DOI: 10.1002/ev.20444

ORIGINAL ARTICLE

WILEY

Introduction to Relational Systems Evaluation

Jennifer Brown Urban[1] | **Thomas Archibald**[2] | **Monica Hargraves**[3] | **Jane Buckley**[4] | **Claire Hebbard**[3] | **Miriam R. Linver**[1] | **William M. Trochim**[3]

[1] Montclair State University, Montclair, NJ, USA

[2] Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

[3] Cornell University, Ithaca, NY, USA

[4] JCB Consulting, Rochester, NY, USA

Correspondence
Jennifer Brown Urban, Montclair State University, Montclair, NJ, USA.
Email: urbanj@montclair.edu

Funding information
John Templeton Foundation, Grant/Award Number: 60483; National Science Foundation, Grant/Award Numbers: 0535492, 0814364, 1811214

Abstract

In this paper, we introduce Relational Systems Evaluation (RSE), the focal topic of this volume. RSE is a framework for program planning and evaluation that is theoretically grounded, empirically tested and focused on building evaluation capacity. Theoretically, RSE is rooted in an evolutionary approach to program development and evaluation informed by systems thinking. The application of RSE involves a collaborative partnership approach pairing evaluators or researchers with program practitioners, the use of collaborative modeling and planning tools, and an emphasis on nurturing an Evaluative Thinking mindset. The goal of RSE is to produce more thoughtful and useful evaluations and, ultimately, better programs.

Relational Systems Evaluation (RSE) is a framework for program planning and evaluation that is theoretically grounded, empirically tested, and focused on building evaluation capacity. Although RSE covers the full evaluation cycle (planning, implementation, and utilization), it emphasizes the importance and value of evaluation planning. In this introductory paper, we introduce the components of RSE which include: (1) its theoretical basis, an evolutionary approach to program development and evaluation; (2) a collaborative partnership approach; (3) the use of collaborative modeling and planning tools (the Systems Evaluation Protocol [SEP] for Evaluation Planning and the Netway); and (4) the importance of nurturing an Evaluative Thinking (ET) mindset (see Figure 1). Subsequent papers in this volume will provide more detail on the theoretical grounding and philosophical orientation of RSE (Trochim & Urban, this volume), practical information on how to apply RSE (Trochim & Urban, this volume; Urban, Hargraves, et al., this volume; Urban, Linver, et al., this volume), and empirical findings on the efficacy of the framework (Buckley, Hargraves, & Moorman, this volume; Chauveron et al., this volume; Hargraves et al., this volume). The volume concludes with commentary from a funder who has both funded and participated in RSE (Bollinger, this volume).

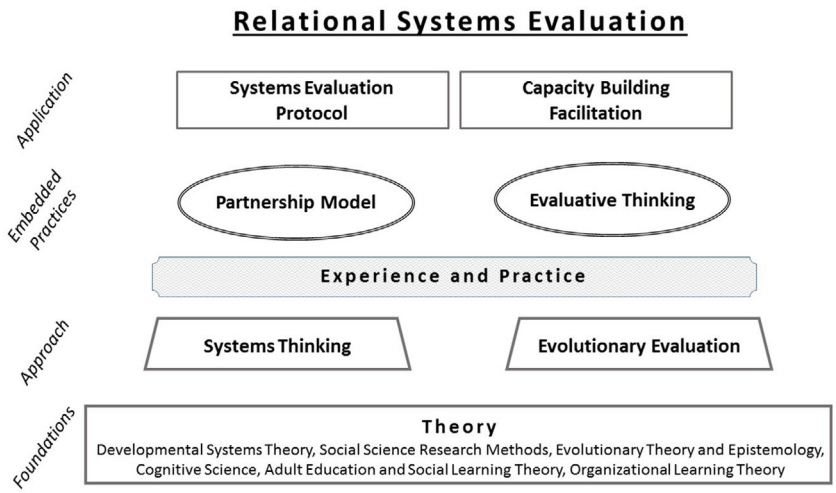

FIGURE 1 Overview of relational systems evaluation framework

High-quality evaluation necessarily begins with good evaluation planning. All too often, there is a rush to measurement without putting in the careful thought and attention to a program's underlying theory of change, the evaluation questions that are important to answer, and the more extensive system within which a program is embedded. This can lead to a waste of resources when the data collected fail to address the question(s) of interest and/or the evaluation design is not appropriate for the stage of development of the program (e.g., using a randomized controlled trial with a program that is being implemented for the first time; Urban, Hargraves, & Trochim, 2014). Good evaluation planning yields more thoughtful and useful evaluations, higher quality data, and ultimately a better understanding of program effects and needs. This is both the intent and ambition of RSE.

AN OVERVIEW OF RSE

Theoretical foundations

Our work on RSE began with a grant from the National Science Foundation in 2005 when our focus was on applying systems thinking to evaluation. As systems thinking was applied to evaluation, the overarching approach naturally integrated a breadth of evaluation theories. RSE is grounded in formal research and theory from developmental systems theory (e.g., Lerner, 2006; Overton, 2006, 2010), evolutionary theory (Darwin, 1859; Mayr, 2001), evolutionary epistemology (Bradie & Harms, 2006; Campbell, 1974, 1988; Cziko & Campbell, 1990; Popper, 1973, 1984), and systems theory (Bertalanffy, 1995; Laszlo, 1996; Midgley, 2003; Ragsdell, West, & Wilby, 2002). It is also grounded in experience-based learning and testing with hundreds of educational science, technology, engineering, and math (STEM) programs across the globe. Trochim and Urban (this volume) present a detailed description of the theoretical principles of RSE. Out of these theoretical foundations, we developed an approach that combined Evolutionary Evaluation (Urban et al., 2014) and a Partnership Model.

Approach

Evolutionary Evaluation considers the complex factors inherent in the larger systems within which a program is embedded (Trochim et al., 2016; Urban & Trochim, 2009; Urban et al., 2014; Urban, Hargraves, Hebbard, Burgermaster, & Trochim, 2011). Just as we characterize human development into broad phases (e.g., infancy, childhood, adolescence, early adulthood), we can similarly discuss the development of programs in terms of broad phases.

Every program—like an organism—has a lifecycle and proceeds through different phases: they are initiated (born); they typically go through cycles of rapid change and growth; they may stabilize and become more "settled"; they may be disseminated widely, and at any point, they may be retired or replaced. This view of program evolution is operationalized in the SEP by articulating four phases of program lifecycles: Phase 1—Initiation, Phase 2—Development, Phase 3—Stability, and Phase 4—Dissemination. Importantly, it is not simply the passage of time that marks a program's evolution but rather a substantive progression that includes refinement and stabilization of program content and approach—a process that is enormously strengthened by appropriate evaluation. This progress is an evolution in the "state of the program" over the course of the program's lifecycle.

It is useful to recognize a parallel process occurring on the evaluation side: the scope and strength of the claims one can make—the "state of knowledge"—is also evolving. Just as the state of the program evolves over its lifecycle, so too does the state of knowledge about a program increase over the lifecycle of evaluation. This progression is operationalized by defining four phases of evaluation lifecycles: Phase 1—Process and Response, Phase 2—Change, Phase 3—Comparison and Control, and Phase 4—Generalizability. For any given program lifecycle phase, there is an appropriate type of evaluation work to be done—that is, a corresponding evaluation lifecycle phase. Alignment between program and evaluation phases is essential for ensuring that a program obtains the kind of information that is most needed at that point in the life of the program, and that program and evaluation resources are used efficiently. When program practitioners, program managers, and evaluators conceptualize program evaluation from this evolutionary perspective, better decisions can be made about whether to keep, change, or retire a program and about what kinds of evaluations to conduct and fund (Urban et al., 2014). In summary, Evolutionary Evaluation is foundational to RSE. It provides important guidance in selecting evaluation methodology during the evaluation planning stage, and also at the utilization stage, by directing attention to how to interpret results in light of the program lifecycle phase, and determining what the program needs in order to evolve and allocate program resources most effectively.

RSE also draws heavily from systems thinking and uses a set of systems thinking metaphors to draw parallels to evaluation. Several systems thinking principles are highlighted in RSE, including part-whole relationships, static and dynamic processes, scale, multiple perspectives, boundaries, and causal pathways. The RSE philosophical orientation is discussed in Trochim and Urban (this volume).

Embedded practices

Partnership model

RSE recognizes there is often a false divide between research and practice that can hinder the application of research knowledge in real-world settings and the application of practitioner knowledge to evaluation and research agendas. Evaluation Partnerships among

researchers and practitioners are used to integrate evaluation into organizational practices, enhance capacity building, and strengthen both program and research outcomes (Tseng, Easton, & Supplee, 2017; Urban & Trochim, 2009). RSE relies on a collaborative partnership model defined by mutual respect and shared goals between program and evaluation professionals. The primary purpose of these relationships is to effectively use evaluation methodology and research knowledge, together with practitioners' expertise and insights about program context, capacity, and community realities, for program improvement. To establish productive and meaningful collaborative partnerships, evaluation and program professionals must rely on relational principles including flexibility, responsiveness (to each other, the program, the context, the organization, and other systems), open communication, and mutual benefit (see Buckley et al., this volume; Urban et al., 2014). Recognizing the value of contributions made by both researchers and program professionals as well as being flexible and responsive to the context, systems, and individuals surrounding a program are both critical to research-practice integration, high-quality evaluation and, ultimately, program evolution. Buckley et al. (this volume) provide more detail on RSE's partnership approach.

Evaluative Thinking

RSE is focused on developing evaluation capacity in part by facilitating the adoption of an Evaluative Thinking mindset. Often, descriptions or definitions of Evaluative Thinking liken it to reflective practice (Argyris & Schön, 1978; Schön, 1983). Buckley, Archibald, Hargraves, and Trochim (2015) define Evaluative Thinking as:

> critical thinking applied in the context of evaluation, motivated by an attitude of inquisitiveness and a belief in the value of evidence, that involves identifying assumptions, posing thoughtful questions, pursuing deeper understanding through reflection and perspective taking, and informing decisions in preparation for action. (p. 378)

The development of Evaluative Thinking is infused in the Evaluation Capacity Building (ECB) activities that comprise RSE to increase the likelihood of realizing ECB's ultimate goal: improved program outcomes. Evaluative Thinking integrates the same skills that characterize good evaluation throughout an organization's work practices (Baker & Bruner, 2012); it is characterized by "a willingness to do reality testing, to ask the question: how do we know what we think we know? … It's an analytical way of thinking that infuses everything that goes on" (Patton, 2005, para. 10). Buckley et al. (2015) have sought to integrate knowledge and experience from across fields, including ECB, critical thinking, and education. They have worked to further refine the definition of Evaluative Thinking which is operationalized through the partnership work described in this issue. These researcher-practitioner partnerships are flexible and systems-focused. The ECB work described in this volume is not limited to the impact on the individual, nor does it take on every organization as a whole. Instead, the focus is on responding to the needs of each researcher-practitioner partnership individually, strategically promoting Evaluative Thinking and evaluative practices in ways that maximize learning and program improvement for each partner. Additional detail about Evaluative Thinking in RSE can be found in our prior publications (Buckley et al., 2015; McIntosh, Buckley, & Archibald, 2020). Chauveron et al. (this volume) will present empirical findings regarding the adoption of Evaluative Thinking behaviors as a result of engaging in RSE.

Application

Systems Evaluation Protocol

The Systems Evaluation Protocol (SEP; Trochim et al., 2016) is a step-by-step guide for program evaluation, covering the three phases of evaluation planning, implementation, and utilization. The SEP is grounded in systems thinking and builds on the premise that both evaluators and program professionals bring unique expertise and perspectives to program evaluation, and that effective integration of these strengths offers a powerful and essential basis for planning and conducting a high-quality evaluation. The SEP is a detailed protocol that includes three phases: Phase I: Evaluation Planning, Phase II: Evaluation Implementation, and Phase III: Evaluation Utilization. The SEP Evaluation Planning phase has three stages: (1) Preparation, (2) Modeling, and (3) Evaluation Plan Development. Several products result from completing the SEP Evaluation Planning steps, including a: Partnership Agreement or Memorandum of Understanding, Stakeholder Map, Logic Model, Pathway Model, and comprehensive Evaluation Plan. These products embody the SEP's distinctive partnership approach, systems thinking, and Evolutionary Evaluation principles. The SEP Implementation phase guides users through the process of putting the evaluation plan into action, with attention to practical considerations and keeping a focus on the plan's overarching goals in the face of surprises that are inevitable in real-world circumstances. The SEP Utilization phase focuses on ensuring that the evaluation effort and results are put to their best use. Urban, Hargraves, et al. (this volume) will provide a detailed description of the SEP for Evaluation Planning and how it can be implemented. Urban, Linver, et al. (this volume) will present six case study examples of Phase I SEP implementation.

Technology platforms for collaborative modeling and evaluation planning can enhance the partnership's work. The Netway—a web-based tool for developing program models and evaluation plans—is a cyberinfrastructure developed in tandem with the SEP. As such, the Netway supports collaboration and encourages community networking. Additional benefits include context-sensitive access to the SEP steps and information, instructional videos, worksheets, and handouts. The embedded "Help" system is also available to address technical questions. The Netway system, now encompassing hundreds of programs, was made available to the public in the summer of 2015. At present, any user can create a free account at www.evaluationnetway.com.

Capacity building facilitation

RSE is accessible to evaluators with a basic level of evaluation skills. The SEP provides step-by-step instructions, including worksheets and activities, which guide evaluators through how to plan and implement evaluations. RSE is also deeply ingrained with a collaborative ECB approach to evaluation; as such, it requires facilitation skills that may be considered secondary for many evaluators.

Although the SEP is highly accessible, RSE may not be appropriate for evaluators who prefer to operate from a neutral, external, or uninvested position. RSE requires a mindset most appropriate for evaluators who are at least marginally interested in ECB and/or who are actively working in ECB. The partnership model that forms the foundation of RSE requires evaluators and program professionals to see each other as equally knowledgeable, albeit with distinct areas of expertise.

LIMITATIONS OF RSE

Some of the strengths of RSE in practice include its contribution to individual and organizational evaluation capacity, the depth of thinking it promotes and instills, the emphasis on grounding evaluation decisions in the detailed articulation of the program's theory of change, and on aligning evaluation plans with the program's lifecycle stage and stakeholder priorities. However, these strengths also contribute to its limitations. In particular, using RSE can be an intensive and sometimes lengthy process. This is particularly true the first time a team undertakes an evaluation using this approach, as it is an investment in new ways of thinking and in new skills. This investment pays off in future evaluations as the process becomes more familiar. Still, the demands can be difficult for staff who have limited time for engaging in evaluation, or where there is insufficient support for and commitment to evaluation at higher levels of the organizational system. For programs facing immediate and substantial funding uncertainties, this investment can be particularly difficult to make. For programs whose evaluations are highly prescribed, with little room for pursuing evaluation needs that may not align with a funder's mandate or other imperatives, the ground-up decision-making inherent in the RSE approach may be out of place. Finally, since RSE invokes a partnership model in which evaluators and program professionals work together to integrate their respective expertise, there can be challenges in ensuring that the two sides of that partnership are well matched, and have a shared understanding of what the work will entail and how it will be valuable. Experiences with these kinds of real-world challenges will be discussed in depth in Buckley et al. (this volume) and Urban, Linver, et al. (this volume).

ACKNOWLEDGMENTS

This work supported in part by the John Templeton Foundation, grant # 60483, and the National Science Foundation, grant #s 0535492, 0814364, and 1811214.

ORCID

Jennifer Brown Urban https://orcid.org/0000-0001-8405-3078
Thomas Archibald https://orcid.org/0000-0002-3567-9143
Monica Hargraves https://orcid.org/0000-0002-6074-4445
Jane Buckley https://orcid.org/0000-0003-4206-2341
Claire Hebbard https://orcid.org/0000-0001-6429-7788
Miriam R. Linver https://orcid.org/0000-0002-1470-5779
William M. Trochim https://orcid.org/0000-0003-0369-2922

REFERENCES

Argyns, C., & Schon, D. (1978). *Organizational learning: A theory of action approach*. Reading, MA: Addison Wesley.

Baker, A., & Bruner, B. (2012). *Integrating evaluative capacity into organizational practice*. Cambridge, MA: The Bruner Foundation. Retrieved from http://www.evaluativethinking.org/docs/Integ_Eval_Capacity_Final.pdf

Bertalanffy, L. V. (1995). *General system theory: Foundations, development, applications* (Revised ed.). New York: Braziller.

Bradie, M., & Harms, W. (2006). Evolutionary epistemology. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2006 ed.). Stanford, CA: The Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University.

Buckley, J., Archibald, T., Hargraves, M., & Trochim, W. M. (2015). Defining and teaching evaluative thinking: Insights from research on critical thinking. *American Journal of Evaluation, 36*(3), 375–388. https://doi.org/10.1177/1098214015581706

Campbell, D. T. (1974). Evolutionary epistemology. In P. A. Schilpp (Ed.), *The philosophy of Karl Popper* (pp. 413–463). LaSalle, IL: Open Court.

Campbell, D. T. (1988). Evolutionary epistemology. In E. S. Overman (Ed.), *Methodology and epistemology for social science: Selected papers of Donald T. Campbell*. Chicago, IL: University of Chicago Press.

Cziko, G. A., & Campbell, D. T. (1990). Comprehensive evolutionary epistemology and bibliography. *Journal of Social and Biological Structures, 13*(1), 41–82. https://doi.org/10.1016/0140-1750(90)90033-3

Darwin, C. (1859). *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. London, UK: John Murray.

Laszlo, E. (1996). *The systems view of the world: A holistic vision for our time*. Creskill, NJ: Hampton Press.

Lerner, R. M. (2006). Developmental science, developmental systems, and contemporary theories of human development. In R. M. Lerner & W. Damon (Eds.), *Handbook of child psychology: Theoretical models of human development* (6th ed., Vol. *1*, pp. 1–17). Hoboken, NJ: John Wiley & Sons.

Mayr, E. (2001). *What evolution is*. New York: Basic Books.

McIntosh, J. S., Buckley, J., & Archibald, T. (2020). Refining and measuring the construct of evaluative thinking: An exploratory factor analysis of the Evaluative Thinking Inventory. *Journal of MultiDisciplinary Evaluation, 16*(34), 104–117.

Midgley, G. (2003). *Systems thinking*. Thousand Oaks, CA: Sage.

Overton, W. F. (2006). Developmental psychology: Philosophy, concepts, methodology. In R. M. Lerner (Ed.), *Theoretical models of human development* (6th ed., Vol. *1*, pp. 18–88). Hoboken, NJ: Wiley.

Overton, W. F. (2010). Life-span development: Concepts and issues. In W. F. Overton & R. M. Lerner (Eds.), *Handbook of life-span development: Cognition, biology, and methods* (Vol. *1*, pp. 1–29). Hoboken, NJ: Wiley.

Patton, M. Q. (2005). In conversation: Michael Quinn Patton. Interview with Lisa Waldick, from the International Development Research Center. Retrieved from https://www.idrc.ca/en/research-in-action/conversation-michael-quinn-patton

Popper, K. (1973). *Evolutionary epistemology*. Paper presented at the Sections I-VI of "the rationality of scientific revolutions" given at the Herbert Spencer lecture, University of Oxford, Oxford.

Popper, K. (1984). Evolutionary epistemology. In D. M. Miller (Ed.), *Popper selections* (pp. 78–86). Princeton, NJ: Princeton University Press.

Ragsdell, G., West, D., & Wilby, J. (2002). *Systems theory and practice in the knowledge age*. New York: Kluwer Academic/Plenum.

Schön, D. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.

Trochim, W. M., Urban, J. B., Hargraves, M., Hebbard, C., Buckley, J., Archibald, T., … Burgermaster, M. (2016). *The guide to the Systems Evaluation Protocol (V 3.1)*. Ithaca, NY: Cornell Digital Print Services. Retrieved from https://core.human.cornell.edu/research/systems/protocol/index.cfm

Tseng, V., Easton, J. Q., & Supplee, L. H. (2017). Research-practice partnerships: Building two-way streets of engagement. *SRCD Social Policy Report, 30*(4)

Urban, J. B., Hargraves, M., Hebbard, C., Burgermaster, M., & Trochim, W. M. (2011). Evaluation in the context of lifecycles: 'A place for everything, everything in its place'. Paper presented at the annual meeting of the American Evaluation Association, Anaheim, CA.

Urban, J. B., Hargraves, M., & Trochim, W. M. (2014). Evolutionary evaluation: Implications for evaluators, researchers, practitioners, funders, and the evidence-based program mandate. *Evaluation and Program Planning, 45*, 127–139. https://doi.org/10.1016/j.evalprogplan.2014.03.011

Urban, J. B., & Trochim, W. M. (2009). The role of evaluation in research-practice integration: Working toward the "golden spike". *American Journal of Evaluation, 30*(4), 538–553. https://doi.org/10.1177/1098214009348327

AUTHOR BIOGRAPHIES

Jennifer Brown Urban is professor of Family Science & Human Development, and co-director of the Institute for Research on Youth Thriving and Evaluation, both at Montclair State University. Her research focuses on positive youth development and Evaluation Capacity Building in youth-serving organizations.

Thomas Archibald is an associate professor and extension specialist in Virginia Tech's Department of Agricultural, Leadership, and Community Education, where his research and practice focus primarily on Evaluation Capacity Building and Evaluative Thinking in community development contexts.

MONICA HARGRAVES is the associate director for Evaluation Partnerships in the Cornell Office for Research on Evaluation, at Cornell University, where her work focuses on a collaborative approach to evaluation and Evaluation Capacity Building using the tools of Relational Systems Evaluation.

JANE BUCKLEY of JCB Consulting is an evaluation planning facilitator and Evaluation Capacity Building consultant whose focus is on fostering a culture of Evaluative Thinking in partner organizations.

CLAIRE HEBBARD is the project coordinator at the Cornell Office for Research on Evaluation (CORE), where she collaborates on research on evaluation methods, assists in facilitating Evaluation Capacity Building, and helps implement evaluations in small outreach programs.

MIRIAM R. LINVER is professor of Family Science & Human Development, and co-director of the Institute for Research on Youth Thriving and Evaluation, both at Montclair State University; her research focuses on diverse contexts of youth development.

WILLIAM TROCHIM is professor of Policy Analysis and Management at Cornell University who specializes in research methodology and program evaluation.

How to cite this article: Urban, J.B., Archibald, T., Hargraves, M., et al. (2021). Introduction to Relational Systems Evaluation. *New Directions for Evaluation*, 1–8. https://doi.org/10.1002/ev.20444