Check for updates

American Journal of Evaluation
33(3) 384-394
© The Author(s) 2012
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1098214012440026
http://aje.sagepub.com
SAGE

Evaluation of Natural Resource Interventions

Andy Rowe[1]

Abstract

This article provides a frame for evaluation of natural resource interventions, which necessarily involves both human and natural systems. Two-system evaluands require us to adapt evaluation methods for comparison and attribution and to address differences in time and space occurring across the systems as well as potentially very different values among stakeholders. While two-system evaluands can be challenging, it does not follow that evaluation in these settings is necessarily more difficult than evaluations located solely in human systems settings. This article suggests that if we are concerned about use and influence, then our responses to challenging evaluation settings should not automatically favor additional rigor; they should prioritize salience and legitimacy through joint knowledge production processes with decision makers and stakeholders.

Keywords

evaluation, natural resource, environment, conservation

This article addresses evaluation of natural resource interventions where the distinguishing characteristic is an evaluand that always involves both human and natural systems and is located at the intersection of those systems. To conceptually frame evaluation of natural resource interventions, I avoid a distant Google earth view where the distinction between human and natural systems is blurred and adopt a more "close-up" view where evaluations are actually conducted with real evaluands featuring multiple connected and distinct human and natural systems. I refer to this as a two-system evaluand. The purpose of the article is to introduce to general evaluation audiences this framing for evaluation of natural resource interventions and illustrate some of the resulting methodological issues. By conceptually framing evaluation in natural resource settings, the article also provides some general guidance for evaluation practitioners working in these venues.

Assessments of the evaluation literature addressing natural resource settings are unusual; a current World Wildlife Foundation facilitated review of the evaluation literature in conservation appears to be a rare undertaking (Pailler et al., 2012). Currently, there is no systematic direction for evaluation specific to natural resource settings nor are there efforts to distinguish evaluation in these settings from other settings such as education or health where much of the conceptual and applied evaluation work has been located.

[1]ARCeconomics, Camden, SC, USA

Corresponding Author:
Andy Rowe, ARCeconomics, 1917 Carriage Hse Lane, Camden, SC 29020, USA
Email: andy.rowe@earthlink.net

This article posits a two-system evaluand as the characteristic that distinguishes evaluation in natural resource settings and suggests that this requires some adjustments to evaluation methods. For example, counterfactuals need to address both systems. Moreover, attribution to the intervention is challenging because each of the systems brings separate and important external outcome-affecting factors. Human systems have different temporal and spatial frames than natural systems, and individual systems within the natural system can have radically different temporal and spatial frames. Human interests and cultures differ in how they use and value natural resources, and sometimes an evaluation of a natural resource intervention will need to consider multiple and very different human values for the same resource. While evaluation of natural resource interventions must address additional challenges from two-system evaluands, it does not follow that evaluation in these settings is necessarily more difficult than evaluations located solely in human systems settings.

Natural Resource Interventions

Many natural resource interventions can be classified as either specifically targeted or more broad and encompassing. Examples of highly targeted interventions include U.S. Environmental Protection Agency (EPA) Superfund cases addressing toxic problems at particular sites, and protective interventions focused on addressing a particular high value at-risk species or habitat. Examples of broader, more encompassing interventions intended to address a class of problems include the design and implementation of voluntary standards for environmental protection; or strategies for designating protected areas in marine and terrestial settings or ocean zoning. Evaluation is a means of assessing the efficacy and effectivness of these interventions, regardless of scope.

The focus of this article is evaluation of human interventions directly targeting natural systems. This includes interventions authorizing or supporting fishing, mining, and other primarily commercial use of natural resources (resource use interventions); interventions that attempt to prevent or remediate damage to natural systems caused by human use either to protect human health (environmental interventions) such as those authorized in the United States under the Clean Air and Water Acts; or interventions to conserve species, habitat, and ecosystems (conservation interventions). Interventions to promote sustainable development worldwide have also gained prominence in view of the 1992 UN Conference on Environment and Development (Rio Process) and from growing and intensifying corporate concerns about sustainability of supply chains, increasing corporate social responsibility and from public awareness and concern for the environment. These and other forces are introducing sustainability as a new element into evaluations of how humans use natural resources (primarily resource use interventions) and will likely have growing influence on how we approach evaluation in all sectors. Many other human interventions such as urban and rural planning regulations could be variously considered as resource use, environmental, or conservation interventions. These illustrate potential extensions of the concepts addressed here, but none are directly addressed in this article.

Ecology is the natural science that directly considers the relationships that living organisms have to each other and to their natural environment. Tansley (1935) was one of the first to incorporate humans into concepts of ecosystems, both as a source of disequilibrium and destruction and also as source of new systems. Following Liu et al. (2007), I use the concept of coupled human and natural systems to acknowledge the complex interactions among human and natural systems at diverse spatial, temporal, and organizational scales. However, complex interactions in the totality of an ecosystem do not mean that evaluation of an intervention in that ecosystem is complex.

Why a Two-System Evaluand?

Readers are invited to consider a relatively simple natural resource intervention to establish a management structure and new rules for using off-road vehicles (ORV) on a national seashore such as Cape Cod in the northeast United States. The seashore provides habitat for year-round and seasonal wildlife, insects, and birds, including the Piping Plover, which is listed as endangered under the Endangered Species Act (ESA) affording it special protection. The dunes and shoreline provide important habitat for many species and are strongly affected by wind, shoreline wave, and tidal action as well as by human use including ORV. The National Park Service (NPS) is mandated to protect habitat and species and also to provide for human use of the park. The initiating actions for the new rule came from issues associated with ORV use, including a potential lawsuit under the ESA from an environmental group and aggressive responses of ORV users to Cape Cod Seashore permits and actions to protect the Plover. This illustrates an evaluand with coupled human and natural systems, where positive and negative effects can occur in either or both systems. Evaluation of the negotiated rule making for ORV use at Cape Cod National Seashore revealed that most of the incremental gains attributable to the rule-making process lie wholly in the human systems of park management including ORV rules and other recreational uses of the Seashore. Within the natural system, the NPS was already aggressively protecting the Plover and habitat critical for various resident and visiting species.

Without the human system, there is no intervention to evaluate, and study of the natural system alone without human interventions would not be an adequate evaluation. Thus, evaluation of these interventions involves examining issues, practices, and developments in both the human and the natural systems.

In natural resource interventions, there can be large disconnects in the worldviews of those undertaking different types of interventions. Those whose vocation is resource extraction or processing (e.g., fishing, mining, and logging) often view those who act to protect the natural system with suspicion. Resource extractors and processors may vigorously oppose modifying their access to resources or their extraction or processing methods. Likewise, some conservationists view the resource extractors with disdain and are often in adversarial and litigious disputes with them as well as with environmental protection agencies, who they think should do more to protect the environment under existing legislative authorities. Environmental agencies also draw the ire of industries that claim environmental regulations undermine their viability. These are often highly charged and politicized settings where representatives of the different constituencies have deep and difficult-to-negotiate differences with one another. Consequently, it is often difficult for evaluators to include all affected populations in the evaluation process as directed by **Utility Standard 2 of** the Program Evaluation Standards: "Evaluations should devote attention to the full range of individuals and groups invested in the program and affected by its evaluation" (Yarbourough, Shulha, Hopkins, & Caruthers, 2011). The challenge is illustrated by the current evaluation literature addressing conservation interventions where it is conventional to regard the human system as a threat to natural systems (Salafsky & Margoluis, 1999) and to evaluate only effects in the natural system (Ferraro, McIntosh, & Ospina, 2007). However, in the example (Rowe, 2009) of Cape Cod National Seashore, a singular focus on the natural system would have led to an erroneous evaluation conclusion of no effect since all effects that did occur were in the human system, and these were substantial.

The remainder of this article illustrates some of the implications of the two-system evaluand for methods to evaluate natural resource interventions and in doing so it also further introduces the premise of a two-system evaluand as the distinguishing characteristic for evaluation of natural resource interventions. The approach is largely descriptive because of the intent to introduce a new conceptual framing for evaluation in settings unfamiliar to most evaluators.

Does a Two-Evaluand System Affect Evaluation Methods?

An evaluation embedded in two-coupled systems faces unique issues related to the approach to the evaluation as well as the choice of methods. While the issues discussed below apply to all forms of evaluation—developmental, formative, and summative—this article uses summative evaluation settings to illustrate how

- comparison to an alternative requires a control or match to a two-system evaluand,
- attribution to the intervention must consider a wide range of differentiated contextual and intervening factors,
- judgments of worth can be complicated and the ethics and standards for judging worth ambiguous, and
- time and location differ conceptually and practically between the two systems.

Evaluators need to be attentive to erroneously assuming that methods-as-usual will consistently work without adaption for the two-system evaluand in natural resource settings. Evaluators working with two-system evaluands need to attend carefully to whether their methods choices address the unique characteristics of these evaluands.

Comparison to Identify the Incremental Contributions (Merit) of the Intervention. Many evaluators believe that comparison is critical to summative evaluation because without a comparison one cannot tell if observed effects were in fact attributable to the intervention in question. Typical evaluation designs using comparisons include: experimental and quasi-experimental designs, interrupted time series, and case studies (Government Accountability Office [GAO], 2009).[1] Others have suggested a much wider range of possible candidates for counterfactuals (Bamberger, Carden, & Rugh, 2009) and methods to construct a counterfactual (Hind, 2010) and (Rowe, 2011). While the relative merits of these methods are periodically a contentious matter, evaluators can usually identify a reasonable comparison that stakeholders and decision makers regard as salient and legitimate; for example, a population with similar characteristics that does not receive the intervention.

The challenge of finding suitable comparisons in two-system evaluands is often considerable. For example, the Marmot Round Butte hydro dam in Oregon was removed in 2007 as part of decommissioning the Bull Run hydro facility, an economic decision for the utility but one that also involved issues associated with the ESA, tribal rights and responsibilities, and water rights. Agreement to decommission the dam was the result of a highly successful mediated decision-making process, sponsored by the utility Portland General Electric (PGE) and including tribal, environmental, recreational, and commercial interests; engineering studies concluded that blowing up the dam was the most effective form of removal (Portland General Electric [PGE], n.d.). While the prospect of randomly blowing up dams to implement an experimental evaluation design in service of "high-quality" evaluation is appealing it is not a feasible or ethical option. More realistically, one would seek another means of comparison such as a matched control. This is also very challenging because matching the characteristics of the Marmot using other forms of dam removal would require matching on a long list of features including, similar geological (streambed and shore) and water (flow, seasonality, passage, sedimentation, and temperature) conditions; similar fish populations with similar physical conditions and affected by similar threats; areas that have undergone similar resource extractive use by the human system on the shoreline areas (e.g., similar agriculture and forestry), with similar passage and other issues affecting salmon mortality above and below the Marmot dam; and, similar human uses of the water and aquatic resources (commercial, recreational, subsistence, and spiritual/traditional). However, the human and natural systems are tightly intertwined in the case of the Marmot dam because most of the same interests (e.g., environmental, commercial, Tribal, and government) have a lengthy history and in particular most were part of another PGE licensing effort, also mediated, that began prior to and concluded after Marmot. This effort altered the level and distribution of social capital among the parties impacting the Marmot decision. Thus, connectivity among the interests active in the Marmot case also needs to be considered for the comparison. All in all, Marmot was a challenging undertaking suggesting that comparison can be tricky with evaluands located at the intersection of human and natural systems.

Attribution to the Intervention. Attribution requires distinguishing effects caused by the intervention from effects whose cause lies elsewhere. All natural resource interventions start in the human system and target immediate and long-term changes in one or both of the natural and human systems. External factors from both systems can complicate or confound attribution of observed changes in targeted outcomes to the intervention.

The life cycle of contemporary salmon illustrates the range of external factors making attribution more challenging and also shows the many points where human and natural systems couple and why both systems need to be included in an evaluation such as for Marmot. Salmon hatch in gravelly freshwater rivers and streams where they spend the first portion of their life before swimming out to sea as smolts. This requires them to migrate down rivers that have been dramatically altered over the past century. Numerous dams constructed for irrigation and hydro generation create barriers to fish passage and have significantly reduced water flow by drawing large amounts of river water for agriculture and the needs of large urban areas, especially those in arid locations. The young fish must successfully navigate the many physical barriers to their passage and those portions of the rivers where flow has been reduced to levels that are problematic for migrating salmon, where water temperatures have risen to risk levels and where the cooling and protective benefits of riparian growth has been removed through agriculture, forestry, or residential development. These challenges increase mortality of migrating smolt and impair the health of those who successfully reach the sea. Once at sea, the salmon may encounter drift nets, legal and illegal commercial harvests, changes in ocean chemistry and temperature, and diminished food sources during 1–7 years they spend in the oceans, depending on species. Those individuals that survive return to complete the reproductive cycle, encountering the same barriers that they successfully navigated years earlier except now they are much larger so their needs for water flow and open passage are greater, and of course, they are now the target of commercial fishers and recreational anglers. In this example, specific human and natural systems intersect at many temporal and spatial points and we know that mortality caused at successive intersection points is cumulative (Lee, 1993).

Evaluators must be able to separate the contribution of removing a single dam to the relevant portion of the river's salmon population from all of these other factors; a difficult undertaking that evaluators must address in natural resource settings. Doing this requires a range of tools such as simulation models, or triangulation with expert judgments and using fish stock monitoring data and other information sources, or by moving down a level to where it is possible to attribute changes to the removal of the dam by considering less-confounded outcomes that strongly influence salmon survival such as water flow or passage. Attribution in challenging natural resource settings such as the Marmot decommissioning has stimulated development of a new approach for counterfactuals and enhancement of existing approaches. These are reported elsewhere (Rowe, Colby, Niemeyer, & Hall, 2004).

However, the evaluation of natural resource initiatives does not always challenge the evaluator seeking attribution. For example, at many U.S. EPA Superfund sites, the problem is a toxic substance such as heavy metals in soil and their removal is carefully recorded. Environmental health economists are able to convert this to a measure of reduced exposure and link this to estimates of the effects on human health. Attribution and measurement of the typical removal decision is complicated, as is making the connections between removal of the toxic substance and public health effects, but attribution and measurement are less problematic than for a mobile species such as salmon with many cumulative transactions with human systems.

Specifying Time and Location for Two-System Evaluands. It is always necessary to locate any evaluand in time and space, but a two-system evaluand introduces at least two different temporal and spatial structures that complicate the evaluation undertaking.

Time. Evaluation always occurs within the time frame relevant to the human system, often within a few years after the intervention is implemented. Changes in the human system are generally observable earlier than those in the natural system. For example, interventions such as specific site actions or policy and regulatory decisions seek to modify human behavior, and thereby limit, mitigate, or end adverse human system effects on the natural system. Many of the behavioral changes in the human systems can be observed or reliably predicted at the time of the evaluation and so do not require any special methodological attention.

Changes in the natural systems can take much longer. For example, it takes 15 or more years before substantial numbers of species recruit into tropical forest restoration areas, and twice as long before many of the unique values of tropical forests are realized (Ghazoul & Sheil, 2010). Another example from marine systems is provided by efforts to restore populations of Northern Cod in eastern Canadian waters, once the most productive fishery in the world. Although it has now been almost 20 years since the moratorium on harvesting was imposed, populations have not recovered and, as recently as 2011, all populations of Northern Cod were listed as endangered by the Canadian equivalent of the ESA. While natural systems have the potential to recover even after sustained periods of deprivation and damage, evaluation often occurs well before these natural systems have fully recovered, and even before they are showing any systematic or early signs of recovery (Fisheries and Oceans Canada, n.d.). And unfortunately, as with Northern Cod, expected recoveries do not always occur.

Generally, we can assume that human patience expires within 10 years or so, while changes will emerge in the natural system over a long time and over that period it is likely that human pressures on and interests in the natural system will also change.

Variants of program theory approaches enable evaluators to stretch the temporal reach of their judgments beyond what can be observed when the evaluation is undertaken (GEF Evaluation Office & Foundations of Success, 2007). This has proven to be a useful approach for many challenging evaluation settings. However, the challenge in natural systems is not just longer time frames, but the mix of different time horizons and spatial scales between and within natural and human systems included in an evaluation of a natural resource intervention. Often there is also considerable contingency and uncertainty and little if any agreement on the causes of observed problems. For example, the decline of Northern Cod stocks and their failure to recover is variously attributed to rapacious Canadian or Canadian-licensed fishing companies inside, and foreign fleets outside the 200 mile management zone; oceanographic conditions; destructive fishing technology and practices; and greedy gray seals.[2] Science knowledge provides insufficient guidance for formulation of a credible program theory. However, program theory is still a valuable approach in less complicated interventions such as local fishery-based sustainable development efforts, especially in very productive waters where regeneration of stocks can be rapid; it is also an important element in the mix of methods required for more complicated settings.

The shorter time frames of human systems means that evaluations are often undertaken when human behaviors can be observed or at least forecast with reasonable reliability and before all but the earliest targeted changes in the natural systems are observable. This complicates the conceptual and measurement aspects of evaluation in these settings and introduces the possibility that evaluations will have more credibility for those interested in the human system compared to those whose interests are in the natural system where observations and evaluation judgments can often and by necessity be more tentative with a weaker empirical basis. This can affect decision maker and stakeholder perceptions of the utility of evaluation for many natural resource interventions and impair prospects for use and influence of evaluation, especially for conservation interventions where the focus of the intervention is primarily the natural system.

Space or location. Natural resources are not necessarily contained within a single political boundary and scarce resources such as water, gold, diamonds, and ivory often cause conflict within and between nations and may result in boundary changes. Ownership and rights of access may be poorly defined or contested, and our management structures usually follow human geographies and rarely align with the territory of the natural system where ecosystem is the applicable spatial category. Thus, some evaluands with concepts of space fixed in the human system require different spatial boundaries for the natural system. Nowhere is this truer than with migratory and mobile species, especially birds, fish, and marine mammals but also wide-ranging species such as wolves and elephants and resources located beyond demarcated and agreed marine boundaries such as current debates over ownership and access to Arctic resources that will become accessible as the polar ice melts.

Many natural resource interventions can be tied to specific locales, such as treating decades of deposits of toxic substances such as polychlorinated biphenyls (PCBs). Deposits on land can be fairly well defined such as at the manufacture site or where PCB-laden soil from the site has been moved, for example, to school playgrounds or for housing construction as was the case with the General Electric (GE) Pittsfield site. Deposits are more mobile in water because of different rates of settling, again for GE Pittsfield, the removal and treatment from the Housatonic River was different for the first half mile from the site, the next one and a half miles and for the remainder of the river (U.S. Environmental Protection Agency [EPA], n.d.). Terrestrial impacts are usually over a much narrower spatial frame than impacts in water or air, but even on land-targeted species such as wolves and elephants roam wide areas without regard to human boundaries. This is not unique; effects in human systems can also be spatially specific. The benefits of school improvement are largely experienced by those associated with the improved schools but can radiate outward with more dispersed social and economic effects.

For some evaluands, it is important to specify the neighboring landscape as well as the landscape specific to the evaluand. For the Marmot dam removal, it was important to specify neighboring uses on the river banks so that it was possible to include important contextual factors such as sedimentation and nutrient leakage from agriculture and changes in water acidity and flow resulting from upstream logging operations. Most of these affect the same fish habitat as removal of the dam; while not directly associated with the decision being evaluated, they do affect outcomes. Interventions targeting water or air have stronger prospects for outcome-affecting upstream or upwind externalities compared to terrestrial interventions; it appears plausible that the need to consider externalities from neighboring locales is similar between terrestrial natural resource interventions and interventions targeting the human system alone. After all, humans are a terrestrial species.

Spatial boundaries of the evaluand usually reflect the interests and structures of the human systems and there is no reason to expect natural system boundaries to align with the boundaries of the human systems, including resource management boundaries. This can be a significant challenge, especially since resource management organizations are frequent commissioners of evaluations and have a strong preference for structuring the evaluation to align with the spatial areas they are accountable for managing.

Two-System Evaluands and Value. Many of the techniques available for evaluating two-system evaluands were developed by economists addressing natural resource issues. As a result, valuing in natural resource settings might be considered less challenging than for human settings and for settings where natural and human systems connect. Sensible application of economic techniques such as contingent valuation and benefit transfer (Kirchhoff, Colby, & LaFrance, 1997) can address some of the valuation requirements for evaluations of natural resource interventions, especially where the systems are not highly connected. For example, there are good possibilities of monetizing the value of observed changes in fish stocks, recreational opportunities, forests, mining, and other resource outcomes. There are also techniques for estimating the value of changes in human health status resulting from environmental interventions, judging the value to humans of food security, livelihoods, and ecosystem services, and assigning values to conserving coral reefs and similar market and quasi-market species. In the many instances where monetary valuation is not possible, we can still talk about, for example, changes in the number of bird years for Piping Plover and qualitatively describe human and nonhuman benefits associated with the intervention.

Differences in human cultural values present a particular challenge to valuing in natural resource settings. Scarce and endangered species often have a place of prominence in one or more cultures. For example, the aphrodisiac value of parts of threatened or endangered species, or the relative indifference of many tribal cultures to wild and wild-interbred-with-hatchery salmon compared to the adamant preference for wild salmon by recreational anglers of European origin. Allen Putney (International Union for Conservation of Nature [IUCN], 2008) provides a useful description of the spiritual values of landscapes. The value of ground rhino horn in eastern medicinal markets is higher than world prices for gold, diamonds, or cocaine (Huffington Post, 2011). Bears provide another good example: some male human populations value bear gall bladders highly as an aphrodisiac, for some North American tribal and aboriginal nations the bears have great spiritual significance,[3] while some North Americans enjoy killing bears and others are dedicated to protecting them as important members of ecosystems. Evaluating any initiative targeting bears will encounter serious issues when valuing observed changes in bear populations. Similarly, cessation of commercial whale hunting would be judged a positive effect by Western values, negative for some Eastern and far northern. This raises challenges of ensuring culturally appropriate valuation of effects in the natural system for the affected human systems, not just those applicable to the culture from which the intervention originates.

Evaluating natural resource interventions often directly engages competing values and radically different values. Ultimately, it can require difficult evaluative judgments between competing sets of culturally appropriate values, or at least careful enumeration of the effect that different values have on evaluation judgments.

Achieving Use and Influence. A commonplace assumption is that more difficult problems require more rigorous methods. Recent research about use and influence of science knowledge suggests that increasing rigor is not a promising approach to improved prospects for use and influence. Bill Clark and colleagues (Clark, Mitchell, & Cash, 2006) examined factors affecting use of Global Environmental Assessments (e.g., the Intergovernmental Panel on Climate Change, the Global Biodiversity Assessment, and Millennium Ecosystem Assessment) and observed that prospects for use and influence are strongly enhanced when the knowledge generation is undertaken as a social process with a focus on joint production of knowledge by domain experts and users (stakeholders, decision makers, affected entities, and communities). Their work aligns with research on evaluation use and influence (Kirkhart, 2010; Kelli Johnson, 2009; Mark & Henry, 2004) and provides succinct guidance toward enhancing use from sources familiar and credible to those from natural and physical science domains. Clark, Mitchell, and Cash identify three attributes observed to strongly influence use: salience (relevant, timely, and on point), legitimacy (addresses concerns and processes, regarded as fair), and credibility (good quality knowledge and methods). Only the latter is associated with increased rigor, the first two are associated with knowledge production processes. Working with Kai Lee of the Packard Foundation Science Program, we applied the work of Clark and his colleagues to develop a theory of change of science knowledge use, termed Linking Knowledge with Action (LKwA). The LKwA theory of change implies very different approaches from the implicit default in natural and physical applied research where routes to use are presumed to pass first through peer-reviewed publication followed by promotion and dissemination by the knowledge producer or others (David and Lucile Packard Foundation, 2010).

Applying this theory of change about use of science knowledge to evaluation of natural resource interventions suggests that evaluators should address the additional challenges arising from the two-system nature of these settings with a social knowledge process that promotes salience and legitimacy improving prospects of use and influence rather than prioritizing credibility by ramping up rigor. This is a perspective from science that can help evaluators improve prospects for use and influence in general, and especially for harder settings where we might elevate rigor over use. The guidance of John W. Tukey is also useful here (Tukey, 2011).

"Far better an approximate answer to the right question, than the exact answer to the wrong question."

Conclusion

This article has proposed that evaluation of interventions in natural resource settings is distinguished by an evaluand comprised necessarily of coupled human and natural systems. In these coupled systems, evaluators encounter the problem of judging value within diverse spatial, temporal, and organizational scales. Thus, evaluation of human interventions in natural resource settings should always consider the evaluand to consist of both human and natural systems and address both systems in the evaluation, described as a two-system evaluand. While strong differences in the worldviews of key human interests in the natural world (e.g., conservation and resource extractors) might promote opposition to this view, this article suggests that evaluators addressing natural resource interventions interpret the Utility Standards for evaluation as providing direction to devote attention to all affected human and natural system interests.

The article has described why evaluations need to be adapted to address two-system evaluands in natural resource settings. Focusing on summative evaluation approaches, I have suggested that the two-system evaluand introduces some additional considerations in selecting counterfactuals and that attribution to the intervention can sometimes be very challenging. I have also described why an evaluation needs to address multiple temporal and spatial settings as well as the challenge of dealing with multiple and very different human values for some species and other natural resources. The argument that evaluation approaches and methods need to be adapted for particular settings is not unique to natural resource interventions, nor do I want to suggest that natural resource settings are necessarily more complicated than all other settings where evaluation is applied. However, it is a useful reminder that our methods need to be tailored to the evaluand and that sometimes this requires attention to very basic questions such as what are the appropriate temporal and spatial frames for this evaluand, and whose values count.

Recent work on use and influence of science knowledge provides a useful insight into how, when evaluators are faced with challenging evaluands, we balance rigor against use and influence. It suggests that if we are concerned about use and influence, then our responses to challenging evaluation settings should not automatically favor additional rigor but should prioritize salience and legitimacy through joint knowledge production processes with decision makers and stakeholders.

The intention of this article has been to introduce a conceptual frame for evaluation in natural resource settings suggesting that they are distinguished by a two-system evaluand with coupled human and natural systems. A secondary purpose has been to remind evaluators that the character of the evaluand matters and that we always need to be attentive to how this can and should affect our methods.

Acknowledgments

The author thanks the editor, reviewers, and Eric Trum for their helpful comments and suggestions. The author also thanks Kai Lee and Jack Putz for their inputs.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. GAO also identifies expert opinion as an appropriate method for summative evaluation; however, it does not necessarily use comparison.
2. As this article was being revised, Canada announced a plan for a 5-year annual cull of 140,000 adult gray seals to help restore Northern Cod stocks. The predictably acrimonious discourse that followed this announcement nicely illustrates the challenge of locating cod and seals in a theory of change, especially identifying seals as a mechanism of change in cod populations. In fact, the Canadian Science Advisory Secretariat, a branch of the federal agency that will presumably approve the cull states "... science evidence needed to justify a cull is highly uncertain and indirect" (Bowen & Lidgard, 2011, p. 16).
3. For example, the Spirit Bear in British Columbia (see the Sprit Bear Youth Coalition http://74.54.116.9/).

References

Bamberger, M., Carden, F., & Rugh, J. (2009, November). *Alternative approaches to the counterfactual*. Retrieved August 16, 2011, from Real World Evaluation: http://www.realworldevaluation.org/Counterfactuals.html

Bowen, W. D., & Lidgard, D. (2011). *Vertebrate predator control: Effects on prey populations in terrestrial and aquatic ecosystems*. Maritimes Region: Canadian Science Advisory Secretariat.

Clark, W. C., Mitchell, R. B., & Cash, D. W. (2006). Evaluating the influence of global environmental assessments. In R. B. Mitchell, W. C. Clark, D. W. Cash, & N. M. Dickson (Eds.), *Global environmental assessments* (pp. 1–28). Cambridge: MIT Press.

David and Lucile Packard Foundation. (2010). *Linking knowledge with action*. Retrieved November 5, 2010, from http://www.packard.org/assets/files/conservation%20and%20science/Linking-Knowledge-with-Action.pdf

Ferraro, P., McIntosh, C., & Ospina, M. (2007). The effectiveness of the US endangered species act: An econometric analysis using matching methods. *Journal of Environmental Economics and Management*, *54,* 245–261.

Fisheries and Oceans Canada. (n.d.). *A fishery decision-making framework incorporating the precautionary approach*. Retrieved August 30, 2011, from Fisheries Renewal: http://www.dfo-mpo.gc.ca/fm-gp/peches-fisheries/fish-ren-peche/sff-cpd/precaution-back-fiche-eng.htm

GEF Evaluation Office & Foundations of Success. (2007). *GEF impact evaluation: Final report on a proposed approach to GEF impact evaluation*. Washington, DC: Evaluation Office Global Environment Facility.

Ghazoul, G., & Sheil, D. (2010). *Tropical rain forest ecology, diversity, and conservation*. Oxford, England: Oxford University Press.

Government Accountability Office. (2009). *Program evaluation: A variety of rigourous methods can help identify effective interventions*. Washington, DC: United States Government Accountability Office.

Hind, J. (2010). Additionality: A useful way to construct the counterfactual qualitatively? *Evaluation Journal of Australasia*, *10*, 28–35.

Huffington Post. (2011, September 25). *Huff Post*. Retrieved September 25, 2011, from Rhino Horns Worth More Than Gold, Leads To Wave Of Thefts In Europe: http://www.huffingtonpost.com/2011/08/26/rhino-horns-worth-more-th_n_938278.html

International Union for Conservation of Nature and Natural Resources, Gland, Switzerland. (2008). *Sacred dimensions: Understanding the cultural and spiritual values of protected areas*. Author.

Kelli Johnson, L. O. (2009). Research on evaluation use: A review of the emperical literature from 1986 to 2005. *American Journal of Evaluation*, *30*, 377–410.

Kirchhoff, S., Colby, B. G., & LaFrance, J. T. (1997). Evaluating the performance of benefit transfer: An empirical inquiry. *Journal of Environmental Economics and Management*, *33*, 75–93.

Kirkhart, K. (2010). The usefullness of utilization-focused evaluation (fourth edition) by Michael Quinn Patton. *American Journal of Evaluation*, *31*, 588–596.

Lee, K. (1993). *Compass and gyroscope: Integrating science and politics for the environment*. Washington, DC: Island Press.

Lee, K., & Rowe, A. (2012). Linking knowledge with action: An approach to philanthropic funding of science for conservation.

Liu, J., Dietz, T., Carpenter, S. R., Folke, C., Alberti, M., & Redman, C. L., Schneider, S. (2007). Coupled human and natural systems. *AMBIO: A Journal of the Human Environment*, *36*, 639–649.

Mark, M. T., & Henry, G. T. (2004). The Mechanisms and outcomes of evaluation influence. *Evaluation*, *10*, 35–57.

Mitchell, R. B., Clark, W. C., Cash, D. W., & Dickson, N. M. (2006). *Global Environmental Assessments: Information and influence*. Cambridge: MIT Press.

Pailler, S., Thieme, M., Rowe, A., Pullin, A., Burgess, M., Danielsen, F., et al. (2012). Measuring success in conservation. Manuscript submitted for publication. *Conservation Biology*.

Portland General Electric. (n.d.). *Sandy River*. Retrieved January 19, 2012, from http://www.portlandgeneral.com/community_environment/initiatives/protecting_fish/sandy_river/default.aspx

Rowe, A. (2009). *Environmental and Economic Effects of ORV Rules at Cape Cod and Fire Island National Seashores*. (Unpublished report). Department of the Interior, U.S. Government.

Rowe, A. (2011). *Getting past the stale RCT debate: A new approach for impact evaluation*. Retrieved September 10, 2011, from SAMEA Virtual Conference: http://wpeg.wits.ac.za

Rowe, A., Colby, B., Niemeyer, M., & Hall, W. (2004). *Evaluating environmental and economic effects of collaborative decisions*. (Unpublished report). Hewlett Foundation, Menlo Park, CA.

Salafsky, N., & Margoluis, R. (1999). Threat reduction assessment; A practical and cost-effective approach to evaluating conservation and development projects. *Conservation Biology*, *13*, 830–841.

Tansley, A. (1935). The use and abuse of vegetational concepts and terms. *Ecology*, *16*, 284–307.

Tukey, J. (2011, July 28). *John Tukey*. Retrieved September 20, 2011, from Wikipedia: http://en.wikipedia.org/wiki/John_Tukey

U.S. Environmental Protection Agency. (n.d.). *GE/Housatonic River Site in New England*. Retrieved September 20, 2011, from http://www.epa.gov/region1/ge/index.html

Yarbourough, D., Shulha, L., Hopkins, R., & Caruthers, F. (2011). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks, CA: SAGE.