estimate impacts for multiple **subgroups**, examine whether and how intervention impacts appear to vary for these different subpopulations.

Tip #1: Look for consistent patterns of favorable findings from multiple rigorous evaluations of the same intervention.

The most effective ways to determine whether findings from a single evaluation are likely to be generalizable is to determine whether other evaluations of that program or program component generated similar findings, even if carried out in different contexts.

The Department of Labor’s Clearinghouse for Labor Evaluation and Research (CLEAR) provides several tools for identifying whether consistent patterns of favorable findings exist for a reemployment intervention. CLEAR’s [Reemployment Synthesis](#) provides an overview of select reemployment interventions and summarizes findings from multiple evaluations. CLEAR also rates the effectiveness of RESEA interventions (see Box 1). RESEA interventions with high or moderate [RESEA intervention effectiveness ratings](#) indicate that the intervention has a consistent pattern of favorable impacts on reemployment outcomes. A high intervention effectiveness rating indicates that more than one rigorous study has found the intervention to be effective. This rating indicates a higher-level of confidence that implementation of this intervention may produce favorable impacts.

Box 1. RESEA Intervention Effectiveness Ratings in CLEAR



High: At least two credible impact studies have found favorable effects on UI duration and employment with strong statistical significance ($p < .05$).



Moderate: At least one credible impact study has found favorable effects on UI duration and employment with modest statistical significance ($p < .10$).



Potentially Promising: At least one impact study (irrespective of study rating) found favorable impacts on UI duration or employment with modest statistical significance ($p < .10$).



No Rating: No impact studies of the intervention exist or studies that exist do not show favorable impacts.

Tip #2: Consider specific ways that program features or context might be different from your own.

Sometimes, it is not possible to find multiple sources of evidence regarding the effectiveness of a program or program component. What can you do when you are only able to identify a single rigorous evaluation of a program or program component?

In some cases, a single evaluation may have been conducted across many sites or states. Such evaluations often have a higher degree of external validity. That is, because such studies are designed to examine program impacts across a variety of contexts, it is reasonable to have more confidence in the generalizability of the evaluation’s findings.

In other cases, however, the only rigorous evaluation may have been conducted in a single study site. In these cases, it is important to examine the specific ways that the context of that study might differ from the context in which your program operates and consider whether any of those differences might have implications for the effectiveness of the intervention. For example, you might examine whether a study was carried out in a location where the labor market, including the general economy and prevalent industries, differed significantly from the labor market in your own state. Or you might examine whether the demographics of the population served by the program under study differed from the population served by your state’s RESEA program.

If you do decide to implement a new program or program component that has not been the subject of multiple rigorous evaluations, you may want to consider running your own evaluation to learn about whether that intervention is effective in your state’s context. The evidence that this evaluation generates

could be useful not only for your own state, but also for other states that operate similar programs or program components.

For impact evaluations that CLEAR has reviewed, relevant details about a study's context are typically contained in the **study profile**. CLEAR study profiles are short, one-page summaries of information about the program or program component being evaluated, the study's research design, key findings from the study, and features of the research site and study sample. Box 2 provides a *hypothetical* example of how you might consider contextual differences in a situation where only a single-site evaluation of a program is available.

Box 2. Considering the Implications of Program Context, A Hypothetical Example

Your state's RESEA program serves a largely rural population and relatively few UI claimants live within reasonable driving distance of an American Job Center (AJC) or workforce development agency. Your RESEA program currently provides job search assistance via in-person sessions but is considering a transition to fully-remote services, where case managers would provide job search assistance exclusively via online meetings.

To ensure that your program is making an informed, evidence-based decision about whether to proceed with this change, you use evidence on the effectiveness of remote job search assistance in CLEAR. Using the **Search for Studies tool** on the CLEAR website, you find one rigorous evaluation of an online job search assistance intervention that was conducted in a densely populated urban area. For the purposes of this hypothetical example, let's say that this evaluation found that the job search assistance intervention was less effective at improving outcomes for UI claimants when services were delivered online than when services were delivered in person. What should you do with this information?

In this hypothetical situation, there is no existing evidence to indicate that switching from in-person to online job search assistance services will improve reemployment outcomes for UI claimants. However, the study was conducted in a densely populated urban area, where short drive times and accessible public transportation made it comparatively easy for UI claimants to attend in-person RESEA meetings or get to a job. In this case, it might be reasonable to anticipate that online job search services might be more effective in rural areas, where accessing in-person services can be a significant challenge due to transportation issues.

If after reviewing the details of the existing evaluation, you still believe that it is possible that online job search services could improve outcomes for UI claimants in rural settings, consider whether it is feasible to implement **and** evaluate a version of this strategy in your state. Conducting your own evaluation will enable you to obtain new evidence about whether online job search assistance services are effective in your state and provide valuable new evidence for other states about the effectiveness of these services in *rural* areas.

Tip #3: When possible, consider how impact estimates vary for different subgroups.

In some cases, an existing RESEA evaluation will report impacts not only for the entire study sample, but also for specific population subgroups. For example, an evaluation might compare program impacts for men and women, by education or employment history, or for participants who reside in urban, suburban, and rural areas. In such cases, looking carefully at these subgroup estimates can give you a sense of the extent to which intervention impacts appear to vary across different populations, as well the extent to which the intervention produced favorable impacts for specific population groups that your state's RESEA program serves.

Impact estimates for subgroups are not always summarized in CLEAR study profiles and you may need to consult the evaluation report directly in order to review these findings. For publications that are

available online, you can always find a link to the original publication in the CLEAR study profile. Because impact estimates for specific population subgroups are generated using smaller samples, these estimates are less precise. When reviewing subgroup estimates, consider whether the study included statistical tests to determine whether differences between impact estimates for a particular subgroup and impact estimates for the full sample (or from other subgroups) were statistically significant.

KEY TAKEAWAYS

When reviewing evidence to learn about the effectiveness of an intervention, to the extent possible, draw on rigorous evaluation findings that are likely to be relevant to *your* program context.

The best way to be confident about the external validity of a finding is to consult multiple evaluations of the same intervention. Doing so can help establish whether there is a consistent pattern of favorable evidence across a variety of contexts. CLEAR’s RESEA intervention effectiveness ratings and reemployment synthesis are valuable tools for identifying those RESEA interventions that have been the subject of past evaluations. When available, meta-analyses and literature reviews can also help you learn about programs or interventions that have consistently improved reemployment outcomes.

In cases where only a single evaluation of an intervention is available, pay close attention to ways that the context of the evaluation might differ from the context in which your own RESEA program operates. If you decide to implement a program or program component in a context where it has not previously been evaluated, consider running your own evaluation of the intervention to help build the RESEA evidence base.

In cases where evaluation reports include impact estimates for population subgroups, examining how intervention impacts vary can provide information about the extent to which the intervention produced favorable impacts for population groups that your state’s RESEA program serves.

ABOUT THIS SERIES

State agencies’ evaluations of their RESEA interventions (i.e., the whole program or aspects of it) are integral to expanding the evidence base of effective service delivery strategies. This tip sheet is part of a series developed to serve as a quick-reference tool for states planning, launching, and carrying out evaluations of their RESEA programs. For additional information on this topic, please also see the [Informed Use of Evidence for RESEA Evaluations Webinar](#)

Suggested citation: Travis, A. (2023). *Considerations for Assessing the Applicability of Evaluation Evidence*. RESEA Evaluation Technical Assistance Tip Sheet. Prepared for the U.S. Department of Labor, Chief Evaluation Office. Rockville, MD: Abt Associates.

Additional Resources and Information

DOL’s RESEA Helpline: RESEA@abtassoc.com

Additional RESEA evaluation technical assistance resources are available here:

https://rc.workforcegaps.org/resources/2019/07/30/17/32/RESEA_Evaluation_Evidence_Resources

Study reports, publications, and other materials are available here:

<https://www.dol.gov/agencies/oasp/evaluation/completedstudies/RESEA-Evidence-Building-Implementation-Study>

Chief Evaluation Office
U.S. DEPARTMENT OF LABOR



This document was prepared for the U.S. Department of Labor (DOL), Chief Evaluation Office by Abt Associates under Contract # 1605DC-18-A-0037. The views expressed are those of the authors and should not be attributed to DOL, nor does mention of trade names, commercial products, or organizations imply endorsement of same by the U.S. Government.