Evaluation of research data services: What things should we evaluate? What are the proper metrics for judging success?

Charlotte L. Flynn
Syracuse University
charlflynn5@gmail.com

Problem & Solution
- An emerging role for information professionals in higher education is supporting researchers in managing data.
- Evaluation of data support services and tools is limited.
- This includes local project improvement and larger scale decision-making.

Solution: The TIER (Teaching Integrity in Empirical Research) project is a model for developing an infrastructure that builds local evaluation capacity and a research base that connects local activities to long-term outcomes. Urban and Trochim’s (2009) Systems Evaluation Protocol (SEP) is one such approach.

We need a better way of conceptualizing evaluation than asking, “What metrics should we use?”

This question presumes:
- Evaluation = measurement
- Uniformity of measurement and evaluation questions across projects evaluation as an afterthought.
- Measurement involves passive data collection that is the same for all projects at all phases of project development.
- Measurement does not involve collecting structured data in the context of an evaluation study.
- It is possible to measure project impact outside the context of an evaluation study and without taking into account the logic of how program activities lead to long-term outcomes.

This approach to evaluation of EDM interventions is a blunt instrument that lacks mechanisms for informing improvement or meaningful comparison of interventions.

If we hope to build an infrastructure of research data support, more robust evaluation is in order. In the professional program evaluation community, measurement choices depend upon context. Central considerations are stakeholders questions and the intended use of evaluation findings.

Example: TIER Teaching Integrity in Empirical Research
- A protocol for teaching students who are doing original analysis of quantitative data to document the steps of data management and analysis to facilitate replicability.
- Students learn documentation experientially in the context of learning research methods.
- For use in situations where:
  Students do research projects using real data to engage in statistical inquiry (not prescribed analysis using assigned data).
  Students receive instructor feedback and mentor over the course of their project.
  Developed and implemented collaboratively over the past 10 years by Richard Ball, an economics faculty member, and Norm Medeiros, a librarian, at Haverford College.

TIER Logic Model
- Inputs: TIER protocol (procedures for assembling data to document analysis & data management)
- Activities: Student research outputs are documented for replicability.
- Outputs: Students gain more in-depth understanding of research methods.
- Short-term outcomes: Students adopt documentation practices.
- Medium-term outcomes: Professional research is documented.
- Long-term outcomes: Professional research is of higher quality.

TIER Pathway Model
- Students do research projects using real data to engage in statistical inquiry (not prescribed analysis using assigned data).
- Students receive instructor feedback and mentor over the course of their project.
- Developed and implemented collaboratively over the past 10 years by Richard Ball, an economics faculty member, and Norm Medeiros, a librarian, at Haverford College.

Factors to take into account before deciding what to measure
- Description
  - Key activities/elements participants experience that contribute to outcomes.
  - Informed by differences in information or values that shape stakeholder perceptions of the program.
- Phase of development
  - Life cycle analysis
- Theory of change
  - Expresses stakeholder understanding of how program brings about change
- Logic model = pathway model
- Program-system links
  - Evidence that links local activities to long-term outcomes.
- Evaluation scope
  - Feasibility, credibility, accuracy, usefulness

Phases of Development
- Programs have life cycles
- Less stability early on, trial and error
- Tend to stabilize as they mature
- Program phases: initiation, development, stabilization, dissemination
- Evaluation should have a lifecycle
  - Early on: rapid feedback for program improvement
  - Once program stabilizes and is disseminated, evaluation should entail greater social science rigor.
- Over the course of a program, evaluation questions attend first to process, then change, then comparison, then generalization.

References

Teaching Integrity in Empirical Research
http://www.haverford.edu/TIER/