Chapter 21: Evaluating progress with a transdisciplinary science framework

21.1 Introduction

This text has described a method for addressing environmental problems that brings in complexity, uncertainty, and values. This is necessary because most of the problems that we are dealing with as a society are "wicked". These problems contain uncertainty and non-coherent values between individuals and the community, which are simultaneously interacting and. A key characteristic of "wicked" problems is that they are never solved; there is no stopping rule that tells us we are done. Wicked problem example 1. Wicked problem example 2. Thus, from the beginning, we should not expect clear outcomes that signal success and completion. Instead we must rely on a constant process of evaluation and iteration. There may be pieces of the projects that can be addressed with traditional scientific hypothesis testing, but for the larger flow of the projects we will have to rely on a more reflective epistemology. We have to learn from our efforts and make adjustments while we continue to work on the problem.

21.2 Defining a scientific evaluation process

The evaluation will be scientific in that it is systematic, rigorous and verifiable. We need to use a restricted definition of science that does not assume everyone involved agrees on what a "fact" is or how to verify if a fact is true. Outside bench science, and in any enterprise that includes the public, the assumption that there is a single method to verify what a fact is just doesn't hold. Our modified definition of science also needs to avoid the implication that the use of technology is required or any biases that science will lead to progress. Instead we can define science as:

> Science is a rigorous, systematic and iterative activity that builds knowledge through seeking empirical evidence and making testable predictions followed by evaluation and revision. The activity should build knowledge that can be reliably used by others.

This definition can be applied to assessing activities that are creating new types of knowledge. A key characteristic of this knowledge is that it is created and shared by scientists, professionals and the public. We will also be able to use this definition of science to describe quality and measures of success that will lead to identifying good practices.

It is important that values are considered as part of the evaluation. We can stay within objective, fact-driven, decision processes by creating objective statements about values. For example, "stakeholder group X values biodiversity more than it values an efficient and large water treatment plant" can be treated as a fact and can be verified with members of stakeholder group X. This is a statement about values, not a value judgment on the part of the observer. Extending this to include the judgment or members in stakeholder group X, you might state that they favor biodiversity more than a sewage treatment plant "because they feel there is ample evidence that the threatened biodiversity loss can't be replaced, and they don't want to make the trade-off to lose any species to this proposed project". Again, the judgment criteria are described as they hold for this group. This is not a statement that it is a fact that there will be a loss of biodiversity or that biodiversity is more important than the sewage treatment plant. Thus observers and coordinators can make objective statements about values, but for stakeholders to be involved they have to make their own statements.

In order for stakeholders and participants to inject their values into scientific judgments, they must make statements that are based on evidence for the problem at hand and are not allowed to introduce non-negotiable demands, or pre-experiential beliefs of dogma. For example an involved citizen might make a statement such as "Based on the evidence that I've seen and my analysis, I believe that we should create a reserve for endangered White-Tailed Deer." It would not be useful in this scientific evaluation process for them to say "I consider preserving these deer to be a sacred trust and I cannot discuss any project that would compromise their chances of survival to any extent." The first statement is what we referred to in Chapter 19: Scientific Adaptive Management as a "considered value", i.e. the person is basing the value on evidence pertinent to this particular decision and willing to consider changing their belief if different evidence were made available or if they were presented with a different analysis of the problem. The second example is what we called a "held value", in which the holder of this value will not consider any other information. Strongly held values are important for civilization and are handled by social and political mechanisms. Scientific evaluation cannot reconcile conflicts that arise between these beliefs and must limit its focus to the region of facts and considered values. One of the powerful aspects of scientific evaluation comes from drawing the line between considered and held values; doing so centers the discussion in a situation where everything (including values and beliefs) must be based on pertinent evidence. A "litmus" test for stakeholders is that they should be able to describe evidence or analysis that would make them revise their beliefs. Applying these criteria for evidence at the beginning of a decision process should improve the flow of the deliberation and allow that process to be rigorous and systematic without discarding important information about how participants' values and beliefs.

21.3 Evaluation of personal progress

Thoughtful and deliberate citizens should always be evaluating if their effort to learn about a problem has been valuable; i.e. to ask the question, "Has my effort been worth it?" Answering this question should involve examining the progress made but also assessing whether you think you're on the right path. Will this approach to learning and acting on an environmental problem meet your goals? At some level this is second nature to all of us, but the intentional self-evaluation should include more than an itemization of the specific tasks completed. Based on what you have learned so far in addressing the problem, you need to ask yourself if the goals that you set are still appropriate. It may be your engagement with the problem has changed your understanding or values and you need to restate your goals. For example, you may have been working on cleaning up a streambed with the goal of creating an attractive natural area, but in doing so you realized that removal of some barriers downstream would make this whole area accessible to native fish. In this case, the engagement refined your goals to focus on a stricter definition of what a natural area should entail. Or you might have been cleaning up a streambed only to realize that the sources of pollution and litter upstream were un-controlled. You might shift your focus to addressing that problem or, if you believe it is an insurmountable problem, you might pick another stream to volunteer on. A very challenging re-evaluation and re-adjustment involves considering the level of uncertainty and complexity of the problem as you first imagined it and how that might have changed with your increased knowledge. As you learn more about any wicked problem and become personally involved, your level of uncertainty is bound to go up and even call into question your personal values and beliefs. You should not dismiss this because this level of re-evaluation is the most valuable form of learning; however, you do have to take the longer view, as described elsewhere in this text, one that embraces the uncertainty that will eventually be valuable.

On a procedural level, an evaluation of your personal involvement in a problem should examine which approaches and tools you have brought to bear and their effectiveness. You should look at the discovery and diagnostic tools that were employed and how much effort was assigned to each (informally or deliberately). This should lead to re-allocating effort between approaches or adding additional approaches that now seem potentially effective.

The personal reflection described above is scientific because it is systematic, rigorous, iterative, and includes values. It is systematic because one must evaluate all of the inputs and efforts in order to gain new knowledge. The rigor comes from testing each portion of new understanding down to the level of questioning original assumptions to see if they still hold, and, in the event that they do not, creating new goals. This evaluation needs to take place concurrently with approaches to solving the problem so that adjustments can be made or a whole a new set of objectives can be iterated if necessary. Finally, personal values are stated with respect to whether progress is made toward intended goals and whether or not efforts have been worth it. The ability to evaluate your own progress without becoming paralyzed by the uncertainty of whether you are doing the right thing is learned through experience and perseverance.

21.3 Multiple-participant project evaluation.

The evaluation of projects uses the same basic template as self-evaluation. The process includes examining progress on tasks and objectives, re-evaluating goals, assessing the value of the knowledge gained and coming to grips with the uncertainty that has been created through the creation of new knowledge.

One major difference for evaluating environmental projects is that the problems are situated in authentic communities that have varied social, economic and scientific issues. For example, addressing the progress on establishing fishing quotas and a marine reserve would have to start by acknowledging where the community was socially, economically and environmentally at the beginning of the project and working from there. This can be challenging because participants may have very different and conflicting descriptions of the previous state of the resource. The evaluation process will be different than a strictly technical project

*** in five key ways. First, the project should be creating more knowledge and this knowledge should include new types of information that might not have been predicted at the start. Thus the evaluation of a possibly successful project has to expand its original definition of knowledge. Second, the evaluation needs to be contextualized in the community, not in the participating academic disciplines. This includes using everyday language as the dominant form of communication and avoiding silos of expertise within the project. Third, different participants may have different and non-converging definitions of success. The goal of the evaluation should be to accurately state the range of definitions, not to force convergence. Fourth, the ultimate solution may require a paradigmatic shift in the community. This means that the evaluation would document the discontinuity from one way of doing business to another disconnected method. Such paradigmatic shifts are often un-predictable and can't be described in terms of cause and effect. In essence, the shift in paradigm may be supported by many contributing factors but no one set would force the change. This fourth condition is very similar to the fifth, which is that the path to success may not be deterministic but may rely on some emergent behavior of the system. For example, a public campaign to clean up a stream may drag along for quite awhile until a critical threshold of participants and social connections is met and then progress takes a leap forward. There is no way to engineer getting to that threshold or even replicating it. Global sustainability may be the most important instance of emergence. We might have to all be doing all the right sustainable "things" and then, by some stroke of luck which we don't understand, there could be a global paradigm shift and the condition of sustainability would emerge. In these five descriptors, it will be necessary to document the different requirements that each stakeholder group brings to the project and maintain broad language that acknowledges contradictory values. This is important because the purpose of this evaluation is to re-evaluate approaches and goals. Remember that with most interesting and challenging environmental problems, we are in an infinite, iterative loop. There will never be a final report.

21.4 Engaging in the solution of environmental problems produces new knowledge

One of the major differences between traditional science and the transdisciplinary approach to science that we have adopted here is that our approach creates different types of knowledge that can't be reviewed and assessed very easily through peer review. A major strength of traditional science is that the peer review process is a robust mechanism for both improvement and building trust. However, environmental projects create many types of knowledge that may be inaccessible for anyone outside the project to assess. For example, a project that is restoring wetlands may develop and test hypotheses that lead to publications and presentations. These products can be peer reviewed. However, there is additional knowledge created by the wetland managers and the staff that did the restoration work. Some of this could be captured with written narratives of the processes, but some of it is tacit knowledge that allows the teams and team members to remove invasives and plant natives in just the right way. This leads to two major differences between "traditional" and what we call "mode 2" science (ref ***). First, the full team that is responsible for the project will be diffuse and ephemeral. The project is planned, carried out and then the team disperses to work on other projects. The people involved in the project are probably trained in a wide variety of disciplines, which complicates the analysis. Assessment of a successful project must include how well the members of the team worked together and whether the final "product" is illustrative of the team meeting its goals and objectives, not the narrative or summative evaluation. The assessment of an unsuccessful project would be even more problematic. Is there evidence to determine that the reason for failure was based on unrealistic objectives, poorly applied principles, applying the wrong principles or ineffective implementation? Although sorting this out would be very valuable, many failed projects seem to erode away with no clear statement of failure that would trigger an evaluation. Fortunately there are two important characteristics of a project that can be evaluated with "mode 2" scientific approach and can help establish a high degree of rigor and reliability.

Successful solutions will build the participants capabilities. The traditional conception of technology transfer is that information becomes available and is used in new instances. In Mode 2 science, the technology is transferred through the people who are involved. They learn information, techniques and skills that allow them to perform tasks and analysis required by for project. The "technology" is the human capital that develops, not specific knowledge products (such as publications) or machinery. The test of the quality of these capabilities is whether the participants join subsequent projects and contribute to other successful efforts. Thus, instead of judging the quality of a project by the production of a static and reliable publication (as in traditional science), quality is judged on the value added to a dynamic network by the diffusion of innovation.

Just as good traditional science has activities that are considered good practice, Mode 2 / transdisciplinary / project based science has characteristics that indicate good practice. In both cases, good practice is a necessary condition and does not guarantee high quality. There are three categories of good practice. First, there needs to be a high communication density. Information needs to flow back and forth between all of the participants and into each social and economic sector that is involved. The connections can be characterized using network descriptors. In particular, the structure of the communication network should have relatively high connectivity across the entire community, but there may be interesting brokerage and holes that help define information flow within the community. Another parameter that can be used to track the value of the network is "ascendency". This parameter is a measure of whether the right information got to the right person in a timely manner. High ascendency is desirable but requires infrastructure and investment in communication and social networks. Second, the number of sites where this approach is adopted indicates good practice. The participants with experience in best practices will use those in other venues and subsequent projects. Finally, tracking the diffusion of key innovations will demonstrate a valuable outcome of the originating project. These three qualities combined can be used to describe the quality of a project, i.e. it should have high communication density during the project, participants involved should use a similar approach in other successful projects, and key innovations should appear in subsequent successful projects.

21.5 Shift from accumulating knowledge to designing solutions

Transdisciplinary, problem-based science is similar to applied science or consultancy (Funtowicz and Ravetz ****). The focus is on the specific issue and its context. Solutions must have particular structures to deal with aspects of the problem, just as control systems need to have the same level of complexity as the system being controlled, the so-called Ashby's Law of Requisite Complexity. Therefore it is crucial to focus on the design of the solution and how all of the partners and their actions work together. There are approaches, such as the one EDA described in Chapter 1, that help identify the structure of knowledge and action. Applying design principles is particularly applicable to entrepreneurial solutions (see Chapter 20) because the entrepreneur is essentially attempting to remedy structural mismatches between resource allocation and the problem. Paul Polak provides a good example of focusing on design of a product and the context. He re-envisioned the cause of poverty as "people are poor because they don't make enough money" *** check actual quote *** (ref). His solution for rural farmers was to design a treadle-style footpump that would be able to irrigate shallow wells and provide enough water so that the farmer could grow enough excess produce to easily pay back the cost of the pump. This entrepreneurial approach provided a structural solution to poverty that worked in the context of sub-surface water, local produce markets and availability of human power. There have even been museum exhibits on all the human-scale designs that are aimed at "the other 90%" of the population who need clean water and extra produce more than they need an iPod (Smithsonian Institution 2007). Most of these designs work with existing components or components that can be fabricated locally to create sustainable solutions to the environmental and economic problems facing the world's rural poor. The value of this product is in the combination and the usefulness of that particular product in that particular situation. The quality control over the product is embedded in the community of users. The characteristics of these design processes can be assessed using the same Mode 2 science framework, because the products are put together from existing components but in a different fashion in each case.

21.6 Challenges for evaluation of complex environmental projects

Communication, the flow of information and the connection of meaning are essential for evaluating environmental projects. However, it can be extremely challenging to get the public engaged in a dialog centered around the many interacting parts (complexity) of a project, especially when the project does not lend itself to a definable successful outcome (uncertainty). The public may be underprepared to hear or deal with this message. As Wolfgang Sachs laments (ref ****) about American, how can we talk about sustainability when people are so busy trying to drive their cars a little bit faster of the freeways? There is also the temptation to leave complex questions to the technocrats who, in matters of public resource allocation, present a significant challenge to our democracies. Another source of resistance is the view that traditional science has been so successful in creating progress that we should not want to replace it with Mode 2 science. The response to this is that we are trying to provide for both traditional and Mode 2 when appropriate, which is a subtle distinction at best. To many it may seem as if Mode 2 science is just a cover for our inability as an environmental community to reach any consensus on how to deal with these complex socio-econo-environmental issues. It's difficult to argue with people who believe that there is a single objective and that we can arrive at an optimal solution if we just study the problem enough. They see the discussion of multiple possible viewpoints as an erosion of the objective approaches that have made so much headway in the last centuries. All of these challenges converge to form a situation in which organizations that employ the situational and transient nature of Mode 2 evaluative methods are unlikely to be backed up by the stable institutions that are currently successfully practicing traditional evaluations of quality (Gibbons et al. 1994). Even though these are substantial challenges, the main point is that Mode 2 / transdisciplinary scientific approaches can be used to reliably assess quality and reliability, and that these evaluations will be extremely useful to all parties involved.

21.7 Summary

Ongoing evaluation is a critical element in any enterprise. A scientific evaluation must be systematic, rigorous, based on evidence and verifiable. Dealing with "wicked problems" requires a level of community stakeholder involvement, a commitment to allowing values to be incorporated into management from the beginning, and a respect for the inherent uncertainty of any specific outcomes. These characteristics undermine the utility of the traditional scientific modes for evaluating quality. Mode 2 science is an appropriate approach for transdisciplinary issues and is probably the best method for dealing with wicked problems. Mode 2 involves evaluation of the new forms of knowledge that have been gained by members of the community (not just scientists) and assesses how these people use their newly acquired capabilities to solve the current problem and how they disseminate and employ these capabilities in subsequent projects. One major strength of traditional science is that the quality of freestanding, timeless knowledge products is judged through stringent peer review. While these products are very valuable to science in general, Mode 2 science instead focuses on the ephemeral increase in capacity to solve the problems.

Probably the major challenge to adopting Mode 2 approaches is that people feel that the success of traditional science can be extended to cover these situations. However, in attempting to extend traditional science to meet these needs, they use a narrow definition of objectivity that reduces the importance of incorporating values and essentially casts the entire evaluation onto a single dimension. As with all of the intellectual tools presented in this text, more use of the approach leads to more skill and better outcomes. It may take more practice and experience to be able to effectively employ this approach on wicked problems with enough expertise to outperform the more tried-and-true traditional evaluative techniques.

