A Map Of Uses Of Impact Evaluation Approaches

Journal of MultiDisciplinary Evaluation
Volume 20, Issue 48, 2024

JMDE
Journal of MultiDisciplinary Evaluation

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

Thomas Delahais
Quadrant Conseil

François Jégou
Strategic Design Scenarios

Agathe Devaux-Spatarakis
Quadrant Conseil

Karen Rousseau
Agence Française de développement

Background: Not all evaluation practitioners are experts in the diversity of impact evaluation approaches. How can these approaches be made clear and accessible? How can practitioners be better able to choose the relevant one? A collaborative project was initiated in 2021 to address these questions and involved the Agence Française de Développement (AFD) evaluation officers, external evaluators, and public-sector codesigners.

Purpose: This article presents the development of a visual tool designed to facilitate meaningful discussions between evaluation officers and project managers regarding an array of available impact evaluation approaches. In our view, the positive reception of the map of uses of impact evaluation approaches within and beyond the AFD underscores the significance of anchoring discussions on evaluation scope and methodologies in real-world evaluation uses.

Setting: Not applicable.

Intervention: Not applicable.

Research Design: Not applicable.

Data Collection and Analysis: Not applicable.

Findings: Not applicable.

Keywords: *visualization; evaluation use; impact evaluation; concept mapping; codesign.*

Over the past 20 years, the evaluation community has developed a wide range of ways to evaluate the impact of policy interventions. We define impact evaluations (IEs) as approaches rooted in different paradigms of causality, meaning different perspectives on the world or the role of evaluation, which all aim to estimate the effects of interventions through specific causal inference processes. The broad range of IE designs has the potential to make evaluations more relevant and useful to various stakeholders (Stern et al., 2012). However, our experience has been that despite a number of available resources, in France and possibly other countries, only a fraction of evaluation practitioners might be able to tell the difference between realist evaluation and process tracing, two IE approaches, let alone which one would be best in a given context. If evaluators do not know of these various approaches, not to mention their attributes and differences, how can they use them?

This question was the starting point of a project launched in 2021 by the Evaluation and Learning Department of the Agence Française de Développement (AFD), which resulted in a map of uses of impact evaluation approaches (Delahais et al., 2022). [1] In this article, we present the development of this project, consider some of the visualization challenges associated with the tool, and discuss the map's reception and current uses. Finally, we outline some of the implications of the process and its output.

Motivation for the Map of Uses of Impact Evaluation Approaches

In the early 2020s, the AFD's evaluation unit was in the process of rethinking how it supported and managed its (mostly external) evaluations. In the previous decade, the term "impact evaluation" (IE) had been reserved, within AFD, to refer mainly to quasi-experimental designs, and impact questions were formulated accordingly (Naudet et al., 2012). The potential of the diversity of IE approaches that had emerged in recent years had been left mostly untapped. Therefore, the initial aim of this project was to help the evaluation team explore and, where appropriate, prescribe these approaches. In the longer term, the hope was to make these new IE approaches more credible and visible. The means to do so were open.

Development of the Product: Initial Attempts

The evaluators and codesigners [2] involved in the project had a longstanding commitment to the diversity of IE approaches, and to working together in policy-making and evaluation (Delahais et al., 2019). They had previously co-elaborated a card game to address evaluation barriers, *Evalophobia* (Sage et al., 2017). The initial attempt was to develop a new game in which the players would start with all possible IE approaches, listed on a set of cards. Depending on the characteristics of the intervention being evaluated, evaluation questions, and context, IE approach cards would be discarded until only the most relevant ones remained. The proposal was quickly rejected by the team at AFD, who was concerned that a game—even a serious one—would work against the long-term objective to reinforce the credibility of alternative IE approaches.

The second attempt focused on the needs of the project managers (PMs) who commission evaluations. PMs are responsible for the preparation, contracting and monitoring of AFD-funded projects at headquarters or in country offices. Initial interviews conducted by the project team had revealed that PMs knew much less about what various IE approaches implied than did evaluation officers (EOs). The idea was to facilitate a conversation similar to that of a customer with a salesperson selling household appliances; that is, where the customer has only a rough idea of their needs, but the salesperson presents the main features of the products and helps them refine their needs. To do so, the team developed a prototype, a one-page narrative of a fake evaluation process rich with distinct features of IE approaches.[3] Reactions of AFD's EOs were mixed. They saw a pitfall of getting caught up in a methodological exchange rather than a conversation about the objectives and uses of the exercise. Clarifying the origin of discomfort with the tool revealed that the EOs saw themselves, not the PMs, as its primary users. They wanted a tool that would not be prescriptive but

[1] The map is available on the AFD website under a free license: https://www.afd.fr/en/ressources/impact-evaluation-map-uses

[2] Codesigners are professionals trained in industrial design who apply a user-driven approach to public sector challenges (Bason, 2014).

[3] Visuals of the project at different steps are available on the Open Science Framework (OSF): https://osf.io/mt7ns

would help them have substantive discussions with PMs about IE designs.

A Concept Mapping Approach to Identify Categories of Uses

The final attempt aimed at helping EOs suggest a range of IE approaches started no longer from methodologies but instead from the uses that PMs could make of them. The first challenge was therefore to identify these uses. AFD staff routinely relied on a dichotomy between *accountability* and *learning* uses. However, initial interviews had shown that these concepts were not useful in making choices in terms of evaluative designs, and rather amounted to obfuscating "aidspeak" (Cornwall & Eade, 2010). We did not want to replace them with other concepts of use, which would still sound abstract to those directly concerned. Rather, we wanted to have categories of use that would be meaningful to AFD staff. To do so, we followed a structured process inspired by concept mapping, through which it is possible to develop conceptual categories from a series of statements (Trochim, 1989). Here, these statements were formed on concrete stories of use we had collected in our interviews, which we supplemented with examples illustrating theories of use not represented in the initial collection (Delahais & Devaux-Spatarakis, 2022). We reached a list of 50 prompts, then gradually reduced it to 31.[4] A workshop was then organized, in which 12 out of 17 EOs participated. They rephrased and then sorted those prompts freely into as many groups they deemed relevant, and labeled them. PMs proved unavailable for an in-person workshop, but 11 of them agreed to do a similar activity online.

The analysis was performed using a multidimensional scaling approach and k-means clustering to identify the prompts that had been grouped together most frequently. A number of clusters emerged initially, but this first classification was nuanced by the differences between the 12 EOs and 11 PMs who had distinctly sorted evaluation uses differently and used different wordings to label their groupings. For instance, EOs were more likely to group uses related to project management and the long-term improvement of interventions under the label of "operational use," whereas the PMs in charge of operations kept them apart and would not use such a term to describe their uses.

Several groupings were tested before negotiating the final clustering, emphasizing meaningfulness over statistical distribution for boundary items (Euréval & Delahais, 2010). The final formulation of each category of use was done collectively, using the language that participants used in labeling groups in the process. This ultimately resulted in 4 groups of uses (strategic, management support, continuous improvement, and dialogue) and 10 groups of sub-uses. This final grouping prioritized the views of the PMs over those of the EOs.

Visualizing the Results of the Concept Mapping

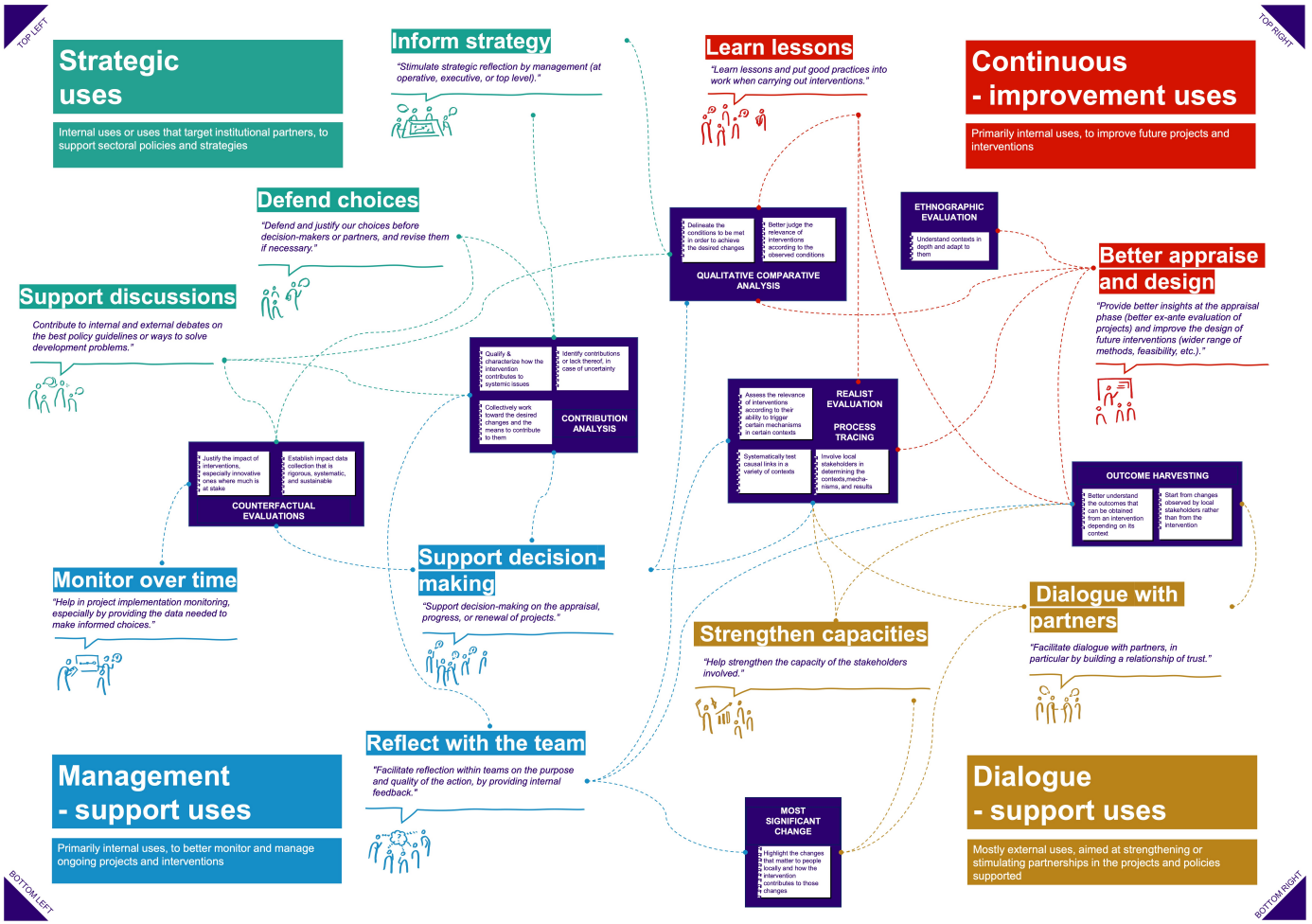
The next challenge was to move from statistical representation to an actual tool. In concept mapping, coordinates and axes have no value in themselves. Our choice was then to display the categories and subcategories of use into four quadrants. Seven major IE approaches (Delahais, 2017) were associated with the uses through double-ended arrows in our first map prototype.

This prototype was tested with the EOs in a workshop. The participants were divided into subgroups and given a fake evaluation request in the form of a short email outlining a PM's expectations. The aim was to identify the potential uses that emerged from the analysis of the email using the tool, and then to propose a range of IE approaches relevant in that case. The immediate reaction of the EOs was that the map, with its numerous blocks and connections, was difficult to read. However, once they had got past this initial stage, they were able to use it relatively easily. Post-workshop feedback was quite positive and focused on simplifying the map and using it for conventional evaluations.

The final map is simpler to read and use. It is provided as a deck of PowerPoint slides and organized in concentric circles. It comes in two versions, one with and one without the IE approaches (Figure 1 and Figure 2 respectively). The decision to create two versions was taken to accommodate a wider range of utilizations of the tool, such as discussing the uses of evaluation without specifying an approach, and collaboratively exploring the best balance between the expected uses and the complexity or cost of the approach. Each approach is also described separately in a short handout.

[4] The list of prompts is available on OSF: https://osf.io/mt7ns

Figure 1. The Map of Uses of Impact Evaluation Approaches



Note. From "Impact Evaluation. A Map of Uses" by Agence Française de Développement, 2022, p. 3. https://www.afd.fr/en/ressources/impact-evaluation-map-uses . CC BY-NC-SA.

Figure 2. The Map of Uses Without Impact Evaluation Approaches

Note. From "Impact Evaluation. A Map of Uses" by Agence Française de Développement, 2022, p. 3. https://www.afd.fr/en/ressources/impact-evaluation-map-uses . CC BY-NC-SA.

Benefits and Limits of the Visual

We see several benefits of the maps of uses—both with and without impact evaluation approaches included. Visuals that show the diversity of IE approaches are scarce, even in recent guides (e.g., Befani, 2020; Vaessen et al., 2020). Our map is an exception. Moreover, the primary objectives were to allow users to grasp all the content at once and use the tool intuitively, and several design features support this. Its two-dimensional environment, similar to a geographical map, facilitates orientation in a complicated topic. The hierarchy of blocks and text sizes and the use of colors make it easy for non-specialists to immerse themselves in the map. The proximity between uses and approaches provides an intuitive understanding of the different families of practices. The connectors facilitate navigation, both by showing a range of approaches relevant to a given expected use, and by displaying those alternative uses for which the selected IE approach was particularly good. Lastly, the point was not to weigh up the attributes of each approach to arrive at the most rational choice, as in conventional guidance, but to clarify the uses and agree on the range of IE solutions that could be deployed. The tool facilitates this non-technical discussion in which all parties can have their say. The choice of the most appropriate approach is left to further discussion between EOs, PMs and other stakeholders.

The visualization also shows its limitations. In the graphic design, we struggled to produce a simple map that still respected the complex system of relationships between evaluation uses and approaches. The choice of a simple, low-tech format—with the same content in both the self-printing tool and the electronic whiteboard for remote collaboration— also placed more constraints on the graphic design. However, it was positive in terms of adoption by users, who are not always comfortable with the latest digital tools. Yet, it only includes the language and views of internal users, excluding potential users outside AFD, such as those in national or local authorities, in Non-Governmental Organisations, in companies, etc. Whether positive perceptions and use consistently hold true for users outside of AFD is an open question.

Actual Use of the Map

Some intended uses have been realized; others have not. Within AFD, the map of uses of impact evaluation approaches was instrumental in positioning and lending credibility to the diversity of existing IE approaches. In general, EOs who were involved in the process now share a similar framework for discussing uses and IE approaches. It is now included in the training provided for new EOs. The four types of uses from the map are listed in the standard terms of reference (ToRs) for project evaluation. What the map has not fully achieved, however, is a change in the interaction between PMs and EOs, which was the initial objective. EOs, rather, use it as a visual reminder of uses and approaches to think about.

The tool was published online in October 2022 under an open-access license in French, English, and Spanish. AFD agreed to allow this only if the map was presented as a generic tool, even though the uses presented were quite specific to AFD. By the end of 2023, the map had been downloaded around 3,200 times in French, 3,000 times in English, and 2,100 times in Spanish, making it one of the most downloaded evaluation documents on the AFD website.

The map was presented at various French and international evaluation events, where it was well received by evaluation practitioners. The format of the map and the choice of simple words to describe the uses, making it accessible to evaluation neophytes, were generally praised. Attendees often appreciated that it helped them discuss expected uses as a way of choosing an IE approach, rather than the other way around. In fact, the map presenting the uses without the IE approaches generated more interest than its more complete counterpart.

Concluding Thoughts

The initial aim of this project was to help the evaluation team explore the diversity of IE approaches. Our initial strategy was to describe the characteristics of these approaches better to facilitate choice. The final product, however, reveals and highlights evaluation uses as an entry point to discuss IE. This was made possible by a collaborative trial-and-error process typical of codesign, but not of the evaluation market, as framed by procurement rules.

The process of starting from actual uses rather than from theoretical or institutional categories is part of a larger effort to evaluate research, evaluation, and other knowledge products in a way that is closer to what knowledge and knowledge-generation processes do (Delahais & Devaux-Spatarakis, 2022; Delahais & Lacouette-Fougère, 2019; Delahais & Quadrant Conseil, 2022; Delahais & Toulemonde, 2017). But in the context of this

project, the inclusion of codesign professionals was invaluable in lending credibility to this approach. The designers acted as embedded dummies, defending a user perspective and usability concerns and guarding potential users against evaluation jargon and bureaucratic language. In the end, the language of PMs and EOs was used as much as possible. The map does not include the dreaded words "accountability" and "learning," nor does it contain the language of evaluators, which, in our view, is no less incomprehensible to project managers. There was also no normative consideration of what constitutes a good use or a misuse (Alkin & King 2017).

Ultimately, this map contributes to the view that making evaluation known and useful requires specific processes of knowledge brokering, which are separate from the usual evaluation activities, and which acknowledge the logics of policy making outside of evaluation (Mc Sween-Cadieux et al., 2019). In a way, this work can be seen as a manifestation of E. Chelimsky's remarks:

> Looking at evaluation only from an evaluator's perspective might cause us to underestimate, misinterpret, or rule out purposes for evaluation that we would recognize as valid if we saw them from a different, broader perspective. We're not unlike those ants, asked to write a zoology paper, who divided the animal kingdom into two classes: the kind, gentle beasts such as the lion, tiger or jackal, and the ferocious ones like the chicken, duck or goose. (Chelimsky, 2006, p. 34)

References

Alkin, M. C., & King, J. A. (2017). Definitions of evaluation use and misuse, evaluation influence, and factors affecting use. *American Journal of Evaluation*, *38*(3), 434–450. https://doi.org/10.1177/1098214017717015

Bason, C. (2014). *Design for policy*. Routledge.

Befani, B. (2020). *Choosing appropriate evaluation methods: A tool for assessment & selection (version 2)*. Centre for the Evaluation of Complexity Across the Nexus. https://www.cecan.ac.uk/wp-content/uploads/2020/11/Final_Choosing-Appropriate-Evaluation-Methods-1.pdf

Chelimsky, E. (2006). The purposes of evaluation in a democratic society. In I. Shaw, J. Greene, & M. Mark (Eds.), *The Sage handbook of evaluation* (pp. 34–55). Sage Publications Ltd. https://doi.org/10.4135/9781848608078.n1

Cornwall, A., & Eade, D. (Eds.). (2010). *Deconstructing development discourse: Buzzwords and fuzzwords*. Practical Action Publishing and Oxfam. https://oxfamilibrary.openrepository.com/bitstream/handle/10546/118173/bk-deconstructing-development-buzzwords-010910-en.pdf?sequence=1

Delahais, T. (2017). *How can impact be evaluated? An impact evaluation approach tree*. Quadrant Conseil. https://www.quadrant-conseil.fr/ressources/impacttree.html

Delahais, T., & Devaux-Spatarakis, A. (2022). *Qu'attendre de la recherche pour éclairer l'action publique?* [What can we expect from research to inform public action?] (Papiers de recherche 229). Agence Française de Développement. https://www.cairn.info/qu-attendre-de-la-recherche-pour-eclairer-l-action--1000000148971-page-1.htm

Delahais, T., Devaux-Spatarakis, A., & Jegou, F. (2022). *Se redonner du choix dans les approches d'évaluation d'impact: Rapport de recherche-action* [Regaining choice in impact evaluation approaches: Action research report]. Agence Française de Développement.

Delahais, T., Gouache, C., & Vincent, S. (2019). Le design de l'action publique: Vers une hybridation entre culture design et culture de l'évaluation [The design of public action: Towards a hybridization between design culture and evaluation culture]. *Sciences du Design*, *10(2)*, 83–89. https://doi.org/10.3917/sdd.010.0083

Delahais, T., & Lacouette-Fougère, C. (2019). Try again. Fail again. Fail better. Analysis of the contribution of 65 evaluations to the modernisation of public action in France. *Evaluation*, *1(3)*, 1–18. https://doi.org/10.1177/1356389018823237

Delahais, T., & Quadrant Conseil. (2022). *External evaluation of MOPAN 2015–2021*. Multilateral Organisation Performance Assessment Network. https://www.mopanonline.org/aboutus/evaluations/items/MOPANEvaluation_Final_August_web%20(1).pdf

Delahais, T., & Toulemonde, J. (2017). Making rigorous causal claims in a real-life context: Has research contributed to sustainable forest management? *Evaluation*, *23(4)*, 370–388. https://doi.org/10.1177/1356389017733211

Euréval & Delahais, T. (2010). *Étude portant sur la contribution de la traduction à la société multilingue dans l'Union européenne* [Study on the contribution of translation to multilingual society in the European Union].

Commission Européenne – Direction Générale de la Traduction. Publications Office of the EU. https://op.europa.eu/fr/publication-detail/-/publication/7d0a5d03-5c81-44df-9861-430a3d1c535f

Mc Sween-Cadieux, E., Dagenais, C., Somé, D. T., & Ridde, V. (2019). A health knowledge brokering intervention in a district of Burkina Faso: A qualitative retrospective implementation analysis. *PLOS ONE*, *14(7)*, 1–21. https://doi.org/10.1371/journal.pone.0220105

Naudet, J.-D., Delarue, J., & Bernard, T. (2012). Évaluations d'impact: Un outil de redevabilité? Les leçons tirées de l'expérience de l'AFD [Impact evaluations: An accountability tool? Lessons learned from AFD's experience]. *Revue d'économie du développement*, *20(4)*, 27–48. https://www.cairn.info/revue-d-economie-du-developpement-2012-4-page-27.htm

Sage, K., Jegou, F., Gouache, C., & Honoré, V. (2017). *Evalophobia*. http://www.sustainable-everyday-project.net/evalophobia/

Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., & Befani, B. (2012). *Broadening the range of designs and methods for impact evaluations: Report of a study commissioned by the Department for International Development* (Working Paper No. 38). Department for International Development. https://www.oecd.org/derec/50399683.pdf

Trochim, W. M. (1989). An introduction to concept mapping for planning and evaluation. *Evaluation and Program Planning*, *12(1)*, 1–16. https://www.sciencedirect.com/science/article/pii/0149718989900165

Vaessen, J., Lemire, S., & Befani, B. (2020). *Evaluation of international development interventions: An overview of approaches and methods*. World Bank. https://doi.org/10.1596/34962