

Improving the Evaluation of Conservation Programs

DEVRA G. KLEIMAN,* RICHARD P. READING,† BRIAN J. MILLER,†
TIM W. CLARK,‡ J. MICHAEL SCOTT,§ JOHN ROBINSON,** RICHARD L. WALLACE,‡§
ROBERT J. CABIN,†† AND FRED FELLEMAN‡‡

*Department of Zoological Research, National Zoological Park, Washington, D.C. 20008, U.S.A.,
email dkleiman@nzp.si.edu

†Department of Conservation Biology, Denver Zoological Foundation, 2900 East 23rd Ave., Denver, CO 80218, U.S.A.

‡Yale University School of Forestry and Environmental Studies, 205 Prospect St., New Haven, CT 06511, U.S.A.

§U.S. Geological Survey, Department of Fish and Wildlife, University of Idaho, Moscow, ID 83844, U.S.A.

**Wildlife Conservation Society, 2300 Southern Blvd., Bronx, NY 10460, U.S.A.

††USDA Forest Service, 23 East Kawili St., Hilo, HI 96720, U.S.A.

‡‡Conservation Consultants, Inc., 3004 NW 93rd St., Seattle, WA 98117, U.S.A.

Abstract: *The evaluation of conservation programs is rare but increasingly important in improving their effectiveness. Regular evaluations of conservation programs and the implementation of recommendations resulting from such assessments are infrequent because of resistance by participants and lack of funding. Evaluations may be internal or external, depending on the purpose of the review and how broadly it is focused. We strongly recommend external peer review of long-term complex conservation programs every 5 years, supported by more frequent (annual) internal reviews. Criteria for success must encompass both biological and social measures and include learning and the application of new knowledge to management. Evaluations must also go beyond monitoring to assess the value of the program. We emphasize the need to include the organization and function of a conservation program (the process) in any evaluation in addition to substantive criteria for success, which usually involve biological measures (numbers). A dysfunctional program organization and process can as effectively cripple a conservation effort as can a major biological catastrophe. We provide examples of different types of conservation program evaluations, including moderated workshops and case-study analysis, and provide advice on the logistics and organization of the review, emphasizing the importance of the evaluation process itself to a successful outcome. One important aspect of an evaluation is having an individual with leadership ability and considerable expertise to organize the format and oversee the review process itself. Second, it is essential at the outset to ensure agreement among the program participants and the review committee on the goals and objectives of the conservation program, what is to be evaluated, and the criteria for defining success. Finally, the best evaluations are inclusive and involve all participants and stakeholders.*

Mejoramiento de la Evaluación de Programas de Conservación

Resumen: *La evaluación de programas de conservación es poco común, pero cada vez más importante para mejorar la efectividad de dichos programas. Tanto las evaluaciones de los programas de conservación como la puesta en operación de las recomendaciones resultantes de dichas evaluaciones se realizan con poca frecuencia debido a la resistencia de los participantes y a la falta de apoyo financiero. Las evaluaciones pueden ser de carácter interno o externo, dependiendo del propósito de la evaluación y de la amplitud del enfoque. Recomendamos enfáticamente que colegas del mismo ámbito evalúen aquellos programas de conservación considerados complejos y de largo plazo cada cinco años, y que estas evaluaciones sean reforzadas por evaluaciones internas realizadas con mayor frecuencia (anualmente). Los criterios que se aplican para*

§ Current address: Eckerd College, 4200 54th Ave. South, St. Petersburg, FL 33701, U.S.A.
Paper submitted November 30, 1998; revised manuscript accepted May 26, 1999.

lograr el éxito deben comprender parámetros tanto de carácter biológico como social, e incluir, asimismo, el aprendizaje y la aplicación de este nuevo conocimiento a la administración de los recursos. Las evaluaciones también deben ir más allá de la inspección o revisión que se realiza para determinar el valor del programa. Resaltamos la necesidad de incluir en cualquier tipo de evaluación la organización y el funcionamiento (o proceso) de un programa de conservación, además de los criterios fundamentales considerados para lograr el éxito, la mayoría de los cuales incluyen parámetros (números) biológicos. Un programa estructurado con una organización y proceso débiles puede paralizar un esfuerzo de conservación con la misma facilidad con que lo haría una catástrofe biológica de dimensiones mayores. Proporcionamos ejemplos de diferentes tipos de evaluaciones de programas de conservación, incluyendo talleres y análisis de estudios de casos, y proporcionamos asesoramiento con relación al procedimiento y a la organización de la evaluación, poniendo énfasis en la importancia del proceso de evaluación en sí para poder obtener un resultado satisfactorio. Un aspecto importante es el contar con un individuo con habilidades de liderazgo y experiencia considerable para organizar el formato de la evaluación y supervisar el proceso de revisión. En segundo lugar, es esencial que los participantes y el comité revisor estén de acuerdo en cuanto a las metas y los objetivos del programa de conservación, así como los criterios para definir el éxito del programa. Finalmente, las mejores evaluaciones son inclusivas e involucran a todos los participantes e interesados.

Introduction

Program evaluations determine how well a program has performed and assign responsibility and accountability for success or failure (Clark 1996a). Although practitioners in almost any field agree that periodic review and evaluation is vital to improved program performance, conservation programs rarely receive comprehensive, in-depth, external, peer-reviewed evaluations. The paucity of such reviews (Backhouse et al. 1996) exists because they are costly and difficult (Clark 1996a), they may provide unwanted suggestions, and their recommendations may be difficult to implement. Even when external evaluation takes place, assessments of success can vary greatly, depending on the time scale used, the range of activities and issues examined, and the standards and criteria employed (Reading & Miller 1994).

We compared the purpose, structure, and function of internal and external program evaluations for a range of conservation programs. Although many of our specific examples are derived from endangered-species recovery programs, the issues are similar for programs focusing on conserving biodiversity, habitats, or ecological and evolutionary processes. We argue that the results of evaluations at a variety of levels and time frames can permit the refinement of parts of conservation programs, the alteration of whole programs, or even a change in the entire approach to conservation problems (i.e., paradigm shifts; Kuhn 1970). We provide examples of how evaluations can be structured and suggest improvements to the evaluation process for conservation programs. Finally, we discuss some of the issues that impede and deter conservationists from seeking programmatic evaluations on a regular basis and implementing the resulting recommendations.

Definition of Evaluation

To define what we mean by *evaluation*, we support Clark and Brunner's (1996:4) suggestion, with Lasswell (1971) as a guide, that evaluations should "Assess success and failure in terms of goal achievement and accountability for outcomes. Gather information on how past decision functions worked. Assess quality of performance. Disseminate findings and recommendations to appropriate people and publics."

Evaluation differs from simple program monitoring and assessment. Evaluations introduce values into our determination of what constitutes success based on the current social norms and principles of twentieth-century western culture. Evaluations require starting with a comprehensive definition of the problem that includes an understanding of the associated context (Dery 1985; Weiss 1989; Margoluis & Salafsky 1998). Good evaluations go beyond assessing whether goals were reached to assess the adequacy of the goals and the reasons for success or failure. Many programs have poorly defined goals—or none at all—because goal definitions are either controversial or implicit (Clark 1996a). Alternatively, goals may be developed within political and socioeconomic constraints but then may be inappropriately presented as biologically based (Scott et al. 1995).

Evaluations of conservation programs should include assessment of both the substance of a program (what the program accomplished) and the process used (how effectively and efficiently the program functioned). For example, a program may reach its scientific goals but may do so inefficiently or with negative secondary effects, such as loss of local support, interorganizational conflict, or negative effects on nontarget species (Miller et al. 1996; Clark 1997). Similarly, a program may operate smoothly but fail to reach its stated biological goals.

Evaluations of process should review how well a program operated during all phases. Social scientists (Lasswell 1971) have divided programs into phases: *initiation*, identifying the problem; *estimation*, defining the problem; *selection*, choosing a prescription or plan of action; *implementation*, acting on the plan; *evaluation*, appraising the program; and *termination*, ending or changing the program (Brewer & de Leon 1983; Clark 1996a; Clark & Brunner 1996). Most conservation problems should be viewed as a continuous feedback loop instead of linearly, with a clear start and finish, because they involve living systems and constant change.

Thus, strategic planning for conservation programs needs to include a stepwise process with multiple milestones and many opportunities for incorporating feedback (Margoluis & Salafsky 1998). Conservationists also need to realize that there may be more than one route to the final destination and many disparate hurdles to be overcome before success is achieved.

Internal and External Evaluations

Many program evaluations are internal, informal affairs (Backhouse et al. 1996), precipitated by the preparation of progress reports or new proposals for funding. These reviews monitor activities and performance in the short term but usually do not question the overall values or long-term focus of a program.

Regular, more formal internal evaluations, in which participants review all or some aspects of their program and change their activities as new knowledge and understanding are acquired, should be standard for most conservation programs. Internal evaluations should include specific questions concerning the adequacy of the program goals and the process. They should also monitor the effect of management practices so that the program may be adjusted continuously in the face of new information (i.e., adaptive management; Holling 1978).

External evaluations tend to be less frequent, to be broader in focus, and to involve peer review. They are also usually highly structured (Backhouse et al. 1996). External evaluations have the potential to benefit a program greatly because peer reviewers, if well chosen, are less constrained by (1) rewards or sanctions associated with program success or failure, which removes pressure to distort findings; (2) dominant paradigms of thought or even "group think," which may influence program participants (Janis 1972); (3) fear of negatively influencing personal or working relationships; (4) conflict of interest; or (5) peer pressure. In a discussion of the importance of peer review of research for implementation of the Government Performance and Results Act (GPRA) of 1993, Kostoff (1997) comments that "Peer review properly designed to support GPRA would provide credible indication to the research sponsors of intrinsic program

quality, program relevance, management quality, and appropriateness of direction, and has the potential to improve the quality of the research program as well."

The Society for Conservation Biology officially supports independent scientific review of decisions and policy related to the environment (Meffe et al. 1998).

There is great value in integrating internal and external evaluation (Backhouse et al. 1996) and in using structures for evaluation that will reveal the more complicated aspects of a program. Program participants often have in-depth understanding of a program, many important aspects of which are complex, difficult to express, poorly realized or understood, and/or highly sensitive. Much of this information is not included in written reports or even expressed during interviews and other less formal review processes. Other program participants may have a narrow view of the purposes of a conservation program. A well-structured review, whether internal or external, can extract and integrate these separate components.

The decision over whether to use an internal or external evaluation depends on several factors, including the purpose and focus of the evaluation, available resources, and frame of reference. An evaluation may cover a month, a year, or several years. It may appraise all aspects of the program, selected parts, or a few more isolated actions. More frequently conducted evaluations with a narrower focus should probably be internal and informal, including participation only from those individuals directly involved. Indeed, resource constraints often require this. Broader, less frequent, and more formal evaluations (conducted, for example, every 5 years) usually justify spending resources to bring in high-quality external reviewers.

Evaluation Scale and Focus

Both internal and external evaluations can focus on different levels within a program. For example, an evaluation may examine only one aspect of an endangered species conservation program (e.g., numbers of animals existing in the wild or outreach efforts), or it may look at the entire program (Yaffee 1994; Miller et al. 1996). Evaluations may focus on limited time scales (1–2 years) or over the life of a program (15–20 years). Evaluations may review the strategic plan (e.g., a recovery plan), the budget (Kleiman et al. 1991), or the performance of individual participants. Especially for the latter, many agencies have regular performance-appraisal systems in place for staff; as such, participants in conservation programs may receive separate internal evaluations relating to their activities within these programs.

Conservation programs involving several geographic or political areas (different states, different nations, different continents) may be especially difficult to review

because the more agencies and actors involved, the more complex the organization and process. Broader evaluations look across programs to assess particular approaches within one aspect of conservation (e.g., Gunderson et al. 1995), such as general approaches to endangered species recovery (Kohm 1990; Clark et al. 1994). Other approaches may focus on a single species or specific methodologies. For example, an evaluation could appraise the usefulness of a single tool used in several conservation efforts, such as captive breeding or reintroduction (Griffith et al. 1989; Snyder et al. 1996; Wolf et al. 1996) or it could pass judgment on entire recovery efforts (Tear et al. 1993, 1995; Schemske et al. 1994).

Criteria for Evaluation

Evaluations must develop criteria for defining success based on the goals defined by the conservation program. Ultimate goals usually specify when a population (or habitat) is "recovered" or how much loss of biodiversity, genetic heterogeneity, or population size is acceptable. Two examples are the use of (1) a species-level goal, such as the criteria of the World Conservation Union for both endangerment and recovery, modified from the Mace-Lande criteria (Mace & Lande 1991), and (2) a biodiversity-level goal, such as the United Nations' widely cited 12% rule for preserving ecosystems. But these standards are only occasionally adopted. More commonly, each conservation program develops its own criteria, often based on subjective, albeit scientifically informed, opinions. Organizations, especially bureaucracies, may be under heavy pressure to demonstrate success (e.g., by delisting endangered species). T. W. Clark (1996a:4) suggests that a "major challenge of appraisal is to set up a vehicle capable of examining the complex issues at hand while remaining loyal to common goals and minimizing the temptation to distort findings."

We have already mentioned that criteria for success must cover more than simply biological considerations (Clark & Wallace 1998). Because socioeconomic and political forces are at the root of most conservation problems, these important considerations must be addressed (Kellert 1985; Brussard 1991). Truly interdisciplinary approaches to conservation incorporating both biological and social science considerations have a better chance for success. Definitions of success should reflect this by explicitly including interdisciplinary criteria (Margoluis & Salafsky 1998).

Moreover, because a program's ultimate goals often cannot be reached for decades, measures of success must be defined in a stepwise fashion. Intermediate-level criteria permit managers and reviewers to evaluate progress toward ultimate goals and can illuminate the in-

terdependency of biological and social measures of success (Margoluis & Salafsky 1998).

Substance Versus Process Criteria

Substantive criteria for evaluating program success are more easily assessed than are process criteria. It is easier to measure whether a conservation program has accomplished its objective of achieving 500 individuals of a species or 20,000 ha of habitat than to estimate the degree of reduction in conflict between scientists, government officials, and developers. Substantive criteria are crucial if programs are to remain focused on specific goals, but they can be both biological and social.

Important biological criteria in endangered-species recovery programs include indices of population abundance, demography, and density; size of a species range; habitat quality and protection; degree of genetic diversity; and trends in all these indices over time. Broader biological goals, such as maintaining biological diversity, ecological processes, ecological communities, unique biological phenomena (e.g., migration), and evolutionary potential are also important but are more difficult to measure.

Some substantive biological measures of success may be inadequate because they do not consider appropriate temporal and geographic scales. For example, a species-reintroduction program that assesses success by looking at survivorship and dispersal requires biologically meaningful measures of these variables, such as number surviving to reproduction and dispersal that maintain contact with conspecifics. For golden lion tamarins (*Leontopithecus rosalia*), the changing and decreasing ratio of reintroduced tamarins to tamarins born in the wild from the reintroduction program might be interpreted negatively, but it reflects successful reproduction of offspring born to reintroduced tamarins or their offspring (B. Beck, personal communication).

Programs can occasionally appear to make clear progress but may do so at a cost to the ultimate objectives of the conservation effort. For example, the black-footed ferret (*Mustela nigripes*) recovery program is considered a success on the basis of numbers of ferrets currently in the wild. Nevertheless, the continual decrease and fragmentation of the prairie dog ecosystem upon which the ferret depends may ultimately lead to its extinction in the wild (Miller et al. 1996).

Substantive social criteria refer to indices such as public support; interorganizational relations; the support of conservation-program practitioners and involved agency personnel; relevant values, attitudes, and knowledge of key stakeholders; and trends in each of these variables (Reading & Kellert 1993). Conservation programs that address biological issues but fail to assess and address the attitudes of the local public may ultimately fail or have a negative effect on future conservation opportunities (Reading & Miller 1994; Margoluis & Salafsky

1998). Clark and Wallace (1998) describe the social conflict and possible long-term negative outcome resulting from the translocation of monk seal (*Monachus schauinslandi*) males in Hawaii without adequate public discussion.

In contrast, there was a more positive attitude toward wildlife, the forest, and golden lion tamarins within the local municipalities surrounding the tamarin's Poço das Antas Biological Reserve, as well as a significant increase in knowledge about nature and conservation after a 2-year community conservation education program (Dietz et al. 1994). The overall success of the conservation efforts for this species would have been reduced without the focus on community education.

Too many programs fail to educate the public or confront the sociopolitical roadblocks to conserving biodiversity in general. The black-footed ferret program may demonstrate a lost opportunity for using ferrets as a potential flagship species for conservation of the entire prairie dog ecosystem (Miller et al. 1996). Currently, one branch of the U.S. government finances recovery of the ferret while another pays for the destruction of its habitat. Comparable situations occur wherever actors have different objectives even when they have overtly accepted common goals. In the United States, inattention to socioeconomic issues and unresolved sociopolitical conflicts have resulted in state agencies, ranchers, developers, and resource extractors lobbying intensively for weakening environmental legislation such as the U.S. Endangered Species Act (Mann & Plummer 1995).

Process refers to the organization and function of conservation programs. All programs that achieve their goals are not equally sound in their methods of operation. Also, conditions are constantly changing and resources are too scarce for the overabundance of urgent conservation problems. Thus, there should also be concern over the efficiency and speed with which objectives are reached.

Understanding organizational structures is critical to evaluating the conservation process properly. Conservation programs can make substantial progress despite poor, inadequate, and ineffective organization. Problems with process are more obvious and more easily described by external peer review. Issues to focus on include how information is shared among involved stakeholders, the frequency and quality of communication among all parties, the management skills of the program leader, the decision-making procedure (consensus, majority vote, etc.), and the standards or norms for the function of the group. Kleiman and Mallinson (1998) discuss some of these issues for the lion tamarin recovery programs.

Different conservation programs often involve the same key actors. Dysfunctional programs can make progress, but if the organizational design is poor and conflict develops between actors, the strain between agencies or indi-

viduals can undercut future decisions involving other species and programs in the same jurisdiction.

In addition to competition, power differentials and conflicting mandates, policies, ideologies, priorities, and underlying values may inhibit successful coordination and cooperation in conservation programs involving multiple participants from different stakeholder groups (Salwasser et al. 1987; Ernst 1990). Evaluations should consider how well formal and informal coordinating mechanisms manage conflict and improve, or at least maintain, communication and relations between key participants (Harvey 1987; Salwasser et al. 1987).

One of the major challenges facing the conservation movement today is the negative effect of conflict among conservation scientists, government agencies, and non-governmental organizations (NGOs) all competing to "own" a conservation problem or solution. Marketing departments of conservation NGOs depend upon the uniqueness of their product to raise money, and productive collaborations often remain unacknowledged to the detriment of all. Similar pressure affects state and federal government offices and individuals. Professional conservationists, regardless of employer, are trained to be competitive. Given that teamwork is frequently not valued in western or developed cultures, program leaders may require special training in organizational development, team building, dispute resolution, and the leverage value of cooperative efforts.

A program's organizational structure should be well matched to the conservation challenge (Clark & Westrum 1989; Clark et al. 1989). Because many conservation problems are characterized by uncertain, urgent, complex, and rapidly changing task environments, organizational structures should usually be more flexible and decentralized, with little hierarchy, a strong task orientation, and fast and effective communication (Gordon 1983; Perrow 1986).

Conservation programs are often expensive; they should strive to decrease consumption of resources through increased efficiency and reduced time commitment. Paradoxically, the personal and professional career investment and long-term commitment to a conservation program by scientists and other stakeholders may militate against solving the conservation problem promptly. By changing the standards or the targets, legitimately or not, conservationists can prevent the timely termination of a conservation program. Conservationists also may not clearly recognize their accomplishments. Finally, a key measure of a conservation program's success in process is its ability to develop a political strategy to mobilize financial, technical, and other resources.

Learning

Indices of success in terms of a substance or process should include measures of the degree to which learning

has occurred (Lee 1993; Clark 1996b). Recovery programs may progress without rigorous hypothesis testing, experimental design, or adequate documentation. The number of animals may increase with little understanding of the mechanisms involved, so there may be no scientific basis on which to make adjustments when problems arise (Miller et al. 1996). A solid knowledge base provides opportunities to improve efforts and to apply techniques to other species. For example, results from the long-term monitoring of reintroduced golden lion tamarins enabled scientists to alter their methodologies for preparing tamarins for release (Beck et al. 1991; Castro et al. 1997). Managers, especially in bureaucratic cultures, might not risk applying new methods because the fear of failure is stronger than the potential rewards of increased efficiency and effectiveness (Miller et al. 1996).

At any stage in the development and implementation of an action plan for solving a conservation problem, scientists face uncertainty over the precision and accuracy of their data and scientific techniques. Conservation biologists must constantly redefine their criteria, goals, and measures of success because scientific standards and expectations will change over time. Science will advance our understanding of biological processes, technical innovations will permit deeper or fundamentally different analyses (e.g., interpreting genetic relatedness through DNA sequencing versus starch gel electrophoresis), new paradigms will be developed in the life sciences (e.g., small-population models), and new data will become available.

Within conservation science, as in all aspects of human life, there are different levels of learning (Clark 1996b; Lee 1993). Individual learning is important and is the first step in adjusting strategies toward better results. Each time individuals in a program share their new knowledge with team members and convince colleagues to act on the new information, it becomes easier for everyone to make the next cognitive leap. Participants can then move beyond the dominant paradigm and question the basic assumptions underlying a problem (Argyris 1992). We need to ask not only "how well are we doing?" but also "does it make sense to do it, even if it is being done well?" (Leeuw et al. 1994:9).

Clark (1996a) advocates the broader, interdisciplinary approach of focusing on case studies for evaluating endangered-species recovery programs. In contrast to the case-study approach, which provides a history of a program and thus usually includes at least a description of nonbiological issues, most assessments are "science-based" and largely ignore nonbiological aspects.

Clark (1996a) further suggests the use of prototypes as a method of developing "model" conservation programs that proceed by constant evaluation and learning. Similar to adaptive-management approaches (Holling 1978; Gunderson et al. 1995), prototyping represents a conscious effort to learn and improve performance

through relatively small-scale, exploratory approaches that are systematically and regularly evaluated (Clark et al. 1995b). Flexibility, innovation, and creativity are key to successful prototypes. Lessons obtained from such model programs often can be applied to other programs because, despite their differences, the problems facing conservation programs are often similar (Clark 1996a).

Evaluation Methods

A good program evaluation begins with an understanding of context and the problem the program addresses. Therefore, participants in an evaluation—reviewers and members of the conservation program—should agree on a comprehensive definition of the problem and the aims, objectives, and context of the program (Lasswell 1971; Dery 1985; Weiss 1989). The problem definition should include determining the magnitude and direction of past and recent trends, analyzing the conditions that caused those trends, and projecting future directions and developments (Lasswell 1971; Clark 1996a). Such analyses then facilitate developing, comparing, and selecting the strategies and general direction of the program (Lasswell 1971).

The broadest evaluations should consider conservation science, conflict management, organizational issues, professional behavior, sociological issues, and policy science (Clark et al. 1994). Evaluations should include a mechanism for setting future direction so that after the review participants can, through a strategic planning process, work for consensus on future goals that are clear and measurable, with associated timetables and designated responsibilities and accountability (Backhouse et al. 1996).

The structure of program evaluations varies greatly in both form, what is included in the review, and procedure, the way the review is conducted. Formal evaluations may be conducted by individuals (internal or external) reviewing the program independently or under contract with the originating agency; they may use moderated workshops or case-study analyses (e.g., Hutchins et al. 1996). Some examples of peer review evaluations are the National Research Council's review of the status of seven species of salmon in the Columbia River Basin (National Research Council 1996) and of the recovery program for the Hawaiian crow (*Corvus hawaiiensis*; National Research Council 1992) and the National Audubon Society and American Ornithologists Union's review of the recovery program for the California condor (*Gymnogyps californianus*; Ricklefs 1978). Recently, the American Institute of Biological Sciences (AIBS) oversaw an independent review of the habitat conservation plans developed under the U.S. Endangered Species Act with the aid of students in university programs in conservation biology (Pyke et al. 1998).

During such appraisals, program participants should be encouraged to think about how their activities contribute to the overall program goal(s). In many instances this means not only assessing how an activity contributes to a subgoal or objective but also how those objectives themselves contribute to the overall conservation goal. The sum total of these narrower reviews can then be evaluated within the context of the overall program. Backhouse et al. (1996) refer to this as "net assessment," and it permits evaluation of the relative importance and contribution of each activity to resolving the problem.

Program evaluation can also proceed as a case study that tracks and reviews all or parts of a program. Individuals occasionally may conduct independent evaluations of conservation programs outside officially sanctioned channels (e.g., Snyder & Snyder 1989; McFarlane 1992; Walters 1992; Alvarez 1993; Wallace 1994; Yaffee 1994; Miller et al. 1996; Clark 1997). There are few examples of evaluation across programs, in which several case studies are examined together (Kleiman & Mallinson 1998), or the usefulness of a particular technique, such as reintroduction, is appraised (Griffith et al. 1989; Wolf et al. 1996). Such broad evaluations look for common themes, both with respect to problems facing conservation programs and methods for addressing problems (Kohm 1990; Clark et al. 1994; Gunderson et al. 1995). Gunderson et al. (1995) refer to these as "barriers" and "bridges," respectively. Several types of evaluation are usually appropriate, depending on the time scale, frequency, and goal of the evaluation.

The Logistics of Evaluation

The first step to conducting a good external evaluation is the choice of a leader of the review process. He or she should have excellent leadership qualities, should have participated before in peer review activities, and should be known for fair and impartial assessments. Kostoff (1997) recommends that reviewers include not only individuals competent in all relevant subject areas or disciplines, but also in the program's highest-level objectives to ensure that the right questions or issues are being pursued. Peer reviewers should not be direct competitors and should have a wide range of backgrounds and experience.

Participants then need to agree on what is to be evaluated and to develop a schedule for the review. All involved need to determine the best format for gathering information about different elements of the program and the kinds of materials requiring preparation and distribution. Information can be obtained through compiling annual reports, distributing written surveys and questionnaires, and conducting workshops, focus groups, and individual interviews. Interviews with outside individuals and agencies, other stakeholders, and peers may need to be organized.

The results of program evaluations should be properly documented and, where possible, subjected to peer review and submitted for publication (Backhouse et al. 1996). This permits improved evaluations across programs, continued learning, and, ultimately, improvements to future conservation initiatives. The participants also need to agree on what type of report to prepare, the report's audience, and how it will be published or distributed. A schedule for the completion of the process needs to be developed and adhered to, because both peer reviewers and program participants must achieve closure on the evaluation process.

The third step is the development of standards or criteria. These criteria should be comprehensive, interdisciplinary, and based on the goals of the program. To help avoid later conflicts, all participants in the review—program participants and peer reviewers—should agree to the criteria prior to initiating the evaluation.

It is crucial to develop and maintain a good working environment for all evaluations. Appraisals are best if they take place in a supportive, constructive environment that permits an open exchange of ideas and information (Backhouse et al. 1996). Information should be available to all interested parties, and participants should be encouraged to speak freely, without fear of reprisal. Participants therefore should be properly buffered from outside influence (Gordon 1983), in some cases by ensuring anonymity.

The best external evaluations are inclusive, involving all participants and all available information (Backhouse et al. 1996). Conflicts among participants may arise, but conflict is negative only if left unmanaged (Gordon 1983). Broad participation not only ensures that the evaluation will be comprehensive and representative, but, as Backhouse et al. (1996:174) state, "Participating in reviews may be the only way for some individuals working on part of a program to fully understand and appreciate the entire program."

It is important for external reviewers to understand that in evaluating a program they are doing more than reviewing a grant application or manuscript. For some participants, the program represents *the* major investment of their career. A heavy-handed focus on negative aspects of a program can discourage the major players and have the opposite effect intended by evaluators. Although academics are trained to be critical, the level and type of criticism needs to be attuned to the time and energy expended by participants, the constraints within which they have worked, and the time frame involved. Also, reviewers should bear in mind that conservationists rarely receive rewards for their efforts. Thus, the best reviews are constructive and balance the positive aspects of a program with those needing improvement or change. Finally, the evaluation process and final reports need to consider national, regional, and cultural differences. There will be resistance to accepting criti-

cism or implementing recommendations if they are not presented in a culturally sensitive manner.

Resistance to Evaluations

Individuals in an organization may learn and think creatively, whereas their parent organization remains mired in standard operating procedures. Thus, an organization's culture, values, and norms may impede its acceptance of evaluation as appropriate. Yet for conservation programs to advance, organizations must also learn and change. Improving the process of learning should occur at individual, organizational, and (ideally) successively higher levels of society.

Resistance to the concept of evaluation, whether individual or organizational, often arises from fear of negative input and exposure as well as fear of change. Such resistance can result in a lack of sufficient resources being budgeted to carry out an evaluation or to implement the recommendations therein, or in ignoring the evaluation report itself.

A lack of resources often constrains the frequency and type of evaluation possible, whether from obstruction or thoughtlessness. Kostoff (1997:652) suggests that peer review costs can be a "non-negligible fraction of total program costs, depending on review frequency, the level of technical detail desired, and whether the programs are labor or hardware intensive." Few conservation programs now include regular review costs as a percentage of the total budget, but several conservation organizations are developing formal guidelines and mechanisms for the internal and external evaluation of programs that they support and fund.

Conclusions

Evaluation is critical to improving the effectiveness and efficiency of conservation programs. Without proper evaluation, learning and change are difficult. Periodic evaluations followed by program alteration can result in adaptive conservation management programs that continually improve prospects for success. To be comprehensive, evaluations should address not only whether the primary biological goals were met, but also (1) how well science was employed; (2) how efficiently resources were used; (3) the degree to which public support was garnered for the program (and for conservation in general); (4) how well the program was organized and functioned to address the conservation challenge; (5) the degree to which the program was characterized by innovative problem solving and individual and organizational learning; and (6) to what extent economic, biological, and social considerations were distinguished when goals and objectives were established (Reading &

Miller 1994; Scott et al. 1995). Programs that improve in function as a result of periodic reviews may then be used as prototypes to guide future conservation efforts and their evaluation (Clark et al. 1995a; Clark 1996a).

No matter what specific procedure is used, evaluations should be institutionalized in every conservation program. Backhouse et al. (1996) discuss the need and means for integrating mechanisms of internal and external appraisals into conservation programs. Less formal internal evaluations should occur regularly and frequently among all groups involved in the program as a part of routine operations. These evaluations should take place at all levels, from the individual to the working group to the programmatic level. More formal external reviews should occur less frequently (probably at 5-year intervals). Such reviews, however, should be formally mandated by governments, NGOs, and funding agencies.

Acknowledgments

This paper was an outgrowth of a special session on evaluation held at the 1997 annual meeting of the Society for Conservation Biology in Victoria, British Columbia. We thank the people who participated in the session, either as speakers or during the moderated discussion. Many of the major ideas discussed in this paper were a direct outgrowth of those discussions. D. Johns provided helpful additions to the manuscript. The anonymous reviewers and editorial office suggested many useful organizational and stylistic changes.

Literature Cited

- Alvarez, K. 1993. Twilight of the panther: biology, bureaucracy, and failure in an endangered species program. Myakka River Publishing, Sarasota, Florida.
- Argyris, C. 1992. On organizational learning. Blackwell Science, Cambridge, Massachusetts.
- Backhouse, G. N., T. W. Clark, and R. L. Wallace. 1996. Reviewing recovery programs for endangered species: some considerations and recommendations. Pages 170-179 in S. Stephens and S. Maxwell, editors. Back from the brink: refining the threatened species recovery process. Surrey Beatty & Sons, Chipping Norton, New South Wales, Australia.
- Beck, B. B., D. G. Kleiman, J. M. Dietz, I. Castro, C. Carvalho, A. Martins, and B. Rettberg-Beck. 1991. Losses and reproduction in reintroduced golden lion tamarins, *Leontopithecus rosalia*. Dodo, Journal of Jersey Wildlife Preservation Trust 27:50-61.
- Brewer, G. D., and P. de Leon. 1983. The foundations of policy analysis. Dorsey Press, Homewood, Illinois.
- Brussard, P. F. 1991. The role of ecology in biological conservation. Ecological Applications 11:6-12.
- Castro, M. I., B. B. Beck, D. G. Kleiman, C. Ruiz-Miranda, and A. L. Rosenberger. 1997. Environmental enrichment in a reintroduction program for golden lion tamarins (*Leontopithecus rosalia*). Pages 113-128 in D. J. Shepherdson, J. D. Mellen, and M. Hutchins, editors. Second nature: environmental enrichment for captive animals. Smithsonian Institution Press, Washington, D.C.
- Clark, T. W. 1996a. Appraising threatened species recovery efforts:

- practical recommendations. Pages 1-22 in S. Stephens and S. Maxwell, editors. *Back from the brink: refining the threatened species recovery process*. Surrey Beatty & Sons, Chipping Norton, New South Wales, Australia.
- Clark, T. W. 1996b. Learning as a strategy for improving endangered species conservation. *Endangered Species Update* 13(1&2):5-6, 22-23.
- Clark, T. W. 1997. *Averting extinction: reconstructing endangered species recovery*. Yale University Press, New Haven, Connecticut.
- Clark, T. W., and R. D. Brunner. 1996. Making partnerships work in endangered species conservation: an introduction to the decision process. *Endangered Species Update* 13(9):1-5.
- Clark, T. W., and R. L. Wallace. 1998. Understanding the human factor in endangered species recovery: an introduction to human social process. *Endangered Species Update* 15(1):2-9.
- Clark, T. W., and R. Westrum. 1989. High-performance teams in wildlife conservation: a species reintroduction and recovery example. *Environmental Management* 13:663-670.
- Clark, T. W., R. Crete, and J. Cada. 1989. Designing and managing successful endangered species recovery programs. *Environmental Management* 13:159-170.
- Clark, T. W., R. P. Reading, and A. L. Clarke, editors. 1994. *Endangered species recovery: finding the lessons, improving the process*. Island Press, Washington, D.C.
- Clark, T. W., G. N. Backhouse, and R. P. Reading. 1995a. Prototyping in endangered species recovery programmes: the eastern barred bandicoot experience. Pages 50-62 in A. Bennett, G. N. Backhouse, and T. W. Clark, editors. *People and nature conservation: perspectives on private land use and endangered species recovery*. Transactions of the Royal Zoological Society of New South Wales. Surrey Beatty & Sons, Chipping Norton, New South Wales, Australia.
- Clark, T. W., R. P. Reading, and G. N. Backhouse. 1995b. Prototyping for successful conservation: the eastern barred bandicoot program. *Endangered Species Update* 12(10/11):5-7, 10.
- Dery, D. 1985. *Problem definition in policy analysis*. University of Kansas Press, Lawrence.
- Dietz, J. M., L. A. Dietz, and E. Y. Nagagata. 1994. The effective use of flagship species for conservation of biodiversity: the example of lion tamarins in Brazil. Pages 32-49 in P. J. S. Olney, G. M. Mace, and A. T. C. Feistner, editors. *Creative conservation: interactive management of wild and captive animals*. Chapman and Hall, London.
- Ernst, J. P. 1990. Federalism and the Act. Pages 98-113 in K. A. Kohm, editor. *Balancing on the brink of extinction: the Endangered Species Act and lessons for the future*. Island Press, Washington, D.C.
- Gordon, J. R. 1983. *A diagnostic approach to organizational behavior*. Allyn and Bacon, Boston.
- Griffith, B., J. M. Scott, J. W. Carpenter, and C. Reed. 1989. Translocation as a species conservation tool: status and strategy. *Science* 345:447-480.
- Gunderson, L. H., C. S. Holling, and S. S. Light, editors. 1995. *Barriers and bridges to the renewal of ecosystems and institutions*. Columbia University Press, New York.
- Harvey, A. H. 1987. *Interagency conflict and coordination in wildlife management: a case study*. M.S. thesis. University of Michigan, Ann Arbor.
- Holling, C. S. 1978. *Adaptive environmental assessment and management*. Wiley, London.
- Hutchins, M., R. J. Wiese, and J. Bowdoin. 1996. *Black-footed ferret recovery program analysis and action plan*. American Zoo and Aquarium Association, Bethesda, Maryland.
- Janis, I. L. 1972. *Victims of groupthink: a psychological study of foreign-policy decisions and fiascoes*. Houghton Mifflin, Boston.
- Kellert, S. R. 1985. Social and perceptual factors in endangered species management. *Journal of Wildlife Management* 49:528-536.
- Kleiman, D. G., and J. J. C. Mallinson. 1998. Recovery and management committees for lion tamarins: partnerships in conservation planning and implementation. *Conservation Biology* 12:27-38.
- Kleiman, D. G., B. B. Beck, J. M. Dietz, and L. A. Dietz. 1991. Costs of a reintroduction and criteria for success: accounting and accountability in the golden lion tamarin conservation program. Pages 125-142 in J. H. W. Gipps, editor. *Beyond captive breeding: re-introducing endangered species to the wild*. Oxford University Press, Oxford, United Kingdom.
- Kohm, K. A. 1990. *Balancing on the brink of extinction: the Endangered Species Act and lessons for the future*. Island Press, Washington, D.C.
- Kostoff, R. N. 1997. Peer review: the appropriate GPRA metric for research. *Science* 277:651-652.
- Kuhn, T. S. 1970. *The structure of scientific revolutions*. 2nd edition. University of Chicago Press, Chicago.
- Lasswell, H. D. 1971. *A pre-view of the policy sciences*. Elsevier, New York.
- Lee, K. N. 1993. *Compass and gyroscope: integrating science and politics for the environment*. Island Press, Washington, D.C.
- Leeuw, F. L., R. C. Rist, and R. C. Sonnichsen, editors. 1994. *Can governments learn? Comparative perspectives on evaluation and organizational learning*. Transaction Publishers, New Brunswick, New Jersey.
- Mace, G. M., and R. Lande. 1991. Assessing extinction threats: toward a reevaluation of IUCN threatened species categories. *Conservation Biology* 5:148-157.
- Mann, C. C., and M. L. Plummer. 1995. *Noah's choice: the future of endangered species*. Knopf, New York.
- Margoluis, R., and N. Salafsky. 1998. *Measures of success: designing, managing and monitoring conservation and development projects*. Island Press, New York.
- McFarlane, R. W. 1992. *A stillness in the pines*. Norton, New York.
- Meffe, G. K., P. D. Boersma, D. D. Murphy, B. R. Noon, H. R. Pulliam, M. E. Soulé, and D. M. Walker. 1998. Independent scientific review in natural resource management. *Conservation Biology* 12:268-270.
- Miller, B., R. P. Reading, and S. Forrest. 1996. *Prairie night: black-footed ferrets and the recovery of endangered species*. Smithsonian Institution Press, Washington, D.C.
- National Research Council. 1992. *Scientific bases for the preservation of the Hawaiian Crow*. National Academy Press, Washington, D.C.
- National Research Council. 1996. *Upstream: salmon and society in the Pacific Northwest*. National Academy Press, Washington, D.C.
- Perrow, C. 1986. *Complex organizations: a critical essay*. 3rd edition. McGraw-Hill, New York.
- Pyke, C. R., B. Bierwagen, and N. Goldstein. 1998. Strange bedfellows: observations on the current relationship between recovery plans and habitat conservation planning. *Endangered Species Update* 15(6):97-104.
- Reading, R. P., and S. R. Kellert. 1993. Attitudes toward a proposed black-footed ferret (*Mustela nigripes*) reintroduction. *Conservation Biology* 7:569-580.
- Reading, R. P., and B. J. Miller. 1994. The black-footed ferret recovery program: unmasking professional and organizational weaknesses. Pages 73-100 in T. W. Clark, R. P. Reading, and A. L. Clarke, editors. *Endangered species recovery: finding the lessons, improving the process*. Island Press, Washington, D.C.
- Ricklefs, R. E., editor. 1978. *Report of the advisory committee on the California condor*. Audubon conservation report 6. National Audubon Society and American Ornithologists Union.
- Salwasser, H., C. Schonewald-Cox, and R. Baker. 1987. The role of interagency cooperation in managing for viable populations. Pages 159-173 in M. E. Soulé, editor. *Viable populations for conservation*. Cambridge University Press, New York.
- Schemske, D. W., B. C. Husband, M. H. Ruckelshaus, C. Goodwillie, T. M. Parker, and J. G. Bishop. 1994. Evaluating approaches to the conservation of rare and endangered plants. *Ecology* 75:584-606.
- Scott, J. M., T. H. Tear, and L. S. Mills. 1995. Socioeconomics and the recovery of endangered species: biological assessment in a political world. *Conservation Biology* 9:214-216.
- Snyder, N. F. R., and H. A. Snyder. 1989. *Biology and conservation of the California Condor*. *Current Ornithology* 6:175-267.

- Snyder, N. R. F., S. R. Derrickson, S. R. Beissinger, J. W. Wiley, T. B. Smith, W. D. Toone, and B. Miller. 1996. Limitations of captive breeding in endangered species recovery. *Conservation Biology* 10:338-348.
- Tear, T. H., J. M. Scott, P. H. Hayward, and B. Griffith. 1993. Status and prospects for success of the Endangered Species Act: a look at recovery plans. *Science* 262:976-977.
- Tear, T. H., J. M. Scott, P. H. Hayward, and B. Griffith. 1995. Recovery plans and the Endangered Species Act: are criticisms supported by data? *Conservation Biology* 9:182-195.
- Wallace, R. L. 1994. The Florida manatee: organizational learning and a model for improving recovery programs. Pages 131-155 in T. W. Clark, R. P. Reading, and A. Clarke, editors. *Endangered species recovery: finding the lessons, improving the process*. Island Press, Washington, D.C.
- Walters, M. J. 1992. *A shadow and a song*. Chelsea Green Publishing, Post Mills, Vermont.
- Weiss, J. A. 1989. The powers of problem definition: the case of government paperwork. *Policy Sciences* 22:97-121.
- Wolf, C. M., B. Griffith, C. Reed, and S. A. Temple. 1996. Avian and mammalian translocations: update and reanalysis of 1987 survey data. *Conservation Biology* 10:1142-1154.
- Yaffee, S. L. 1994. *The wisdom of the Spotted Owl: policy lessons for a new century*. Island Press, Washington, D.C.

