

Rethinking Evaluation Methodology

Michael Scriven

Claremont Graduate University

Medicine, engineering, and evaluation have a highly significant common feature: they completely ignored the ban on evaluation that controlled the social sciences through most of the last century. What is it that doctors do in their core practice? They diagnose *disease* and *malfunction*, they *recommend* treatment, they encourage *good health*. What do engineers do? Amongst other tasks, they work out why the bridge *failed*, why the plane *crashed*, and how to correct the underlying *errors* and build *better structures* thereafter. And evaluators do the same with programs or policies or products or personnel—find the *best*, *improve* the *flawed*, report on the *worst*. It is the core nature of these essentially practical enterprises to be evaluative; they were not just describing or explaining or predicting how the world is, but trying to improve it. They simply didn't take seriously that the essential nature of science had to be 'value-free.' Are there any lessons to be learnt from the methods used by our fellow-practitioners in these highly evaluative disciplines?

A very important recent book contains an answer that makes it essential reading for evaluators interested in methodology. (It doesn't hurt that the story it tells, and the way it is written, is fascinating and entertaining as well as extremely educational; it's on the *New York Times*

top ten list at the moment.) The book is by a surgeon and it proposes a methodological technique for medicine that it bases in part on a study of engineering, in particular civil and aeronautical engineering. The methodology is proposed as a way to handle a serious problem in medicine—the problem that bad medical practices are killing more than 100,000 US citizens a year. (In fact, 99,000 of them are killed by avoidable infections picked up in intensive care units alone.) The solution proposed by the author was inspired by the technique that aeronautical engineers developed to reduce plane crashes due to pilot error, and that civil engineers use to reduce mistakes in the construction of large buildings. The book is *The Checklist Manifesto: How to Get Things Right* by Atul Gawande (2009).

Reading it made me go back and revise an article I first posted on The Evaluation Center's website in 2000, as part of the Checklist project that Dan Stufflebeam and I kicked off then and have nurtured a little since then, with much help from the many contributors of their own checklists and Lori Wingate's editorial supervision. The article is called "The Logic and Methodology of Checklists" and can be found at: <http://www.wmich.edu/evalctr/checklists/papers/index.html>. I've sent in the fourth edition of this and hope it will

be posted shortly. The first three editions began as follows: “The humble checklist, while no one would deny its utility in evaluation and elsewhere, is usually thought to fall somewhat below the entry level of what we call an evaluation methodology, let alone an evaluation theory. But many checklists used in evaluation incorporate a quite complex methodological theory...” (Scriven, 2000, p. 1) which I try to outline there.

My present thinking on this pushes a little harder; what I’d now say is that: (i) checklists are an important strand in evaluation *model* development—not quite the same as evaluation theory—led by Dan Stufflebeam’s start with the first distinct effort at an evaluation model, the CIPP checklist (2007), which he improved on all the way through to the mighty Program Evaluation Standards. (ii) They are also the structure for scores of evaluation *tools*, many of them at our web site, many more to be found elsewhere (e.g., in the thousand-odd pages of the invaluable *Modern Personnel Checklists* by Richard J. Melucci [1982]). (iii) The logic of checklists is quite complex—for example, there is a big difference between merely *taxonomic* ones (‘laundry lists,’ purely mnemonic), *procedural* ones (‘recipes,’ e.g., ‘you should make these nineteen checks in order to prevent infections in the operating room’), and *evaluative* ones (‘decathlons,’ e.g., ‘the quality of a training program is the weighted sum of its ratings on these eleven dimensions’). The logic of checklists is part of the logic of evaluation, although not the heart of it—which is the logic of how to combine empirical data with values to generate evaluative conclusions. Checklists are an essential part of evaluation, as valuable for the improvement of evaluation as they have been and will continue to be in medicine and engineering.

I hope readers will use, and find time to add to, our repertoire of these online, and to send criticisms and suggestions to me at mjscriv1@gmail.com.

References

- Gawande, A. (2009). *The checklist manifesto: How to get things right*. New York, NY: Henry Holt and company, LLC.
- Melucci, R. J. (1982). *Modern personnel checklists, with commentary*. Boston, MA: Warren, Gorham & Lamontby.
- Scriven, M. (2000). *The logic and methodology of checklists*. Retrieved February 2, 2010 from http://www.wmich.edu/evalctr/checklists/papers/logic_methodology.pdf
- Stufflebeam, D. L. (2007). *CIPP evaluation model checklist* (2nd ed.). Retrieved February 2, 2010 from http://www.wmich.edu/evalctr/checklists/cippchecklist_mar07.pdf