Rowe, A. (2019). Sustainability-ready evaluation: A call to action. In G. Julnes (Ed.), *Evaluating Sustainability: Evaluative Support for Managing Processes in the Public Interest*. *New Directions for Evaluation*, *162*, 29–48.

2

Sustainability-Ready Evaluation: A Call to Action

Andy Rowe

Abstract

Evaluation is at the cusp of two urgent challenges: indigenous evaluation and sustainability. How we respond to these challenges can dramatically affect the future of evaluation. A sustainability-ready evaluation will be transformative. It will be an evaluation that recognizes that human and natural systems are coupled, and that evaluation portfolios are now and will increasingly be affected by our connections to natural system forces including climate. Sustainability-ready evaluation will be an evaluation that reaches well past the intervention to important public policy goals and to key sustainability challenges. Evaluating coupled human and natural systems will be challenging. Fortunately, technical barriers do not prevent us from starting to infuse sustainability into evaluation; the barriers are social and associated with the worldview and vision of evaluation. To facilitate the development of sustainability-ready evaluation, this paper provides an initial checklist and references to useful resources. Absent transformations to become sustainability-ready evaluation will lack relevance for many of the current and future key issues of our times. Fields lacking relevance are themselves not sustainable.

Introduction

Evaluation faces two urgent challenges, each with the potential to radically reshape the ethical, theoretical, and practical underpinnings of the field. The character of efforts underway to address indigenous lives and worldviews will emerge over time and could lead evaluation into very different places. If evaluation were to address sustainability, this too would be transformative. Addressing the challenges of indigenous and sustainability-ready evaluation provides a unique opportunity for evaluation to learn, refresh, invigorate, and engage in transforming the evaluation of today to a knowledge and practice that is fit for the deepening challenges we face.

Sustainability and climate are deep global challenges that reach to every level and facet of human activity. Following a period of growing awareness and concern about the environment and seminal publications such as Rachel Carson's Silent Spring (Carson, 1962), the report of the World Commission on Environment and Development provided the contemporary framing of sustainability as meeting the needs of the present without compromising the ability of future generations to meet their needs (Brundtland Commission, 1987). This intergenerational framing of sustainability has subsequently been enhanced to encompass social and economic as well as environmental pillars of sustainability (Uitto, 2016). These pillars have always been regarded as highly and dynamically coupled. Indigenous worldviews recognize this, and evaluation approaches that truly incorporate indigenous worldviews would have strong prospects of also being sustainability-ready.

Sustainability-ready evaluation cannot be bounded by contemporary partitioned and sectoral evaluation approaches. The evaluation we have today treats human and natural systems as unconnected and rarely considers the natural system (Rowe, 2018). A sustainability-ready evaluation will be a *connected* evaluation that reaches to public policy goals (Chelimsky, 2012; van den Berg, 2011). Adapting evaluation to a systems orientation that considers human and natural systems as actively and dynamically coupled is a significant challenge that will transform evaluation and can position it to be a useful contributor to thought and action in sustaining life on Earth.

Evaluators question and fret about the use and influence of our work; if we do not develop sustainability-ready evaluation, present levels of use and influence could be the high point. Absent transformations to become sustainability-ready, evaluation will lack relevance for many of the current and future key issues of our times. Fields lacking relevance are themselves not sustainable.

Though the global Sustainable Development Goals (SDGs) and commitments provide the appearance of a growing worldwide evaluation agenda on sustainable development, this is more a matter of appearance

than substance (Uitto, this volume). The reality is worldwide and especially in North American evaluation, there is only an initial awareness of sustainability, or the urgency and level of challenge, or the opportunities. Evaluation is currently missing in action on sustainability.

This chapter aims to sound a call to action for a sustainability-ready evaluation. First, I describe some core features of a sustainability-ready evaluation, building on the foundations laid in recent papers (Rowe, 2018; Uitto, 2016). Then, I provide an overview assessment of the readiness of the intellectual foundations and capacities of evaluation, explore the role of dominion in shaping this, and conclude with an initial checklist for developing sustainability-ready evaluation.

Sustainability Is a Cross-Cutting Issue for Evaluation

Like gender and race, sustainability is an issue that evaluation must address systematically as core, crosscutting evaluation criteria and questions.

To illustrate, consider any intervention located in a city. In 2010 (NOAA, 2013), over half of the U.S. population lived in coastal watershed counties where the population is aging most rapidly, is most dense, and where the already deteriorating physical infrastructure is highly vulnerable to extreme weather events as well as a sustained rise in water levels.

A recent International Resources Panel (IRP)[1] report (Swilling, 2018) illustrates how human and natural urban systems are unavoidably coupled and how the limits of the natural system are being reached. Population increase, continued urban flows and dedensification of cities is forecast to increase urban settlement to two-thirds of the world population, requiring significant expansion of existing cities and construction of new cities. Construction and operation of urban settlement requires billions of tons of raw materials, from fossil fuels, sand, gravel and iron ore, to biotic resources such as wood and food. To support the forecasted 12% increase in urban populations to 2050, necessary consumption of materials will more than double (from 40 billion tons in 2010 to 90 billion tons in 2050). This greatly exceeds what the planet can sustainably provide. The growth in cities and continuing dedensification of cities will increase global urban land use from 1 to 2.5 million km^2 over the same period, reducing agricultural land and threatening food supplies worldwide (Westhoek, 2016).

Many evaluators work on important human matters located in cities. The interventions that are currently evaluated rely on the foundation of natural capital (European Environment Agency, 2008) provided by ecosystems (air, water, plants, and trees, for example) termed biotic natural capital and from abiotic natural capital sources such as fossil fuels, minerals and metals, wind and solar (Uitto, forthcoming). These evaluations cannot continue

[1] The IRP is an important knowledge source. See http://www.resourcepanel.org/.

to ignore the connections to the natural system; the rapidly approaching limits of natural capital with more extreme natural conditions will assert the coupling to natural systems and prove disruptive to human system interventions. Interventions will, as a matter of urgent necessity, incorporate adaptation; and are likely to turn to evaluation to help assess and improve the coupled adaptation and programmatic efforts. A few illustrations are as follows:

- Public health: Adapting public health systems to address arrival of tropical and highly communicative zoonotic diseases (infections shared between humans and animals) (Mills, Gage, & Kahn, 2019) to the north, outbreaks of diseases such as cholera associated with frequent storms and flooding, health effects of extreme temperatures; broadening scope to incorporate climate change (Enkelejda, Butzbach, & Brousselle, 2019) and adapting to global health approaches (Whitmee, 2015)
- Education: Educating in the context of more frequent disruptions from extreme weather, increasing share of budgets allocated to addressing costs of sustainability issues, more frequent school closures; adapting curricula to changed realities, addressing political influences
- Disabilities, mental health, homelessness: Providing safety, mobility and services to already vulnerable and heavily urban populations made more vulnerable by climate change; addressing increased vulnerability to sanitary risks (Nicholas, Breakey, & Winter, 2015)
- Transportation, urban governance, infrastructure and environmental services: Disruptions and physical instability of many current transport, sewage and water, food storage and distribution systems, responding to extreme weather and natural disasters
- Public safety and climate risk: Destabilising dynamics from extreme weather events, wildfires, and flooding; insurance and recovery programmes.
- Economic and community development: Incorporating ecosystem services, waste, recycling (UNEP, 2019)

A changing climate affects the prerequisites of economic activity, population health, urban settlement, food production, and most other aspects of human life. Sustainability is an evaluation issue that potentially intersects with and affects all aspects of human and natural systems. Incorporating sustainability into evaluation is no longer a choice or moral issue but an imperative.

Where Are We Now?

In a recent paper, I employed the example of U.S.[2] school siting and construction policies as a simple illustration of how the thought, practice, and

[2] The school siting policies refer to United States and Canada; school siting is a coupled system evaluand everywhere.

structures of contemporary evaluation are far from ready for sustainability (Rowe, 2018). Here, I advance that illustration, showing that the focus of conventional evaluation on a single (human) system creates bias toward positive assessments and restricts the potential contributions and relevance of the evaluation endeavor. Some program managers and evaluators will regard this as evaluation-creep and too challenging to enact; the initial checklist below provides some guidance on how evaluators can manage this challenge.

Briefly recapping the illustration, many U.S. school siting policies give strong preference to new construction over rehabilitation of schools when the cost of new construction is not more than 165% of the costs of rehabilitation (Cohen, 2010; Dalbey, 2012). In addition, the siting policies require quite large areas for new schools (Weighs, 2016). As a result, many or most new schools are sited on the outskirts of urban areas and located on inexpensive land.

Inexpensive land is relatively unproductive of agriculture, often wet and located outside current commercial development areas. Prior to construction of the school, the site likely[3] provided important and valuable ecosystem services (BISE, 2016; European Environment Agency, 2008; Zaidi, Dickinson, & Male, 2015) such as water quality management through filtration of harmful nutrients and retention of sediments as well as reducing water flows in extreme weather events; contributed to biodiversity by providing habitat for wildlife, birds, and insects; sequestered carbon; and so on[4].

The human actions of building and landscaping convert the site from being a provider of ecosystem services to a generator of ecosystem problems: grassed and fertilized playing fields add to nutrient loading, removal of mature trees and their sediment-retention services; water is shed from new impervious surfaces such as parking, roads, and roofs and now transports harmful chemicals from car and bus discharge directly into the watershed; fencing installed with the school and playgrounds obstruct essential wildlife movement. These are but some of the *direct effects* of the school siting policies and their implementation through design, construction, and landscaping. There are also demonstrated indirect and induced effects of these school-siting policies contributing to obesity, reduced family function,

[3] Likely rather than certainly because some sites might have generated minimal ecosystem services such as brownfield or reclaimed heavily polluted lands that can be rendered suitable for uses such as schools (US EPA, 2003). In effect, environmental damage is attributed to prior and redress provided by subsequent human activity.

[4] Including, and usually overlooked, inclusion of cultural services *. . . nonmaterial benefits people obtain from ecosystems," and specifically lists "cultural diversity, spiritual and religious values, knowledge systems, educational values, inspiration, aesthetic values, social relations, sense of place, cultural heritage values, recreation and ecotourism"* (Daniela et al., 2012). See also (Gregory & Tousdale, 2009)

neighborhood decline, reduced air quality and associated diseases, reduced water quality, increased vulnerability to flooding, and so on[5].

A contemporary evaluation of school siting would likely focus exclusively on human system issues lying largely within the accountability frame of the School Board and units responsible for the decisions about design and construction of the new school and landscaping, addressing questions such as safety (e.g., student and staff safety, traffic flow), the learning environment and efficiency and effectiveness. The signs of the evaluation findings for these human system issues will most certainly be positive because the main direct outcomes are all performance criteria incorporating professional standards for design and construction. Compliance with the standards for design, construction, and costs is legally required and inspection and approval processes promote compliance.

Absent remediation all of the signs for effects on the immediate natural system will be negative; and the effects are directly attributable to the school, for example, water quality and storm flow management will all be worse; the school, roads and parking will increase CO_2 emissions; air quality at the site will be worse from the slow-moving traffic and diesel particulates from school buses; carbon sequestration lessened and the habitat for species will have been impaired. A strong positive bias is thereby rendered to the evaluation through exclusion of natural system effects even though these are *directly caused* by the school siting decision and its construction and landscaping.

Mitigation of adverse natural system effects can provide on-site learning labs for students, teachers, and the community and provide school-community opportunities to collaborate through maintaining and improving the mitigation measures. Examples of actions that could provide valuable forward-looking learning opportunities include installing rain gardens to mitigate the effects of impervious groundcover transporting substances through runoff; protect habitat function by retaining priority wetland and natural habitat portions of the site; ensure fencing provides safety without impairing wildlife passage; utilize climate-compliant roofing and heating/cooling systems to reduce CO_2 release; and retain mature shade-providing, sediment-retaining, and habitat-providing trees.

Evaluations usually take a single system perspective (Patton, 2011, pp. 117–120) regardless of whether the evaluation was launched from a human or natural system concern (Uitto, 2016). For example, evaluations addressing conservation concerns in the natural system tend toward regarding the human system as the cause of, and barrier to addressing the problem and are likely to be unconcerned with effects in the human system; and as illustrated by school siting but true generally, evaluations initiated from a human system concern do not consider even the direct effects in the natural

[5] The National Clearing House for Educational Facilities is a valuable source http://www.ncef.org/.

systems. The work of the CHANS (Coupled Human and Natural Systems) Network[6] (Liu et al., 2007; Ostrom, 1990) and of the GEF IEO (Independent Evaluation Office of the Global Environment Facility)[7] illustrates how evaluations can be conducted from a connected natural and human system platform using mixed methods approaches. We have the technical capacity to conduct evaluations in the human and natural systems but systematically choose an either/or approach, either the human or the natural system. The barrier to evaluating sustainability can be understood as a social barrier and one that is solidly within the remit of evaluation.

When evaluation treats a school siting evaluand as a human system intervention and only addresses human system effects, it is unable to contribute to improvement and to establishing more sustainable school policies and decisions. When the direct effects on the natural system are excluded from the evaluation or assessed more lightly as indirect or unintended results then evaluation reinforces the performance management structures that seek to narrow accountability; whereas evaluation that includes the direct effects in both systems will point to the need and potential of broadening the accountability structures and contribute to establishing a sustainable school. In effect, a sustainability-ready evaluation will contribute to repositioning evaluation from silo-reinforcing to silo-busting.

The evaluation we have today is systematically biased toward positive assessments, serves to reinforce silos, is unlikely to address sustainability, and systematically underfills the evaluation mandate.

Intellectual Infrastructure of Evaluation for Sustainability-Ready Evaluation

Sustainability-ready evaluation requires capacities in biophysical and social sciences and evaluation working together to develop and test approaches and undertake evaluations including those set in complex coupled system settings where the systems have important spatial and temporal boundary differences. The focus on coupled systems has been long part of ecology and most recently emerged in the social sciences associated with the work of the Eleanor Ostrom, winner of the 2009 Nobel Prize in Economics for her work on shared use and governance of natural resources (Ostrom, 2009) and the work of William Clark on Sustainability Science (Clark, 2007).

Evaluation of coupled systems can present especially onerous challenges, largely because of the complex dynamics of coupled systems, and because the temporal and spatial boundaries of natural systems are almost always different than the boundaries of human systems (Rowe, 2012); and because the unit of account in human systems is usually more homogeneous than for natural systems.

[6] http://chans-net.org/about-us
[7] http://www.gefieo.org/

The spatial scales that we use to manage resources reflect political and jurisdictional boundaries. Separate responsibilities at national and subnational and even regional levels for fish, wildlife, minerals, forests, and so on; separate responsibilities for mental health, homelessness, poverty and urban issues. None aligns at all with ecosystem boundaries that are the relevant spatial category for most natural system matters. Contemporary thinking shows the need to turn to ecosystem-based-management (Barnes, Bozi, & McFadden, Unkown; Millennium Ecosystem Assessment, 2005; Wondolleck & Yaffee, 2017; DeLauer et al., 2014).

Temporal scales also vary widely; there is no foundation for assuming that the temporal scales relevant to the human system will be relevant in the natural system; for example, water management is seasonal, intergenerational effects are often used to assess impacts on fauna and insects, and for some species a new generation occurs in a matter of days while in others it can be decades and longer. The temporal scales for an evaluation are usually framed by the temporal scale relating to the intervention and are unlikely to align at all with the temporal scales relevant for the natural system.

The unit of account for natural systems is often one or more ecosystems with multiple coupled species whereas human system evaluation is single species. In the natural system of the school siting example, there are multiple species belonging to several Kingdoms (e.g., flora and fauna) and even domains (e.g., including bacteria important to water quality)[8]. The wetland was habitat for mammals, insects, bacteria, birds, reptiles, and so on. Some, such as frogs and flora, are multigenerational residents, others such as deer or fox are wetland visitors (seasonal, daily for water or for food).

However, diversity in the unit of account does not necessarily imply a complex evaluation. Biophysical sciences have developed a knowledge base and methods to assess environmental effects such as best management practices (National Research Council, Committee on Reducing Stormwater Discharge Contributions to Water Pollution, 2009) as an example, use of indicator species for assessing habitat quality (Government of Canada, 2008), satellite imagery enabling assessment of changes in ground and forest cover over large geographies (IEO: GEF, 2017), models for data poor settings such as marine areas (NOAA National Marine Fishery Service, 2016) and ecosystem services and natural capital approaches (BISE, 2016; Millennium Ecosystem Assessment, 2005; TEEB, 2016).

Capacities exist for evaluation in natural systems as well as the strong evaluation capacities for human systems. The natural system work is largely communicated outside the field of evaluation, often in conservation, natural resource management and climate venues. The intellectual infrastructure in North America for sustainability-ready within the evaluation field includes private practitioners, evaluation units in philanthropic organizations, natural resource and environment government departments in Canada and

[8] For a quick overview of biological classification, see (Wikipedia, 2018)

some in the United States and environmental non-government organisations; university-based infrastructure with research and training capacities specific to evaluation of natural systems or coupled human and natural systems are rare. Thus, the intellectual infrastructure for evaluation in the natural system lacks those whose mission includes research and publishing about evaluation. The effect is limited institutional resources in North America for development and testing of methods and approaches, syntheses, and training for new evaluators and for the evaluation field in general. Examples of important and existing capacities (mostly) outside of the evaluation field per se include:

- Capacities for application of specific evaluation methods: systematic review, see Collaboration for Environmental Evidence (CEE)[9], the journal Environmental Evidence and publications by Andrew Pullin (Pullin & Stewart, 2006); for impact evaluation see, for example, the work of Paul Ferraro (Ferraro & Pattanayack, 2006; Ferraro & Hanauer, 2014; Wunder, 2015). An encouraging development has been progress in addressing coupled human and natural systems with mixed evaluation methods and creative use of big data (Blackman, 2012; Charreire et al., 2017; IEO: GEF, 2017; Ndyeshumba, 2000);
- Citizen science (University Library, University of Illionois at Urbana-Champaign, 2018) as in bird counts and in other applications (see The Cornell Lab of Ornithology[10]), evaluation of community-based conservation and coalitions (Jenkins-Smith & Sabatier, 2008; Lubell, Leach, & Sabatier, 2009; Wondolleck & Yaffee, 2017);
- Where sustainability is an important policy area (Mickwitz, 2013), or sustainable development (Larsson, 2015; Raggamby, 2012; Christopher, 2012);
- Some important European evaluators working on natural systems matters now lead significant government agencies (Hans Bruyninckx is the Executive Director of the European Environment Agency, Per Mickwitz is Research Director at Finnish Environment Institute)[11];
- Evaluation units such as in the Stockholm Environment Institute and the London-based International Institute for Environment and Development have been established[12];
- As well as capacities developed for "other" reasons such as organizations expectations and accountability (Conservation International

[9] Collaboration for Environmental Evidence http://www.environmentalevidence.org/
[10] http://www.birds.cornell.edu/citscitoolkit/evaluation/instruments
[11] European Environment Agency https://www.eea.europa.eu/, Finnish Environment Institute http://www.syke.fi/en-US.
[12] http://www.environmentalevidence.org/, Stockholm Environment Institute https://www.sei.org/, International Institute for Environment and Development https://www.iied.org/about

[McKinnon, Mascia, Yang, Turner, & Bonham, 2015]) or the World Wildlife Fund[13]; and professional interests and priorities (Mascia et al., 2014) of those working within conservation organizations.

To put the intellectual infrastructure for sustainability-ready evaluation in perspective a quick search of contents of the *American Journal of Evaluation* and *New Directions in Evaluation* identified forty articles for the entire 1998–2016 period that contain in their title or abstract key words such as natural resources, climate, sustainability, or conservation; only eighteen of the forty were deemed to address some aspect of the natural system such as evaluation of an agricultural intervention or the social aspects of water quality improvement[14]. An average of one article per year in the two publications of the AEA combined addressing some aspect of evaluation involving the natural system suggests that the intellectual infrastructure in North America for evaluation in the natural system is weak compared to what exists for the human system. By comparison, and over a shorter period, Clark notes a 15–20% annual growth in sustainability science publications 1997–2007 (Clark, 2007).

The extent to which existing evaluation approaches and methods can be adapted to address coupled two-system evaluands and incorporate multiple sciences is not currently known. The question is likely best addressed as an empirical matter[15] to be approached inductively. There is no doubt that evaluation will be better off from the effort. In addition, it is hard to ignore the need for coupled evaluation to address the sustainability issues we face today.

Getting Started

Sustainability-ready evaluation does not exist, nor is there much in the way of intellectual infrastructure for evaluation in coupled two-system evaluands, particularly in North America. This section provides some initial thoughts on how evaluators might start to adjust current evaluation approaches to systematically engage sustainability in evaluation. As noted above sustainability-ready evaluation will address coupled human and natural systems and two-system evaluands; connectivity is the key mechanism for sustainability-ready evaluation approaches and methods and evaluation[16]. The following items are the first steps in developing a sustainability-ready checklist, a beginning and by no means complete. The first element

[13] http://wwf.panda.org/knowledge_hub/endangered_species/marine_turtles/lac_marine_turtle_programme/projects/climate_turtles/planning/m_e/

[14] Review of the two journal contents was undertaken by Eric Trum.

[15] For example the article (van Mierlo, Arkestejn, & Leeuwis, 2010)

[16] It is useful to read the GEF publication providing similar advice for GEF project design (GEF IEO Scientific and Technical Advisory Panel, 2018).

in the checklist considers if the evaluation is addressing a coupled two-system evaluand and if so, the remaining items are concerned with using sustainability-ready approaches to conduct the evaluation.

1. Given current evaluation capacities and outlook, it is important to start with the assumption that you are evaluating a two-system evaluand. Only if initial inquiry (potentially folded into an evaluability assessment) says otherwise should an evaluator proceed with a human-system only evaluation. For example, school-siting evaluands will always be two-system under current policies because the resulting siting of new schools of necessity disrupts ecological functions.
2. Connectivity is the key mechanism when addressing a two-system evaluand. To illustrate, open up the theory of change creating an *unconstrained theory of change* that extends the temporal and spatial reach of the intervention, articulating natural system as well as human system inputs, assumptions, mechanisms, and effects Caballero (2015), Sustainability Consortium 2017.
3. A sustainability perspective is very unlikely where the evaluand is primarily or exclusively framed within the stated goals and descriptions of the intervention or abstracts from context. Most intervention managers will quite rightly resist the resulting extension of the evaluand on the basis that they were not authorized or resourced to address the expanded framing. A negotiation needs to occur to enable the evaluation to adopt a sustainability perspective while not placing unfair expectations on the intervention. Consensus building / (ADR) theory and practice is instructive for this (Susskind, 2012; Islam & Susskind 2018; McBride et al., 2017).
4. Engage as core members of the evaluation team relevant natural science expertise from the outset and throughout. Evaluating two-system evaluands requires knowledge of the two systems and research capacities relevant to the systems. Having expertise from both systems in the evaluation team will provide assurance to external and internal interests that the team is legitimate and the methods credible. It is important that the representatives of natural system knowledge are integrally involved with all aspects of the evaluation including determining the evaluation questions and research design (Rockwell & Buck, 1995).
5. Assume that human and natural systems have different scales and units of account. An evaluation needs to be clear on this from the outset—it is part of defining the evaluand. Be aware that interests' perspectives of scales and units of account will differ, and of the need to achieve agreement across interests that the scales and units of account are appropriate.
6. Consider key stakeholders to include representatives of all interests who can influence success of the intervention and representatives of all interests affected by the intervention—including representatives

from the natural system. Failing to include any can be an important source of bias; by including all interests, the biases of particular interests (e.g., program, industry, environmental groups, indigenous peoples, commercial developers) are balanced by the biases of other interests.[17]

7. In addition to the types of challenges to use sketched in the evaluation literature, sustainability-ready evaluation adds factors such as interests many of whom will be unfamiliar or skeptical of evaluation methods; evaluators working on territories where many will not be fluent or even functional in the language of the subject matter or sciences; involves sciences and the accompanying methods and approaches that are not naturally inclined to attribution; and through pursuit of connectivity takes human system interests outside the familiar narrow boundaries defined by program scope and accountability. The literature on science knowledge use emphasizes the knowledge process over the knowledge product, and that the knowledge process should be jointly engaging researchers (evaluators) and users. This promotes an assessment that addresses the right questions including those that are relevant to potential users; and one that potential users regard as legitimate including being undertaken by appropriate researchers, credible in terms of the methods and analysis and timely in terms of the assessment knowledge being available when there are openings for use, even if this means curtailing the work or reporting before it is completed. Attend to communications knowing how, when and where to identify and communicate with the right audiences. (Clark, Mitchell, & Case, 2006; Committee on the Science of Science Communication National Academy of Sciences, 2017; Courage & Baxter, 2005; Jacobs et al., 2007; Clark et al., 2016; Committee on the Science of Science Communications National Academy of Sciences 2017).

These points are but the starting point of a checklist for conducting a sustainability-ready evaluation, and by no means an adequate enumeration on their own. The key is that we need to start to systematically adapt evaluation to identify and incorporate elements required for assessment of coupled human and natural systems. Some current methods will find this more challenging than others, new capacities and partnerships with other sciences will need to be developed, and importantly a developmental approach to the developing sustainability-ready evaluation itself is required.

[17] (Rowe, under review) introduces the concept of interest-based approaches to stakeholders in evaluation.

Why Is Evaluation So Doggedly Unisystem: Worldviews, Coupled Systems, Public Policy Goals, and Accountability

Both human and natural systems can be complex; coupled human and natural systems are normally complex, sometimes "profoundly complex" (Clark, Dickson, & Matson, 2008, p. 4572), meaning that evaluating from a sustainability perspective will be more challenging (Mermet, Bille, & Leroy, 2010). The issue is largely where we set the frame for the evaluation; nominally narrow frames such as local effects in a single system are relatively uncomplicated evaluations, single system evaluations reaching to public policy goals will have a broader frame with more moving parts and so will be more complicated, coupled systems with more and highly and dynamically connected moving parts tends toward profoundly complex. (Liu et al., 2007) identify observed characteristics of coupled systems to include reciprocal effects and feedback loops, nonlinearity and thresholds, surprises, legacy effects and time lags, different resilience levels and heterogeneity.

That the two systems are understood with sciences having significant differences in methods of inquiry does not explain why evaluation is so doggedly monosystem. As I have illustrated it is technically possible to evaluate the direct local effects in each system with the systems coupled, and have posited that the barrier is social. An explanation for this social barrier is that our worldview systematically ignores the natural system when making most decisions. Natural system resources are regarded as having little or no value unless commercially harvested, privately owned, or iconic. As a result, they are not considered when making decisions. The worldview of Indigenous peoples is very different and regards natural things as equal and all part of the whole, understanding the human–nature relationship as one of stewardship with responsibility to respect and sustain other species and natural things.

The rationale for disconnecting the evaluand from its system setting is connected to accountability structures that Alkin and Christie have proposed as one of two rootstocks for evaluation (Alkin & Christie, 2004). I have observed the logic playing out that results-based accountability and performance management/measurement carry incentives for managers and programs to narrow the frame of their accountability to reduce risk of falling short or being unable to achieve commitments because of insufficient resources and limited control / influence over the larger outcomes. For example, the school siting and design decisions are accountable for creating a physically good learning environment; and while they contribute to the social learning environment and actual learning these lie outside the accountability frame. This is consistent with the truncation of the public policy goals from interventions noted by Chelimsky (Chelimsky, 2012; Elvidge, 2012; van den Berg, 2011) and the lack of direction from evaluation toward meeting important public policy goals observed by Williams

(under review) who notes that the implication of many evaluations is to expand services, yet even when government has done so conditions kept worsening and now absolute limits on even maintaining current funding have been reached. This cycle has led Savoie (2013, Chapter 7) to comment that evaluation is "like turning a crank that is not connected to anything."

The presumption of dominion accompanying Western European colonization and embedded in the western social sciences influences decisions about what is included in evaluation (Rowe, 2018). The presumption of dominion of humans over all other species and natural things is a force much broader than evaluation, we do not value air, water, and other natural inputs unless there is an ownership right to these resources. Dominion has contributed to the exclusion from evaluation of biophysical systems as causal forces and as affected systems illustrated above with reference to evaluation focusing on human system matters in urban settings. It is an open question whether dominion-embedded western social science will remain in ascendancy for indigenous evaluation or if indigenous worldviews, knowledge, and practices will provide the foundation. An indigenous evaluation built on indigenous worldviews is potentially already sustainability-ready. Sustainability-ready evaluation directly challenges an evaluation founded on dominion. Evaluation absent dominion is a transformed evaluation.

Dominion refers to the presumption of ascendancy over other peoples, species, and things and is a key concept in biblical accounts of the relationship between humans and the rest of material creation. The Christian God was held as the source of man's authority over the world, when proposing to create man said, "Let us make man in our image, after our likeness. And let them have dominion over the fish of the sea and over the birds of the heavens and over the livestock and over all the earth and over every creeping thing that creeps on the earth." (Gen. 1:26). The concept was advantageous to monarchs and other elites and is now enshrined in western law through claims of ownership and rights and is embedded in western social science. Crowshoe and Manneschmidt (2002) explain the role of dominion in justifying colonization and in ignoring Indigenous protocols and mechanisms, and the perceptive Annie Proulx has described dominion as experienced by western European settlers in North America and by the resident indigenous peoples. The fictional settler Père Crème writes his sister:

Dear Sister Marguerite.

While I have great sympathy for the Indians, they are difficult. The sorest Point is their Refusal to grasp the Fact and Land belongs to the Man who improves it as Scriptures show. They only fish (an idler's occupation) and wander through the Forest taking Animals and Plants for Sustenance, but when a White Man comes and cuts the oppressive encroaching Forest, builds a House for his Family and Shelter for his Beasts; the Indians complain that he takes

their Land, Land they have done nothing to improve, but rather have allowed to thicken with more and more Trees. They do not understand that the White Man who struggles and strives to reduce the Forest's grip has exerted his God-given Right to claim the cleared Land as his own. By virtue of the suffering of Indian Attack and severe Labor as well as the adversities of removing from their Homelands to that up a Place in the Wilderness it is the Destiny of the French to hold this Land as they have earned moral Title to it from God.

(Proulx, 2016, pp. 179–180)

A worldview with dominion will comfortably separate the school siting decision from its natural place and assume that it is appropriate to disregard natural system effects; a stewardship worldview cannot separate these. Alkin and Christie (2004) might well be correct in asserting that the rootstock of contemporary evaluation is comprised of accountability and social inquiry. However, rootstocks are the mechanisms by which the tree draws nutrition from the soil; rootstock also provides the structural scaffold for the tree. The soil and scaffold from which the roots of evaluation draw their nourishment is a soil of dominion; *the terroir of evaluation is dominion*.

Summary

This chapter is a call to action for sustainability-ready evaluation. Two stark facts capture the current status of evaluation as a sustainability-ready undertaking. First, sustainability and climate are everywhere affecting every aspect of daily lives and activities including institutional arrangements, physical infrastructures, and international agreements. Worldwide rapid adaptation is ongoing from how we feed and shelter our families to the conduct of war; adapting is essential and not a choice. It is inconceivable that evaluation can be relevant without systematically incorporating and responding to this fact by providing an evaluation that expects to and is ready and able to address sustainability. The second fact is that contemporary evaluation is monosystem, the capacity to work with natural systems is weak, and to work with coupled human and natural systems even weaker. Yet evaluation is one of the few fields that explicitly addresses incrementality and potentially could ask questions such as the difference the intervention makes to sustainability in both human and natural systems. Evaluation addresses questions that are essential to improving sustainability. If only evaluation would address sustainability.

Developing a sustainability-ready evaluation will be transformational because it requires incorporation of different worldviews that regard human and natural systems as coupled and each important. This challenges some important foundational elements in evaluation that rest on dominion, and notions that social inquiry and accountability are an appropriate rootstock for evaluation. Indigenous evaluation approaches that incorporate indigenous worldviews could prove to be the polar star for sustainability-ready evaluation.

The barriers to evaluating coupled human and natural systems are social and, initially at least, not technical. This chapter has provided some initial elements for a sustainability-ready evaluation checklist.

Absent transformations to become sustainability-ready evaluation will lack relevance for many of the current and future key issues of our times. Sustainability is and will continue to be evaluated; the question is whether the evaluation field wants to contribute or if assessment of sustainability will continue to be undertaken by those with more sustainability-ready approaches but lacking the special attributes of evaluation.

Acknowledgements

Many have contributed to my thinking, most importantly Elder Reg Crowshoe and my partner Linda King on the concept of dominion and Elder Albert Marshall on two-eyed seeing, as well as comments from those attending my plenary presentation at the 2017 Australasian Evaluation Society meetings, a plenary panel at the 2018 Canadian Evaluation Society meetings and at New Zealand venues sponsored by ANZEA, especially to Kate McKegg who recognized this as a call to action, Juha Uitto a fellow traveller on the broad pathway of evaluation of coupled human and natural systems, George Julnes, editor of this issue, and many evaluation colleagues, especially Mel Mark, Kai Lee, Jim McDavid, and Eleanor Chelimsky. I also respect and value the many expressions of concern and requests for guidance from younger evaluators.

References

Alkin, M. C., & Christie, C. A. (2004). An Evaluation Theory Tree. In M. Alkin (Ed.), *Evaluation roots: Tracing theorists views and influences* (pp. 12–65). Sage.

Barnes, C., Bozi, L., & McFadden, K. *Exploring and ecosystem approach to management: A review of the pertinant literature*. Retrieved from https://pdfs.semanticscholar.org/2851/8c5e8055560f5fa35908c3ff4504ba3c5da6.pdf

BISE. (2016). Ecosystem services. Retrieved from http://biodiversity.europa.eu/topics/ecosystem-services

Blackman, A. (2012). Ex-post evaluation of forest conservation policies using remote sensing data: An introduction and practical guide. *Environment for Development*, *39*, 1–16.

Brundtland Commission. (1987). *Our common future*. Oxford, UK: Oxford University Press.

Caballero, P. (2015). Connecting the dots in 2015 for sustainable development. Retrieved from http://blogs.worldbank.org/voices/connecting-dots-2015-for-sustainable-development?CID=ENW_TTEnvironmentEN_D_EXT

Carson, R. (1962). *Silent spring*. Washington DC: Houghton Mifflin.

Charreire, H., Mackenbach, J., Ouasti, M., Lakerveld, J., Compernolle, S., Ben-Rebah, M., … Oppert, J.-M. (2017). Using remote sensing to define environmental characteristics related to physical activity and dietary behaviours: A systematic review (the SPOTLIGHT project). *Health and Place*, *25*, 1–9.

Chelimsky, E. (2012). Public-interest values and program sustainability: Some implications for evaluation practice. *American Journal of Evaluation*, *35*, 527–542.
Clark, W. C. (2007). Sustainability science: A room of its own. *Proceedings of the National Academy of Sciences*, *104*, 1737–1738.
Clark, W. C., Dickson, N. M., & Matson, P. A. (2008). Knowledge systems for sustainable development. *PMAS*, 4570–4578.
Clark, W. C., Mitchell, R. B., & Case, D. W. (2006). Evaluating the influence of global environmental assessments. In R. W. Mitchell (Ed.), *Global environmental assessments: Information and influence* (pp. 1–28). Cambridge, MA: MIT Press.
Clark, W. C., van Kerkhoff, L., Lebel, L., & Gallopin, G. C. (2016). *Crafting usable knowledge for sustainable development*. Washington, DC: National Academy of Sciences.
Cohen, A. (2010). Achieving healthy school siting and planning policies: Understanding shared concerns of environmental planners, public health professionals, and educators. *New Solutions: A Journal of Environmental and Occupational Health Policy*, *20*, 49–72.
Committee on Analysis of Global Change Assessments, National Research Council. (2007). *Analysis of global change assessments: Lessons learned*. Washington, DC: National Academy of Sciences.
Committee on the Science of Science Communication National Academy of Sciences. (2017). *Communicating science effectively: A research agenda*. National Academy of Sciences.
Courage, C., & Baxter, K. (2005). *Understanding your users*. San Francisco, CA: Morgan Kauffman.
Crowshoe, R., & Manneschmidt, S. (2002). *Akak'stiman: A blackfoot framework for decision-making and mediation processes*. Calgary, Canada: University of Calgary Press.
Dalbey, M. (2012). School siting and community building. Washington, DC: EPA Office of Sustainable Communities.
Daniela, T. C., Muharb, A., Arnbergerb, A., Aznarc, O., Boydd, J. W., Chane, K. M., & Costazaf, R. (2012). Contributions of cultural services to the ecosystem. *Proceedings of the National Academy of Sciences*, *109*, 8812–8819.
DeLauer, V. G., Rosenberg, A. A., Popp, N. C., Hiley, D. R., & Feurt, C. (2014). The complexity of the practice of ecosystem-based management. *Integral Review*, *10*, 4–28.
Elvidge, S. J. (2012). *The enabling state: A discussion paper*. Dunfermline: Carnegie Trust UK.
Enkelejda, S.-R., Butzbach, C., & Brousselle, A. (2019). Planetary health: countering commercial and corporate power. *The Lancet*, *3*, 30241–30249.
European Environment Agency. (2008). European Environment Agency. Retrieved from https://www.eea.europa.eu/publications/briefing_2008_2
Ferraro, P. J., & Hanauer, M. M. (2014). Quantifying causal mechanisms to determine how protected areas affect poverty through changes in ecosystem services and infrastructure. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 4332–4337.
Ferraro, P. J., & Pattanayack, S. K. (2006). Money for nothing? A call for empirical evaluation of biodiversity conservation investments. *PLoS Biology*, *4*, 482–488.
GEF Scientific and Technical Advisory Panel. (2018). *Integration: To solve complex environmental problems*. Washington, DC: GEF IEO.
Government of Canada. (2008, July 8). *Biobasics: Bio-indicators*. Retrieved from http://www.biobasics.gc.ca/english/View.asp?x=740
Gregory, R., & Tousdale, W. (2009). Compensating aboriginal cultural losses: An alternative approach to assessing environmental damages. *Journal of Environmental Management*, 2469–2479.

IEO: GEF (Independent Evaluation Office: Global Environment Program). (2017). *Measuring environmental outcomes using remote sensing and geospatial methods*. Washington, DC: Global Environment Facility.

IRP. (2019). *Global Resources Outlook 2019: Natural Resources for the Future We Want*. Oberle, B., Bringezu, S., Hatfield-Dodds, S., Hellweg, S., Schandl, H., Clement, J., and Cabernard, L., Che, N., Chen, D., Droz-Georget, H., Ekins, P., Fischer-Kowalski, M., Flörke, M., Frank, S., Froemelt, A., Geschke, A., Haupt, M., Havlik, P., Hüfner, R., Lenzen, M., Lieber, M., Liu, B., Lu, Y., Lutter, S., Mehr, J., Miatto, A., Newth, D., Oberschelp, C., Obersteiner, M., Pfister, S., Piccoli, E., Schaldach, R., Schüngel, J., Sonderegger, T., Sudheshwar, A., Tanikawa, H., van der Voet, E., Walker, C., West, J., Wang, Z., Zhu, B. A Report of the International Resource Panel. United Nations Environment Programme: Nairobi, Kenya.

Islam, S. S., & Susskind, L. (2018). Using complexity science and negotiation theory to resolve boundary-crossing water issues. *Journal of Hydrology*, *562*, 589–598.

Jacobs, J., Brasseur, G. P., Barron, E. J., Benedick, R., Chameides, W. L., Dietz, T., … Richels, R. (2007). *Analysis of global change assessments: Lessons learned*. Washington DC: National Academy of Sciences.

Jenkins-Smith, H. C., & Sabatier, P. A. (2008). Evaluating the advocacy coalition framework. *Journal of Public Policy*, *14*, 175–203.

Larsson, M. (2015). Effects on sustainable development from large environmental programs: A review of 16 evaluations. *Journal of Integrative Environmental Sciences*, *12*, 85–105.

Liu, J., Dietz, T., Carpenter, S. R., Alberti, M., Folke, C., Moran, E., … Taylor, W. M. (2007a). Complexity of coupled human and natural systems. *Science*, *317*, 1513–1516.

Liu, J., Dietz, T., Carpenter, S. R., Folke, C., Alberti, M., Redman, C. R., … Kratz, T. (2007b). Coupled human and natural systems. *Ambio*, *36*, 639–649.

McBride, M. F., Lambert, K., Huff, E. S., Theoharides, K., Field, P., & R. Thompson, J. R. (2017). *Increasing the effectiveness of participatory scenario development through codesign*. Ecology and Society, *22*, 848–858.

Lubell, M., Leach, W. D., & Sabatier P. A. (2009). Collaborative watershed partnerships in the epoch of sustainability. In D. A. Mazmanian & M. E. Kraft(Eds.), *Toward sustainable communities: Transition and transformations in environmental policy*[Online publication], MIT Press.

Mascia, M. B., Pailler, S., Thieme, M. L., Rowe, A., Bottrill, M. C., Danielsen, F., … Burgess, N. D. (2014). Commonalities and complementarities among approaches to conservation monitoring and evaluation. *Biological Conservation*, *169*, 258–267.

McKinnon, M. C., Mascia, M. B., Yang, W., Turner, W. R., & Bonham, C. (2015). Impact evaluation to communicate and improve conservation non-governmental organization performance: The case of Conservation International. *Philosophical Transactions of the Royal Society B*, *370*, 1–9.

Mermet, L., Bille, R., & Leroy, M. (2010). Concern-focused evaluation for ambiguous and conflicting policies: An approach from the environmental field. *American Journal of Evaluation*, *31*, 180–198.

Mickwitz, P. (2013). Policy evaluation. In A. Jordan (Ed.), *Environmental policy in the European Union: Actors, institutions and processes* (pp. 267–286). London, UK: Routledge.

Millennium Ecosystem Assessment. (2005). *Ecosystems and human well-being: Health—Synthesis report*. New York: Island Press.

Mills, J. N., Gage, K. L., & Kahn, A. S. (2019). *Potential influence of climate change on vector-borne and zoonotic diseases: A review and proposed research plan*. Retrieved from https://ehp.niehs.nih.gov/doi/10.1289/ehp.0901389

National Research Council, Committee on Reducing Stormwater Discharge Contributions to Water Pollution. (2009). *Urban stormwater management in the United States*. Washington, DC: National Academies Press.

Ndyeshumba, P. (2000). The use of remote sensing for environmental impact assessment and determination of the area affected by refugees in ngara district north western tanzania. *International Archives of Photogrammetry and Remote Sensing*, Proceedings XXXIII, 981–984.

Nicholas, P., Breakey, S., & Winter, S. (2015). Climate change, climate justice and environmental health issues. In S. Breakey, I. Corless, N. Meedzan, & P. Nicholas (Eds.), *Global health nursing in the 21st century*. New York: Springer.

NOAA. (2013). *National coastal population report: Population trends from 1970 to 2020*. Washington, DC: NOAA.

NOAA National Marine Fishery Service. (2016). *NOAA Fisheries Toolbox version 3.1*. Retrieved from http://nft.nefsc.noaa.gov/index.html

Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge, UK: Cambridge University Press.

Ostrom, E. (2009). A general framework for evaluating sustainability of social–ecological systems. *Science*, *325*, 419–422.

Patton, M. Q. (2011). Systems thinking and complexity concepts for developmental evaluation. In M. Q. Patton (Ed.), *Developmental evaluation* (pp. 111–151). New York: The Guilford Press.

Proulx, A. (2016). *Barkskins*. New York: Scribner.

Pullin, A. S., & Stewart, G. B. (2006). Guidelines for systematic review in conservation and environmental management. *Conservation Biology*, *20*.

Raggamby, A. v. (2012). *Sustainable development, evaluation and policy-making: Theory, practise and quality assurance*. Cheltenham, UK: Edward Elgar Publishing.

Rockwell, S. K., & Buck, J. S. (1995). An interdisciplinary and interagency evaluation team: Benefits for other-discipline specialists. *Evaluation Practice*, *16*, 239–245.

Rowe, A. (2012). Evaluation of natural resource interventions. *American Journal of Evaluation*, *33*(3), 384–394.

Rowe, A. (2018). Ecological thinking as a route to sustainability in evaluation. In R. Hopsin, & F. Cramm (Eds.), *Evaluation in complex ecologies*. Stanford: Stanford University Press.

Rowe, A. (under review). *Rapid impact evaluation*.

Savoie, D. (2013). *Whatever happened to the music teacher?* Montreal: McGill-Queen's University Press.

Susskind, L. T. (2012). Building concensus for sustainable development. In K. Zoeteman (Ed.), *Sustainable development drivers: The role of leadership in government, business and NGO performance*. Cheltenham: Edward Elgar Pub.

Sustainability Consortium. (2017). Sustainability consortium. Retrieved from https://www.sustainabilityconsortium.org/about/

Swilling, M. H. (2018). *The weight of cities: Resource requirements of future urbanization*. Nairobi, Kenya: International Resource Panel. United Nations Environment Programme.

TEEB (The Economics of Ecosystems and Biodiversity). (2016). *The economics of ecosystems and biodiversity*. Retrieved from http://www.teebweb.org/

Uitto, J. I. (2016a). The environment–poverty nexus in evaluation: Implications for the sustainable development goals. *Global Policy*, *7*, 441–447.

Uitto, J. (forthcoming). *Evaluation for sustainable development: Understanding the nexus of natural and human systems*. In G. Julnes.

Uitto, J. I. (2016b). Evaluating the environment as a global public good. *Evaluation*, *22*, 108–115.

Elkins, P., Hughes, N. Brigenzu, S., Clark C. A., Fischer-Kowalski, M., Graedel, T., … Westhoek, H. (2016). *Resource efficiency: Potential and economic implications. a report of the international resources panel*. United Nations Environment Programme.

University Library, University of Illionois at Urbana-Champaign. (2018). Citizen science: Evaluating citizen science. Retrieved from https://guides.library.illinois.edu/citizen-science/evaluation

US EPA (United States Environmental Protection Agency). (2003). *New middle school will reduce classroom crowding and bring better educational opportunities and open green spaces to the heart of neglected neighborhood*. Retrieved from https://archive.epa.gov/region1/brownfields/web/html/lowell_ksms_agp.html

van den Berg, R. D. (2011). Evaluation in the context of global public goods. *Evaluation*, 405–415.

van Mierlo, B., Arkestejn, M., & Leeuwis, C. (2010). Enhancing the reflexivity of system innovation projects with system analyse. *American Journal of Evaluation*, *31*, 143–161.

Weible, C. M., Heikkila, T., deLeon, P., & Sabatier, P. A. (2012). Understanding and influencing the policy process. *Policy Sciences*, *45*, 1–12.

Weighs, J. (2016). *School site size: How many acres are necessary*. Retrieved from http://old.cefpi.org/pdf/state_guidelines.pdf

Westhoek, H. I. (2016). *Food systems and natural resources*. Nairobi, Kenya: UN Environment.

Whitmee, S., Haines, A., Beyrer, C., Boltz, F., Capon, A. G., de Souza Dias, B. F., … Yach, D. (2015). Safeguarding human health in the Anthropocene epoch: Report of the Rockefeller Foundation–Lancet Commission on planetary health. *The Lancet*, *386*, 1973–2028.

Wikipedia. (2018, June 27). *Taxonomy (biology)*. Retrieved from https://en.wikipedia.org/wiki/Taxonomy_(biology)

Williams, R. (under review). Evaluation in dangerous times. *Evaluation Matters*.

Wondolleck, J. M., & Yaffee, S. L. (2017). *Marine ecosystem-based management in practice*. Washington, DC: Island Press.

Wunder, K. B.-R.-d.-B. (2015). Mainstreaming impact evaluation in nature conservation. *Conservation Letters*, *9*, 58–64.

Zaidi, A., Dickinson, T., & Male, T. (2015). *Incorporating natural infrastructure and ecosystem services in federal decision-making*. Retrieved from https://obamawhitehouse.archives.gov/blog/2015/10/07/incorporating-natural-infrastructure-and-ecosystem-services-federal-decision-making

ANDY ROWE is a former president of the Canadian Evaluation Society, has been awarded for his service to evaluation in Canada and is a Fellow of the Society. He has a PhD from the London School of Economics and has worked in universities, with government, a public/private oceans research corporation, and globally as a consultant.