. . . restoration ecology now faces two major conceptual challenges. One is to decide what we really mean by our goals when we pretend that we are restoring natural, self-sustaining communities—which we rarely, if ever, are. The second is to decide, given limited amounts of time and money, what we most urgently need to know in order to achieve our goals.
—Jared Diamond, 1987

CHAPTER 8

MONITORING AND ADAPTIVE MANAGEMENT

Jeffrey L. Kershner

By now the reader should be convinced that restoring degraded watersheds is important. Millions of dollars have been spent to reclaim our aquatic and riparian resources and millions more probably will be. Intuitively, we know that restoration can improve these resources, but clearly restoration dollars must be spent wisely.

Monitoring is the measure of success of any restoration. Well-designed monitoring should (1) indicate whether the restoration measures were designed and implemented properly, (2) determine whether the restoration met the objectives, and (3) give us new insights into ecosystem structure and function. Monitoring should help us reexamine our understanding of aquatic and riparian ecosystems and provide information needed to adapt the goals for restoring those systems. Significantly, as much or more is learned about systems by monitoring and reporting failure as is learned by reporting success.

If monitoring is so important, why is so little effective monitoring undertaken in proportion to the number of restoration projects? Probably foremost is the lack of funding for monitoring, an institutional problem that persists (Noss and Cooperrider 1994). Notable exceptions exist (e.g., Toth et al. 1997, this volume), but many resource management agencies are reluctant to commit funding for monitoring, particularly for the long-term monitoring of restoration.

Part of the problem is that much restoration implemented today may not yield significant benefit for years or even decades. Watershed restoration practitioners need to understand the trajectory of recovery and how to adapt our management to changing environmental conditions. This requires a long-term approach to funding by agencies, foundations, industry, tribal groups, and others. Unfortunately, this long-term perspective is currently missing from much restoration monitoring.

A second, more pervasive reason that monitoring gets short shrift is that restoration practitioners are intimidated by monitoring. They are intimidated because the job looks so large: Did the treatments work? How long did they work? Did they make a difference? Were the restoration objectives chosen correctly? How much money and time should be spent on monitoring?

Although increasing the funding for monitoring may not be possible, we can develop monitoring that is well-designed, that reflects realistically the available personnel and funding, and that answers questions about restoration measures.

The purposes of this chapter are (1) to identify the key components of watershed restoration monitoring and (2) to aid the monitoring practitioner in designing credible monitoring for watershed restoration, given varying levels of funding and personnel.

TYPES OF MONITORING

The restoration practitioner must understand what types of monitoring fit a particular project (MacDonald et al. 1991). Three types are particularly useful for restoration: *implementation monitoring*, *effectiveness monitoring*, and *validation monitoring*.

Implementation monitoring asks: Was the restoration implemented properly? This monitoring should be part of every restoration project and is normally performed during or shortly after restoration. During the project, implementation monitoring continually evaluates the project design to determine its appropriateness in the field. Midcourse corrections are often necessary because field conditions may make the original design unworkable. Implementation monitoring during the project can identify these problems early and suggest workable solutions. This is particularly important where restoration contracts identify specific restoration measures, materials to be used, and designs to be followed. (Nothing is more frustrating than to follow contract specifications in the field, only to realize that the project is doomed because the design was inappropriate!)

Implementation monitoring sets the stage for other types of monitoring by demonstrating that the restoration treatments were done correctly and followed the design. The practitioner can then concentrate on identifying and correcting design problems if failures occur.

Effectiveness monitoring asks: Was restoration effective in attaining the desired future condition and in meeting restoration objectives? Effectiveness monitoring is more complex than implementation monitoring and requires understanding of the physical, biological, and sometimes the social factors that influence aquatic ecosystems. This understanding is translated into quantifiable objectives or benchmarks that describe the function of healthy aquatic systems. The primary purpose of effectiveness monitoring is to measure whether objectives are met by restoration.

Trend monitoring is a less rigorous form of effectiveness monitoring; it often involves visual estimates or photographs of changing resource conditions over time.

Validation monitoring is more specialized and primarily has a research focus. Validation monitoring verifies the basic assumptions behind effectiveness monitoring. For example, until 20 years ago, large woody debris in streams was removed to facilitate fish movement (Sedell et al. 1988). More recent research has shown that woody debris is important in structuring stream communities in many areas of the country and is an important link between the physical and biological functions of streams (Maser and Sedell 1994). Validation monitoring is a research tool with which to examine the basic scientific understanding of how aquatic systems work. Effectiveness and validation monitoring are necessary steps to evaluate adaptive management prescriptions.

A RESTORATION MONITORING PROCEDURE

When multiple restoration activities are underway, the additional task of monitoring can appear overwhelming, so practitioners may neglect it. A common lament is that ''it will take too much time or money,'' or ''I don't have the statistical background,'' or ''it's not a priority for my supervisor.'' Understanding the starting point and taking the first steps will alleviate this anxiety and get monitoring in motion. The following steps provide a template for sound restoration monitoring:

1. define participants;
2. establish clear goals and objectives;
3. design monitoring to detect change to (a) distinguish treatment effects from other variations, and (b) take replicate samples over space and time;
4. prioritize monitoring activities;
5. implement field prescriptions and techniques;
6. analyze data and report results; and
7. adapt goals and objectives to new information.

Further explanation of each step is provided next.

Define Participants

Watershed-scale restoration may involve myriad resource specialists, agency personnel, and nonagency partners. All should develop ownership in the monitoring. Interdisciplinary development of monitoring goals and objectives usually is best. For example, fisheries objectives may have hydrologic or geomorphic (landform) components that will determine the success of the restoration. Thus, hydrologists or fluvial geomorphologists (specialists in stream patterns and stream-related landforms) should actively participate in setting objectives, study design, analysis, and other appropriate phases of the project. Other specialists may be required to successfully complete restoration monitoring.

State and federal agencies, private landowners, tribal groups, and citizen groups often have ownership in watershed restoration projects and are interested in their success or failure. These groups can provide labor, financial support, and technical assistance. Their involvement can extend monitoring resources and allow expanded monitoring that otherwise would be impossible.

However, use of nontechnical personnel requires caution. Monitoring tasks that demand high technical expertise may require extensive training. It is best to question nontechnical participants to determine their expertise and desired level of participation before committing them to monitoring.

Establish Clear Goals and Objectives

Monitoring often fails because the overall mission, goals, and objectives are not clearly articulated. Establish the purpose of monitoring by developing clear goals and objectives (see Appendix Examples 1 and 2 at the end of the chapter). A successful monitoring plan has clear objectives to use as benchmarks for the analysis. These objectives define the project's purpose and determine the type and extent of restoration. Objectives should come from analyses of limiting factors for the species or community of interest and normally are determined during a full watershed analysis (Kershner, in press; Ziemer 1997, this volume).

Objectives are typically measurable and quantifiable and represent some desired future condition, within the constraints of resources. It is important to understand the objectives within the actual spatial and temporal scales that are operating on the subject landscape (Frissell et al. 1986; Wissmar 1997, this volume; Ziemer 1997). Thus, the spatial and time scales of interest need to be described specifically (Conquest et al. 1994). In doing so, the actual study design and sampling protocol can be more easily defined.

Objectives may represent standards and guidelines from broad planning or policy documents, or they may represent benchmark conditions from ''healthy'' aquatic systems. In all cases, these broader objectives must be modified for local conditions. Watershed-specific objectives should consider local disturbance processes and geomorphic conditions. For example, a watershed goal might be to reduce the total sediment input from roads. Thus, an effectiveness monitoring objective might be to quantify change in residual pool depths in low-gradient, unconfined channels that are sensitive to sediment inputs, or to quantify change in fine sediments in spawning areas.

In any case, one should select monitoring objectives that are the best indicators of change and measure them in the appropriate areas that are responsive to change. For example, it may be difficult to determine whether bank stabilization is reducing fine sediment if fine sediment inputs are measured in

TABLE 8.1.—Ten principles for conducting environmental field studies. (Green 1979.)

1. State concisely to someone else the research question. (The results will be as coherent and comprehensible as the initial conception of the problem.)
2. Take replicate (multiple) samples within each combination of time, location, and any other controlled variable. (Differences among sample groups can be demonstrated only by comparison to differences within the groups.)
3. Take an equal number of randomly allocated replicate samples for each combination of controlled variables. (Putting samples in ''representative'' or ''typical'' places is not random sampling.)
4. To test whether a condition has an effect, collect samples both where the condition is present and where the condition is absent, but all else is the same. (An effect can be demonstrated only by comparison with a control.)
5. Conduct preliminary sampling to provide a basis for evaluation of sampling design and statistical analysis options. (Those who skip this step because they do not have enough time usually end up losing time.)
6. Verify that the sampling device or method is sampling the population it is supposed to be sampling, and with equal and adequate efficiency over the entire range of sampling conditions to be encountered. (Variation in the efficiency of sampling from area to area will bias among area comparisons.)
7. If the area to be sampled has a large-scale environmental diversity, break the area into relatively homogenous subareas and allocate samples to each in proportion to the size of the subarea. If estimating abundance over the entire area, make the area allocation proportional to the number of organisms in the subarea (by adjusting the numbers of sample units).
8. Verify that the sample unit size is appropriate to the size, density, and spatial distribution of each organism being sampled. Then estimate how many replicate samples are required to obtain the desired precision.
9. Test the data to determine whether the error variation is homogenous, normally distributed, and independent of the mean. If it is not, as will be the case for most field data, then (a) appropriately transform the data, (b) use a distribution-free (nonparametric) procedure, (c) use an appropriate sequential sampling design, or (d) test against simulated H_o data.
10. Having chosen the best statistical method to test the hypothesis, stick with the result. (An unexpected or undesired result is not a valid reason for rejecting the method and hunting for a ''better'' one.)

high-gradient riffles. These areas typically transport sediment and may not be responsive to changing inputs of fine sediment.

Design Monitoring to Detect Change

Monitoring must detect change. While this in itself is a simple concept, it actually places an enormous burden on how monitoring is conducted and how the data are analyzed and interpreted. Often it is difficult to distinguish the effect or result of a particular activity among many interacting factors and great natural variability. It is useful to recall the ten principles described by Green (1979) for designing environmental field studies, presented in Table 8.1.

Implementation monitoring is relatively straightforward, and may involve simple yes–no answers. It may help to develop a checklist that identifies the project prescription and indicates whether each element was completed properly. This obvious step often is overlooked, but it is nearly impossible to conduct effectiveness monitoring unless it is clear that the project was implemented correctly.

If the project was implemented correctly, effectiveness monitoring is conducted to determine whether project objectives are being met. The following two general design principles are appropriate for effectiveness monitoring studies (Armour et al. 1983).

Distinguish treatment effects from other variations.—The study design must allow treatment effects (explained sources of variation) to be distinguished from all other sources of variation. This can be done by isolating ''treatments'' or restoration from other ''untreated'' sites. In this way, it can be determined whether restoration caused the change or whether similar changes would have occurred naturally without restoration.

It may be difficult to detect change from site-specific restoration in larger watersheds. An important factor is natural variability, which can make difficult the detection of change between restored and untreated areas. It is critical to minimize this variability by choosing areas of similar size, geology, morphology, stream discharge, and other characteristics. Ideally, they should vary only in the extent of restoration and should consist of one or more pairs of treated–untreated smaller subwatersheds within a larger watershed.

Realistically, practitioners may encounter other environmental factors that prevent the ''ideal'' situation. Although comparisons between partially and wholly treated smaller watersheds can be used, the contrast in restoration extent should be as great as possible, given the inherent difficulties in detecting change. Untreated areas can serve as the ''control'' or ''reference,'' while restored areas can serve as the ''treatment.''

It is useful to collect pretreatment data in areas being considered for restoration (House 1996). Pretreatment data are often useful as a benchmark where suitable control areas may be unavailable. If possible, information should be collected on physical and biological characteristics of sites before treatment so that change resulting from restoration can be measured and documented. Pretreatment data may not be available over a long period, but whatever can be acquired is useful.

Take replicate samples over space and time.—Replication refers to the number of sites or samples needed over space and time. How many are needed depends on natural variation in the variables measured and the precision and accuracy desired. Two things are important in determining the sample quantity: (1) level of significance and (2) the probability of detecting a difference when one exists (statistical power).

Selection of the level of significance generally depends on the values and risks associated with the variable being measured (MacDonald et al. 1991). The level of significance must be appropriate for the project. For example, many field studies set a high level of significance of 0.05 to detect changes caused by management. (This level indicates that only 5% of the observed difference is due to chance.) Although the reliability of the answer might be quite high, more samples are generally required to reach this higher level of statistical significance. Picking a lower level may still provide reasonable certainty but may require fewer samples and be more cost-effective over the long term.

Another consideration when choosing sample size is being able to detect a difference when it exists, which is called statistical power (Sokal and Rohlf 1981; Peterman 1989). Similar to determining the level of significance, an increase in sample size will generally increase the statistical power of the test. When identifying the level of significance, it is important to remember that decreasing the level of significance will generally increase the statistical power. This may have advantages where funding and personnel limitations reduce how many samples can be obtained.

If the practitioner is interested in monitoring the direct, localized effects of a spatially limited activity, then the appropriate study design might use paired sites (see Appendix Example 1). This may involve selecting pairs of monitoring stations upstream and downstream or perhaps side by side.

If one is monitoring the combined effectiveness of varied restoration activities within a watershed, then the paired control–treatment approach should be used for a number of objectives and integrated throughout the watershed (see Monitoring Example 2). In these cases, an expanded study design would include comparisons between whole tributary watersheds or even entire watersheds.

Because multiple activities must be monitored in a watershed, single-activity monitoring cannot distinguish all the sources of change. For example, a decreasing ratio of stream channel width to depth could be the combined result of three different events: (1) decreased hillslope sediment delivery from a landslide stabilization project, (2) culvert and road fill removals in headwater areas, and (3) vegetation enhancement projects at the site to reduce bank erosion and resulting sedimentation. A combination of implementation monitoring and effectiveness monitoring will be needed to assess overall restoration effectiveness in the watershed. The watershed where these activities are taking place should be compared against a watershed or smaller subwatershed where similar management exists, but without similar restoration.

Because multiple activities are involved, multiple measurements are necessary on a variety of variables over an extended period. Perhaps the ''best'' variables are those that integrate a variety of processes and that serve as appropriate biological and physical monitoring surrogates (see Monitoring Example 2). Biological indices such as the index of biotic integrity (Karr 1991, 1993; Angermeier 1997, this volume) or aquatic invertebrate metrics (Plafkin et al. 1989) may be particularly suited to monitoring when restoration emphasizes the recovery of aquatic communities.

The data and analysis results are greatly influenced by spatial and temporal factors and by observer error. For example, changes in flow and channel morphology (shape) along the stream affect many variables commonly used for monitoring. Our inability to consistently identify features can be a very large source of error. Standardizing and using quantitative methods helps to improve repeatability in monitoring.

Prioritize Monitoring Activities

A well-designed restoration plan should identify all of the key elements to be monitored, but should prioritize them according to importance and availability of resources. The plan should estimate the time, money, personnel, and equipment required to implement each plan element. This information is essential to the monitoring practitioner and to supervisory personnel when planning a work force and budget.

Rarely can practitioners do all of the restoration monitoring they would like. During times of decreasing personnel and budgets, some monitoring elements may be deferred or dropped. An effective strategy is to prioritize program elements so that at least some level of monitoring can be accomplished in any year. For example, implementation monitoring of a riparian (streambank) planting project may be done annually for the first few years of the program, and deferred to every 3 years if it appears that vegetation has been successfully reestablished. This may make money and personnel available for other aspects of the monitoring plan.

Another approach is to share personnel and funds among administrative units. Multiple agencies and groups may be involved in the monitoring of large watersheds. It is often useful to combine resources and develop monitoring teams where similar types of monitoring activities are being implemented. Teams can be dedicated to limited work assignments, and once training in field techniques and data collection has been accomplished, these teams are often more efficient at collecting information. Typically, this improves the reliability and precision of data and should save time and money. The only down side is that it may be difficult to establish priorities for these teams in a multiple group or agency setting.

Implement Field Prescriptions and Techniques

Monitoring analyses can be only as good as the data gathered. Because different personnel often perform the same monitoring in different years, it is important to establish consistent field protocols to collect field measurements in the same way, year after year. This step is missing in many monitoring programs. It may be useful to develop a set of field protocols for each phase of the monitoring plan. A simple checklist, like the one mentioned above for implementation monitoring, can help ensure that each critical element of a restoration is evaluated. Trend monitoring often requires that photos be taken in successive years to evaluate changing conditions. A narrative outlining the date, time of day, compass direction, and location on an aerial photo is useful to minimize the variation from differing field conditions.

Graphical and quantitative analyses require consistent, replicated measurements to reduce variation from field measurement error. Practitioners should write detailed narratives that establish sampling location, frequency, technique, and equipment to be used. They should try to avoid field variables that require subjective judgment. For example, habitat classifications that require considerable professional judgment but which lack consistent, repeatable field protocols should be avoided. In general, the more complex a field variable, the greater the chance for error.

It is advisable to establish a training program before the field season to train

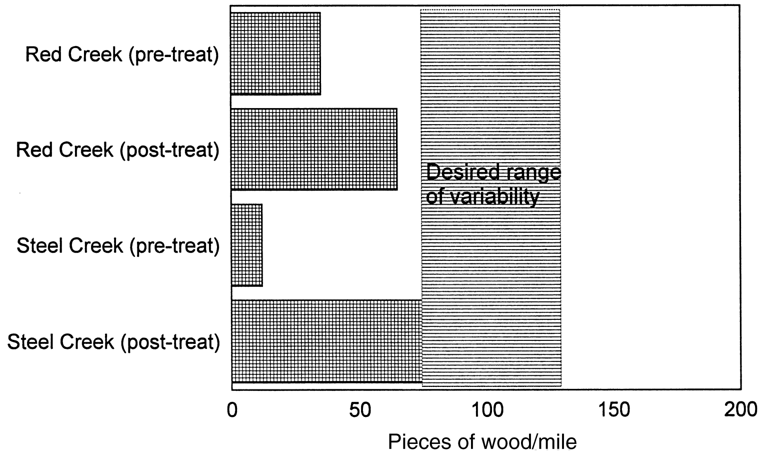

FIGURE 8.1.—An example of a graphical data display comparing frequency of large woody debris before and after restoration to a hypothesized desired range of variability.

field personnel in data collection protocols and techniques. Only when it is clear that monitoring personnel are collecting measurements consistently should they be sent into the field. Quality control checks should be made throughout the field season.

Analyze Data and Report Results

Several useful ways exist to display and analyze monitoring data. Most importantly, a variety of very powerful tools are available if the sample design considerations described earlier have been followed carefully. For some variables, a qualitative approach may be more appropriate, given the complexity or great variation associated with them. For example, it is difficult to accurately quantify bank erosion, but a photograph provides vivid and unambiguous documentation.

Comparative analyses can be simple graphical displays of how data in control and treatment areas differ in the same year and over time. These simple comparisons are often the most visually powerful evidence that change has occurred. It may also be appropriate to compare data from these controls and treatments to objectives from policy documentation. Photographs, bar graphs, and line graphs are particularly powerful ways to show analytical results to nontechnical people and decision makers (Figure 8.1). Graphical analyses should be the first step in any statistical report.

Statistical interpretation may provide further insight that may not be readily apparent from graphical analyses alone. If the study is designed using the principles outlined in Table 8.1, then other statistical options will be available for the analysis. Statisticians should be consulted often during the study design, pilot study, and analysis phases of the project to keep on track. Few statisticians are willing or able to help analyze data from a poorly designed study (Green 1979).

Restoration often fails because data are not analyzed and reported. Monitoring reports can build support for restoration by demonstrating positive environmental change. The results may lead to greater support for restoration and restoration monitoring from supervisory personnel and the public. It is as important to report restoration failure as it is to report success. Determine whether failure was caused by poor design, poor implementation, catastrophic conditions, or poor planning. Practitioners should share information with other restoration professionals, document what worked and what did not, and adapt restoration goals and objectives as new information becomes available.

Technical note on statistics.—Parametric statistics are used to measure data against a set of ''assumptions'' regarding their use. These assumptions have to do with the distribution of the data, whether the error variation is homogeneous, normally distributed, and independent of the mean. Random or stratified-random samples are a condition for using these analyses as well as nonparametric analyses described below. If data do not meet the assumptions outlined above, then it may be appropriate to statistically transform the data before performing further analyses. Statisticians should be consulted to help with the appropriate transformations of data. Parametric tests that can be used for monitoring data include *t*-tests, analyses of variance, and means tests. The questions asked during the study design phase will guide which of these tests are appropriate.

Nonparametric statistics are used when the assumptions described above are violated for some reason. Random or stratified random samples are still a condition for the use of these tests. Nonparametric tests are not a panacea for poor study design; they are to be used only when the assumptions for parametric tests are violated and cannot be corrected by transformation.

Adapt Goals and Objectives to New Information

Adaptive management is the process whereby management is initiated, evaluated, and refined (Holling 1978; Walters 1986). It differs from traditional management by recognizing and preparing for the uncertainty that underlies resource management decisions. Adaptive management is typically incremental in that it uses information from monitoring to continually evaluate and modify management practices. It promotes long-term objectives for ecosystem management and recognizes that people's ability to predict success is limited by knowledge of the system. Adaptive management uses information gained from past management experience to evaluate both success and failure and to explore new management options.

Monitoring provides the information needed to evaluate management. Monitoring may suggest new approaches or goals for watershed restoration and management. An adaptive management strategy for the restoration practitioner may be to experiment with different types of restoration throughout a watershed while continually monitoring the performance of the measures. By learning from both successes and failures, the restoration practitioner can see which techniques may be most useful and gain insight into which practices best promote recovery. It may be advantageous to conduct several smaller experimental restorations to gauge their effectiveness and system response as a prelude to large-scale projects.

CONCLUSIONS

Watershed restoration monitoring is an important component of aquatic ecosystem management. Resource stewards cannot determine the success or failure of watershed restoration without well-designed and properly implemented monitoring. The call for more efficient government and wise use of public funds will place restoration under increased scrutiny from the public and legislative bodies. The accountability that comes from proficient monitoring will be essential to continued restoration funding.

The challenge to restoration practitioners is to move beyond project implementation into careful analyses and reporting of restoration results. The lack of published restoration monitoring results indicates either that restoration generally is not followed by careful analysis or that restoration practitioners have not widely shared their analyses and findings, or both. Any of these cases is unacceptable. To move forward with well-conceived and well-supported watershed restoration, we must look backward to lessons of the past.

APPENDIX: MONITORING EXAMPLES

Example 1: Introducing large woody debris to improve habitat variety, restore deep pool habitat, and increase juvenile steelhead (rainbow trout) numbers.

Goal: Improve juvenile steelhead habitat to restore runs of summer steelhead.

Key limiting factor: Complex pool habitats created by large woody debris.

Objectives: (1) Increase structurally complex rearing habitat for juvenile steelhead as measured for deep pools and woody debris frequency in the current administrative policy. (2) Increase the numbers of juvenile steelhead to meet downstream migrant numbers defined as optimal in state management plan.

Question to be addressed: Was summer rearing habitat for juvenile steelhead restored?

Implementation monitoring: Number of deep, complex pools.

Measurements.—Numbers of pools per mile; frequency of woody debris after restoration.

Effectiveness monitoring: Was the restoration effective in creating quality summer rearing habitat that produces more juvenile steelhead?

Measurements.—Residual pool depth; cover index; numbers of juvenile steelhead.

Study design:

Control section(s).—No treatment upstream or downstream of treated section(s); no influence from treatment. Sampling unit is two riffle pool sequences, randomly selected in a low-gradient stream segment. Separated from treatments by long distance (300–1,000 yards).

Treatments.—Introduce large woody debris into a randomly selected section, which is two riffle pool sequences in a low-gradient stream segment.

Sampling.—Annual field monitoring of residual pool depth, percentage of complexity, and juvenile steelhead numbers in both the controls and treatments until recovery is established. Each treatment site should have a matched control and replicates of both controls and treatments. Field measurements should be quantifiable and repeatable. (Remember, structural enhancements may move or deteriorate over time.) Monitoring during periods of low flow during summer will allow for the detection of changes in the working efficiency of the structure and suggest when maintenance or further improvement is needed.

Analyses:

Graphical.—Graphical comparison of residual pool depths, complexity values, and juvenile steelhead in the control and treatment sections. Graphical comparison of treatment values and policy standards (if quantitative). Multiyear comparison graphs showing the recovery trajectory of effectiveness-monitoring variables.

Statistical.—Tests to determine whether data meet the assumptions for parametric statistical methods. Transform data to meet assumptions where possible. If transformations do not work, determine appropriateness of nonparametric methods. Parametric methods that may be appropriate include *t*-tests (two-sample paired sites), analysis of variance, and means tests. Consult statisticians to determine appropriate statistical methods.

Example 2: Watershed restoration in a small watershed that provides spawning and rearing habitat for an important anadromous fish stream.

Goal: Restore spawning and rearing habitat for summer steelhead in the subwatershed. Specific restoration goals include reducing fine sediment from roads, increasing complex pool habitat with woody debris, and providing for long-term woody debris inputs from riparian stands that consist currently of mixed hardwoods and conifers of subcommercial size.

Key limiting factors: complex pool habitats created by large woody debris; spawning substrates having minimal fine sediments.

Objectives: (1) Increase structurally complex rearing habitat for juvenile steelhead as measured for deep pools and woody debris frequency in policy standards. (2) Decrease the percentage of fines in spawning gravel to less than 10% during spawning and incubation. (3) Increase the numbers of juvenile steelhead to meet downstream migrant numbers defined as optimal in the state steelhead management plan.

Question to be addressed: Was rearing and spawning habitat for summer steelhead restored in the subwatershed?

Implementation monitoring: Wood additions; miles of roads closed; riparian stand conifer stocking.

Effectiveness monitoring: Was the restoration effective at improving rearing and spawning conditions for summer steelhead?

Measurements.—Numbers of complex pools per mile (residual pool depth, complexity rating); percentage of fines in spawning gravel; emergence survival; numbers of juvenile summer steelhead; long-term woody debris inputs.

Study design (three parts):

1. Wood Introduction

Control sections. No treatment upstream or downstream of treated sections; no influence from treatment. Sampling unit is two riffle pool sequences, randomly selected in a low-gradient stream segment. Separated from treatment by long distance (300–1,000 yards).

Treatments. Introduction of large woody debris into site in low-gradient stream segment (measurement area is two riffle pool sequences). Treatment may include different debris inputs to vary the complexity.

Sampling. Annual field monitoring of residual pool depth, percentage of complexity, and juvenile steelhead numbers in both the controls and treatments until recovery is established. For each treatment site there should be a matched control and replicates of both controls and treatments. Field measurements should be quantifiable and repeatable. (Remember, structural enhancements may move or deteriorate over time.) Annual monitoring during periods of low flow during summer will allow detection of changes in the working efficiency of the structure and suggest when maintenance or further improvement is needed. Monitoring frequency may be decreased once structures are stabilized.

2. Road Obliteration

Assumptions. Road obliteration typically involves obliterating the existing road grade, removing culverts, regrading to native surface, and replanting. The assumption is that primary sediment delivery occurs via small side drainages that were crossed by the former road and that these drainages delivered the majority of sediment to the spawning tributary.

Controls. Controls exist above the road obliteration and are above the side drainages that delivered sediment from the road. Controls should be placed in spawning riffles under geomorphic conditions in the upstream areas that are similar to those in sites below treatment areas.

Treatments. Treatments are sites below confirmed sources of sediment that were linked to the road. Sites can be placed below side drainages in spawning riffles throughout the treatment area or below the last side drainage that delivered sediment to the spawning stream. To separate treatment effects, sites below the treatment area should be placed in spawning areas that do not have confounding effects from other management sources. It is important to have replicate controls and treatments for each site.

Sampling. Establish sampling protocols to measure fine sediment in spawning riffles that will allow detection of change during the critical spawning–incubation period. Protocols should be measurable and repeatable. Uncalibrated visual estimates of fine sediment are normally unacceptable for detecting change. Preferred are methods that provide repeatable quantitative estimates of sediment size distribution (such as freeze-core sampling or McNeil samplers).

To understand sample variance at the site, a sample protocol needs to be established that takes multiple samples within the riffle in random locations. A small pilot study should help determine the variation between samples. A power analysis will help determine how many samples to take in the riffle to reasonably estimate the particle-size distribution (Green 1979; MacDonald et al. 1991). A lone sample at each control and treatment normally is insufficient to detect a change between control and treatment.

A second part of this sampling may require emergence traps (traps that capture young fish as they emerge from the gravel) to understand embryo survival in the changing gravel conditions. Typically, these are placed behind known nesting locations to estimate the emergent fry that are surviving from each nest. Similar sample-size guidelines to those described for sampling fine sediments may be used when estimating emergence survival.

The temporal component of this monitoring will vary depending on the size of the treatment and climatic conditions. Sediment-rich drainages that have high sediment loads may take decades to respond to treatment. In these cases, it is important to understand the trajectory of recovery and to note whether fine sediment in spawning gravel is decreasing as a result of treatment. Having control sites is particularly important in this case to compare recovery rates between treated and untreated areas. Changes between treatment and controls may be detectable before long-term, large changes become evident in treated sites alone. Monitoring annually for the first few years may be necessary to understand their rate of recovery. Then monitoring frequency can be decreased to every three years or longer, depending on the rate of change.

3. Riparian Silvicultural Restoration

Assumptions. Silvicultural restoration of riparian forest stands refers to planting trees along streambanks to enhance woody debris production, and it may take decades or centuries to fully realize the benefits. In this example, the measure of effectiveness is the amount of wood produced by the stand that will be available to the stream and riparian zone as coarse woody debris.

Controls. Control areas are larger-scale sections (small watersheds and larger) that have not been treated with tree plantings. These may include old-growth forest or other regenerated stands that have not been subject to timber harvest.

Treatments. Treatments are riparian stands subject to planting.

Sampling. Sampling can be a combination of remote sensing (aerial photography) and field methods. The ultimate measure of effectiveness is the contribution of woody debris to the stream channel. However, since time frames will be long, short-term surrogates for woody debris input may be necessary. These could include stand stocking rates (how many trees), canopy closure (how much shade), and stand growth rates. Aerial photography and videography could be used to establish stand densities and canopy closure over intervals of 10–30 years to document stand conditions in the treatment areas. Field measurements of stand stocking rates and canopy closure can be used to verify and supplement remote sensing methods.

Over the long term, woody debris in streams and along their banks can be resampled using basinwide protocols from inventory methods. If full-basin surveys are used, statistical comparisons are unnecessary because the full population (*N*) is being sampled. Consequently, the total difference between control and treatment areas will be known. If control and treatment sites are subsampled, then it will be necessary to randomize the sample sites and follow sample design recommendations from Table 8.1. Sampling frequency will be relatively infrequent, given the rate of temporal change in vegetative composition (10–30 years).

Analyses:

Graphical.—Graphical comparisons of variables in control and treatment sections. Graphical comparisons of treatment values and policy standards (if quantitative). Multiyear comparison graphs showing the recovery trajectory of effectiveness-monitoring variables.