

Assessment Burden: When does the Burden of Assessment and Evaluation Outweigh the Benefits?

James P. Van Haneghan
University of South Alabama
jvanhane@usouthal.edu

Presentation at the National Evaluation institute,
Wilmington, NC
October 9, 2008

Context

- Experiences with four types of evaluation systems
 - College of Education Assessment System Implementation
 - Experience with a criterion referenced testing system in a large school system
 - Evaluation of educational programs in the context of schools
 - Work with our faculty evaluation system

Problem of Assessment Burden

- Assessments that add to the work that program participants and stakeholders otherwise carry out as part of their program add burden
- Question is when is that burden too much?
- What are the conditions that suggest that the added value of the evaluation does not merit the burden required in an assessment system?

Program Evaluation Standards

- **F1 Practical Procedures** The evaluation procedures should be practical, to keep disruption to a minimum while needed information is obtained.
- **F2 Political Viability** The evaluation should be planned and conducted with anticipation of the different positions of various interest groups, so that their cooperation may be obtained, and so that possible attempts by any of these groups to curtail evaluation operations or to bias or misapply the results can be averted or counteracted.
- **F3 Cost Effectiveness** The evaluation should be efficient and produce information of sufficient value, so that the resources expended can be justified
- **U3 Information Scope and Selection** Information collected should be broadly selected to address pertinent questions about the program and be responsive to the needs and interests of clients and other specified stakeholders
- Note new standards modify F2 and F3. Add one on project management in feasibility section

The Joint Committee on Standards for Educational Evaluation (James Sanders, Chair) (1994). *The program evaluation standards*. Thousand Oaks, CA: Sage.

Changing Face of Evaluation

- Evaluation in the past was something that was done to a program or organization
- **Now Evaluation is something that is intrinsic to most organizations**
- **Because more people are involved in evaluation, need to consider the burden on their time**

Metaphor of Cognitive Load: Organizational Assessment/Evaluation Load

Paas, Renkl, and Sweller (2003)

Intrinsic Cognitive Load: Task intrinsic complexity

Extraneous Cognitive Load: Unnecessary additions to cognitive load that do not help learning but add to load of instruction

Germane Cognitive Load: Load associated with processes of schema development

Cognitive Organizational Load of Evaluation

- Use the metaphor of cognitive load theory to examine how assessment systems work within organizations
- Intrinsic cognitive load of the assessment system
 - Cognitive difficulty of interpreting system
 - Cognitive complexity of the measure
 - Example of Rasch model-based test that district could not interpret simply was dropped after a year of use as part of an incentive package
 - Coordination of classroom and program level assessments

Organizational Cognitive Load

- Extraneous Cognitive Load
 - Data from assessment is not in a useable form.
 - Assessment could be set up for automated or more easily interpretable scoring but is not
 - Parts of the assessment are useful in the evaluation process, but other elements are not.
 - There are too many different types of scores that need to be interpreted from too many different angles and this adds to the load associated with interpretation

Organizational Cognitive Load

- Germane Cognitive Load
 - Load associated with collaborative examination of scores to make sense of the test information
 - Load available to learn from assessment and evaluation information
 - Reduction in extraneous and perhaps task intrinsic cognitive load can influence germane cognitive load available for purposes of using assessment information

Challenge

- Design system that reduces extraneous load and leaves some ability to interpret and use evaluation information (germane load)
- Load is only one issue

Multiple Accountability Demands

- Demands for data from a variety of sources
 - Local
 - State
 - National
 - Accreditation Agencies
- Perception sometimes greater than the reality (find ways to serve multiple purposes)
- When setting up a system need to find ways to avoid building separate systems for each accountability agency
- If setting up several different data collection systems for several different stakeholders than may be duplicative and may overburden users of system

Implementation of Electronic Assessment System in College of Education

- 160 indicators of standards need to be assessed (across a program)
- Faculty may assess only a few (foundations courses) or many (student teaching)
- High intrinsic load to the task? Maybe
- We surveyed faculty in terms of the beginning implementation
- Found that there was not much change in the number or kinds of assessments they were using.
- Primary added burden was putting assessment into the electronic system and we have tried to lessen that burden by tying it into the e-learning system

Implementation of E-Assessment System Continued

- Lack of change in assessments not surprising
 - Should be teaching indicators of standards in courses anyway
 - May be validity issues down the line (to avoid more work assessments faculty may sometimes use existing but less valid assessments—we are now to the stage where we will be looking at this by sampling assessments each year)

Perceptions of e-assessment system

- Some view new system as add on that uses up time but does not accomplish anything
- We are one year in and will be using the data we have collected for program feedback.
- Interestingly, the state that required the system will not be collecting the data, but only maybe asking for it on site visits or at accreditation time
- We need to make it useful to change the perceptions that it is not just an added burden in their already busy professional lives
- Perhaps change the perceptions as they relate to the value of evaluation can change the perceived burden of the assessment and evaluation system

Two Views of Evaluation

- Add on—something that adds work
 - Recent evaluation of professional learning teams weakest part of process is evaluation of what they are doing
 - This perception leads to perception of burden
- Part of the process of continuous improvement—something that is intrinsic to the organization
 - Reflection on data and progress part of the process of continuous improvement
 - This perception leads to the belief that evaluation is just part of the job and does not create extra burden

Evaluation Use and Feedback

- Perception of burden when there is no information fed back from data
- Even if system is potentially useful, the system may be viewed as burdensome because there is no information about how assessment is used
 - Student interview assessments. Could not get the district feedback quickly enough, so decided to stop doing

Burden as Unintended and Negative Change in Evaluand

- CRT system
 - Two weeks a quarter on CRT
 - 8 weeks in year lost to instruction preparing for and taking CRTs
 - Add time for standardized assessment preparation
 - Assessment system changed what people do
 - Could be in good way (more accountable teach better, but evidence from recent dissertation found no change in state test pass rates after implementation)
 - But sometimes in an effort to “game” the system teaching changes to test preparation (which changes instructional programming in a negative way)
 - If program of testing does not lead to positive change, may view testing system as just another burden on time

So What are the Signs of Assessment Burden?

- Cognitive Organizational Overload
 - Examine system for cognitive load
 - Try to remove extraneous load
- A unsystematically designed assessment system
 - Separate systems for each accountability agency
 - Lack of coordination of those systems
 - Perception that cannot do job because of so many requests for data and information
- View of evaluation as an add on
 - Assessment and evaluation will always be seen as a burden if the culture of the organization does not see it as a part of the way one does business.
- Collection of data without feedback about what it means and for whom
- Lack of knowledge of how to use data (either because of other demands or lack of training on how to use data)
- Changes in the activity that the assessment was meant to measure that refocus it in a nonproductive way (note can change it in productive ways when there is accountability, but much change is toward teaching to test)
- Evaluation Standards serve as a useful Guide

References

- The Joint Committee on Standards for Educational Evaluation (James Sanders, Chair) (1994). *The program evaluation standards*. Thousand Oaks, CA: Sage.
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*. 38, 1-4.