

Comments
on
Field Evaluation of Behavioral and Cognitive
Science-Based Methods and Tools for
Intelligence and Counter-Intelligence

Paul Lehner
The MITRE Corporation
September 2009

A Field Evaluation Short Story

- An apparent success
 - Study examined a current system for evaluating how much analysts should rely on a certain type of information
 - Developed a simple objective method to retroactively measure the real value of this type of information
 - Small but rigorous study performed
 - Results suggested that current evaluation system was very inaccurate
 - Replicating this study, to obtain definitive results, would have been very cheap
- But
 - Original sponsor of study had moved to a new position
 - New sponsor wouldn't release study unless approved by another organization
 - The organization that produces this type of information
 - The senior expert in that organization disagreed with our results
 - Never officially responded to request
- Study was not released

What went wrong?

- In the IC we need permission to release results
 - Given our domain, this is how it should be.
- Practitioner overconfidence
 - Research in other fields shows that many “tried and true” expert practices don’t work
 - Experience is a surprisingly poor discriminator of “what works”
 - Objective evaluations often show that common expert practices don’t work
 - Scientific results are often discounted because “they haven’t been there”
 - Practitioners are sure the science is wrong
 - Why would the expert reviewer approve release of a study that they “knew” was wrong?
- Bureaucracy changes
 - Field research usually requires a government “champion”
 - Who will probably move before study is complete
 - Why would new manager boldly push something they didn’t champion?
 - Or even believe?

There is a strong institutional bias against
obtaining or reporting negative results

It's not political

It's not “protecting turf”

Everybody is sincerely trying to do “what's right”

Some other sources of this bias

- Methodologist overconfidence
 - Methodologists believe in *their* expert practices
 - Quite sure that good field evaluation will validate their practices
 - So fielding of new methods shouldn't wait for full validation.
 - Once fielded, lots of validating field experiences will occur even if method doesn't work
 - By the time rigorous field evaluation would occur, new methods are already “tried and true” → see earlier story
- Urgent need
 - New methods/tools are fielded in response to true operational need
 - At most new method/tool receive preliminary testing
 - Once fielded, lots of validating field experiences ... → see earlier story

Why the bias against negative results is bad

- Most good ideas just don't work
 - Edison tested about 2000 prototypes before “discovering” the light bulb
- Even many ideas supported by lots of validating field experience (extrapolating from other fields)
- Expert Judgment based on field experience is a surprisingly poor discriminator of “what works”
 - This simple fact is the basis of “evidence-based” movements in policy, education, medicine, management, forecasting, analysis, sports, ...
- We should expect that most practices/methods/tools don't work, including many that are tried and true
- The bias against negative results makes it exceedingly difficult to find the small subset of things that really do work
 - We may be trapped in an illusion of success

My Suggestions

- For the Intelligence Community (field *then* test)
 - Its impractical to wait for full scientific validation before fielding new methods and tools
 - Needs are urgent and good science is slow
 - Foster a culture that is open to negative evaluation of current practices
 - Promote the need to ask “Does this really work?”
 - Encourage rigorous field evaluation after deployment
 - Reward discovery of negative results (e.g. funding for new methods/tools)
 - This workshop is an encouraging first step
- For the Scientific Community (test lessons learned process)
 - Most “evaluations” will always be based on field experience – get over it
 - Testing everything is simply impractical
 - The process of deriving “lessons learned from experience” should itself be an active research topic
 - Develop and rigorously test systematic processes for developing lessons learned
 - Qualitative Research Methods (as a systematic process) may be a good starting point
 - Help minimize the extent to which false lessons are learned from experience