

Rethinking rigor: Increasing credibility and use

Jewlya Lynn (Spark Policy Institute) & Hallie Preskill (FSG), January 2016

Traditionally, evaluation has focused on understanding whether a program is making progress against pre-determined indicators. In this context, the quality of the evaluation is often measured in part by the “rigor” of the methods and scientific inquiry, such as the use of experimental and quasi-experimental methods. However, evaluations of complex systems change strategies or adaptive, innovative programs cannot use this same yardstick to measure quality: experimental designs are hard to apply when a strategies’ success is not fully defined up front and depends on the strategy being responsive to the environment.

Below is a new definition for rigor, one that applies in any setting, but particularly in complex, adaptive ones. The definition is based in the idea of balancing four criteria, each of which is critical, though some will be a greater priority in some settings than others.

1. Quality of Thinking

The quality of thinking that guides an evaluation’s design and implementation should...

- Engage in deep analysis.
- Focus on patterns, themes, and values, e.g., systems thinking.
- Seek alternative explanations and interpretations,
- Situate findings within the research literature (theories, frameworks, models), and
- Look for outliers that offer different insights (perspectives and experiences).

2. Credible and Legitimate Claims

Ways of ensuring the trustworthiness of data include considering...¹

- Credibility, or your confidence in the 'truth' of the findings.
- Transferability, the ability to show the findings have applicability in other contexts.
- Dependability, being able to demonstrate the findings are consistent and could be repeated.
- Confirmability, which is the degree of neutrality or the extent to which the findings of a study are shaped by the respondents and not evaluator bias, motivation, or interest - whose voices are serving to create that credible frame.

3. Cultural Context and Responsiveness

Ways of ensuring responsiveness to the cultural context include...²

- Evaluation questions that reflect stakeholders' values and context.
- Outcomes that reflect stakeholders' definitions of success, which may necessitate incorporating multiple definitions of success.
- Culturally-appropriate data collection methods, procedures, and instruments.
- Diverse stakeholders, who help interpret data and generate insights about the implications, leading to recommendations generated through an inclusive process.
- Conclusions that reflect stakeholders' values and perspectives of the quality and effectiveness of systems/programs.

4. Quality and Value of the Learning Processes

The quality of the learning process that includes...³

- Engaging the people who need the information when they most need it,
- In a way that allows single and double loop learning (reflection, dialogue, testing assumptions, and asking new questions), and
- Contributes to decision-making that help improve processes and outcomes.

Balancing the Criteria

The concept of balancing the four criteria is at the heart of this redefinition of rigor. Regardless of its other positive attributes, an evaluation of a complex, adaptive program that fails to take into account systems thinking will not be responsive to the needs of that program. Similarly, an evaluation that fails to provide timely information for making decisions lacks rigor even if the quality of the thinking and legitimacy of the claims is high.

References:

¹ Lincoln, YS. & Guba, EG. (1985). [Naturalistic Inquiry](#). Newbury Park, CA: Sage Publications.

² CDC. (2014). [Practical Strategies for Culturally Competent Evaluation](#). Atlanta, GA: US Dept of Health & Human Services.

³ Preskill, H & Torres, RT. (1999). [Evaluative Inquiry for Learning in Organizations](#). Sage.

