Evaluation is not just for NSF: Tools and strategies for using evaluation to empower, learn and grow successful programs (and get funded)

Jason Ravitz
Evaluation by Design
jason@evaluationbydesign.com

Abstract

Over the years the Society for Information Technology and Teacher Education (SITE) has brought many creative and ambitious educators and organizations together to create and share best practices. There is still a question, however, about how we really know when we have best practices, and once we are convinced, do we know how to share evidence with each other? This paper provides access to insights, guidance and tools for educational leaders and organizations based on years of educational program evaluation and research. Some of this is based on work for the National Science Foundation (NSF), but much of it applies more broadly. The tips are intended to help those who are in the process of developing offerings and want to make evaluation work for them. The free resources are helpful, but what matters most is the ideas they carry that can help you design an evaluation that will help you learn, improve and grow more impactful (and fundable) programs.

Purpose

Investigating and developing new educational technologies and programs is fraught with challenges. There are some strategies, however, that seem to work over and over. These are basic ideas for how to deliver professional development courses and workshops. In terms of theory, what we are seeing follows years of research and development in the areas of change management (Ravitz & Hoadley, 2005), project-based learning (Ravitz & Blazevski, 2014), teacher professional growth (Ravitz, 2002), technology- supported assessment (Ravitz, 2004) and program and empowerment evaluation (Fetterman & Ravitz, 2017). While some of these ideas may be tried and true, new thinking about evaluation capacity building -- and building these ideas into innovations in online tools -- can help participants produce more worthwhile evaluations.

The common "best practice", that each of the following tools carries forward, is using evaluation design thinking to produce more credible evidence and impactful programs. This paper reviews tools that allow you to accomplish the following, in order:

- A tool for tracking evaluation progress
- An evaluation planning worksheet for individual evaluations (or cohorts)
- An empowerment evaluation exercise to promote buy-in
- A shared rubric feedback model (several iterations) for PD or instruction

Honorable mentions that represent more current directions include application of these ideas to a 21st Century Skills survey and rubrics, and a system for linking skills with performances for career-technical education, as well as automated scoring and reliability testing. There has been considerable cross-pollination. Ideas have emerged and re-emerged as useful, resulting in footnotes that sometimes point backwards

Many colleagues have helped keep a spotlight on inspirational cases where evaluation incorporated into design thinking has results in substantial opportunities. Telling each other these stories is important. As Tiffany Decker, formerly of MIT (now at Columbia University) reported from working with professors before a summer camp: "Once they saw how useful the data were, they asked for more." A team leader from Google volunteered this insight: "I'd always considered evaluation as an afterthought, but I've come to learn that a good evaluation strategy is a central component of an effective research and development cycle. Moving forward, I can't imagine developing new programs without it."

Tracking Evaluation Progress

A key best practice is keeping track of all the ways evaluation allows programs to gather better evidence, tell more credible stories, and make sure they are heard. Tool #1 helps solve the problem of how to measure evaluation as its own programmatic effort, indicating if progress is being made on numerous teams, in a way that can also help program management.

What does progress look like?

Stage 1	You have some measures or you are talking about evaluation already
Stage 2	Metrics and methods are mapped to evaluation **questions**
Stage 3	**Analyses** help answer key evaluation questions
Stage 4	Results are used to **influence** the program
Stage 5	Findings are **reported** to key stakeholders and leadership (or more widely)
Stage 6	Evaluation is **institutionalized** into new programs (budget, staff, etc.)

Source: Presentation to American Evaluation Association (Ravitz, 2018).

Evaluation Worksheet

Another key best practice is setting time aside for careful and strategic evaluation planning. Without this, many teams will essentially stall out at Stage 1 (see above), declaring "we have metrics" and so "mission accomplished". Going through a systematic planning process makes sure you are asking yourself questions you need to ask to make your study can still fly, in the face of stronger headwinds. This evaluation tool originated at Google while working with CS and STEM education nonprofits and educators, and has been incorporated for use in nonprofit evaluation capacity building cohorts (National Girls Collaborative Project, 2017), workforce training, a mobile app to learn coding, a graduate school self-study using empowerment evaluation (Fetterman & Ravitz, 2017), and many others. The evaluation planning worksheet helps teams create detailed plans in four stages:

1. describe the program -- including its purpose, who it serves and how it works
2. define the evaluation -- its particular purpose, focus, audience and context
3. plan the evaluation -- with specific questions, procedures, analyses and reporting
4. manage the evaluation -- with defined responsibilities, time frames and budget

Current: tinyurl.com/evalworksheet-2019   Earlier: tinyurl.com/evalworksheet-Google

To help make this resource useful, it includes a video (Ravitz, 2016) and slide deck clarifying the context and intended use. Intangible benefits or working through this process with an evaluator (or evaluator's mind) include clarifying the stage of development (e.g., are you even ready for an impact study?); tightening up how everyone talks about your work; focusing on what matters most; and, making sure you can get to your most valuable sources of evidence and knowledge.

Empowerment Exercise
Another best practice is getting "buy-in" or making sure everybody is on the same page. Empowerment evaluation (EE) builds capacity and fosters program improvement by helping programs evaluate themselves in a way that gives all key stakeholders a voice. As noted above, this can be combined with the worksheet to great effect. Key concepts include having critical friends, cycles of reflection and action, and a community of learners. These all can enhance efforts to produce measurable outcomes and demonstrate accountability. Increasingly, we also see how technology can play a key role. Essential EE steps that can be done online include helping to do the following:

1) establish a mission
2) take stock – systematically rate and discuss priorities and performance
3) plan for the future – establish goals and strategies to accomplish objectives

Online (or blended) use of this tool, along with online polls and videoconferencing, has been demonstrated successfully with National Science Foundation evaluators (Haynie, 2017), broader CS and STEM education evaluators (Fetterman, Ravitz & Haynie, 2018), and a graduate evaluation course (Fetterman & Ravitz, 2017). This process is best when managed by a skilled facilitator who can maintain transparency and help balance all perspectives; the numbers are non-deterministic, for example, with trading or combining votes often encouraged. After a group discussion, a dashboard can be used to set up and monitor quarterly and annual goals, for example. The evaluator is a critical friend and facilitator, working with program staff and stakeholders to guide them on the path they set forward. Check out the online EE tool here: tinyurl.com/eeblank

Collaborative Feedback (shared rubrics)
The final best practice is sharing feedback to advance and link both individual and organizational learning (Kimmerle, Cress & Held, 2010). Evaluating educational technology innovations has been an exciting field since the early days of the web, and longer, per Saettler (1990) or Cuban (n.d.). As much as the landscape has evolved, the challenge (at least in online learning) has remained. How do we know when learners (or organizations) have demonstrated expertise or produced quality work? From the beginning, it was clear answering this required a new kind of interactivity at the project level (Ravitz, 1997) and to promote more meaningful individual learning. (Ravitz, 1998)
There have been multiple iterations of this "shared rubric" strategy with a variety of discussion - review - database tools tested out and applied to educational websites (Ravitz & Lake, 1996), handheld assessment of group collaboration (Yarnall, et. al. 2008), project-based learning workshops (Ravitz, et. al., 2004) and evaluation planning courses (Fetterman & Ravitz, 2017). In all these cases, with illustrations below, technology tools and processes supported learning in community with peers and instructors. Identifying and talking about quality work engendered persistent and meaningful use of shared criteria with the associated form of critical thinking (Case,

2005).

        The first re-usable online review form for professional development purposes emerged from the NSF-funded National School Network Testbed II project for the Online Internet Institute (Ravitz & Serim, 1997) and was reused at LeMoyne College (Ravitz & Lake, 1996). The next test involved an extended 6-week online course that collected valuable, but relatively unstructured, feedback (Ravitz, 2004). The Buck Institute for Education created an online database with rated examples of projects and shared rubrics for teacher professional development sessions (Ravitz, et al., 2004) and there was extensive use of rubrics with three rounds of project design reviews in our West Virginia study (Hixson, Ravitz & Whisman, 2012). In recent years, a graduate level evaluation course incorporated self-, peer- and instructor-level reviews with an online shared rubric to guide development of projects, and facilitate instruction and grading (Fetterman & Ravitz, 2017).

        Finally, research on school technology coaching has harnessed a shared rubric along with surveys representing another step forward for the West Virginia-derived measures (Bakhshaei, 2019). Looking ahead there will be new tools emerging that incorporate automated dashboards from these items, as well as tools for automated rater agreement tests, as developed for observing nursing school simulations (Sweeney, et al., 2020).

Conclusion

        Regardless of which of the above tools are used, the methods they convey are likely to have clear benefits for programs. There are many reasons to evaluate, as noted in this video by Jeanne Century (2015), Director of Outlier Research & Evaluation at the University of Chicago, in this video from the Google office in Cambridge, MA. Some of these benefits are more directly linked to communication needs involving funders. Others are more internally-valuable to make sure programs shape how their own story is told, or collect stories from participants more easily.

        Going forward the vision is to put as many of these practices into use as possible, and in new ways. One experiment involves collecting opportunities to learn, using the West Virginia surveys, and linking these to performance assessment reviews in career-technical education (Serim & Ravitz, 2020). As far as we know, this would be the first time linking validated surveys from students and teachers about what is taking place in classrooms to actual examples of student work with multiple ratings.

        What this paper has tried to show is that technology, and the transparency and flexibility it provides, can help you to formalize and facilitate evaluation processes in order to streamline your work, increase your credibility and utility, and produce better and more fundable programs. Not all of these strategies are directly tied to funding, or meeting the requirements of funders, although there are tools for that too. No one should feel that evaluation is a burden that hurts program operations. Designing for evaluation as outlined here can make sure the resources and efforts that go into this work are worthwhile.

REFERENCES

Bakhshaei, M. (2019). Sharing tools for measuring impactful technology use. Blog Post: November 19, 2019. https://digitalpromise.org/2019/11/19/

Case, R. (2005). Moving critical thinking to the main stage. *Education Canada, 45*(2), 45-49.

Century, J. (2015). Video. *Why evaluate?* Interactive session for Global RISE Summit attendees at Google in Cambridge, MA. https://www.youtube.com/watch?v=NNcK2zEClLs

Cuban, L. (n.d.). Blog. Larry Cuban on School Reform and Classroom Practice https://larrycuban.wordpress.com/

Fetterman, D., Ravitz, J. & Haynie, K. (2018). Using Google Sheets and Forms and Zoom to Remotely Facilitate an Empowerment Evaluation Exercise. https://tinyurl.com/zoomeval

Fetterman, D.M. & Ravitz, J. (2017). A Google-enhanced empowerment evaluation approach in a graduate school program, Chapter in *Collaborative, Participatory, and Empowerment Evaluation Approaches: Stakeholder Involvement Approaches*, Guilford: NY, 105-117

Haynie, K. (2017). Expanding our evaluation practice: Community development in 2017. Computer Science Outcomes Networked Improvement Community (CSONIC) blog. https://csonic.org/2017/11/03/

Hixson, N., Ravitz, J. & Whisman, A. (2012). The impacts of extended PBL professional development on West Virginia teachers. Charleston, WV: West Virginia Department of Education. http://eric.ed.gov/ED565466.pdf

Kimmerle, J., Cress, U., & Held, C. (2010). The interplay between individual and collective knowledge: technologies for organisational learning and knowledge building. *Knowledge Management Research & Practice, 8*(1), 33-44.

National Girls Collaborative Project (2017). CS OPEN: Computer Science Outreach Program Evaluation Network. https://ngcproject.org/sites/default/files/ngcp_cs_open_summary.pdf

Ravitz, J. (2018). Learning from evaluations and the programs we fund. In Eric Barela (Chair), *Evaluating social impact at the intersection of the technology and nonprofit sector*s. American Evaluation Association. Atlanta, GA. November 2, 2018.

Ravitz, J. (2016). Edu on Air: Empowering leaders: Introducing best practices in program evaluation https://www.youtube.com/watch?v=Ets8-3qKv1E

Ravitz, J. (2004). A doorway to new tools and practices: Supporting teacher education, research, and development with an online netcourse. In N. Sabelli and R. Pea. (Eds.). *Uniting people, technology and powerful ideas for learning. Six years of knowledge networking in learning sciences and technologies,* 94-101.

Ravitz, J. (2002). The enthusiasm effect, the reality effect and other things to know when evaluating

professional development**.** *National Educational Computing Conference*. Seattle, WA.

Ravitz, J. (1998). Emerging practical theory of assessment: Interactive portfolios. Third Annual Conference on Computer-Supported Collaborative Learning (CSCL). In G. Stahl (Chair). *Workshop for collaborating on the design and assessment of knowledge-building environments*. Palo Alto, CA.

Ravitz, J. (1997). Evaluating learning networks: A special challenge for Web-based Instruction? In B. Khan (Ed.)., *Web-based Instruction*. Englewood Cliffs, NJ: Educational Technology Publications, 361-368.

Ravitz, J. & Blazevski, J. (2014). Assessing the role of online technologies in project-based learning. *Interdisciplinary Journal of Problem-based Learning, 8*(1).

Ravitz, J. & Hoadley, C. (2005). Supporting change and scholarship through review of online resources in professional development settings. *British Journal of Educational Technology, 36*(6), 967-974.

Ravitz, J. & Lake, D. (1996). An authentic tool for teachers: The OII WWW site evaluation form. *Proceedings of the FSU/AECT Distance Education Conference*, Tallahassee, FL. June 20-23.

Ravitz, J., Mergendoller, J., Markham, T., Thorsen, C., Rice, K., Snelson, C., & Reberry, S. (2004). *Online Professional Development for Project Based Learning: Pathways to Systematic Improvement*. Novato, CA: Buck Institute for Education. Available: eric.ed.gov/ED486531.pdf

Ravitz, J., & Serim, F. (1997). Summary of first year evaluation report for the Online Internet Institute. In *Proceedings of the Edward F. Kelly Evaluation Conference*. State University of New York, Albany. Retrieved from: tinyurl.com/oii-evaluation

Serim, F. & Ravitz, J. (2020). NextREADY Learning Model – A Platform with Evidence-Based Tools for Career Readiness. Paper presented at SITE 2020, New Orleans, LA (virtual). April 9, 2020.

Sweeney, N.L., Rollins, M.C., Gantt, L., Swanson, M. & Ravitz, J. (2020). Development and reliability testing of the Sweeney-Clark Simulation Evaluation Rubric©. *Clinical Simulation in Nursing, 41*, 22–32. https://doi.org/10.1016/j.ecns.2019.04.002

Yarnall, L., Penuel, W., Ravitz, J., Murray, G., & Means, B .(2003). Portable assessment authoring: Using handheld technology to assess collaborative inquiry. *Education, Information & Communication, 3(*1), 7-55