



Innovations in Program Evaluation to Improve School Leadership: How Can We Perform One?

Richard DLC Gonzales
Inno-Change International Consultants, Inc.

How to cite

Gonzales, R. D. L. C. (2020). Innovations in Program Evaluation to Improve School Leadership: How Can We Perform One? *Asia Pacific Journal on Curriculum Studies*, 3(1), 22-28. <https://doi.org/10.53420/apjcs.2020.3>

Abstract. This paper describes a number of innovative approaches in conducting a program evaluation for school leadership. It also discusses the factors driving program evaluation innovations and the factors enabling innovations to occur in program evaluation. Furthermore, the paper defines what qualifies as innovations in program evaluation and the criteria that should be met to be considered innovative. The article ends with a detailed discussion on some of the program evaluation innovations and how to apply program evaluation innovations.

Keywords: program evaluation, innovations in programs evaluation, innovative approaches

Introduction

Evaluation is broadly defined as the “examination of the worth, merit, or significance of an object” (Chen, 2015; Scriven, 1998). Rossi, Lipsey, and Freeman (2004) and Schwandt (2015) also defined evaluation as the systematic use of scientific approaches to assess the design, implementation, improvement, or outcomes of a program. The World Health Organization (2013) also defines evaluation as an assessment that follows a systematic and impartial approach to an activity, project, program, strategy, policy, topic, theme, sector operational areas, and institutional. The WHO based its evaluation policy on UNEG definition of evaluation (UNEG, 2012b).

Program is defined as “any set of structured activities with particular short- or long-term goals”. Programs are usually supported with human and financial resources to achieve specific and intended results (U.S. Department of Health and Human Services Centers for Disease Control and Prevention, 2011). Programs may include organized activities such as educational services, social interventions, media advocacy campaigns, service provisions, public policies, research projects, and training programs. Programs are also defined as high-effect activities based on packaged resources that delineate and define a series of specified undertakings.

Program evaluation is a systematic method for gathering, analyzing, and using data and information to examine the worth, merit, or significance of a project, policies, and programs (Chen, 2015; Mertens & Wilson 2012; Stufflebeam & Shinkfield, 2007). Generally, program evaluation usually occurs in an organized setting that looks into the effectiveness and efficiency of any set of organized activities and their impact on its intended users (Pancer & Westheus, 1989; Scriven, 1998). In many cases, program evaluation's primary purpose is to demonstrate program effectiveness to funders and document program development and activities to ensure successful replication (Olnet & Barnes, 2013; Shadish, Cook, & Leviton, 1990). Program evaluation also purports to document accomplishments. Hence, evaluation looks at programs and project objectives and determines whether they have been achieved, judges the worth of on-going programs, decides upon the usefulness of new or recurring programs or projects (Boulmetis and Dutwin, 2005, p. 3). Furthermore, it can also be viewed as a systematic approach in assessing the current state of a program or project and identify any discrepancies of the current state and what the intended state is supposed to be. Overall, program evaluation is also regarded as useful to any organization in several ways (Preskill, Parkhurst & Juster, 2014; Wholey, Hatry & Newcomer, 2010; WHO, 2013).

In schools and educational settings, programs are developed, introduced, reinforced, and evaluated more often. A program evaluation in schools is more frequently focused on assessing its students' achievements and educational interventions' efficiency

(Burkhauser et al., 2012). A program evaluation is an indispensable tool for school managers and teachers to find ways and approaches to strengthen their programs' quality and improve students' learning outcomes (Wright & Wallis, 2019). Among schools, program evaluation responds to fundamental questions about a program's effectiveness and gathers data that can be used to improve any school or educational program. Wiseman (n.d.) argues that recognizing the value of program evaluation in many sectors and the changes occurring in the field where methods and strategies have been evolving, and the prominent role of big data, it is necessary to respond to these changes by considering innovations in the program evaluation process. Innovation as suggested by Perrin (2002) is defined as a novel means of performing things better or differently, often by quantum leaps instead of incremental gains. This definition is anchored on the European Commission's Green Paper on Innovation (1995), which defines innovation as a successful production, assimilation, and exploitation of novelty in the economic and social spheres. Perrin (2002) further argues that innovations can be on a small scale, such as a school principal trying a new way to monitor teachers. But it can be on a large scale, like determining a new program approach to implement online learning or competency-based curriculum.

Hence, in this paper, innovations in program evaluation methods and approaches are presented, considering the need for program evaluation to respond to higher standards, fast-changing environment, and technology impact. Specifically, this paper describes a number of innovative approaches in conducting program evaluation to improve school leadership. It also looks into the factors driving innovations in program evaluation of school leadership and the factors enabling innovations to occur in program evaluation.

Why innovation in program evaluation necessary in improving school leadership?

Schools are progressively taking innovative approaches to improve school leadership. Efforts to improve school leadership are critical in improving student outcomes (Burkhauser, Pierson, Gates & Hamilton, 2012; Pancer & Barnes, 2013). Undeniably, principals and other school leaders and managers are essential in schools' effectiveness and student learning. School leaders monitor and evaluate the performance of schools and delivery services. Various schools' stakeholders are interested in knowing whether programs in improving school leadership would show student achievement results (Burkhauser et al., 2012). However, the conventional and standard approaches in evaluating these school leadership programs may not significantly capture student achievement outcomes. Therefore, schools and educators are concerned with having program evaluations that respond to school leadership programs' changing needs. The desire to achieve global targets in education and the use of technology has necessitated schools' stakeholders to be more innovative in evaluating student achievement due to school leadership programs. As technology is moving forward, opportunities for innovation in evaluating school leadership is opening up.

Essentially, innovation signifies pioneering something new into any approach in program evaluation. Evaluators believe that the current tools to collect and analyze data for program evaluation are overwhelming with a large amount of data from disparate sources, incompatible with one another requiring time-consuming manual entry to put data to use, inconsistent in their ability to report data and the level of details provided, and too slow to provide information in time to modify instruction in meaningful ways (Olney & Barnes, 2013; Melinda & Bill Gates Foundation, 2014). Because of these barriers, evaluators perceive that data are often siloed and challenging to work with, inflexible, and unable to track programs' progress over time.

What factors are driving innovation in program evaluation of school leadership?

In the study done by Burkhauser et al. (2012) on evaluating efforts to improve school leadership, they recognized that evaluation is critical for effective use of resources, but it poses challenges. They provided recommendations for evaluators and policymakers as follows: a) give time for improvements to be seen, mainly because the indirect relationship between principals and students may affect the immediate changes in students' learning outcomes; 2) employ several evaluation measures that can determine the differences in student characteristics and outcome measures, principal characteristics, and school contexts; and 3) interpret the evaluation findings carefully before applying them to other school, districts, or states.

Using a single program evaluation method is not sufficient anymore (Patton, 2002; Rossi, Lipsey & Freeman, 2004; Scriven, 1998). Yarbrough, Shulha, Hopson, and Caruther (2011) argues that the complexities of determining outcomes necessitate multiple and multi-level methods because outcomes are more challenging to measure. They count more than the activities of a program (Wright & Wallis, 2019). Similarly, schools' stakeholders and civil societies demand accountability from school leaders, which single method and traditional evaluation approaches may not capture more efficiently. Innovations in program evaluation in school leadership are also required because of flexible and faster data collection and program evaluation. Lastly, innovation is imperative to avoid cognitive bias and any other social biases.

The boom of ICT, the rise of “big data,” and having more mature schools’ stakeholders and civil societies also enable innovations to place in program evaluation. Undeniably, the technology and existence of the concept of “big data” transform the way evaluators conduct program evaluations.

What are some examples of innovative program evaluation?

As a rule, whether evaluators use the traditional or innovative approach in program evaluation, it is still essential that they adhere to the four primary phases of program evaluation – planning, implementation, completion, and dissemination, and reporting (Boulmetis & Dutwin, 2005; Chen, 2015; Pancer & Westhues 1989; Wholey et al., 2010). Stufflebeam and Shinkfield (2009) emphasized that these phases should also complement program development and implementation stages. Each of these basic program evaluation phases has specific issues, methods, and procedures (Boulmetis & Dutwin; 2005; Patton, 2002).

This section of the paper discusses at least six innovative approaches to program evaluation. The techniques discussed in this paper are considered innovative because they offer significant impact or clear potential to change, merely a better, faster, and cheaper way of data gathering, and can be replicated and potentially scalable across different contexts (Mertens & Wilson 2012). They are also considered innovative because they introduce something novel, a new idea or method, or the use of a device. Big data and data exhaust in gathering data through social media or geo-tagging qualify them as innovative approaches.

However, while approaches are regarded as an innovation, they still use and borrowed ideas and methods from the traditional program evaluation (Chen, 2015; Olney & Barnes, 2013; Preskill et al., 2014; Schwandt, 2015). They are also benchmarked from other disciplines and professions. Still, some of these approaches are a patchwork of commonly used evaluation techniques but used more systematically.

These approaches are selected as examples of innovative program evaluation because they contribute, at least indirectly, to:

1. Increased independence – implies the freedom from undue political influences of schools’ stakeholders and/or organizational and community pressure;
2. Credibility – refers to the degree of transparency of the evaluation process, the quality of outputs, the balances of reporting successes and failures in school leadership, and the level of participation of the schools’ stakeholders and beneficiaries; and
3. Utility – refers to the relevance and usefulness of the findings to improve, enhance, and expand school leadership scope.

Table 1
Innovations in Program Evaluation for School Leadership

Method	Description	Innovative Element	How and When to Best Use it?	Advantages and Disadvantages
1. Multi-level mixed evaluation method	This approach includes the deliberate, massive, and creative use of mixed (quantitative and qualitative) methods on multiple levels for complex evaluation. It is used mainly for instructional delivery systems.	Provides a paradigm shift in evaluation methodology which focuses on impact. Allows experimentation with different approaches and concretizing evaluation methods into one evaluation design.	Best for instructional delivery systems, face-to-face, online, distance mode of learning. Appropriate in very complex and potentially high-stakes evaluations.	Advantages <ul style="list-style-type: none"> • Leads to more validity, reliability, and diversity of findings and insights into the evaluation objectives. • Reveals unexpected findings with policy implications. • Allows multiple options for triangulation. Disadvantages <ul style="list-style-type: none"> • Requires careful and deliberate planning for an appropriate methodological mix to be credible. • It requires a team of evaluators with experience in quantitative and qualitative methods

Method	Description	Innovative Element	How and When to Best Use it?	Advantages and Disadvantages
				and how to combine them at multiple levels.
2. Participatory Statistics	This approach uses local people (teachers, parents, students) to generate statistics. Then, this approach is replicated to a large number of groups to produce robust quantitative data.	<p>Uses decentralized statistics data collection (from top-down) to empowering citizens, which makes it catalytic.</p> <p>Allows the possibility of collecting statistics on sensitive topics that are mostly inaccessible to traditional surveys.</p> <p>Enables data gathering approach to be increasingly tested and applied.</p>	Particularly suitable for social and census mapping for students, household enumeration and scoring, well-being ranking, trends and change analysis, and causal-linkage analysis. Best when the empowerment of school leaders is part of the policy, service, or program.	<p>Advantages:</p> <ul style="list-style-type: none"> • Produces more valid, reliable, and accurate data when carefully aggregated and triangulated. • Empowers citizens and schools' stakeholders. <p>Disadvantages</p> <ul style="list-style-type: none"> • It can be time-consuming if school stakeholders are asked to collect the necessary data. • Requires to be built-in into a school policy, service delivery, or educational program from the very beginning.
3. Outcomes harvesting	This is an evaluation approach that does not measure progress towards predetermined outcomes. However, it collects evidence of what has been achieved and works backward to determine whether and how the school leadership program or intervention contributed to the change or improvement of student learning outcomes. The Outcome Mapping approach also inspires this.	<p>It allows evaluating policies or programs where the relations or cause-and-effect are not fully understood and have previously been challenging to evaluate.</p> <p>Warrants the search and identification of unintended results that frequently missed in traditional evaluation methods (Grau, 2012)</p>	<p>Appropriate when the relationship of cause-effect- of school policies or educational services are unknown.</p> <p>Best in situations where complexities are high and outcomes are not well-defined or vague.</p>	<p>Advantage</p> <ul style="list-style-type: none"> • It can be useful when policies, services, or educational programs are complex and not based on a precise results chain or theory of change. <p>Disadvantage</p> <ul style="list-style-type: none"> • Can be time-consuming • A potential bias may occur when the evaluators interpret the expected outcomes, and it might skew findings.
4. Crowd-sourcing	This is an approach when many people actively report on a situation around them, often using social media platforms and mobile phone technology.	<p>Encourages significant participation and civic engagement.</p> <p>Permits data collection on a scope usually not feasible through traditional evaluation tools.</p> <p>A great variety of open-source software platforms and social media platforms are ready to use.</p>	<p>Best to utilize this approach when data collection requirements go beyond the scope, and more traditional program evaluation or quantitative information is required.</p> <p>Best used for sensitive issues (e.g., fraud, corruption) because anonymity can be installed.</p>	<p>Advantages</p> <ul style="list-style-type: none"> • Can collect massive location-specific data in real-time with lower running cost than more traditional tools • Encourages civic engagement <p>Disadvantages</p> <ul style="list-style-type: none"> • Requires incentives for citizens (or netizens) to be involved continuously • Requires tailoring a crowdsourcing platform. • Prone to fake participants
5. Mobile data collection	This approach gathers structured information	Mobile data gathering can	It can be used when mobile data	Advantages

Method	Description	Innovative Element	How and When to Best Use it?	Advantages and Disadvantages
	using mobile phones, tablets, or personal digital assistant (PDA) using special software applications.	incorporate first-hand and different information in designed surveys such as geographic location through auto-geo-tagging, photographs, and videos. More extensive participation because mobile and PDAs are widely available.	collection benefits offset the benefits of more traditional paper-based surveys. Where data collection demands or significantly benefits from audio, video, or geographic information.	<ul style="list-style-type: none"> Enables improvement of timelines and accuracy of the data collection. Mobile platforms allow customized surveys to include photos, video recordings, GP, coordinates not usually captured in paper-based surveys. Disadvantages <ul style="list-style-type: none"> Technology alone will not improve the survey designs or instruments Potential bias of well-education and well-off individuals.
6. Data Exhaust	This approach uses data generated from trails of information from digital or online activities. Whenever individuals use mobile phones or access the internet contents, they leave trails behind in the form of transactional data called “data exhaust.” These comprise storable choices, actions, or preferences such as log files, temporary files, cookies, and even information generated in any digital or online transactions. Data exhaust is massive, passively collected transactional data from individual’s use of digital services.	Students, teachers, parents, and other schools’ stakeholders use the internet extensively. Passive transactional data readiness has amplified exponentially and analyzing data exhaust can radically change how evaluation is conducted. The passive data on the web are always available. Data exhaust provides a lot of value in program evaluation. They can be tapped anytime, even by commercial services, to create novel data.	When data are processed in bulk for program evaluation, they allow to determine the current status of the entire schools’ stakeholders and institutions and identify changes occurring in real-time through web-based and social media search queries. These online transactions can also be used to predict human behavior.	Advantage <ul style="list-style-type: none"> Data already exist and readily collectible. Disadvantage <ul style="list-style-type: none"> Potential bias makes digital data skewed in favor of better educated, well-off individuals while marginalizing those less articulate individuals or with lesser access to digital services.

Conclusions

Innovative program evaluation in school leadership requires a combination of competent evaluators who can introduce methods and approaches that are actually novel. Innovations in program evaluation do not merely mean putting together strategies that integrate technology or relabeling existing approaches but having a good comprehension of the gamut and extent of evaluation practice and theory applied in school leadership. They should be well-grounded with the traditional approaches of program evaluation.

A successful program evaluation, whether traditional or innovative, should have an evaluation team who are:

1. experienced in the type of evaluation needed;
2. competent in quantitative and qualitative data sources and analysis;
3. can work with a broad sample or population of stakeholders and involving representatives of the target populations;
4. able to develop innovative approaches to evaluation while taking into account the realities affecting a program, such as challenges in data gathering and having limited human and financial resources;
5. can merge evaluation into all program activities;
6. have good grips on the potential benefits and risks of the evaluation;
7. build the capacity of the program staffs in designing and conducting the evaluation; and
8. prepare an excellent report and disseminate it to all evaluation stakeholders.

Furthermore, program evaluation innovations for school leadership should be applied when:

1. the needs are identified;
2. schools and educational institutions can come up with informed decisions about innovations;
3. sufficient resources are available to mobilize the evaluation team;
4. the users and audience are ready and capable of accepting change in the process of program evaluation;
5. there is an assurance of thorough data analysis and follow-up activities; and
6. the beneficiaries are encouraged and ready to participate.

Finally, innovation in a program evaluation should generate discussion around evaluation planning and implementation. They should be able to provide data-driven inputs toward improving any program or activities that would improve school leadership and students' learning outcomes. All innovative approaches must ensure credibility that is useful in cultivating any program evaluation initiatives.

References

- Boulmetis, J. and Dutwin, P. (2005). *The ABCs of evaluation: Timeless techniques for program and project managers* (2nd ed.). San Francisco: Jossey- Bass.
- Burkhauser, S., Pierson, A., Gates, S. M., & Hamilton, L. S. (2012). *Addressing Challenges in Evaluating School Principal Improvement Efforts*. Santa Monica, CA: RAND Corporation. Retrieved from https://www.rand.org/pubs/occasional_papers/OP392.html.
- Chen H.T. (2015). *Practical Program Evaluation: Theory-Driven Evaluation and the Integrated Evaluation Perspective*. 2nd ed. Los Angeles, CA: SAGE Publications; pp. 3–33.
- European Commission (1995). Green Paper on innovation. The document drawn up on the basis of COM (95) 688 final, Bulletin of the European Union Supplement 5/95. Luxemburg. Available online at <http://europa.eu.int/en/record/green/gp002en.doc>
- Melinda & Gates Foundation (2014). *Teachers know best: Making data work for teachers and students*. Retrieved from <https://files.eric.ed.gov/fulltext/ED557084.pdf>; <http://k12education.gatesfoundation.org/download/?Num=2335&filename=TeachersKnowBest-MakingDataWork.compressed.pdf>
- Mertens D.M. and Wilson A.T. (2012). *Program Evaluation Theory and Practice: A Comprehensive Guide*. New York, NY: Guilford Press; pp. 3–12
- Olney C.A and Barnes, S.J. (2013). *Planning and evaluating health information outreach projects. Booklet 2: Planning outcomes-based outreach projects* (2nd ed.). Seattle, WA: National Network of Libraries of Medicine Outreach Evaluation Resource Center.
- Parkhurst, M. & Preskill, H. (2014). Learning in action: Evaluating collective impact. *Stanford Social Innovation Review, Fall*. Retrieved from https://ssir.org/articles/entry/learning_in_action_evaluating_collective_impact
- Patton M.Q. (2002). *Qualitative Research & Evaluation Methods*. (3rd ed.). Thousand Oaks, CA: Sage.

- Perrin, B. (2002). How to – and How not to – evaluate innovation. *Evaluation*, 8(1), 13-28.
- Rossi R H., Lipsey, M. W., & Freeman. H. E. (2004). *Evaluation: a systematic approach*. Thousand Oaks. Call.: Sage Publications.
- Scriven, M. (1998). Minimalist theory of evaluation: The least theory that practice requires. *American Journal of Evaluation*, 19, 57-70.
- Shadish W.R., Cook T.D., and Leviton L.C. (1990). *Foundations of Program Evaluation: Theories of Practice*. Newbury Park, CA: SAGE Publications.
- Schwandt T.A. (2015). *Evaluation Foundations Revisited: Cultivating a Life of the Mind for Practice*. Redwood City, CA: Stanford University Press.
- Stufflebeam, D. J., & Shinkfield, A. J. (2007). *Evaluation theory, models, & applications*. San Francisco, CA: Wiley.
- U.S. Department of Health and Human Services Centers for Disease Control and Prevention. Office of the Director, Office of Strategy and Innovation. (2011) *Introduction to program evaluation for public health programs: A self-study guide*. Atlanta, GA: Centers for Disease Control and Prevention.
- Wiseman S.H. (no date). *Evaluating Efforts to Improve School Leadership: Challenges and Recommendations. A research brief describes work done by RAND Education documented in Addressing Challenges in Evaluating School Principal Improvement Efforts*, by Susan Burkhauser, Ashley Pierson, Susan M. Gates, and Laura S. Hamilton, OP-392-NLNS, 2012, 40. Retrieved from http://www.rand.org/pubs/occasional_papers/OP392.html.
- Wholey Joseph S, Hatry Harry P, Newcomer Kathryn E, et al. (2010). *Handbook of Practical Program Evaluation Third Edition*. San Francisco: Jossey-Bass.
- World Health Organization (2013). *WHO evaluation practice handbook*. Switzerland.
- Wright, B. & Wallis, S.E. (2019). *Practical Mapping for Applied Research and Program Evaluation*. SAGE Publications.
- Yarbrough D.B., Shulha, L.M., Hopson. R.K., and Caruthers. F.A. (2011) *The Program Evaluation Standards: A guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks, CA: Sage.