

# How It All Started and How It Evolved: Ray's Whip

Nicoletta Stame

*A Colorni-Hirschman International Institute*

**Background:** This article reconstructs the origin of the concept of “streams of evidence” in the collective work of the INTEVAL group, as proposed in the book *From Studies to Streams* (Rist & Stame, 2006). Contrary to insisting on the (scarce) utilization of single studies, answering the evaluation demands of static organizations based on their “own incentives and reward structures” (p. 284), Ray challenged the group to engage with the new scenario of the learning organizations “that embraced learning and knowledge as fundamental to their own existence and success” (p. 284). Such organizations needed evaluative knowledge that was “timely, proactive and drawing on multiple sources of knowledge” (p. x), that could be provided by different sources, evaluations, and research, internal or external to the organizations.

**Purpose:** The article aims to show the farsightedness of Ray's intuition. He was able to capture the spirit of time and to encompass different converging trends. His intellectual leadership oriented other people's work, creating a wide space for innovation and adaptation. Moreover, in the new current conditions of pandemics and multi-crises, Ray has

been able to revive and update that thrust, adding a multidisciplinary dimension to the idea of streams.

**Setting:** The concept of streams of evidence is a powerful and versatile metaphor. New instances of flowing stuff can be added to the original taxonomy, based on the source of the evidence and on its organization. It has entered many different fields as an autonomous cognitive tool. In evaluation methodology it has contributed to meta-analysis, realist syntheses, and a broad reinterpretation of mixed methods. In the study of evaluation systems, it has provided the basis for organized practices, from well-tempered forms of knowledge management to the elaboration of evaluation demands through learning agendas.

**Intervention:** Not applicable.

**Research Design:** Not applicable.

**Data Collection and Analysis:** Not applicable.

**Findings:** Not applicable.

**Keywords:** *studies; streams; learning organization; evaluative knowledge.*

---

At the beginning the project on studies/streams<sup>1</sup> focused on the typical question, “Do commissioners or policymakers use our evaluations?” If a direct (“instrumental”) use was not in sight, the prospective co-authors were playing with what Kirkhart had called “influence” (and that Weiss had before called “cognitive” use). In any case, this would have always implied starting from our own work (producing the evaluation) and our frustration for its scarce use, in a one-on-one relationship between evaluator and commissioner.

When we realized that we would have ended up saying nothing deserving a new study, Ray turned the whole argument upside down: Evaluations were useless, because they were not up to the expectations of the actor that we were wooing at the time, when evidence and results had become the mantra of the day: the learning organizations that were creating their own knowledge management systems.

Ray defined learning organizations as “high performance organizations, that embrace learning and knowledge as fundamental to their own existence and success” (2006b, p. 284). Learning organizations needed evaluative thinking and information of various kinds: Evaluations could indeed be very useful if only they took up an approach fitted to those expectations.

Hence, we moved from use (of “information” that had already been produced) to “knowledge” (transforming information into learning, action, change). Better, we moved from use of a single study to learning from a multiplicity of information sources.

Ray constructed the argument this way. Single studies are irrelevant, and therefore useless, for the following reasons:

- They are the product of an administrative way of working, that is not interested in getting new knowledge, and where evaluators are just professionals of a single expertise: evaluation amnesia (same questions asked time and again) on the part of the organization, evaluation routine on the part of the evaluator. Ray noted that in such circumstances “the growing disillusionment about the ability of evaluation to contribute to organizational knowledge management gains strength” (2006b, pp. 284–5). These remarks were a whip to challenging the status quo and the comfort level of the evaluation community, used to answering the evaluation demands of static organizations

based on their “own incentives and reward structures” (p. 284).

- They are mostly in the form of ex post evaluations. They look backward. They deliver results when decisions have been taken, and they are not anymore needed. Ray distinguished between single studies that want to “prove” something and evaluation that wants to “improve” the situation or, in this case, an organization.

What would be relevant, and necessary at a time when organizations are interested in results, “is the expectation, indeed, the demand, for streams of real-time evaluative knowledge within high performance organizations as a tool for managing successfully” (2006b, p. 285). The example of the New York City Police Department, with its absorption of daily data to change the deployment of agents, is an example of “an evaluation system that is relevant at all stages of a policy, program or project” (2006a, p. 13).

Such knowledge would emerge from:

- Bringing together different kinds of information, not only evaluations, and not only produced by the organization itself: From performance management to social research, any kind of knowledge may be relevant. This was the object of Ray’s chapter “The ‘E’ in Monitoring and Evaluation – Using Evaluative Knowledge to Support Results-Based Management Systems.”
- Being produced by the time decisions were taken (and not ex post as with evaluations): This was the object of Ray’s other chapter, “Conclusion: A Brief Critique.”

Ray’s intuition that learning organizations needed an evaluative knowledge that was *timely, proactive, and drew on multiple sources of knowledge* was encouraging and motivating for evaluators (the power of leadership!). At the same time, it allowed us taking advantage of new approaches that had already been introduced around the idea of working on a multiplicity of studies: thematic evaluations, syntheses, integrative reviews.

As Perrin showed in his chapter, however, not all the kinds of syntheses stood up to the challenges Ray had evoked. Yes, they were timely (they provided already existing information) and they

<sup>1</sup> The project that became the volume co-edited by Ray Rist and Nicoletta Stame, 2006, *From Studies to Streams: Managing Evaluative Systems*.

could be used proactively. But not all of them would guarantee the efficiency of the learning organization. It depended on the criteria for selection of data, and the aim of the analysis.

While some kinds of syntheses aim at finding the “best practice,” to be “generalized” everywhere, Perrin identified that the “needs and expectations of the learning organization” would be better met by the opposite idea of “good practice.” The worth of syntheses would consist in comparing different situations and contexts to find out “what works better for whom, where and why” (the precepts of the realist synthesis). Hence, learning would be linked to the idea of “adapting” the good practice to the situation of interest (the daily practice): a process to be unfolded rather than a product to sell everybody.

Thus, we reoriented our research by asking what the hydraulic metaphor meant. What do streams look like? The many tributaries to a river, as when evaluations converge with other kinds of information into a river? The continuous flow of information? All of this, and more. Moreover, beyond conceptual differentiation, we inquired about the way it worked in different contexts (kinds of politics, organizations, institutions), etc.

The result was a series of chapters that corresponded to a taxonomy of streams, constructed according to whether they were (Yes/No) channeled (organized into cognitive or organizational tools) and whether they were composed (Yes/No) only of evaluative information or also information coming from other sources.

## How It Evolved: Two Areas of Inquiry

The concept of streams came thus to cover two different domains of evaluation, which correspond to the double inspiration of the book: on the one side, evaluative knowledge management by “learning organizations” and, on the other side, the new approach to evaluation proposed by syntheses. Yet, they took up somewhat different meanings from how they had been discussed there.

### *First, Evaluative Knowledge Management*

<sup>2</sup> The World Bank Open Knowledge Repository (<https://openknowledge.worldbank.org/home>) presentation reads, “The World Bank is the largest single source of development knowledge. The World Bank Open Knowledge Repository (OKR) is The World Bank’s official open access repository for its research outputs and knowledge products.” Such a repository was

Entranced by the concept of evidence-based politics, many agencies redefined themselves as knowledge organizations, knowledge agencies, etc. Foremost is the example of the World Bank. At the time one could read essays titled like “The World Bank as a Knowledge Agency” (Toye & Toye, 2005), or Stone’s (2003) “The ‘Knowledge Bank.’”<sup>2</sup>

Such a casual use of KM would remind us of Patton’s (2001) suggestion to distinguish between “Knowledge Management ... a process that harvests and shares an organization’s collective knowledge to achieve breakthrough results in productivity and innovation (and) Information Management that merely collects, processes, and condenses information.”<sup>3</sup>

Yet, the idea of KM has often been referred to bringing together knowledge from the organization itself, more than from multiple sources of knowledge, missing the innovative idea of the affluents to the streams.

Another institutional reception of the idea of streams can be found in the learning agendas (LA). Started by the U.S. Office of Management and Budget (OMB) but also adopted in other national evaluation systems (like in Poland), LA are ways of collecting the existing knowledge, from whatever source, on given policies in order to improve those systems’ evaluation plans, and commissioning (Newcomer et al., 2022).

### *Second, Syntheses*

In the years following our book the approaches to syntheses multiplied. Yet, the criteria set by Perrin (best practice vs. good practice) is still the main divide.

The field was mostly occupied by systematic reviews, championed by the Campbell Collaboration, and *3ie*, according to whom evidence is assured by the criterion of the hierarchy of methods (randomized controlled trials being at the top). Studies would be included in a review according to protocols that established the scope of the review, criteria for the selection of cases, quality assurance of the study, with the aim at eliminating the studies that do not correspond to the criteria, reducing them to a small number, from where to draw generalizations about best practices, no matter the different contexts.

intended as a place where one could find all the evaluations conducted by the WB, filtered by topics and areas of intervention.

<sup>3</sup> In an article critical of “best practices” and “lessons learned mania,” Patton (2001) was quoting from *Executive Edge magazine* (1998, p. 16).

Other approaches to syntheses were more inclusive. Petticrew and Roberts (2006) listed many different kinds of reviews (meta-analysis, narrative reviews, and other kinds of less “systematic” reviews) that used not only quantitative methods of research but also qualitative, and that were not limited to asking “what works” but also inquired into the meaning of an intervention.

On its own side, realist syntheses became better defined with Pawson (2006). Instead of asking whether the same treatment produces the same effects on different populations, realist syntheses inquire whether the program theories underpinning different environments—policy, disciplines, organizations (p. 178)—can explain what works for whom in what circumstances and why. Since interventions “are never implemented twice in the same way,” and are reinterpreted differently by participants, “understanding how a particular intervention works requires a study of the fate of each of the many, many intervention theories” (pp. 170–1). Realist syntheses are based on these principles: Do not aggregate data but compare theories. Do not base the analysis on positive cases only, but adjudicate between positive and negative cases. Do not impose best practice as models, but understand how the most appropriate example is adopted.

In all these cases, the idea of streams is utilized as a way of avoiding repetition, but they are based on long-term assessment instead of when “decisions are taken.”

### *Evaluative Streams Today*

What has changed in the last few years is the object of the evaluations, and the streams that should flow. When the aim is no more to evaluate a single policy, but to attain the SDGs, the focus is on the need to fight convergent causes (social, economic, and ecological) of multiple crises (wars, climate change, inequality). The streams of knowledge to be pooled together would need to suit this new feeling: to come from different disciplinary areas (no silos), to really incorporate tacit and explicit knowledge, to understand the consequences that any decision could have on other people, domains, contexts. The concept of an “ecology of evidence” conveys this outlook.

This more complex set of knowledge sources meets nonetheless the same concerns and urgency for timely and proactive knowledge, as when the streams were initially conceived. In this vein, Ray has led the study on the answers to COVID-19 by evaluators (Eliadis et al., 2023). One could sense his

feelings and moral commitment. COVID-19 presented an unparalleled challenge that had devastating effects: It needed a quick answer from a set of competencies and disciplines: biology, psychology, health, environment, security, etc. No single study could do; no previous study was available for such a new occurrence.

The results were less encouraging than with the example of the New York City police, but the whip was cracked as well. There were a lot of lessons to learn about the consequences of how the pandemic was being treated, among which was a need for evaluators to be better prepared for future shocks, and to provide immediate answers. The following articles demonstrate how the challenge of achieving streams of knowledge was tackled in the Global Environment Facility (GEF; Van den Berg and Uitto) and offer an exploration (Ling) of how streams of knowledge could nowadays be explored and managed in innovative ways.

### References

- Eliadis, P., Naidoo, I., & Rist, R. (2023). *Policy evaluation in the Era of Covid 19*. Routledge.
- Newcomer, K., Olejniczak, K., & Hart, N. (2022). Learning agendas: Motivation, engagement, and potential. *New Directions for Evaluation*, 173, 63–83.
- Patton, M. (2001). Evaluation, knowledge management, best practices, and high-quality lessons learned. *American Journal of Evaluation*, 22 (3).
- Pawson, R. (2006). *Evidence-based politics: A realist perspective*. Sage.
- Perrin, B. (2006). How evaluation can help make knowledge management real. In R. Rist & N. Stame (Eds.), *From studies to streams: Managing evaluative systems*. Transaction Publishers.
- Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences*. Blackwell.
- Rist, R. (2006a). The ‘E’ in monitoring and evaluation: Using evaluative knowledge to support results-based management systems.” In R. Rist & N. Stame (Eds.), *From studies to streams: Managing evaluative systems*. Transaction Publishers.
- Rist, R. (2006b). Conclusion: A brief critique. In R. Rist & N. Stame (Eds.), *From studies to streams: Managing evaluative systems*. Transaction Publishers.
- Rist, R., & Stame, N. (2006). *From studies to streams: Managing evaluative systems*. Transaction Publishers.

- Stone, D. (2003). The 'knowledge bank' and the Global Development Network. *Global Governance*, 9, 43–61.
- Toye, J., & Toye, R. (2005). *The World Bank as a knowledge agency*. United Nations Research Institute for Social Development.