Does Evaluation Contribute to the Public Good?

Sandra Mathison
University of British Columbia

Abstract

If evaluation is to make a positive contribution to the social, physical and environmental world we need to analyze our theory and practice from a sociological perspective. Dominant socio-political ideologies shape how evaluation is conceptualized, the methods and models used, how and by whom it is funded, and its efficacy in promoting positive social change. Most evaluation occurs in a micro context, a legacy of evaluation practice that serves other disciplines, decision-makers, policy-makers, funding agencies, and beneficiaries. Evaluation practice is local (even when the context is geographically vast) and mostly responsive to particular concerns about programmatic effectiveness.

Evaluators, while continuing to work within programmatic frames, should investigate the frames themselves. We must ask how these frames establish taken-for-granted forms of problem definition, solutions, and indicators of success. These frames are embedded within ideologies that structure human relations and social practices beyond, but including, evaluation.

I will trace the evolution of evaluation theory and practice as influenced by global ideologies from early progressivism (characterized by public funding for much program evaluation) to a possibly waning neo-liberalism (characterized by increased funding by philanthropists, NGOs and entrepreneurs) to a surging populism, and reflect on evaluation's contribution to the public good through these changes.

While perhaps an uncomfortable consideration, we need to ask whether evaluation as framed by these dominant socio-political ideologies contributes to the public good, whether it contributes to positive change. By most accounts, evaluators' work isn't contributing enough to poverty-reduction, human rights, and access to food, water, education and health care. We need also to consider whether formal evaluation practice may be getting in the way of and hindering social change. I will conclude with some tentative thoughts about what we (evaluators, funders, and users of program evaluation) might do to make a positive contribution to the public good through evaluation.

Good morning. I am delighted to be here with fellow evaluators, and I thank the AES conference committee for the invitation to be with you over the next several days. I would like to acknowledge the ancestral domain and land we are on today and throughout Australia, and pay my respect to Elders past and present.

I have been doing evaluation, primarily in education, for over 40 years. I wish I could say evaluation has fulfilled the promise of its ameliorative assumption. That evaluation practice and theory have made things better. But instead I find myself pessimistic, maybe even cynical about the extent to which evaluation has contributed to the public good, by which I mean the well-being of all people, globally, manifest in things such as food security, health care, education, clean water, and housing. Even though I do not assume that evaluation and evaluators alone are responsible for the public good, we suggest the work we do will make things better, lead to better outcomes, solve social and environmental problems. So, we need to own that claim.

Even though I am quite pessimistic, I will end with a couple ideas we might consider if evaluation is ever to fulfill the promise we think it has.

- the world population is 7.5 billion people
- half of the people in the world live in poverty
- 22,000 children die each day due to conditions of poverty
- 805 million people do not have enough food
- 5 million children die each year before they are 5 years old
- 165 million children under 5 are stunted from malnutrition
- 750 million people do not have adequate access to safe drinking water
- 2300 people die each day from diarrhea
- 214 million women lack access to family planning
- 1 in 7 children in NYC are homeless
- this year, in British Columbia 4 people will die each day from opioid overdoses
- 34,500 people flee their homes every day to avoid violence
- 1 in every 113 people on Earth has been driven from their home by conflict, violence, or human rights violations
- two thirds of illiterate people in the world are women

I will outline three reasons I believe evaluation has not and is not contributing enough to the public good.

First, is that evaluation theory and practice (like many social practices) reflects the values, beliefs and preferences of the time. As such, evaluation is constrained by dominant socio-political ideologies.

Second, is that evaluation fundamentally lacks independence, it is a service provided to those with power and money and in that relationship, becomes a practice constrained in its capacity to contribute to the public good.

And third is that evaluation is fundamentally a conserving practice, working within domains established by others, and more often than not maintaining the status quo.

The Influence of Socio-political Ideologies

How we make judgments about the value and quality of programs, policies, interventions, reforms is a function of socio-political ideologies. Dominant socio-political ideologies shape how evaluation is conceptualized, the methods and models used, how and by whom it is funded, and its efficacy in promoting positive social change. All evaluation requires the specification of the desirable qualities of what is being evaluated. And these qualities are socially constructed; therefore, the dominant approaches to evaluation reflect the current socio-political zeitgeist.

In recent history, there have been two dominant socio-political ideologies that frame how evaluation is conceptualized and conducted: progressivism or social democracy, and neo-liberalism. And, we are potentially on the verge of another dominant ideology: populism.

I am using the term progressivism for this first phase even though in reality the progressive socio-political era has a much longer history. I could also have used the term social democracy. The primary idea I am connecting to is the advocacy for social reform, especially a reduction of income inequality, expanding freedoms and rights, and expressions of collective action and humanitarianism… the idea that the human condition would improve through the application of science, technology, and social organization.

In its early days evaluation practice bore the mark of progressivism. Our work was often supported by public funding and defined itself as a public good, in the interest of all. Evaluation reflected progressive values… including efficiency, social justice, and democracy. The late 60s and into the early 1980s were the gold rush days of evaluation. During this time, models of evaluation proliferated and exciting intellectual work was being done across a number of disciplines. Evaluators borrowed from other disciplines (like journalism, jurisprudence, art) to explore how we can and ought to make value judgements; evaluators explored the potential for evaluation to contribute to democracy, fairness, and equity; evaluators were confident their craft would help to make the world a better place.

Evaluation has now labored under neo-liberalism for decades, and in the current state of neo-liberalism, evaluation increasingly reflects those values, including commodification, competition, and privatization. Neoliberalism is a global socio-political ideology like no other… it rejects traditional political partisanship and disregards national boundaries.

This is a time when capitalism trumped democracy, and when the interdependence of capital and government is solidified. A basic tenet of neoliberalism is the pivotal role state governments play in facilitating and fostering the interests of economic elites.

And evaluation is a practice that, on the one hand, is a tool to rationalize and normalize state actions and neoliberal values. We see this manifest in, for example, new public management… performance management… impact measurement. We see this in the rise of philanthrocapitalism, bringing hard-nosed strategy, performance metrics, and an emphasis on

effectiveness to the nonprofit sector. We see this in the rise of social investment, a purposeful conflation of doing good whilst making a profit.

On the other hand, evaluation has become the tool of the state… constantly monitoring and assessing public policies, the conduct of organizations, agencies and individuals, even serving as the final evaluator. Indeed, evaluation is the privileged creator of knowledge for the rational individual, organization and State seeking impartial, up to the minute and comparable data to make rational choices among competing options. Finally, decision making based on facts, rather than custom, preference, community or magic.

A good example of the state as evaluator is the *What Works Clearinghouse* created by the US Department of Education to tell educators and parents which educational programs, products, practices and policies work and which do not. What works is directly connected to narrow conceptions of how one knows what works (in this case, evidence from study designs that include comparison groups is the only evidence that counts, and preference is given to experimental designs).

Neoliberalism therefore creates both an evaluated State and an evaluating State.

In spite of heartfelt, well-argued, compelling conceptualizations of program evaluation as democratizing, empowering, inclusive and transforming, and/or participatory evaluation practice often (maybe usually) reflects an emphasis on efficiency, effectiveness, and short-term measurable outcomes typically reflected in a logic model or theory of change. Evaluation adopts 'market speak' and measures program success in terms of profit; programs are in a competition; outcomes are conceptualized in economic terms; and individuals are to blame for failure.

Evaluators talk about a good program as one that is a good return on investment or gives value for money, one that decreases costs (whether health, housing, welfare) or results in savings, and is the best of its kind. And when programs don't work or fail it is due to service providers' implementation infidelity or lack of willingness or willpower on the part of beneficiaries. Even when evaluation is formative, the role of evaluation often makes the case for what needs to be done so an intervention can be scaled up or replicated or reach more people.

While much of the world continues in the grip of neo-liberalism there is an emerging ideology, although it is unclear what its future scope and strength will be. What the mainstream media calls right wing populism, but more accurately is authoritarianism or even neo-fascism, is sweeping through democratic societies, notably Britain, France, Turkey, and the United States.

All populism, whether of the right or left variety, has in common the idea of drawing people, who feel mis- or under-represented, together around some political idea to shape an us versus them relationship. Populism can be a response to neo-liberalism and a concomitant appeal to centralism and the hope that compromise, the development of a consensus at the centre, will win the day.

I don't know what evaluation will look like in a world of dominant right-wing populism. But what I have seen in my 40 years of doing evaluation is that we can be assured evaluation will be molded within that socio-political ideology in a way that serves its central tenets and values.

Evaluation Fundamentally Lacks Independence

Evaluation itself is a commodity, a service for hire, exacerbated especially within a neoliberal capitalist framework. Evaluation is a service bought and sold and while many evaluators frame their practice within larger principles of our professional organizations, evaluation and evaluators are nonetheless responsive to those who pay for their services. I believe it is difficult for most practicing evaluators to imagine that whoever has commissioned the evaluation does not have the most say in what the evaluation questions and preferred outcomes will be. Those with the money dominate the definition of what matters, what counts as success and how that is demonstrated. The mode of operation in much evaluation practice is top-down, what Michael Scriven called the "managerial ideology." In the big picture, program/project managers serve their funders and evaluators serve the program/project managers and/or their funders.

It is notable that one of the most researched, talked about, worried about topics in evaluation is USE, or more accurately evaluation non-use. We have been consumed for decades about why our evaluations are not used, how we can get people to use them, how we can do evaluation so that decision-makers will see the utility of our work. Carol Weiss initially liberated us from the idea that evaluation findings would lead directly, instrumentally, to decisions and changes. But in so doing she also left us with the uneasy feeling that our work was a small part of the mix, or even worse, a pawn in a game played by completely different rules. Recently we have even created a new kind of use, process use, to convince ourselves and others that what we are doing really does matter.

Any profession that spends so much energy, so much time, so many resources on why people don't attend to what we do and say is by inference one that probably isn't doing as much good as we might think.

Because we are trapped within this managerial ideology, we pride ourselves on standing up to and delivering evidence about what works and what doesn't, often in the face of irrational or at least ideological commitments to policies and programs.

We hear the phrase "speaking truth to power" used to express what we see as our contribution to doing the right thing, contributing to the public good. A valiant, but usually futile act. "Speaking truth to power" is a cliché, used often by leftists and liberals, and it neglects the likelihood the powerful already know the truth and choose to ignore it (or modify it) to suit their interests and already well-developed ideologies. Unfortunately, our reliance on others in our service role and a long standing couching our findings in other understandings, preferences, and assumptions makes speaking truth to power the entirely wrong approach, or at least not very effective.

Evaluation is Conserving

Most evaluators think in a micro context, a legacy of evaluation practice that serves other disciplines, decision-makers, policy-makers, funding agencies, and beneficiaries. Evaluation practice is local (even when the context is geographically vast) and mostly responsive to particular concerns about programmatic effectiveness. The value of evaluation for alleviating conditions of poverty are typically manifest in identifying effective strategies or improvements to strategies for the relief of poverty. While it is entirely appropriate for evaluation practitioners to work in this way, it blunts our attention to big questions about why this intervention, why this strategy, why these people and not others.

Consider this one simple example—food security, a core concept in poverty reduction, a key contribution to the public good. Programs to increase food security are created by agencies (like Save the Children) or governments (like USAID) who are in turn the architects of definitions of program success. USAID declares evaluations should emphasize performance-based management to "strengthen the impact these programs have on the well-being of their intended beneficiaries." While there are many kinds of food security programs, one strategy is 'food for work,' an approach where food aid payment is exchanged for labor in public works programs designed to build and maintain local infrastructure (such as, roads, wells, latrines, or schools).

This program type (a commodity exchange approach to a basic human need, in this case food) privileges local economic growth as the major outcome, and food security is one means to that end. Beneficiaries are both capitalists and people who need food. A response to food security that emphasized use value (rather than exchange value) would lead to different strategies and indicators of success; for example, a use value conceptualization might see food as family maintenance, love, aesthetics, happiness, neighborliness, and community building.

But evaluators' work is conceptualized in this kind of closed system and it would take a Herculean effort to shift the evaluation focus to the underlying conceptions of a program. As a consequence, evaluation practice is a reaction to the ideas of others, and has limited reach to challenge the idea of the policy or program.

Couple this with the likelihood that we do evaluations of programs whose assumptions and intentions are consonant with our own value systems. And, chances of seriously challenging programs or programmatic intentions are further diminished.

Our work therefore tends to conserve what is already there.

Just a Couple Thoughts Moving Forward

We need truly independent evaluations

External evaluation is often characterized as more independent and likely to provide a free from bias judgment about programs, interventions and strategies. Whether this is true is much debated, but at a minimum, evaluations paid for by someone with a vested interest in the program

will be influenced by those interests. I am suggesting that more independent evaluations, ones done with no monetary interest in the program (although intellectual or moral interests) can provide information about the value and consequences of programs and interventions. And so, independent funding of such evaluations allows evaluators the opportunity to step outside the frames of neoliberalism, to look at long and short-term outcomes, and investigate unplanned and unanticipated outcomes. Earlier, I used the example of food security programs framed as commodity exchange and point out that this framing of both the problem (hunger, starvation) and the solution (food for work) is not value-free and is based on capitalism. Independent evaluation that interrogates these programmatic foundations is a substantial contribution to challenging taken-for-granted assumptions that define food security in capitalist terms.

We might do better to speak truth to the powerless

Speaking truth to the powerless may be far more useful than the cliché of speaking truth to power. The powerless may not know the truth or may be confused about it, which helps to make them inactive, unable to pursue their own interests, unable to see their interests are shared with others.

We have a proliferation of approaches to evaluation that promise participation, emancipation, transformation… I am part of the evaluation community that makes these promises. I am not simply advocating that this type of evaluation is better. But rather, for us to consider how evaluation might contribute to the public good if the poor, the homeless, the diseased were our clients. The powerless need evidence and analysis that will allow them to understand blaming the victim, privatization, union busting, and free market principles that sustain policies and ideologies detrimental to their well-being. We need more truth to the powerless, not the powerful, and we need to empower the powerless to speak for themselves… evaluators should consider how we might do this.

These are ideas for which I have no clear plan, no set of procedures or processes needed for their implementation. But, after a lifetime of work I'm not ready to give up on evaluation's potential. And so even though our work has contributed too little to the public good, has sometimes maintained a harmful status quo, has sometimes even contributed to harmful social practices… I believe we can do better. I believe we should try to do better.

Thank you.