
Visualizing Evaluation Theory: Current Reflections and Future Projections

Melvin M. Mark
Penn State University

Journal of MultiDisciplinary Evaluation
Volume 20, Issue 48, 2024

JMDE
Journal of MultiDisciplinary Evaluation

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

Background: Among other potential benefits, evaluation theories are intended to serve as guides to evaluation practice. Scholars have offered alternative visual representations of multiple evaluation theories, their contents, and/or their interrelationships.

Purpose: I offer brief commentary on the evaluation theory visualizations in other articles in this special section of *Journal of Multidisciplinary Evaluation (JMDE)*. I suggest evaluators should be familiar with multiple representations. I also speculate about a future generation of visual representations which might be interactive, allowing evaluators to select the level of detail most useful to them.

Keywords: *evaluation theory; visual representations.*

Setting: Not applicable.

Intervention: Not applicable.

Research Design: Not applicable.

Data Collection and Analysis: Not applicable.

Findings: Not applicable.

While the primary function of theories of evaluation is to offer guidance about evaluation practice, other potentially important benefits exist (Mark, in press). The benefits of evaluation theory arise, or at least are greatly enhanced, when one is “multilingual” with respect to evaluation theory. Comparison of theories highlights practice options and points to issues on which thoughtful evaluators differ. Theoretically multilingual evaluators recognize the wide scope of evaluation practice and are better equipped to advocate against any overly narrow views of evaluation they encounter. In these and other ways, familiarity with evaluation theory can be a key part of the value proposition for evaluators. That is, the broader and deeper view that comes from being theoretically multilingual makes an evaluator poised to make greater contributions than someone with comparable methods skills but without a similar background in evaluation theory. Evaluation theory can also be a key part of our professional identity, as Shadish (1998) suggested with his assertion that “evaluation theory is who we are” (p. 1).

Several instances of what are called evaluation theories were developed in the early days of evaluation as a field (Shadish et al., 1991), and many more have been added since (Christie & Alkin, 2023). The existence of multiple evaluation theories offers advantages, but the multitude of theories can be overwhelming. In principle, visual representations of evaluation theory can help one make sense of an array of evaluation theories. Visual representations can show (dis)similarities across theories, highlight selected attributes of individual theories, suggest individualized directions for continuing professional education, and in other ways aid the study, understanding, and application of evaluation theories.

This brief article offers comments inspired by the visualizations in this issue of *JMDE*. The next section of the current paper contains selected reflections, especially on strengths and weaknesses of these visualizations. Comparing across the representations, I also suggest that just as evaluators should be multilingual with respect to evaluation theories, they should be multilingual regarding the visualizations. In a subsequent section, I speculate about a future generation of visualizations, perhaps more interactive ones that allow evaluators to control the level of detail and even the attributes presented.

Four caveats before proceeding. First, I’ve written this article without seeing the other articles in this special section. While prior information (e.g., previous articles by the authors) should suffice for my brief commentary, I cannot comment on the specifics of the accompanying articles.

Second, given word limits, I’ve had to ignore several important issues here. These include what an evaluation theory is, and the different ways evaluation theory might help guide evaluation practice (but see Mark, in press). Third, given the more inclusive coverage from some of the visualization developers, the term “approach” may at times apply better than “theory.” Fourth, I am not including here the visual representations themselves, which are available in the accompanying *JMDE* articles and elsewhere.

Brief Reflections on Existing Visual Representations

The grandparent of visual representations is the evaluation theory tree discussed by Christie in this volume, developed by Alkin and Christie (2004a) and expanded on in subsequent publications (e.g., Christie & Alkin, 2023). The theory tree has three branches, each corresponding to an important aspect of evaluation: use, methods, and valuing. Evaluation theories (or, in earlier versions, evaluation theorists) are listed, each on one branch, each theory’s placement indicating that the theory gives greater *relative emphasis* to that branch’s topic. This is not to say that a theory on, say, the use branch ignores valuing and methods—only that it gives relatively more emphasis to use.

An early criticism of the theory tree was that the theorists on it were primarily Western, white, and male. Offshoots of such criticism led Mertens to offer a revision of the evaluation theory tree, with four branches rather than three, the development of which she describes in this volume. To the use, methods, and valuing branches of the earlier tree, Mertens added a social justice branch. Alkin and Christie had included a few of Mertens’s social justice theories/theorists on the values branch, more so in the most recent version. In my view, arguments can be made for both the original three-branch and the revised four-branch tree. But that issue need not be adjudicated to comment on the tree as a visual representation of evaluation theory.

The initial motivation for the tree was an interest in the patterns of influence across evaluation theorists, in part because the “family trees” that capture influence patterns in many fields do not work in evaluation (Alkin, 2004). Based on theorists’ comments, Alkin (2004) and Alkin and Christie (2004b) concluded that, while other influences on evaluation theories exist (e.g., ideas from outside evaluation, a theorist’s personal experience), there is greater cross-theorist influence within than across branches. For

example, theorists on the use branch tended to report being influenced by other theorists on the same branch.

Thus, the theory tree, whether with three or four branches, highlights an important aspect of the relationship *across* evaluation theories: the lines of cross-theory influence (and the theory and practice characteristics associated with the relative emphasis of a given branch). The tree provides one mapping of the lay of the land of evaluation theory, which is valuable, and it does so based on one characteristic of the individual theories (relative emphasis on the three or four branch topics). However, the tree provides relatively little about each theory's *internal content*. Information about other aspects of each theory is absent from the display. And even the information on relative emphasis does not indicate *how* the theorist/theory deals with the emphasized factor. For example, Scriven as well as Guba and Lincoln appear on the valuing branch of the three-branch tree, though they take vastly different approaches to valuing.

In sum, the tree shows that a visualization can portray (one view of) the lay of the land of evaluation theory. And the tree broke ground for subsequent visual representations—including some that attend far more to multiple characteristics of each theory. It also would not be fair to criticize the tree (or subsequent visualization) developers for not doing something they did not set out to do. As we think about the other and especially future visualizations, however, we should consider which features should be highlighted, how, and with what (if any) complementarity across visualizations—recalling a word inapt in its roots but apt conceptually: multilinguality.

Azzam and Donaldson's visualization shows various streams of evaluation purposes. The visualization itself does not link to theorists/theories, but identifying such linkages is easy. One could even imagine transforming the river imagery into tree imagery, with one branch corresponding to each river/purpose. But the benefits of the river metaphor would be lost in translation. By showing all rivers flowing into a single body of water, Azzam and Donaldson's display shows that, across all purposes, evaluation practice is intended to contribute to the common, general end state they label societal improvement. The river visualization also clearly shows bifurcation of one purpose into two, such as when the social justice stream splits, adding the transformation of stakeholders as another purpose.

As with the tree and with visualizations in general, Azzam and Donaldson's display highlights

some attributes of the theories and not others. Evaluation purpose is nicely highlighted, but other aspects of the content of evaluation theories are not. This includes the critical matter of *how* to go about trying to achieve the purpose at hand. Of course, textual material or other displays could add such information. Again, displays by their nature tend to be selective, which holds especially for the most visually accessible ones.

Montrosse-Moorhead, Schröter, and Becho offer a more comprehensive but arguably more visually challenging visual representation. Their garden of evaluation approaches represents each evaluation theory as a flower. Various attributes of each theory are shown with parts of the flower: Paradigm is shown by the color; methodological approach (qualitative, quantitative, mixed) by a design in the center (i.e., pistil) of the flower; and eight other attributes (e.g., values, use) by the length of each of eight petals. The flower format shows the feasibility and the value of a more multidimensional visual representation. Moreover, as more theories are represented as flowers over time, this form of visual representation should allow cross-theory comparisons and the clustering of similar theories. But the display of a full garden, or even a sizable portion, will be rather busy visually. The multidimensionality has benefits, but especially when applied across multiple theories, it also offers challenges for the viewer, especially one new to evaluation theory. Still, the display of multiple attributes of each theory is impressive.

Vaca's periodic table of evaluation, inspired by the periodic table of elements, likewise is relatively busy, though it differs in focus from Montrosse-Moorhead and colleagues' flowers. The flowers each display the judged characteristics of a single evaluation theory. In contrast, though it includes some approaches that have also been called evaluation theories (e.g., realist, deliberative democratic), much of Vaca's periodic table includes other elements or features including paradigms, criteria, designs, and more. While not linked to evaluation theories in the table, these elements can come together to make up an evaluation theory (or an evaluation).

Lemire's evaluation metro map looks like a subway map but replaces station names with evaluation approaches and methods. Like other visualizations, the metro map illustrates the creativity that evaluators can apply using a visual metaphor. Moreover, the metro map offers a readable compendium of various evaluation theories and evaluation methods. However, it is not clear whether the ordering of the theory stops is meaningful—which might instead require

clustering similar theories in different lines, perhaps corresponding to the branches of the tree, or evaluation purposes, or another meaningful grouping. And as always with a visualization, some features are highlighted; others not (e.g., evaluation purpose and bifurcation à la Azzam and Donaldson).

Delahais and colleagues' map of evaluation uses explicitly focuses on impact evaluation. At the core of the map are four possible categories of impact evaluation uses (strategic, continuous improvement, management support, and dialogue support). Delahais and colleagues also identify and describe a range of impact evaluation methods, and they note, graphically and in writing, whether and how each method can serve the four evaluation uses. Future visualization developers should consider including and expanding on this kind of linkage, such as between intended evaluation use and method options (or other features).

More could be said about each of the visual representations briefly discussed here. Before turning to a few cross-visualization observations and some speculations about the future, I'll simply say that (a) all the visualization developers deserve thanks for their work, though (b) none of the visualizations to date is the (hypothetical) perfect visualization of evaluation theories, and so (c) just as evaluators should be multilingual with respect to evaluation theories, they should be multilingual in regards these visualizations.

Musings on Future Visual Representations of Evaluation Theory

In this section, I speculate about a future generation of evaluation theory visualizations. These speculations rest in part on the following observations about existing visualizations. First, visual representations can highlight selected characteristics of individual evaluation theories, or display the lay of the land in the sense of meaningful groupings of theories, or both. Second, as shown within some representations (e.g., the flower garden) and across the various visualizations, numerous characteristics of evaluation theories exist that may be of interest. Additionally, it seems likely that the characteristic(s) of greatest interest will vary across viewers (or within viewers over time). Third, as some visualizations (e.g., the map of impact evaluation uses) show, it is possible to display options that are nested within an evaluation approach, such as the method or design options that fit within a theory. This kind of nesting within theories would be quite helpful for some

visualizations, providing a kind of menu. Fourth, representations vary in terms of the number of elements of theory they display. And generally, a display featuring only one element (e.g., relative emphasis or evaluation purpose) is generally easier to make sense of visually than one featuring numerous elements (e.g., flowers or a table of elements), though it provides less information.

Expanding from this set of observations, at least some future visual representations should be designed both to map the lay of the land of evaluation theory and to highlight the key features of the individual theories and subsets of theories. Sequenced or hierarchical displays might help avoid overload, especially for displays with multiple theory characteristics. For instance, for a subcategory of similar theories, a meta-flower might show the common general pattern, with a subsequent multi-flower display showing the variations across the subgroup's theories.

Another way to manage the potential overload from too many features would be to give the viewer direct control. Imagine being able to choose which features are displayed. Or even being able to select an attribute or attributes as the basis for clustering theories into subgroups. For instance, one viewer might want to group theories based on who has decision-making control for the evaluation, and to see individual theories displayed in terms of a wide array of attributes. Another person might want groupings based on the evaluation criteria that theories prioritize, and then see a display with only a few select attributes of each theory.

The idea of giving viewers control over displays might send today's visualization developers into a paroxysm of despair (or hysterical laughter). However, for visualization developers a few years hence, armed with new tools based on what we have come to call artificial intelligence, this and other ideas about future displays might instead seem quite natural.

Imagine a future variation on the metro subway map that has different subway lines, each corresponding to a different subgroup of evaluation theories—perhaps based on characteristics from current visualizations such as the tree or streams, or grouped empirically according to an algorithm. Each subway line could have multiple stops, representing different stages of or variations in evaluation practice for that theory subgroup. Additionally, at each stop storefronts could present menus of options for that stage of evaluation (e.g., the array of research designs that could be used for an impact evaluation for strategic use). Historical markers or other signs could be included to provide more information for those interested. AI might

even make it possible to “ride” the subway and to linger at stops of interest.

Countless variations and extensions can be imagined. Expanding on the metro map, for example, distinct clusters of evaluation theories could be represented by different transportation modes: subway lines, bus lines, bike paths. Or a visualization might instead include multiple railway lines taking different routes across a country, with each line having a late stop that represents one of Azzam and Donaldson’s evaluation purposes. And the railway lines could all head into the neighboring country of Societal Improvement. Alternative visualizations might be explored, such as a building with wings that represent theory subgroups. Or a map might be of a housing development, with apartments, duplexes and houses clustered into distinct areas of the development—and attributes such as their color

and design representing different evaluation features.

AI-assisted visualization programs might make it easy to try out these and other global variations, as well as variations within a visualization (such as adding flower height or type to represent a theory dimension in the theory garden). AI could also generate drafts of text or spoken messages to be included in hybrid visual displays. It is easy to imagine a taped message from Michael Quinn Patton, or a quote from (or about) Stafford Hood or other evaluation theory luminaries. One can even imagine AI-enabled conversations between such individuals and the user riding the train or walking the garden.

In various ways, then, a future representation of evaluation theory may be hybrid, drawing on communication modalities beyond the visual. And a single representation in the future may in some sense be multilingual.

References

- Alkin, M. C. (2004). An evaluation theory tree. In M. C. Alkin (Ed.), *Evaluation roots: Tracing theorists’ views and influences* (pp. 373–380). Sage
- Alkin, M. C. & Christie, C. A. (2004a). An evaluation theory tree. In M. C. Alkin (Ed.), *Evaluation roots: Tracing theorists’ views and influences* (pp. 12–65). Sage.
- Alkin, M. C. & Christie, C. A. (2004b). Evaluation theory tree revisited. In M. C. Alkin (Ed.), *Evaluation roots: Tracing theorists’ views and influences* (pp. 381–392). Sage.
- Christie, C. A. & Alkin, M. C. (2023). An evaluation theory tree. In M. C. Alkin & C. A. Christie (Eds.), *Evaluation roots: Theory influencing practice*. Sage.
- Mark, M. M. (in press). Evaluation theories: Guidance to evaluating in various circumstances. In L. Bickman & D. Rog, (Eds.), *The evaluation handbook: An evaluator’s companion*. Guilford.
- Shadish, W. R. (1998). Evaluation theory is who we are. *American Journal of Evaluation*, 19(1), 1–19.
- Shadish, W. R., Jr., Cook, T. D., & Leviton, L. C. (1991). *Foundations of program evaluation: Theories of practice*. Sage.