Journal of MultiDisciplinary Evaluation
Volume 20, Issue 48, 2024

JMDE
Journal of MultiDisciplinary Evaluation

ISSN 1556-8180
http://www.jmde.com

What Is This Thing Called Evaluation Theory?

Sebastian Lemire
Abt Global

Bianca Montrosse-Moorhead
University of Connecticut

Christina A. Christie
University of California, Los Angeles

Background: Evaluation theory has played and continues to play an important role in the practice and profession of evaluation. Over the years, and as the range of evaluation theories has continued to grow, evaluation scholars and practitioners have developed various visual classification frameworks of evaluation theory, invoking images such as trees, rivers, the periodic table, metro maps, concept maps, and most recently gardens.

Purpose: The purpose of this special issue is to explore various visualizations of evaluation theory and to discuss their benefits and implications for the practice and profession of evaluation.

Setting: Not applicable.

Intervention: Not applicable.

Research Design: Not applicable.

Data Collection and Analysis: Not applicable.

Findings: Not applicable.

Keywords: *evaluation theory; data visualization; research on evaluation; teaching evaluation.*

Evaluation theory is essential to and for evaluation. Knowledge of and ability to use evaluation theories, for example, continues to be of great importance and is explicitly identified in several competency and capabilities frameworks, including in the Evaluation Knowledge domain of *The EES Evaluation Capabilities Framework* (European Evaluation Society, 2011) and The UK Evaluation Society's *Framework of Evaluation Capabilities* (UK Evaluation Society, 2012), and the Professional Practice domain of *AEA Evaluator Competencies* (American Evaluation Association, 2018). Knowledge and use of evaluation theories is also highlighted in the United Nations Evaluation Group's (UNEG's) *Evaluation Competency Framework* (2016). Moreover, requests for proposals (RFPs) and terms of reference (ToRs) sometimes specifically request the use of a particular evaluation approach (UNEG, 2020).

To what extent, how, and to what end evaluation theory is used in practice is still an ongoing question (Christie, 2003; Christie & Lemire, 2019; Coryn et al., 2011). Few studies (e.g., Bledsoe & Graham, 2005) have sought to capture the complexity of using multiple evaluation theories in practice. One suggestion put forth to explain the less-than-optimal link between evaluation theory and practice is that many evaluator education and training efforts do not include serious attention to the subject, and this explanation has some empirical support (Dewey et al., 2008; Galport & Azzam, 2017; Montrosse-Moorhead et al., 2022). And, when evaluator training does include evaluation theories, learning how to use them *practically* (as opposed to learning about one or more of them in theory) is less frequently explored (Boyce & McGowan, 2019; Lee et al., 2007).

To aid others in learning about and using evaluation theories, scholars have developed visual classification frameworks (e.g., Alkin & Christie, 2023; Lemire et al., 2020; Montrosse-Moorhead et al., 2024). These visual frameworks are presented in different places: Some are published in journals and books; others are presented on blogs or have only been presented at professional evaluation conferences. This diffuse state of presentation makes it more difficult than it should be to learn what visualizations are out there and what each visual classification uniquely illuminates about evaluation theory and practice, and it stifles the learning that could come from reading about and looking across the visuals. This special issue aims to remedy this by providing a common language to understand evaluation theories and by presenting seven visual classification frameworks to promote learning, dialogue, and use.

Definitions of Evaluation Theory

As part of his address at the 1998 American Evaluation Association conference, then-President Will Shadish famously declared that "evaluation theory is who we are" (Shadish, 1998, p. 3). The historical context for his remark was that evaluation as a transdisciplinary field was coming of age. Accordingly, there was a collective need for and interest in defining what evaluation is and should be—to draw boundaries around the practice and profession of evaluation. As Shadish went on to discuss, evaluation theory matters for the practice and profession of evaluation in several important ways: by providing a common language (e.g., "formative" vs. "summative"); by identifying important issues (e.g., evaluation use, nature of causation, equity); by distinguishing evaluation from other professions (e.g., education, economics); and by establishing a unique knowledge base (what makes the discipline and "evaluators" unique). As Shadish concluded, evaluation theory "encompasses many of the things in our field about which we seem to care most deeply" (1998, p.3). Decades later, evaluation theory continues to play an important role in defining the practice and profession of evaluation.

Scholarship generally discusses evaluation "theory" in various—yet related—ways (Donaldson & Lipsey, 2006; Montrosse-Moorhead et al., 2024; Scriven & Davidson, 2021). One way, which is the focus of the classifications in this special issue, is as prescriptions or prototypes for practice, or what scholars argue ought to be done in carrying out evaluations. Other ways include how sociological, political, or psychological frames apply to the practice of evaluation or the thing being evaluated; meta-theories about the nature of evaluation as a discipline or transdiscipline; and hypotheses about how the thing being evaluated works in a particular context.

Speaking directly to the important role and purpose of evaluation theory, evaluation scholars have over the years offered different definitions and conceptualizations of evaluation theory. Shadish, Cook, and Leviton—in their classic volume on evaluation theory—define theory as a "body of knowledge that organizes, categories, describes, predicts, explains, and otherwise aids in understanding and controlling a topic" (1991, p. 30). Framed within the broader commitment to social betterment, they propose that evaluation theory serves to "specify feasible practices that evaluators can use to construct knowledge of the value of social of social programs that can be used

to ameliorate the social problems to which programs are relevant" (p. 36).

In line with this thinking, Alkin (2013) suggests that evaluation theories "offer a set of rules, prescriptions, prohibitions, and guiding frameworks that specify what a good or proper evaluation is and how evaluation should be done" (p. 4). Evaluation theory, then, is normative and prescribes how and with what purpose to conduct an evaluation. Thus, we have evaluation theories that are intended to promote the utilization of evaluation results (Patton, 1978), encourage stakeholder empowerment (Fetterman, 1994), or bring attention to context–mechanism–outcome configurations (Pawson & Tilley, 1997), to name a few. While these theories are often informed by several things, including the theorists' practical experiences with evaluation and their paradigmatic worldviews (Alkin & Patton, 2020), they are, in scope and nature, intended for application across a broad—and often unspecified—range of contexts. In this way, and as suggested by Christie and Lemire (2019), evaluation theories are perhaps best viewed as providing a working logic to assist evaluators in designing evaluations, selecting procedures and methods, and providing a rationale for the selected procedures and methods used in an evaluation.

In scholarship on evaluation theory, several terms (theory, approach, model) are used interchangeably to refer to this working logic. We suspect that these terms are used interchangeably because each term highlights something different that is important to understand about this working logic. "Theory," for example, draws attention to the fact that these are "conceptions of something to be done, or of the method of doing it" (Oxford University Press, 2023a). "Approach" is "a way of considering or handling something, especially a problem," and so the use of this term highlights that a problem of practice is how to carry out evaluations and what sorts of things ought to guide that decision (Oxford University Press, 2023b). "Model" brings to the forefront that these logics are offered as "objects of imitation" and, ideally, serve as prototypes that can be emulated in practice by others (Oxford University Press, 2023c). We use these terms in all of these ways.

Regardless of the term used for this working logic, it should be understood as distinct from other concepts used in evaluation. Donaldson and Lipsey (2006) provide a useful distinction between *evaluation* theory (the working logics that are the focus of the classifications in this special issue), *scientific* theory (how sociological, political, or psychological theories apply to the thing being evaluated), and *program* theory (local and context-bound theories aiming to explain how a program works for a given population in a specific setting, time, and context). Evaluation theory is also distinguished from other common components of evaluation methods or designs, such as research methods, sampling methods, data collection methods, analytical methods, and so on (Abbott & Abbott, 2004; Montrosse-Moorhead et al., 2024; Moss & Haertel, 2016).

Evaluation Theory Classification Frameworks

Over the years, and as the range of evaluation theories has continued to grow, evaluation scholars have made various attempts to provide an overview of the broadening array of evaluation theories through classification frameworks. These classifications differ in their frame, organization, format, and assumptions. As noted by Alkin and Christie (2023), some of the earliest classifications were developed by Worthen and Sanders (1973), Popham (1975), House (1978), Glass and Ellett (1980), Alkin and Ellett (1985), Shadish, Cook, and Leviton (1991), and Alkin and House (1992). Some are organized in tabular form or are verbally described, such as Stufflebeam's classification according to coherence with evaluation (Stufflebeam & Coryn, 2014), Fitzpatrick et al.'s (2023) evaluation approach orientation classification, and Shadish et al.'s (1991) evaluation theory stages classification. Some classifications take a historical perspective (Shadish et al., 1991), and others, such as Azzam and Donaldson (2013), who focus on the purposes of different evaluation theories, focus on one key dimension of evaluation theory.

Extending beyond classifications in text and table format, scholars have also developed various visualizations invoking images such as trees, rivers, the periodic table, subway systems, maps, and most recently gardens. These visual frameworks are the focus of this special issue (listed historically from year of first publication): the evaluation theory tree (Christie & Alkin, 2004; Alkin, 2013; Alkin & Christie, 2023), the tree of evaluation approaches (Mertens & Wilson, 2012, 2018), the evaluation river (Azzam & Donaldson, 2012, 2013), the periodic table of evaluation (Vaca, 2017), the evaluation metro map (Lemire, 2020), the map of evaluation uses (Quadrant Conseil, 2022), and the garden of evaluation approaches (Montrosse-Moorhead et al., 2024).

These visuals can be understood in a couple of different ways. They can be seen as different ways of categorizing, describing, and debating the

boundaries of evaluation practice, including but not limited to its methodological, theoretical, sociopolitical, philosophical, and ethical aspects. By mapping the evaluation landscape, the visuals establish boundaries between different types of evaluation approaches, and boundaries between evaluation and other types of knowledge production. The visualizations can also be viewed as different ways of telling the story of evaluation. Each visual contributes in its own way to a collective narrative—a way of making sense—of the evaluation theories and evaluation practice. In this way, the visuals also reflect the authors' backgrounds, relationships, experiences in evaluation, and so on.

While Montrosse-Moorhead et al. (2024) include an appendix summarizing key dimensions of five highly-cited classification frameworks, no single publication presents all of the visualized classification frameworks side by; some have been published, and others have been shared at conferences or in blogs. Each framework serves a unique purpose, and a key motivation for this special issue is that there is much that can be learned from reading about and looking at these visuals individually and collectively.

The Purpose and Scope of This Special Issue

With these reflections as our backdrop, the purpose of this special issue is to explore the various visualizations that have been used to map the evaluation theory landscape, as well as to discuss their benefits and implications for the practice and profession of evaluation. Toward these ends, the special issue includes articles on seven different visualizations of evaluation theory. Each visualization is described by its author(s) in terms of the background and motivation for its development, its intended purpose, and its structure and content. Authors were also asked to reflect on the benefits and limitations of their visual and to consider future directions for visualizing evaluation theory.

In the first article, Christina Christie describes the process for developing the many iterations of the evaluation theory tree (Christie & Alkin, 2004; Alkin, 2013; Alkin & Christie, 2023) and how it developed over time, and explore ways that others might contribute to its continued growth. The authors highlight key features of the evaluation theory tree whose under-recognition has led to questions about how the tree was developed, what informed it, and its underlying organizational tenets.

In her article "Visualizing Evaluation Theory: Tree, Forest, or Ocean Currents?", Donna Mertens reflects on the influence of the paradigm wars in the early 1970s on her motivation for revising Christie and Alkin's evaluation theory tree to include separate branches for social justice and Indigenous evaluation approaches. Her revised five-branch evaluation tree reflects the assumptions of the post-positivist, constructivist, pragmatic, transformative, and Indigenous paradigms. In reflecting on its benefits and limitations, Mertens concludes that while a tree metaphor is useful because it provides a quick and clear way to show that different assumptions lead to different methodological commitments, it is also limited because it does not show the interrelationships among the various paradigms. Looking ahead, Mertens suggests that an ocean current visualization could potentially accomplish that task.

Motivated by their shared interest in visualizing data and finding creative ways to communicate complex concepts to a broader audience, Tarek Azzam and Stewart Donaldson use a river metaphor in their article to represent and show the evolution of different purposes (e.g., utilization, knowledge generation, accountability) of evaluation. Their map of evaluation uses, which was intended as a teaching tool, shows that the end goal of all these evaluation approaches—the ocean which the rivers flow toward—is societal improvement, and highlights the historical context that has influenced their development.

Motivated by the many methodological choices involved in designing an evaluation, the purpose of Sara Vaca's periodic table of evaluation is to provide an accessible overview of the paradigms, evaluation approaches, methods, designs, and criteria used in real-world evaluations. As a visual thinker, Vaca selected the periodic table because it provides a familiar framework for summarizing many of the diverse choices involved in the science and art of designing an evaluation. Originally presented on her website and shared as a poster at the 2012 European Evaluation Society conference, the intended use of the visual is to provide a quick overview of options for those defining terms of reference for evaluations, crafting evaluation designs, or learning about evaluation.

Combining his longstanding interest in evaluation theory with his son's love for riding the DC metro, Sebastian Lemire developed the evaluation metro map, depicting evaluation approaches and study designs as stations on three main metro lines: one for methods, one for theory, and one for use. Reflecting on its benefits, Lemire notes that the map always generates conversation.

Why is this evaluation approach next to this one? Why is this approach included or not included? If included, which line would it fit on? Is it a new line? In generating these questions, he concludes, the map promotes collective reflection on what evaluation is and can be.

To facilitate discussions between evaluation officers and project managers on different types of evaluations, Thomas Delahais, Agathe Devaux-Spatarakis, François Jégou, and Karen Rousseau developed an interactive map of evaluation uses. Based on findings from concept mapping workshops with evaluation commissioners, the map visualizes how impact evaluation approaches and methods (e.g., experimental design, contribution analysis, process tracing) are connected with different types of uses (e.g., strategic use, continuous improvement, dialogue). As the authors conclude, the collaborative trial-and-error process of codesigning the map with evaluation commissioners highlighted the value of evaluation uses as an entry point to discussing the benefits and limitations of various evaluation theories and methods.

Based on a decade of collaborative work, Bianca Montrosse-Moorhead, Daniela Schröter, and Lyssa Wilson Becho provide details on the creation of the garden of evaluation approaches. They describe several interrelated aspects that led to the planting and blooming of the garden, its intended purpose and guiding question (How do evaluation approaches compare in terms of dimensions that facilitate use and application?), its structure and content, the intended users and uses, and they reflect on the benefits and limits of the garden. Importantly, they draw attention to eight dimensions, philosophical orientations, and methodological dispositions that distinguish approaches. They also describe the mixed-methods process that was used to classify approaches (and is currently being used to classify more). Grounded in a commitment to democratizing evaluation knowledge, they make their work visible and available for free on an open science framework (OSF) page.

In the concluding article for this special issue, Melvin Mark reflects on the strengths and weaknesses of existing visual representations of evaluation theory and on the implications of such visualizations. He suggests that evaluators should be "multilingual," familiar with multiple visualizations. He also speculates on a future generation of evaluation theory visualizations, including the possibility of more interactive representations that allow evaluators to choose the level of detail most useful to them.

The purpose of this special issue is not to argue in favor of a single visualization or visual metaphor. The position we hold is that no single visualization is ideal in all contexts. The practice and profession of evaluation is diverse and continues to diversify, and this diversity should be reflected in the use of a broad (even broadening) range of visuals and metaphors. Moreover, classifications and visuals simplify and impose order. Using them critically means developing and using an intellectual lens, being able to sense what is in focus and what is in the background, being able to sense what has been simplified and in what way (and thus distorted, because all simplifications distort), and being able to read between the lines of what the authors of different classification frameworks are saying. Our modest hope is that the special issue will aid practitioners, clients, program participants, funders, educators, researchers, theoreticians, and so on to continue to develop these habits of mind. Perhaps it will also serve to inspire and motivate new visualizations?

References

Abbott, A. D., & Abbott, D. (2004). *Methods of discovery: Heuristics for the social sciences*. W.W. Norton & Company.

Alkin, M. C. (2004). *Evaluation roots: Tracing theorists' views and influences*. Sage Publications.

Alkin, M. C. (2013). *Evaluation roots: A wider perspective of theorists' views and influences*. Sage Publications.

Alkin, M. C., & Christie, C. A. (2023). *Evaluation roots: Theory influencing practice*. Sage Publications.

Alkin, M., & Ellett, F. (1985). Evaluation models and their development. In T. Husén & T. N. Postlethwaite (Eds.), *International encyclopedia of education: Research and studies*. Pergamon Press.

Alkin, M. C., & House, E. (1992). Evaluation of programs. In M. C. Alkin (Ed.), *Encyclopedia of educational research*. Macmillan.

Alkin, M. C., & Patton, M. Q. (2020). The birth and adaptation of evaluation theories. *Journal of MultiDisciplinary Evaluation, 16*(35), 1–13.

American Evaluation Association. (2018). *AEA evaluator competencies*. https://www.eval.org/About/Competencies-Standards/AEA-Evaluator-Competencies

Azzam, T., & Donaldson, S. I. (2012). *The evaluation purposes visualization* [Unpublished manuscript].

Azzam, T., & Donaldson, S. I. (2013). *The evaluation purposes visualization* [Unpublished manuscript].

Bledsoe, K. L., & Graham, J. A. (2005). The use of multiple evaluation approaches in program evaluation. *American Journal of Evaluation*, *26*(3), 302–319. https://doi.org/10.1177/1098214005278749

Boyce, A. S., & McGowan, B. L. (2019). An exploration of two novice evaluation educators' experiences developing and implementing introduction to evaluation courses. *American Journal of Evaluation*, *40*(1), 119–136. https://doi.org/10.1177/1098214018778812

Christie, C. A. (2003). What guides evaluation? A study of how evaluation practice maps onto evaluation theory. *New Directions for Evaluation*, *97*(1), 7–36. https://doi.org/10.1002/ev.72

Christie, C. A., & Alkin, M. (2004). An evaluation theory tree. In M. Alkin (Ed.), *Evaluation roots* (pp. 12–66). Sage Publications.

Christie, C. A., & Lemire, S. T. (2019). Why evaluation theory should be used to inform evaluation policy. *American Journal of Evaluation*, *40*(4), 490–508. https://doi.org/10.1177/1098214018824045

Coryn, C. L., Noakes, L. A., Westine, C. D., & Schröter, D. C. (2011). A systematic review of theory-driven evaluation practice from 1990 to 2009. *American Journal of Evaluation*, *32*(2), 199–226. https://doi.org/10.1177/1098214010389321

Dewey, J. D., Montrosse, B. E., Schröter, D. C., Sullins, C. D., & Mattox, J. R. (2008). Evaluator competencies: What's taught versus what's sought. *American Journal of Evaluation*, *29*(3), 268–287. https://doi.org/10.1177/1098214008321152

Donaldson, S. I., & Lipsey, M.W. (2006). Roles for theory in contemporary evaluation practice: Developing practical knowledge. In I. Shaw, J.Greene, & M. Mark (Eds.), *The handbook of evaluation: Policies, programs, and practices* (pp. 56-75). London: Sage

European Evaluation Society. (2011). *The EES evaluation capabilities framework*. https://europeanevaluation.org/wp-content/uploads/2020/03/The-EES-Evaluation-Capabilities-Framework-leaflet.pdf

Fetterman, D. M. (1994). Steps of empowerment evaluation: From California to Cape Town. *Evaluation and Program Planning*, *17*(3), 305–313. https://doi.org/10.1016/0149-7189(94)90010-8

Fitzpatrick, J. L., Worthen, B. R., Sanders, J. R., & Wingate, L. A. (2023). *Program evaluation: Alternative approaches and practical guidelines* (5th ed.). Allyn and Bacon.

Galport, N., & Azzam, T. (2017). Evaluator training needs and competencies: A gap analysis. *American Journal of Evaluation*, *38*(1), 80–100. https://doi.org/10.1177/1098214016643183

Glass, G. V., & Ellett, F. S. (1980). Evaluation research. *Annual Review of Psychology*, *31*, 211–228. https://doi.org/10.1146/annurev.ps.31.020180.001235

House, E. R. (1978). Assumptions underlying evaluation models. *Educational Researcher, 7*(3), 4–12. https://doi.org/10.3102/0013189X007003004

Lee, J., LeBaron Wallace, T., & Alkin, M. (2007). Using problem-based learning to train evaluators. *American Journal of Evaluation, 28*(4), 536–545. https://doi.org/10.1177/1098214007305613

Lemire, S. (2020). *The evaluation metro map* [Unpublished manuscript].

Lemire, S., Peck, L. R., & Porowski, A. (2020). The growth of the evaluation tree in the policy analysis forest: Recent developments in evaluation. *Policy Studies Journal, 48(S1)*, S47–S70. https://doi.org/10/ghdjc2

Mertens, D. M., & Wilson, A. T. (2012). *Program evaluation theory and practice: A comprehensive guide*. The Guilford Press.

Mertens, D. M., & Wilson, A. T. (2018). *Program evaluation theory and practice*. Guilford Publications.

Montrosse-Moorhead, B., Gambino, A. J., Yahn, L. M., Fan, M., & Vo, A. T. (2022). Evaluator education curriculum: Which competencies ought to be prioritized in Master's and Doctoral programs? *American Journal of Evaluation*, *43*(2), 269–292. https://doi.org/10.1177/10982140211020326

Montrosse-Moorhead, B., Schröter, D., & Becho, L. W. (2024). The garden of evaluation approaches. *American Journal of Evaluation*, *0*(0). https://doi.org/10.1177/10982140231216667

Moss, P. A., & Haertel, E. H. (2016). Engaging methodological pluralism. In D. H. Gitomer & C. A. Bell (Eds.), *Handbook of research on teaching* (127–247). American Educational Research Association.

Oxford University Press (2023a, September). Theory (n., sense 3). In *Oxford English Dictionary*. https://doi.org/10.1093/OED/1108604610.

Oxford University Press (2023b, December). Approach (n., sense 5b). In *Oxford English Dictionary*. https://doi.org/10.1093/OED/3442162913.

Oxford University Press (2023c, December). Model (n., sense II.10). In *Oxford English Dictionary*. https://doi.org/10.1093/OED/3160470348.

Patton, M. Q. (1978). *Utilization-focused evaluation*. Sage Publications.

Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. Sage Publications.

Popham, W. J. (1975). *Educational evaluation*. Hall Press.

Quadrant Conseil, Strategic Design Scenarios (2022). *Impact Evaluation: A map of uses*. Agence Française de Développement. https://www.afd.fr/sites/afd/files/2023-01-04-20-43/impact-evaluation-map-uses.pdf

Scriven, M., & Davidson, E. J. (2021, November 12). *The transscientific theory of evaluation* [Paper presentation]. American Evaluation Association Annual Conference.

Shadish, W. R. (1998). Evaluation theory is who we are. A*merican Journal of Evaluation*, *19*(1), 1–19. https://doi.org/10.1177/109821409801900102

Shadish, W. R., Jr., Cook, T. D., & Leviton, L. C. (1991). *Foundations of program evaluation: Theories of practice*. Sage Publications.

Stufflebeam, D. L., & Coryn, C. L. (2014). *Evaluation theory, models, and applications* (2nd ed.). John Wiley & Sons.

UK Evaluation Society. (2012). *Framework of evaluation capabilities*. https://www.evaluation.org.uk/app/uploads/2019/04/UK-Evaluation-Society-Framework-of-Evaluation-Capabilities.pdf

United Nations Evaluation Group. (2016). *UNEG evaluation competency framework*. https://www.unevaluation.org/document/download/2610

UNEG Working Group on Evaluation Methods. (2020). *Compendium of evaluation methods reviewed (Vol. 1)*. United Nations Evaluation Group. http://www.unevaluation.org/document/detail/2939

Vaca, S. (2017). *Periodic table of evaluation*. https://twitter.com/VisualBrains/status/933239046658187265

Worthen, B. W., & Sanders, J. R. (1973). *Educational evaluation: Theory and practice*. C. A. Jones Publishing Company.