

A Purpose-driven Action: The Ethical Aspect and Social Responsibility of Evaluation¹

Miri Levin-Rozalis

Davidson Institute at the Weizmann Institute of Science

The American Evaluation Association (AEA) conference always makes my head spin: thousands of people running from hall to hall, hundreds of sessions discussing an infinite number of subjects, the sound of masses of people walking, talking, meeting, looking around, seeking a familiar face or immersed in the program and trying to find the hall they want to get to. Amid this tumult I heard a man's voice call my name in a French accent. A dear colleague from Paris.

"It's been a long time," he said loudly, trying to speak over the din. "Shall we get away from the noise?" he shouted. "It's a lovely day outside."

And so I found myself sitting opposite him on a café veranda on Baltimore's beautiful waterfront.

"I've got a problem that's worrying me," he said. "I've just completed a comprehensive evaluation in conjunction with a government ministry, involving activities with migrants. The ministry is investing a great deal of money, energy, and goodwill, and in the short term, they're improving things— the program is working well in the framework of its rationale and goals, it's meeting the schedule, it's within budget, and it's reaching its planned goals in the best possible way. The problem is that I feel this is done at the

¹ This article is based on chapter 9 in my book: "Let's Talk Program Evaluation in Theory and Practice" Wachtman's sons: CA <http://www.amazon.com/Program-Evaluation-Theory-Practice-Lets/dp/1888820632>

expense of the community and family infrastructure. The program staff are building up a young, educated, modern leadership, one that's easier for the authorities to live with, while shunting aside the traditional leadership. Their point of departure is that it's for the community's benefit: the young leadership will lead it toward better integration into the general population. But I think that these young, educated guys can't lead the community. And what will happen is that they will be integrated into the general population, while the majority of the not well educated community, will be unable to follow them. And since the traditional leadership is being eased out as a result of the intervention, and the strong younger people are leaving, what will remain is a weakened, alienated, headless community that can't find its way. And that, I fear, is a long-lasting tragedy. I'm afraid that such a community will provide endless work for a long line of social workers, perhaps even the police, not to mention the vacuum that will allow various kinds of radicals to fill."

"And you are deliberating over the extent to which you, as a **professional**, have the right to push your own worldview at the expense of that of the intervention."

"Yes. I'm asking about the evaluator's boundaries as a professional. And moreover, I'm asking a moral question about the evaluator's right, or even duty to intervene in matters that aren't part of his job description."

"We've learned that evaluation is judgmental," I said, "because it gives value to the things it evaluates. But we've also learned—with great emphasis—that as evaluators, we're not allowed to be judgmental. We can't allow our personal judgment to bias the evaluation."

“That’s precisely what I’m asking. First, about my right, as a professional, not a private individual, to judge or, in other words, to allow my biases or inclinations or values to lead my considerations. And second, about my criteria for valuing my own judgment over the judgment of others.”

“At the first conference on evaluation I attended many years ago, somebody stood up and spoke about the unholy alliance between evaluators and program operators or their funding bodies. He contended that, by the very fact that evaluators examine a program’s inner workings and not its broader context, they form that alliance. His words have remained with me since then, and I always try to examine which unholy alliance I’m serving.”

“It’s quite similar,” he said sadly. “If I keep quiet, I really am forming an unholy alliance with the intervention, and I feel bad about it. If I speak out, I’m exceeding my authority as an evaluator. Perhaps they’re right in what they’re doing and there’s no alternative but to break up the traditional community.”

“So you’re also asking whether, as an evaluator, **you have the right** to judge the intervention’s goals.”

“You’re right in part,” he agreed. “Evaluation usually examines internal goals, which creates the unholy alliance you’re talking about. It sometimes examines the quality of those goals, too. I’m not talking about judging the goals in the technical sense, judgment that asks to what extent they’re clear and exhaustive, and to what extent they express the actual endeavor. But about judgment of their essence: their contribution to the evaluatees, the community, and society.”

“In other words, you’re also asking about judging the fundamental value of the goals. Do the intervention program’s goals see the entire community or just the specific target population? Do the goals cover a sufficiently long period?”

“Yes, and I believe that if the intervention people had looked at the entire community, it should have appeared in their goals, or at least, in the rationale.”

“Let’s summarize the questions that you’ve raised up to now so that we don’t get lost. The first is, what are the boundaries of the evaluator’s role? (or, in other words, what is the evaluator’s professional responsibility)? The second is how broadly and how far the intervention people, and especially the evaluator, should look (whether to examine only the program or also the context,-). And the third question is, does the evaluator have the right to judge (or evaluate) the intervention’s goals?”

“Right. And I want to examine the goals vis-à-vis a worldview, not just their technical merits.”

“Look, if you’re asking about judging the goals, then Friedman and his colleagues addressed that issue, (Friedman et al., 2006) but more in the direction of the fourth generation. They speak about the emergence of the different worldviews held by a program’s stakeholders. They did this in order to create understanding and agreement between the partners. But you’re going even further and asking how you choose between these worldviews. How do we build a professional compass, something that can tell us what’s good and what’s bad? And that’s perhaps the most important question because it also dictates the breadth of the picture we’ll look at as evaluators. So your fourth question is: What is the role of your worldview versus that of the intervention people in the

evaluation process? Or in other words, what compass does the evaluator have to sufficiently answer these questions?" I said.

"You've put it in a nutshell," he smiled.

"Many serious questions?"

"Yes," he nodded soberly. "All these questions are causing a commotion together in my mind. It's as if the more I know, the more experience I have, the more difficult the decisions become because the questions become more difficult too."

"Werner Ulrich (Ulrich, 2001) addressed expertise and has a very interesting approach to it."

"Enlighten me," he said.

"I hope I can. Ulrich contends that expertise isn't the high road to confidence, but the opposite. Expertise is more a matter of questions rather than answers. He also contends that expertise expands, and with that expansion, the responsibility of the expert expands correspondingly on several levels. First of all, the expert must recognize that it is not humanly possible to discuss all the aspects of a problem."

"But we still labor under the delusion that if we choose the questions and means of testing correctly, we'll obtain a correct answer. But correct according to what? According to whom? According to which criteria?"

"That's just it, there are no answers to those questions, at least according to Ulrich," I said. "One of the problems with making any decision and choosing any methodology is selectivity."

"Which means?"

“That something always remains outside our field of vision. And what we must take into account is what remains outside. What we’re losing by following our choice, not what we’re gaining. Qualitative methods, for instance, leave the distribution and power of the findings outside, whereas with quantitative methods it’s the variety. But even without the methodology’s limitations, any focus on one specific question bypasses what wasn’t asked.”

“Well, that’s certainly thought-provoking.”

“And more important,” I went on, “an expert must know how to identify the questions that she or her methodology can’t answer. That’s the difference between an expert and a technician. Experts have to constantly observe themselves and test whether the path they’ve chosen is the best one in the existing circumstances—the best in the sense that it gives them reliable answers.”

“It’s a real paradox,” my colleague said hotly, “the more expertise expands, the more do the responsibility and the doubts. And as evaluators, our doubts are about everything: our basic assumptions, our questions, work procedures, the findings and how we interpret them, and how we derive recommendations for action from them.”

“There’s no rest for the wicked,” I smiled, wickedly.

“There isn’t, or answers either, because this whole discussion of ours still hasn’t shown me a good way of knowing what to do.”

“I suppose it does make things more difficult than it helps,” I agreed. “Being a technician is far easier than being an expert.”

“Now she tells me.”

“Ulrich further contends that, even when the aim of evaluation is to test whether a specific work method is successful, the subjects of the evaluation are people, and the findings and recommendations are addressed to people. Therefore there’s social significance in what we do as evaluators.”

“Right, that’s where we started,” he said with a sigh.

“Okay,” I went on. “On the question of whether a program is successful, Ulrich suggests answering the question ‘What is improvement?’ in a complex manner, particularly while understanding our own values, which define what improvement is for us, and also in connection with the groups of people likely to be influenced by the evaluation process or its attendant side effects. And they’re not only the evaluation clients.”

“It sounds like fourth-generation evaluation.”

“Not really. He doesn’t talk about joint construction of a worldview. He addresses the evaluator’s sphere of responsibility, and it’s up to evaluators to declare which and whose values their evaluation promotes.. He opposes the frequently heard argument that professional evaluation serves all interests equally.”

“But he does introduce into the equation anybody who’s involved in the process in one way or another.”

“Yes, but into the evaluator’s equation. Ulrich argues that the evaluator’s vision of what change the evaluation needs to create and in whom, must be constantly examined. And evaluators must examine their own vision. According to Ulrich, reflection such as this drives the professional thinking process forward more than any other point of observation.”

“For years we’ve been taught about scientific objectivity, about this separation of powers—we can’t allow our personal inclinations, opinions, and beliefs to dictate the outcome of our evaluation.”

“And now,” I smiled, “your definition of yourself as an expert obliges you to consider your own beliefs and values. Because, as experts, the question we must ask ourselves is not only how to conduct a good evaluation, or which methodology to use, but to what end we’re conducting it, and who and what this evaluation serves, and also who and what we want it to serve. That’s what Ulrich contends.”

“Well, if the question is ‘to what end,’ that means that our prime concern is the character and implications of the evaluation results, and not the methodology used,” he said, somewhat depressed.

“And that’s only part of it!” I exclaimed. “Since evaluation is purpose-driven, it’s teleological. It has direction because it’s done in order to reach a result.”

“And as you’re always saying, Research seeks reasons for phenomena, that’s its purpose. Evaluation uses the reasons for a purpose, to create a result about a concrete object.”

“Absolutely,” I agreed, “and a generally accepted evaluation result is, for instance, to help the person who commissioned the evaluation to make practical decisions about the evaluation objects.”

“And also,” my friend continued, brightening up, “research has a subject but evaluation has objects. And since the object of the evaluation always relates to people’s actions, there’s no way of avoiding a discussion—not only on professional questions—but also on the ethical questions. And that brings me back to the question of the compass.”

“Well, right now I’ve got three different lines of thought on responses to your questions.”

“Three’s fine,” he laughed. “It’s a number I can deal with. What are they?”

“On the question of the compass, the first that comes to mind is Kant’s categorical imperative, (Kant, 1991) which provides a partial answer.”

“And what about the legitimacy of my worldview?”

“On that, Professor Hanan Alexander (Alexander, 2006) provides a certain answer with his ‘a life that is worth living’ concept. And on the question of the broad canvas, we’ll go back to Ulrich (Ulrich, 2001) and Churchman, (Churchman, 1968a, 1968b) and I hope that all of them will come together in a concept of a fifth generation of evaluation as I understand it, a concept that will perhaps provide an answer to all your questions.”

My colleague smiled. “Fine, let’s decipher them one by one.”

“Shall we begin with Kant?”

“Let’s,” he agreed. “Kant held a clear and determined position: ‘Act in such a way that you always treat humanity, whether in your own person or in the person of any other, never merely as a means to an end, but always at the same time as an end,’”² he declaimed with French pathos.

“Indeed, and the research objects in evaluation are people. Always. Even when the evaluation’s objective is to examine whether a work method is successful,” I said.

“But Kant bothers me less,” my colleague said. “He was a deontologist. That is, he contended that ethical considerations and moral laws should guide our behavior regardless of its outcome, and that we must behave in accordance with universal moral

² Kant, 1991.

principles such as honesty, fairness, human rights, justice, and respect for others. I conscientiously behave that way, so Kant doesn't influence the methodology I choose."

"Perhaps not the methodology. The relationship with the evaluatees is always a means. It's never a purpose in and of itself. There are evaluation approaches that attempt to bypass this: empowerment evaluation,³ democratic evaluation,⁴ participatory evaluation,⁵ and constructivist evaluation.⁶ Although all these approaches allow lots of room for the evaluatees, and the evaluatees are likely to gain from the interaction with the evaluation, they are still not the evaluation's purpose. They're a means for creating the knowledge the evaluation needs, that the intervention people need. I've attempted to deal with this problem with my cybernetic approach."⁷

"We've crashed at the first of your three lines of thought," my colleague said, frowning.

"I'm not sure we have. Perhaps we should go a little deeper into Kant's words."

"Enlighten me."

"Dr. Niva Arav⁸, for instance, contends that at the basis of Kant's categorical imperative lies the principle of freedom of action. Kant states that humans are free creatures, and freedom, in his opinion, is manifested in our ability to think about our behavior and our ability to change that behavior in a way that meets our needs.

³ Fetterman, 1994.

⁴ Greene, 2000.

⁵ Cousins & Whitmore, 1998.

⁶ Guba & Lincoln, 1989.

⁷ Levin-Rozalis, 2010.

⁸ Arav, 2010.

“Meaning,” he said enthusiastically, “that all the participatory approaches you mentioned, which give the evaluatees tools and knowledge that extend their range of choices, clearly meet the criterion of Kant’s categorical imperative.”

“They do,” I replied, “because Kant contends that freedom of action is dependent on knowledge, and if we expand the evaluatees’ knowledge and allow them to choose, then there’s evidently no problem. There is, however, a big problem in an action that prevents the evaluatees from setting their goals for themselves.”

“Then the question reverts to the intervention program itself,” he remarked thoughtfully. “Doesn’t the intervention program prevent the participants from setting goals for themselves and deciding about events of which they are a part?”

“A good question. According to Kant, as I understand him, that’s something that evaluation should examine.”

“We’re getting into an infinite loop,” he said, shaking his head, “because, as interveners or evaluators, we have no ability to do anything at all. Any decision we make, even a decision to teach people and give them the tools for making a better choice, is one we’ve made for somebody else.”

“You’re right,” I replied. “But it seems that Kant gave us an answer to that with his principle of universality, with his contention that every moral law or principle must be universal, in the sense of being suitable for everyone. And that’s conditional upon a positive answer to the question of whether I’d want everyone to act like me, or act toward me as I act.”

“That also means,” he went on, “that we must be prepared for the choice made by the evaluatees to be a choice that runs counter to our worldview, values, or wishes. And that seems to be against human nature.”

“It’s certainly against the nature of the various do-gooders,” I laughed,

“And what does all that say about my present quandary?” he asked.

“It raises questions,” I replied. “First, about the community and the choices it can or can’t make in the intervention framework.”

“Yes, but that’s less important for me,” he said pensively.

“But still,” I persisted, “from the intervention’s standpoint, it’s important to examine whether it has extended or narrowed the community’s ability to choose. That can be your important contribution to the community, to place—as a criterion for a good evaluation—extending their possibilities of choice.”

“I’m not sure how intervention can narrow it. After all, in our world at least we intervene in order to do good.”

“I’ll give you an example. At one of the international conferences on evaluation, an evaluator spoke about a project she evaluated in Egypt. The project set up a purified running-water system for a series of villages. Although the water reached the houses, the village women continued the traditional way of walking a long way to a well, drawing water, and carrying the heavy cans of water home on their head.”

“Why?”

“The evaluator examined this and found that, for the women, walking to the well together was the only social interaction they had outside the family, and one of the few

possibilities they had of getting out of the house and away from their interminable household tasks.”

“But going back to my problem, Kant is in fact telling me that my opinion on what’s right or not right for the community isn’t important at all. That is, as long as the principle of universality is maintained. What the community thinks about itself is important. But then what? Not to pipe running water to it? After all, isn’t the right to purified water as universal as the women’s right to social interaction outside their four walls?”

I looked at him, unsure about what to say. It seemed that the question was becoming increasingly complicated. “It looks like we need to think up a third solution, or a series of solutions whereby the lion’s hunger is satisfied but the sheep remains whole.”

“Such as?”

“I don’t know. Maybe by separating the water into drinking water in the faucets, and water for household use, from the well that would compel the women to walk to it. Am I a water engineer?”

“Okay, okay,” he said, holding up his hands. “So what Kant in fact contributes to my question of a compass that I can use to examine my worldview vis-à-vis that of the intervention people is the issue of extending the evaluatees’ possibilities of choice, and the issue of universality. Two things that aren’t easy to examine with a slide rule in complex situations. Sometimes one thing comes at the expense of the other.”

He went on, “It’s reasonable to assume that the program I’m evaluating is extending the possibilities of choice for the young leadership it’s fostering at the expense of the

traditional leadership, and also at the expense of the community's weaker parts, which remain without their own talented young leadership."

"It's true that it isn't easy," I agreed. "Maybe what we need to do here is to draw attention to the weakening of the community and suggest to the intervention people that they seek a third way, one that will keep the young leadership in the community."

"I don't know. How do you do that without manipulation?"

"Perhaps by examining the life that the community sees as worth living?"

"Fine. So tell me about Professor Alexander⁹ and life that's worth living."

"Alexander speaks about educational research, but what he has to say is very appropriate for evaluation. He contends that a researcher must have a unique viewpoint stemming from a subjective perception of reality."

"The researcher?" he asked, raising his eyebrows.

"Mainly the researcher. The researcher or evaluator is responsible for the implications of the research. They don't work in a vacuum."

"But we were taught that our responsibility is solely to rigorously and responsibly conduct worthy research, and that the results of such a research process have a life that's detached from the researcher as though research conducted in an appropriate manner creates a truth that's beyond any consideration of the researcher."

"And that's why Alexander talks about 'a view from somewhere': a subject- or context-dependent viewpoint. And he doesn't agree with detaching the researcher's viewpoint from the research, the research subjects, and the effects of the research."

⁹ Alexander, 2006.

“He’s actually telling me that, as an evaluator, I must examine the wider effects of my evaluation, that I can’t say ‘I’ve done a good evaluation using validated and reliable tools, and all the rest is the intervention people’s responsibility.’”

“Exactly,” I nodded.

“A man after my own heart. What else does he say?”

“He speaks about what he thinks is the unbreakable connection between reality and practice, and theory; between human life experience and objective reality. And in that context he cites Dewey,¹⁰ who the ostensible dualism of practice on the one hand and theory on the other.”

“But if everything is context-dependent, culture-dependent, and dependent on the local perception of a life worth living, don’t we have a problem of relativism? What if the evaluated community thinks that the intervention program money should come to them instead of funding all sorts of programs?”

“First, let’s try to understand what Alexander means by ‘a life that is worth living.’”

“Okay.” He looked at me expectantly.

“Alexander cites Charles Taylor’s¹¹ claim that day-to-day human endeavor is always subject to supreme values, which he calls ‘strong values,’ such as the sanctity of life, loyalty, and friendship—values that indicate for us a life worth living and the nature of a just society in the transcendental sense. It is this transcendental nature that we must research, not the immediate demands or dogmatic concepts of members of society.”

¹⁰ Dewey, 1938.

¹¹ Smith, 2002.

“How do you translate this transcendentalism into action?”

“Alexander doesn’t expand on that, but I think that you simply speak to it in its own language.”

“Speak to it?”

“I’ll give you an example I read not long ago.¹² A researcher examined the use of purified wastewater in the Palestinian Authority. Because of the Muslim laws of ritual purity and also the stigma attached to wastewater, the residents preferred using contaminated water, not purified wastewater. This necessitated long negotiations with the community’s religious leaders, but with much goodwill from both sides. Ideological and religious considerations were discussed, and a compromise was reached: the purified water would be used to irrigate trees, which have a purification system of their own, but not to irrigate vegetables. Although this isn’t about evaluation or research, the notion of going along with the community’s transcendental values and not against them, speaking their language and not using rational reasoning or coercion, seems to me to be well suited to the intention. And most important, this enabled the use of purified wastewater for the benefit of the community without impairing its ideological coherence.”

“Meaning—if I’m managing to follow what Alexander proposes—I must conduct an in-depth study to expose the community’s underlying ideological and normative constructs and go with them? Not my values or those of the intervention people, because they’re irrelevant.”

¹² Nasrallah, 2013.

“Yes. To go back to Ulrich, the values that your evaluation should promote are the underlying ideological constructs of the community within which the intervention is conducted.”

“Is that practical?”

“That’s another question entirely. What Alexander proposes is a matter of principle.”

“And, as with so many in-principle answers, the practical ones only move further away.”

“That’s true,” I agreed.

“Okay, so let’s try to understand the principles.”

“Alexander doesn’t stop at the modern pragmatists,” I said. “He cites ancient authorities, going back to Aristotle, who drew a distinction between two types of knowledge: *sophia*, which deals with theories on how the world works, and *phronesis*, knowledge that enables us to function reasonably and reflectively in the world,¹³ and suggests shifting the emphasis from *phronesis* to *sophia*.”¹⁴

“Why *sophia*?”

“Because it attempts to understand both a physical and metaphysical reality. Even though *sophia* has reasons, it also has a purpose, and that’s what I think is important. For *sophia* attempts to understand the world by means of two processes: *techne*, which attempts to reveal the reasons for things, and *episteme*, which focuses on the purpose of things, their final outcome. But more than that, *episteme* is essentially teleological since it draws things toward their proper natural state, which connects with their essence.”

¹³ Aristotle, 1994.

¹⁴ For an expanded discussion of *sophia* and *phronesis*, see Chapter Two.

“And the essence is always broader than the concrete reason,” he remarked.

“According to Aristotle, it certainly is,” I replied. “In his view, the teleological explanation is more complete, more essential, and more fundamental since it connects things with the complete metaphysical reality—to something wider and bigger than the local reality.”

“In other words, we have to strive not only toward understanding the essence of things, but to allow the essence of things to guide our research. So evaluation has to be teleological? Aimed at a transcendental purpose?”

“As I understand it,” I nodded.

While, in my opinion, the type of knowledge used in evaluation is *phronesis*, but what I do adopt from Alexander is our need as evaluators to examine the implications of our evaluation transcendently, to look forward at the possible implications of our work through the evaluatees’ eyes and their underlying perception of a life worth living.”

“And that’s the compass you use? Is it really possible?”

“Not completely. We’re only human and we have limitations and pressures. But as far as possible, it’s important to try to take these things into account.”

“So what do we do with this transcendental scale? It demands a broad canvas that’s hard to contain.”

“This whole discussion began because you came along with the broad picture, one that’s broader than the evaluation’s goals.”

“Yes,” my colleague sighed, “and now I’ve gotten myself into a bigger tangle.”

“The man who tries to resolve that tangle is again Ulrich.” (Ulrich, 2005, 2012).

“What does he suggest?” my colleague asked with a flicker of hope.

“Ulrich both follows Churchman and takes issue with him. Among other things, Churchman is thought to be one of the fathers of the systems approach in management and social research. (Churchman, 1968b). His most difficult question is ‘How can we design improvement in a large system without understanding the whole system?’” (Churchman, 1968a, 1968b, 1971).

“It’s impossible,” my colleague said.

“So if it’s impossible, how can we understand a system?”

“Good question.”

“And Churchman’s answer is to ‘sweep in’ all the information—of any kind—that can be collected regarding the system in question.”

“No more and no less?”

“Ulrich took Churchman’s concepts of ‘sweeping in’ and ‘whole system judgment’ a step further. He talks about ‘boundary judgment’ instead of judging the whole system.”

“And where does he go from there?” he said.

“What’s interesting is the difference between them,” I said. “Churchman aspired to an overall understanding, whereas Ulrich speaks about means of critical reasoning: ‘the story I mean to tell.’ He contends that since it is impossible to contain the entire system, the question of boundary judgment becomes obvious. Using his systems approach, which he calls ‘critical systems heuristics’ (CSH), he contends that our perception of a system,

its boundaries and contents, is usually heuristic and intuitive, and that's not good enough. In order to be able to discuss a system, we first have to know what we're talking about. What *is* this system? Therefore we must critically examine its boundaries."

"So what Ulrich is telling me," he said, "is in fact this: 'Know what you're talking about.

"Yes, that's what he says," I agreed. "And your decision about what's relevant, what's important, what should be included and what can be left out, should be an informed critical process, not a heuristic and intuitive one."

"So how do you make these decisions? Are there guidelines?"

"There are all kinds of guidelines in all sorts of articles,¹⁵ but the main thing is to go with your central aim. That might be the definition of a problem or of a situation, a suggestion for resolving a problem, a question, and so forth. The aim must be of good quality, where the criterion for quality is its relevance to the players. Theoretical support and precise formulation are far less important and, in and of themselves, are insufficient."

"Like all pragmatists."

"Yes. With this in mind, it's easier for you to decide what's important enough to be included in the system you're examining, and what's less important. Who must be in it and what can be left out. This process enables you to focus without needing to include the whole of France in your deliberations."

"You can laugh," he said, as I smiled. "I accept it with love."

"I'm not laughing; I'm part of your deliberations."

¹⁵ Ulrich, 2005.

“Let’s assume that in my case the part of the community that’s less mobile is important for me because it’s less educated and is employed in work with no possibility of advancement.”

“And the traditional leadership, the educated young people, and the connection between them. The reciprocal effects,” I added. “It’s hard to do it all in one go. Every such group has considerations for and against.”

“At least there’s one thing about which I’m no longer deliberating,” my colleague said with some relief.

“And that is?” I asked.

“The unholy alliance. From everything we’ve talked about, it’s completely clear that as an evaluator, and certainly as an expert evaluator, I don’t have the option of remaining within the limits of the program and examining it from the goals inward.”

“No, you don’t, at least not according to the philosophers whose opinions we’ve discussed.

“Oh yeah...,” my colleague said. “The inevitable conclusion of this entire conversation is that evaluators must take into account where their evaluation is leading. We have to address its implications and the effects it might have in the broadest possible sense, far beyond the question of the nature and quality of the program or its expansion or termination.”

“And we’ve also learned something no less important,” I added, “that the most significant criterion for examining the implications and effects of an evaluation isn’t our personal worldview, nor that of the intervention people. The significant criterion is in the effects

on and worldview of those who'll be affected by those implications and effects: the evaluatees and their environment.”

My colleague concluded. “As evaluators, we bear a very great social responsibility: to look broadly, to look forward, to be responsible for our actions, and as far as possible, to be sure that the purpose is indeed for the benefit of those who will ultimately be affected by our work.”

I nodded silently.

“Now back to my first question,” he smiled. “Do you have a particularly interesting session in mind?”

“I think I've learned enough for one day,” I replied. “I'm going for a walk.”

Bibliography

- Alexander, H. A. (2006). A view from somewhere: Explaining the paradigms of educational research. *Journal of Philosophy of Education* 40(2):205–217.
- Arav, N. (2010). Man solely as a means? Evaluation in light of the Kantian imperative. In Levin-Rozalis, M. & Savaya, R. (eds.), *Evaluation in Israel: Issues and Dilemmas*, Beer Sheva: Ben-Gurion University of the Negev, pp. 333–342. (Hebrew)
- Aristotle. (1994). *Metaphysics*. (D. Bostock, trans.) Oxford: Clarendon Press.
- Churchman, C. W. (1968a). *Challenge to Reason*. New York: McGraw-Hill.
- Churchman, C. W. (1968b). *The Systems Approach*. New York: Delacorte Press.
- Churchman, C. W. (1971). *The Design of Inquiring Systems*, New York: Basic Books.
- Cousins, J. B. & Whitmore, E. (1998). Framing participatory evaluation. *Understanding and Practicing Participatory Evaluation: New Directions for Evaluation* 80:5–24.

- Dewey, J. (1938). *Logic: A Theory of Inquiry*, New York: Henry Holt.
- Fetterman, D. M. (1994). Empowerment evaluation. *Evaluation Practice* 15(1):1–15.
- Friedman, J. V., Rothman, J. & Withers, B. (2006). The power of why: Engaging the goal paradox in program evaluation. *American Journal of Evaluation* 27(2):201–218.
- Greene, J. C. (2000). Challenges in practicing deliberative democratic evaluation, *New Directions for Evaluation* 85:13–26.
- Guba, E. G. & Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. London: Sage Publications.
- Kant, I. (1991). *The Metaphysics of Morals*. (Mary Gregor, trans.) Cambridge, UK: Cambridge University Press.
- Levin-Rozalis, M. (2010). Cybernetics: A possible solution for the “knowledge gap” between “internal” and “external” in an evaluation process. *Evaluation and Program Planning* 33(4):333–342.
- Nagel, T. (1986) *The View from Nowhere*. Oxford, UK: Oxford University Press.
- Nasrallah, W. M., (2013). *Wastewater and Solid Waste Projects in the West Bank – Success Versus Failure*. Unpublished Master’s thesis. Ben-Gurion University of the Negev.
- Peirce, C. S. (1955b). Abduction and induction. In Buchler, J. (ed.), *Philosophical Writing of Peirce*, New York: Dover, pp. 150–156.
- Smith, N. H. (2002). *Charles Taylor: Meaning, Morals, and Modernity*. Cambridge, UK: Polity.
- Ulrich, W. (2001). The quest for competence in systemic research and practice. *Systems Research and Behavioral Science* 18(1):3–28.
- Ulrich, W. (2005). *A Brief Introduction to Critical Systems Heuristics (CSH)*. Milton Keynes, UK: The Open University. Online (retrieved 23 December 2013):
http://wulrich.com/downloads/ulrich_2005f.pdf.
- Ulrich, W. (2012) CST's two ways: a concise account of critical systems thinking. *Ulrich's Bimonthly*, November-December, 2012. Online (retrieved February 2013):
http://wulrich.com/downloads/bimonthly_november2012.pdf.

Miri Levin-Rozalis has conducted, taught, researched and written about evaluation for more than 30 years as a faculty member at Ben-Gurion University of the Negev, the Weizmann Institute of Science, and independently. She is a founding and current member of IAPE – the Israeli Association for Program Evaluation.