

INCREASING THE INDEPENDENCE OF FEDERALLY-FUNDED IMPACT EVALUATIONS

Jacob Alex Klerman (Abt Associates), Jon Baron (Coalition for Evidence-Based Policy),
and Howard Rolston (Abt Associates)

Please address correspondence to Jacob Alex Klerman (Jacob_Klerman@abtassoc.com).

Problem: Federal agencies’ current approach to funding evaluations of program effectiveness can sometimes undercut the evaluations’ independence.

While OMB has emphasized the importance of “rigorous, independent program evaluation [in] help[ing] the Administration to determine how to spend taxpayer dollars efficiently[,]” contracting for such evaluations and assuring that their results are available for the policy debate is an ongoing challenge. Although problems that compromise the independence of federal evaluations do not appear to be common, they do happen (see the examples discussed in the recent “Point/Counterpoint” in the 2008 *Journal of Policy Analysis and Management*; Reingold, 2008; Metcalf, 2008; and in GAO, 2010; Klerman, forthcoming).

Building safeguards against the introduction of bias into science is not just a challenge for the federal government. It is often noted, for example, that the journal publication process may be vulnerable to publication bias – that is, greater likelihood to publish studies with positive findings while leaving studies that confirm the null hypothesis to languish in file cabinets. Researchers seeking publication may therefore “mine” for positive findings. The proposal we make below is addressed to the federal government, because it is in a position to take simple steps that could reduce the likelihood of bias being introduced into the contract evaluations it sponsors, and because of the high value that is placed on government transparency. As discussed below, we believe these steps will help to reduce potential threats to the independence of federally-funded studies from both the funding agencies and the researchers.

Part of the challenge in ensuring such independence is that a funding agency’s ability to control the content and publication of studies that it sponsors is often written into the contracts for what are labeled “independent” evaluations. On the one hand, agencies need considerable discretion – to select a research contractor and then to guide the evaluation. This discretion helps the agency to prevent researcher bias, or mistakes, from leading to erroneous findings. However, that same discretion gives the agency considerable formal and informal influence over the researcher – and therefore has the potential to compromise study independence and lead to evaluation results which do not fully serve the nation’s needs.

Proposed Solution: Two Straightforward Steps to Strengthen Evaluation Independence

Here we propose two straightforward and low-cost first steps which, alone or together, would substantially improve the independence of contract evaluation.

1. Creating a publicly-available Registry on federally-funded impact evaluations

Consistent with OMB’s goal of ensuring “Public Availability of Information on Federal Evaluations” (OMB 2009), a governmental or outside organization might maintain a database — sometimes termed a Registry — of major impact evaluations. Such a Registry could be modeled, in part, on existing study registries – such as NIH’s well-established registry of medical trials (clinicaltrials.gov) and the Institute for Educational Sciences’ new [registry](#) of randomized controlled trials in education – with a few additional features designed to ensure study independence, as outlined below. The Registry would serve as a central repository for information on federally-funded impact evaluations. Specifically, such a Registry might include two types of items.

First, a project timeline – planned and actual – including: (i) date of issue of the Request for Contract (RFC) for each evaluation; (ii) date of contract award; and (iii) original and updated due dates for the Analysis Plan (if any), Interim Reports, and Final Reports.

Second, links to key project products:

- The original RFC.
- A redacted copy of the winning proposal.
- The official Analysis Plan (if any). Current “best practice” (see OMB, 2006, Guidelines 5.1, 5.2, and 5.3; Schochet, 2008) calls for writing a formal Analysis Plan prior to the study’s data collection. Such an Analysis Plan is partial protection against data mining and *ex post* (and therefore potentially invalid) estimates of impacts for multiple outcomes. For our purposes, a public Analysis Plan makes it difficult not to release unfavorable results from the natural analytic approach; i.e., if an analysis was specified in an Analysis Plan, the results of that analysis would be expected in the report. Publication of the Analysis Plan would serve as a counter to both funder and researcher bias.

If no Analysis Plan is required by contract, the Registry might provide a simple form: (i) intervention; (ii) proposed study population; (iii) type of study design (e.g., RCT or, if not, quasi-experimental with a brief summary of how the comparison group will be selected); and (iv) primary outcomes of interest; (v) main subgroups of interest; and (vi) a table shell in which the impact estimates will be displayed.

- The official project Final Report, and any interim reports.
- Any complementary documents released by the contractor (see below).

Publicly available timelines and Analysis Plans make it harder – for the funding agency or for the research contractor – to exclude key results from the study reports or to substantially delay their release.

This proposal for a Registry leaves open the key question of who would maintain it. Possible options include: (i) OMB/Office of Management and Budget, (ii) GAO/General Accountability Office, or (iii) an impartial non-governmental organization. As we move up this list, the ability of the organization to spur agency cooperation increases (e.g., OMB, and to a lesser extent, GAO, are well-positioned to foster agency compliance), but the degree of independence from political interference decreases (e.g., OMB is part of the Executive Office of the President, and could potentially delay release of a report whose findings are inconsistent with an Administration’s positions or agenda).

2. Allowing the contractor to release any draft evaluation report after a certain time period following its submission to the agency.

We propose this idea as an appropriate check on agencies’ discretion as the evaluation funder, to guard against potential misuse of that discretion to modify, delay, or prevent release of results in ways that might compromise study independence.

We propose that a contractor’s submission of a draft evaluation report start a disclosure clock. After a contractually-defined interval – perhaps six months – the contractor would be free to release the draft report and related materials, including text or results (except where doing so would raise human subjects or proprietary data issues). Upon notification by the contractor, the official Registry (see above) would then link to any such “Complementary Disclosure” reports or materials.

The right to Complementary Disclosure would encourage funders to release study reports promptly and avoid making inappropriate changes.

Given that the nation has paid for the evaluation, it is hard to see why the results should not be released to the public. If the funder wants to characterize the results differently than the contractor, the funder is free to release (in its own name, but not using the contractor’s name) its official report – presenting what it sees as the key findings and their interpretation.

Implementing such a Complementary Disclosure policy would likely require changes to contracts for “independent impact evaluations,” and may therefore involve modest additions or revisions to the FAR/Federal Acquisition Regulations language.

Closing Thoughts

We strongly agree with OMB’s goal of ensuring rigorous, independent program evaluations. This paper outlines two concrete ideas that we believe can help the federal government achieve that goal in a cost-effective manner.

Some References

- Klerman, J.A. Forthcoming. “Contracting for Independent Evaluation: Approaches to an Inherent Tension.” *Evaluation Review*
- Metcalf, Charles E. 2008. Threats to independence and objectivity of government-supported evaluation and policy research. *Journal of Policy Analysis and Management* 27 (4): 927–34).
- Office of Management and Budget. 2006. “Standards and Guidelines for Statistical Surveys.”
http://www.whitehouse.gov/omb/inforeg/statpolicy/standards_stat_surveys.pdf
- Reingold, David A. 2008. Can government-supported evaluation and policy research be independent? *Journal of Policy Analysis and Management* 27 (4): 934–41.
- Schochet, Peter Z. 2008. Guidelines for multiple testing in impact evaluations of educational interventions. Princeton, NJ: Mathematica Policy Research.
- U.S. General Accountability Office. 2010. Employment and Training Administration: Increased Accountability Could Improve Research Program. Washington DC. GAO-10-243.
- U.S. National Institutes of Health. ClinicalTrials.gov. <http://clinicaltrials.gov/>
- U.S. Institute for Educational Sciences. Registry of Clinical Trials.
<http://ies.ed.gov/ncee/wwc/references/registries/RCTSearch/RCTSearch.aspx>
- U.S. Office of Management and Budget. 2006. Standards and guidelines for statistical surveys.
http://www.whitehouse.gov/omb/inforeg/statpolicy/standards_stat_surveys.pdf (accessed November 5, 2009).
- U.S. Office of Management and Budget. 2009. “Increased Emphasis on Program Evaluation” (Peter Orszag; October 7).
http://www.whitehouse.gov/omb/assets/memoranda_2010/m10-01.pdf