Biodivers Conserv (2017) 26:2205–2222
DOI 10.1007/s10531-017-1353-5

CrossMark

ORIGINAL PAPER

Assessing the internal consistency of management plans for the recovery of threatened species

Alejandro Ortega-Argueta[1,2] · Greg Baxter[2] · Marc Hockings[3] · Roger Guevara[4]

Received: 2 August 2016 / Revised: 28 March 2017 / Accepted: 18 April 2017 /
Published online: 21 April 2017

Abstract Recovery planning is an important global conservation strategy for threatened species. Despite the existence of international standards for recovery planning, deficiencies and anomalies have been detected in several jurisdictions. This study evaluated the quality of recovery plans based on internal consistency as a measurement of coherent planning. We analyzed 236 plans developed by the Australian Government (1992–2006) using three criteria: (a) consistency of gaps in scientific information with prescribed research actions, (b) consistency of identified threats with prescribed threat abatement actions and (c) consistency of established plan objectives with performance evaluation criteria. These criteria were aggregated in order to calculate an index of plan consistency. We tested two

Communicated by Dirk Sven Schmeller.

Electronic supplementary material The online version of this article (doi:10.1007/s10531-017-1353-5) contains supplementary material, which is available to authorized users.

✉ Alejandro Ortega-Argueta
aortega@ecosur.mx

Greg Baxter
gbaxter@uqg.uq.edu.au

Marc Hockings
m.hockings@uq.edu.au

Roger Guevara
roger.guevara@inecol.mx

[1] El Colegio de la Frontera Sur, Carretera a Reforma km 15.5, Ranchería Guineo, 2ª Sección, 86280 Villahermosa, Tabasco, Mexico

[2] School of Agriculture and Food Sciences, University of Queensland, Gatton Campus, Brisbane, QLD 4343, Australia

[3] School of Earth & Environmental Sciences, University of Queensland, St. Lucia Campus, Brisbane, QLD 4072, Australia

[4] Instituto de Ecología, A. C., Red de Biología Evolutiva, Carretera Antigua a Coatepec 351, El Haya, 91070 Xalapa, Veracruz, Mexico

hypotheses: (1) plans made for single-species would exhibit better consistency than those for multi-species; and (2) plans made under the amended legislation of the Environment Protection and Biodiversity Conservation Act (EPBCA) would exhibit better consistency than those under the rescinded Endangered Species Protection Act (ESPA). In total, over 85% of the plans consistently addressed the research needs. However, the plans addressed threats poorly (66% of all plans exhibited inconsistencies). Moreover, nearly 50% of all plans established inconsistent performance evaluation criteria. Under the ESPA, single- and multi-species plans exhibited equal consistency, but under the EPBCA, single-species plans clearly exhibited higher consistency. Our major contribution is the assessment of attributes of consistency that are paramount for effective recovery planning. Evaluation of these attributes may provide knowledge of universal utility and relevance to other biodiversity conservation efforts.

Keywords Biodiversity conservation · Endangered species · Recovery planning · Policy evaluation · Assessment

Introduction

The formulation and implementation of recovery plans is an important conservation strategy for threatened species. This management instrument is applied worldwide as part of international efforts such as the species conservation programs of the International Union for Conservation of Nature (IUCN). At national level, recovery plans are official instruments of management used in countries such as the United States of America (USA), Canada, Mexico, New Zealand and Australia, among others, under the principles of the Convention on Biological Diversity. These documents comprise management and decision-making guidelines directed towards the restoration and conservation of species at risk, through definition of priority interventions and allocation of human and financial resources. Australia is one of the countries where recovery plans are prepared for species included in a national list of threatened wildlife. The aim of such recovery plans is to maximize the opportunity for the long-term survival in the wild of threatened species or ecological communities (Australian Government 2016a). Recovery planning can be focused on single species, multi-species, ecological communities or landscapes. Production of recovery plans for threatened species is a mandate of Australian legislation under the abrogated Endangered Species Protection Act 1992 (ESPA) and the current Environmental Protection and Biodiversity Conservation Act 1999 (EPBCA). While complementary legislation and policy mechanisms, such as the protected areas systems and the control of the international trade of wildlife, contribute to the conservation of threatened biodiversity, recovery plans remain the most direct conservation scheme carried out by federal and state governments and allied conservation organizations.

Successful natural resource management depends on the application of appropriate and coherent planning, leading to quality plans with a high likelihood of meeting their objectives (Joseph et al. 2008). Scientists and managers normally take the lead in the preparation of recovery plans. However, in practice, the design and development of such plans remain a challenge for governments and conservation organizations. Despite the existence of international standards for planning species conservation (IUCN/SSC 2008) and several national official guidelines (e.g., US: NMFS-USFWS 2010; Australia:

Australian Government 2016b), research has identified deficiencies and anomalies in overall recovery planning. Major issues that undermine the quality of recovery plans are related to poor incorporation of relevant scientific knowledge, poor information with which to identify critical habitat, poor specification of objectives, unclear connection between recovery objectives and performance evaluation criteria, inconsistent definition of actions to address threats, lack of monitoring and evaluation protocols and poor compliance with legislative requirements, among others (Clark et al. 2002; Lawler et al. 2002; Rahn et al. 2006; Ortega-Argueta et al. 2011; Roberts and Hamann 2016).

Effective planning is of paramount importance for conservation action and technical protocols are therefore crucial to the assessment of the quality of recovery plans, based on clear and measurable criteria, prior to adoption and implementation (Ortega-Argueta and Contreras-Hernández 2013). Such protocols may provide the basis to enhance plan design, improving the decision-making process, refining management strategies and allowing an improved allocation of resources (Wallace 2003; Ortega-Argueta et al. 2011). In contrast, deficient planning could have negative repercussions at the implementation phase that may serve to divert attention from priority targets, wasting time, effort and resources (Brigham et al. 2002).

Research into the planning process and the practical application of recovery remains relatively sparse. To date, few efforts have been dedicated to assessing the planning elements of recovery programs at international level (Boersma et al. 2001; Clark et al. 2002; Roberts and Hamann 2016), with much less efforts to formulate general planning principles (Miller et al. 1994; Burbidge 1996). Studies have identified differences in the quality of recovery plans in relation to the taxonomic characteristics of target species (Campbell et al. 2002; Hoekstra et al. 2002; Schultz and Gerber 2002), the periodic revisions of plans (Gerber and Hatch 2002; Lundquist et al. 2002) and whether the plan has either single- or multi-species focus (Boersma et al. 2001; Clark and Harvey 2002). Most research covers recovery planning in the USA; the research has only covered other jurisdictions and key design elements to a limited extent. Learning from other jurisdictions could offer contrasting insights into planning strengths and deficiencies that could be applied to the improvement of existing international standards.

The objectives of our study were: (1) to evaluate the quality of Australian recovery plans, based on three design and planning measures that indicate internal consistency; (2) to examine the likely influence of two planning scopes (single- vs. multi-species plans) and two periods of Australian legislation (ESPA vs. EPBCA) on the consistency of recovery plans. We tested two a priori hypotheses in order to contrast the plans between two pair-wise comparisons. Hypothesis 1 was that recovery plans made for single-species would exhibit higher internal consistency than those developed for multi-species and ecological communities. According to some authors (Clark and Harvey 2002; Hornaday and Bloom 2006), single-species plans are more focused on prescribing management actions, due to the greater availability of technical knowledge and better understanding of biological and ecological aspects at species compared to multi-species level. This is an important issue since there are calls to abandon the single-species scope for conservation planning, with the argument that the ecosystem approach is more effective (Jewell 2000). On the other hand, previous research in Australia (Ortega-Argueta et al. 2011) found that recovery plans exhibited better coverage of key planning attributes following amendments to the EPBCA legislation in 2000 and the introduction of new official planning guidelines that had not existed prior to that year. This change was assessed by the compliance of recovery plans with a checklist of mandatory planning elements. We wished to assess whether the introduction of the new planning guidelines supported better recovery planning according to national standards. Hypothesis 2, therefore, was that these new guidelines and amended legislation would improve the internal consistency of the more recent recovery plans made under the EPBCA (July 2000–January 2006) compared to those that existed under the previous ESPA (1991–July 2000).

Methods

Analysis of recovery plans and assessment of quality criteria

We examined all 236 recovery plans adopted by the Australian government from the first plans in 1992 up to January 2006. We selected this period in order to contrast recovery plans from under the two different periods of legislation. The entire set of recovery plans was readily available from the website of the Australian federal agency (Australian Government 2016a). The complete list of evaluated plans is present in the Online Resource 1.

The assessment framework consisted of three attributes of internal consistency (predefined as quality criteria), a scoring scale of discriminatory categories for each attribute and definitions of the scoring categories. The three criteria of internal consistency were: (1) consistency of identified gaps in the scientific information with prescribed actions calling for research, (2) consistency of identified threats with prescribed threat abatement actions, and (3) consistency of established recovery plan objectives with the performance evaluation criteria. Such criteria are based on the rationale and logical frameworks of planning (Brigham et al. 2002) and the assumptions of desired interventions for the recovery of threatened species (IUCN-SSC 2008). Below, we describe the construct and assumptions of each criterion and how the assessment was conducted (see Online Resource 2 for a detailed explanation of the assessment criteria and scoring system):

Criterion 1) Consistency of identified gaps in the scientific information with prescribed actions calling for research

Recovery plans contain a section that reviews the technical knowledge of target species. According to the official guidelines, plan authors must identify knowledge gaps and prescribe related research that may support improved species management. In each plan, we reviewed the coverage of the scientific information background, and then grouped this into six topics: distribution/habitat, population biology, life history, genetics, behavior/dispersal and general ecology. For the analysis, we considered only the prescription of plans that explicitly mentioned the requirement for further scientific information on these six topics. We subsequently verified whether the recovery plan presented a consistent response by prescribing research actions directed specifically to those topics.

Criterion 2) Consistency of identified threats with prescribed threat abatement actions

We first conducted a review of all of the threats identified in each plan. Since threat nomenclature varied according to the particular plan authorship, we used a classification guide based on the IUCN Threats Classification Scheme (TCS) (IUCN 2012) to standardize this aspect of the assessment. The TCS contains 60 threat categories; these are shown in the Online Resource 3. We searched for a consistent response by quantifying the abatement actions prescribed against each of the cited threats. Only management actions formulated specifically to address each cited threat were considered as a consistent response.

Criterion 3) Consistency of established recovery plan objectives with the performance evaluation criteria

The official guidelines require establishment of measurable objectives and a consistent relationship with performance criteria and monitoring metrics by which progress and achievement of objectives can be measured (Australian Government 2016b). We assessed the expected coherent relationship for at least one recovery objective associated with one performance evaluation criterion and one metric for monitoring in each plan.

Each plan was thus assessed against the three criteria and rated using a predetermined scoring system to systematically and consistently characterize the plans. This method, which assists the evaluation of qualitative attributes and transforms them into quantitative parameters, is known as performance assessment and was designed primarily in education as a method for rating student performance and work quality (Morris et al. 1987; Mertens and Wilson 2012). This approach has been previously applied to the evaluation of recovery planning in the USA (e.g., Boersma et al. 2001; Hoekstra et al. 2002). The quality of recovery planning has been assessed through several attributes, such as the explicit use of science (Boersma et al. 2001), addressing identified threats within plans (Lawler et al. 2002; Schultz and Gerber 2002), the presence of monitoring protocols (Campbell et al. 2002), the adequacy of species status assessments (Boersma et al. 2001) and compliance with mandatory regulations (Ortega-Argueta et al. 2011). In our case, we used internal consistency as a measure of coherent and quality planning (see Brigham et al. 2002) since our three established criteria are mandatory under the Australian legislation (Australian Government 2016b). Our selected criteria can also be universally applied regardless of national context. In this way, the insights obtained from this study can be applied to any context or jurisdiction.

All assessments and associated scores (nominal variables) were stored in an Access® database, using pre-designed computer recording forms as an instrument to facilitate data capture and the assessment of consistency between the different sections of the plans. The three quality criteria were rated in the forms according to three possible responses, as: 1 = 'inconsistent', 2 = 'of limited consistency' and 3 = 'consistent' (see Online Resource 2). Besides the quality criteria scores, queries on the computer forms also solicited complementary information about the plans such as the publication date, in order to discriminate between the two legislation periods, and the planning scope (single- vs. multi-species). We first conducted a trial on a sample of 24 plans (10% of the entire set) to test the validity of the discriminatory criteria for the assessment, suitability of the computer forms and reliability of the data (Morris et al. 1987). Where necessary, adjustments to the database were made. In order to minimize the subjectivity of judgment and inconsistency between plans of differing formats, only the information contained within these recovery plans was used to conduct the assessment. Furthermore, to assure uniformity and reliability of data, the first author undertook all of plan assessments following the pre-defined system of classification, judgment criteria and scoring. For this reason, the issue of inter-evaluator consistency in assessments does not arise. However, we do recognize certain limitations of our study. Internal consistency evaluation requires careful review through several planning elements, at different sections of the plans. We recognize that the inconsistencies identified in plans may be due to many circumstances, from truly conceptual deficiencies in planning to a deliberate decision on the part of plan authors to not prescribe management actions, despite the fact that the evaluated consistency criteria are mandatory requirements of the EPBCA legislation and guidelines. We limited our study to examine the overall internal consistency of recovery plans; the investigation of the motivations of plan authors to respond or not respond accordingly is beyond the scope of our study.

Statistical analysis

We first conducted a disaggregated analysis of scores for each criterion in order to observe its individual patterns, and then an aggregated analysis (that combined the product of three criterion scores) to assess an overall quality index (QI) of recovery plans. We used a contingency table to analyze the frequency of cases with different scores across the three quality criteria reviewed, as well as the overall QI (Table 1). As a second level of analysis, we tested hypotheses 1 and 2 described above. We used polynomial ordered logistic regression (POLR) models to test whether the proportion of cases with different scores was significantly related to planning scope (two levels: single- (n = 200) vs. multi-species (n = 36) plans), legislation period (two levels: ESPA (n = 74) vs. EPBCA (n = 162)) and the interaction between these factors. Statistical analyses were conducted in R (R Development Core Team 2012). Specifically, the POLR models were conducted under the MASS library (Venables and Ripley 2002).

Results

How do recovery plans respond to identified gaps in the scientific information?

The plans obtained the best scores in this criterion; over 85% of the analyzed recovery plans performed well in terms of identifying gaps in the scientific information and formulating consistent research actions (Fig. 1a). We found that the consistency of response varied across the topics of scientific information. Population biology, distribution/habitat

Table 1 Statistical summary of polynomial ordered regression models

		Scientific information gap assessment		Threat response assessment		Performance criteria assessment		Overall QI***	
	Df	X^2	P	X^2	P	X^2	P	X^2	P
*L	1	<0.1	0.915	<0.1	0.973	3.1	0.079	1.4	0.230
**S	1	7.3	0.007	1.4	0.243	5.1	0.025	20.0	<0.001
L × S	1	<0.1	0.949	0.3	0.583	4.4	0.037	9.2	0.002

Legislation: *Endangered Species Protection Act 1992* (ESPA) n = 74, *Environmental Protection and Biodiversity Conservation Act 1999* (EPBCA) n = 162; Scope: Single-species n = 200, Multi-species n = 36

*L legislation, with two levels (ESPA vs. EPBCA periods)

**S scope with two levels (single- vs. multi-species plans), and L × S the interaction between Legislation and Scope (N = 236 Australian recovery plans)

***QI Quality index: product of the three criteria scores

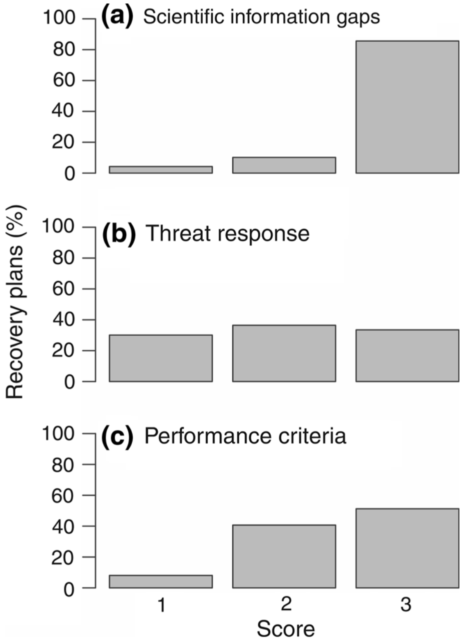

Fig. 1 Frequency distribution of scores (1 = 'inconsistent', 2 = 'of limited consistency' and 3 = 'consistent') across three criteria: **a** of identified gaps in the scientific information with prescribed actions calling for research; **b** identified threats with prescribed threat abatement actions, and **c** established plan objectives with performance evaluation criteria (N = 236 Australian recovery plans)

and general ecology were consistently addressed by over 80% of recovery plans, while life history, genetics and behavior/dispersal were addressed to a lesser extent. Most attention was paid to addressing population biology issues, mainly in the form of research prescription and monitoring activities on demographic parameters and population dynamics.

How do recovery plans respond to the identified major threats?

In contrast to that described above, in this criterion, the plans obtained the lowest scores. Recovery plans performed poorly in identifying and addressing threats, with over 66% of the plans showing 'inconsistent' or 'of limited consistency' scores (Fig. 1b). Plan authors described threats as 'potential' in 73% of plans. Eighteen percent of plans prescribed no threat abatement actions at all. Identification of multiple threats was common to all plans, with an average of six different threats for each target species within a plan. Prescribed threat-related actions ranged from one to eleven across all of the plans.

How consistent are the established plan objectives with the performance evaluation criteria?

In this criterion, the plans produced mixed results; just over 51% of plans had 'consistent' performance criteria for measuring progress in recovery, while around 41% had 'of limited consistency' in defining their performance criteria (Fig. 1c). Limited definition of evaluation performance criteria corresponded primarily to a lack of quantitative measures and thresholds of recovery. In this category, some evaluation criteria were consistent with the objectives, although the established measures lacked a clear definition for performance monitoring (see appendix), impeding the effective assessment of plan progress and species recovery. The remaining 8.1% of plans presented 'inconsistent' performance evaluation criteria.

Quality of recovery plans across planning scopes and legislation

Polynomial ordered linear regression models showed that scope (single- vs. multi-species plans) significantly affected the consistency scores in 'addressing gaps in the scientific information' and 'performance evaluation criteria' (Table 1)—single-species recovery plans performed better than multi-species plans (Fig. 2), thus supporting our hypothesis 1. A higher proportion of multi-species plans were marked as 'inconsistent' or 'of limited consistency' compared to single-species plans ($P < 0.05$) and the frequency of single-species plans marked as 'consistent' was higher than that of the multi-species plans ($P < 0.05$). For the criterion 'addressing identified major threats', no significant difference was detected.

Testing hypothesis 2 showed that only the consistency of the 'performance evaluation criteria' was affected by the interaction between scope and legislation (Table 1), with single-species plans performing better than the multi-species plans, but only under the new EPBCA legislation (Fig. 3). Under the rescinded ESPA legislation, there was no effect of scope, but under the EPBCA a higher proportion of multi-species plans were graded as 'inconsistent' compared to single species plans ($P < 0.05$) and the frequency of single-species plans marked as 'consistent' was higher than that of the multi-species plans ($P < 0.05$). For the criteria 'addressing gaps in the scientific information' and 'addressing identified threats', no significant differences were detected. Overall in the disaggregated analysis, there was no clear trend across the three criteria of quality that could suggest that recovery plans performed better under the EPBCA legislation than the ESPA legislation (Table 1, Fig. 3).

The aggregated analysis of the overall QI score of recovery plans proved hypotheses 1 and 2 (Fig. 4); the QI scores were significantly affected by scope and the interaction between legislation and scope. Under the ESPA legislation, single-species and multi-species plans performed equally, but under the EPBCA legislation a higher proportion of multi-species plans achieved low QI scores compared to the single-species plans ($P < 0.05$), while a higher proportion of single-species plans achieved the highest QI scores compared to multi-species plans ($P < 0.05$).

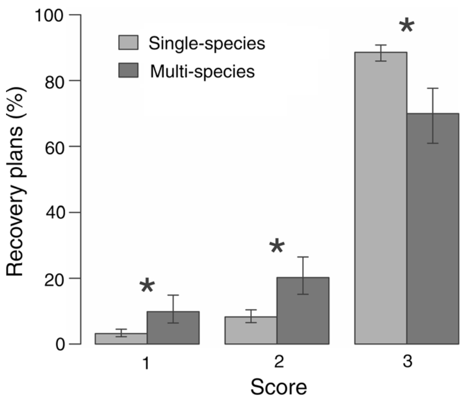

Fig. 2 Frequency distribution of scores (1 = 'inconsistent', 2 = 'of limited consistency' and 3 = 'consistent') of the criterion 'consistency of identified gaps in the scientific information with prescribed actions calling for research', across two planning scopes (single- vs. multi-species) (N = 236 Australian recovery plans)

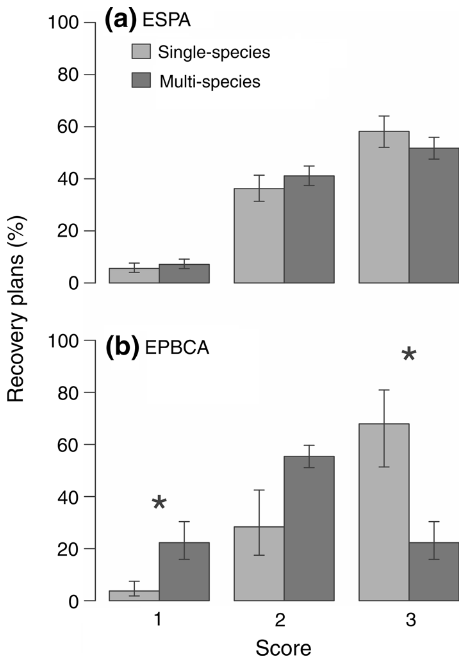

Fig. 3 Frequency distribution of scores (1 = 'inconsistent', 2 = 'of limited consistency' and 3 = 'consistent') of the criterion 'consistency of established plan objectives with performance evaluation criteria', across two planning scopes (single- vs. multi-species) and two periods of Australian legislation (*Endangered Species Protection Act 1992* (ESPA) vs. *Environmental Protection and Biodiversity Conservation Act 1999* (EPBCA) (N = 236 Australian recovery plans)

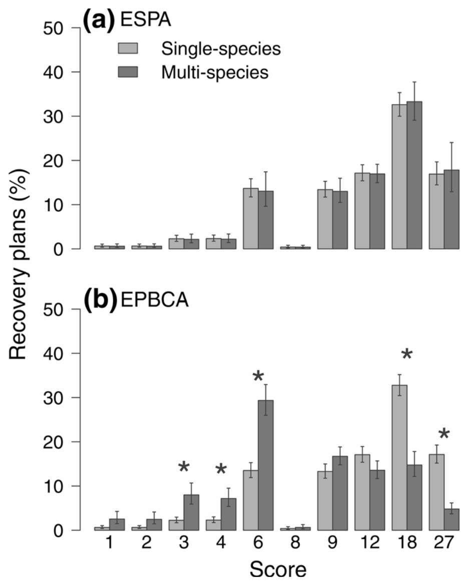

Fig. 4 Overall quality index (product of the three criteria scores: (1) consistency of identified gaps in the scientific information with prescribed actions calling for research, (2) consistency of identified threats with prescribed threat abatement actions and (3) consistency of established plan objectives with performance evaluation criteria) assessed across two recovery planning scopes (single- vs. multi-species) and two periods of Australian legislation (*Endangered Species Protection Act 1992* (ESPA) vs. *Environmental Protection and Biodiversity Conservation Act 1999* (EPBCA) (N = 236 Australian recovery plans)

Discussion

Preparation of recovery plans may appear straightforward for scientists and managers; however, as previous research has shown, there are recurrent design limitations that undermine the quality of planning. Our assessment of the three quality criteria produced mixed results. We detected a good performance of plans in addressing gaps identified in the scientific information but we also observed poor consistency in terms of addressing identified threats and establishing performance evaluation criteria. Below, we discuss the limitations and management implications of this.

Consistency of identified gaps in the scientific information with prescribed actions calling for research

Scientific information is crucial for effective management. Such information is required at various stages of the planning process, including the preparation and adoption of plans, setting of objectives and performance criteria, identification of threatening processes and causes of species decline, identification of critical habitats, formulation of recovery actions, and establishment of monitoring and evaluation (M&E) activities (Carroll et al. 1996; Dickman 1996; Pulliam and Babbitt 1997; Smallwood et al. 1999; Boersma et al. 2001). Nevertheless, given to the scarcity of scientific knowledge regarding most of the threatened species, management decisions are often based on incomplete information and are thus made with a high uncertainty (Tear et al. 1995; Schwartz 2008). Recovery plans fit into this type of situation because they have to provide management guidance on the basis of the best and most up-to-date scientific information available. Thus, identification of gaps in the knowledge is required by the official planning guidelines of several countries in order to indicate where appropriate research and monitoring activities should be directed in order to gather further information (Campbell et al. 2002; Pullin and Knight 2003; Ortega-Argueta et al. 2011). Identification of knowledge gaps is also important for improved recovery planning, since the flow of information between scientists and managers must be reciprocal and continuous. Besides the day-to-day requirement for information for decision-making, new knowledge of target species and their management responses may also be documented for incorporation into an updated future version of the plan. Improved communication mechanisms are needed in order to streamline this process; bringing scientists and practitioners closer together and increasing the mutual exchange of information (Pullin and Knight 2003).

We observed that some topics of scientific information were better addressed than others, with an overall high level of consistency in research prescriptions (>80% of plans). This degree of consistency observed across the range of Australian plans was similar to that observed in plans from the USA (Brigham et al. 2002). The topics most covered by research actions were population biology, distribution and habitat, general ecology and life history. Differences in consistency across topics may reflect not only the priorities of managers and scientists, but also their background disciplines. The more consistently addressed topics, with mainly ecological parameters, may have had a greater relevance for management and for assessing progress towards recovery, since most of the established recovery criteria are based on those measures. Population parameters are often used as proxy indicators of recovery and are established in most monitoring activities (Campbell et al. 2002; Gerber and Hatch 2002). This may be the reason for the strong emphasis given to the population biology topics. On the other hand, previous research identified very poor consideration of social aspects of management in the monitoring protocols, indicators and recovery criteria (Campbell et al. 2002; Ortega-Argueta et al. 2011). This issue has been identified not only in planning for threatened species, but also in other conservation initiatives (Kleiman et al. 2000; Hadlock and Beckwith 2002). One explanation may be the likely bias attributed to disciplinary barriers, given that the background of most professionals involved in the management of threatened species is in the ecological sciences.

Consistency of identified threats with prescribed threat abatement actions

Overall, the response of recovery plans in addressing threats was poor or moderate, with some differences detected among threat types. A similar degree of consistency was observed in recovery plans from the USA (Brigham et al. 2002). This limitation may be related to the poor knowledge of threatening processes across most recovery plans that may affect the prescription of effective management measures (Ortega-Argueta et al. 2011). The knowledge gaps regarding the threats must be addressed through prescribing research to generate understanding of the ultimate and proximal causes of species endangerment. Species facing site-specific and focal-source threats, such as management-related fires, have a better chance of success through the adoption of improved management practices. In contrast, species facing complex and broad-scale threats related to social and economic development, such as mining or coastal development, require greater efforts and novel strategies, which involve serious challenges for management (Lawler et al. 2002). In a previous study, it was found that the species that had improved their conservation status had threats that were more site-specific and easier to address (e.g., a direct cause of mortality), while species facing complex and multi-source threats such as those derived from development activities tended to be in decline (Abbitt and Scott 2001). The inconsistency observed in Australian plans to prescribe management actions against some threats could be generated by: (a) poorer knowledge of complex and broad-scale threats related to socioeconomic development; (b) the 'low-success' management scenario anticipated by wildlife managers who may prefer to invest their efforts in other issues; and (c) political barriers that may be involved in development projects with governmental participation that may also discourage the intervention of wildlife agencies. Given the limitations of the knowledge, in some plans, authors could only identify threats 'for which there was no direct evidence but are highly likely to be causing or contributing to species decline' (Australian Government 2016b) as potential. Perhaps for this reason, threat abatement actions were not prescribed at all in 18% of the plans. This poor understanding of threats is reflected in a weak management response (Lawler et al. 2002). Another critical point is that all species targeted by recovery plans face multiple threats; we found an average of six threats cited in each plan. Some groups of threats may have cumulative impacts across several bioregions, making their management more difficult and expensive (Sattler and Creighton 2002). Analysis of overlapping patterns in similar species, ecologically similar areas, or species facing similar threatening processes, could assist managers to streamline threat management (Foin et al. 1998; Burgman et al. 2006). As an example of a broader perspective, the Australian environmental agency has prepared Threat Abatement Plans (TAP) to fight against the most pervasive natural and anthropogenic processes threatening biodiversity. There are 14 TAP in place for addressing introduced predators (e.g., feral cats, red foxes) and competitors (e.g., rabbits, feral goats, feral pigs), among others (Australian Government 2016c). This broader approach to the management of pervasive and common threats to many threatened species could be aligned and cross-referenced with all of the recovery plans to which those threats are relevant. Thus, technical aspects such as knowledge limitation, complexity, severity and range of threats, as well as management capacity and politics, are factors that may influence the degree of threat response in recovery planning.

Consistency of established plan objectives with performance evaluation criteria

The poor or limited consistency found in setting performance criteria represents another design and planning deficiency. The link between objectives, performance criteria and design of monitoring activities was unclear in nearly half of the plans examined. Other studies have already highlighted this problem in Australia and the USA (Brigham et al. 2002; Campbell et al. 2002; Cork et al. 2006). This weakness may reduce the ability of managers to make decisions involving changes of management direction or priorities, or the reinforcement of successful interventions. Without a clear link between objectives, performance criteria and monitoring, it is impossible to assess progress and the success or failure of recovery efforts (Gerber and Hatch 2002). Poor definition of evaluation criteria may also reduce our ability to protect species or ecological communities against de-listing decisions that may be ecologically unsound (Doremus and Pagel 2001; Brigham et al. 2002).

Quantitative and qualitative criteria are both necessary in a plan. Quantitative criteria linked to population parameters are relevant when accurate information on species trends is available (e.g., 'at least 20 self-maintaining populations should exist with an increasing rate of ten percent within five years, and the minimum size of at least ten populations should exceed 300 adults'). However, given the limited information on the population parameters of most threatened species (Schwartz 2008; Bottrill et al. 2011), definition of criteria and monitoring measures is often poor (e.g., 'the population has increased or decreased'). The current effectiveness of most conservation efforts for threatened species is assessed against population and conservation status and trends. To date, no country has produced complete data sets of population parameters and trends that truly reflect years of monitoring of their threatened species. This is true even of the USA, which has had more than 40 years of recovery programs (Boersma et al. 2001; Gibbs and Currie 2012), and Australia with more than 20 years (Bottrill et al. 2011). In a strict sense, it will be impossible to obtain such a quantity of data for thousands of species; there is simply no governmental or financial capacity to do so. Assessing the effectiveness of recovery efforts based only on population measures is thus unfeasible (Gibbs and Currie 2012). With such an approach, governments and conservation organizations are limited in terms of information for adaptive management and policy accountability. Ortega-Argueta and Contreras-Hernandez (2013) proposed a framework for assessing recovery through multi-criteria indicators, both qualitative and quantitative, for performance monitoring and evaluation, which covers eight aspects of management: (1) research and monitoring, (2) species management, (3) habitat management, (4) threat management, (5) community participation, (6) education and awareness, (7) agency management and (8) policy and legislation. Besides the ecological aspects, it is necessary to examine other elements of the management system for adaptive management and assessment of progress and recovery.

Which plans exhibit better internal consistency?

With regard to our hypothesis 1, which stated that recovery plans developed for single-species would exhibit better consistency than those for multi-species, our results were concordant. We agree with some authors (Clark and Harvey 2002; Hornaday and Bloom 2006; Ortega-Argueta et al. 2011) that single-species plans are more clear and focused on prescribing management actions due to a greater availability of technical knowledge and better understanding of ecological and managerial aspects at species level. Recovery planning entails an intellectual, scientific and creative exercise to identify the major causes of species endangerment and formulate the best likely responses, within given resource and time constraints. This exercise is highly influenced by the amount of available knowledge; there is more scientific information at species level than at the level of ecological communities or ecosystems. This issue may also have had an influence on the other quality criteria, since the multi-species plans exhibit a poorer understanding of the gaps in the scientific knowledge and presented more inconsistencies in terms of setting performance evaluation criteria.

Previous studies contrasting single- vs. multi-species plans highlighted the advantages and limitations of both planning scopes. For example, Jewell (2000) states that the multi-species approach promotes thinking on a broader scale, acknowledging a perspective in population management. In theory, multi-species plans can save time and money; they avoid the need to address habitats and threat management separately, which may ultimately be more cost-effective. Under the umbrella of ecosystem management, multi-species plans can also reduce conflicts between listed species that occur in the same area (Boyer 2001). There are, however, occasions when single-species recovery plans are preferable or when they should be developed in addition to multi-species plans. For example, extremely endangered species may require a more specific and immediate intervention, such as captive management.

While the multi-species approach has a sound ecological basis, once implemented, its management limitations are revealed: multi-species plans are more complex, ambitious and often lack a clear statement of implementation responsibilities and prioritization of action (Boyer 2001). Complexity implies the difficulty of coordinating a large number of recovery activities in parallel, with competing interests and diverse stakeholders involved in implementation, which may demand strong partnerships (Hornaday and Bloom 2006; Joseph et al. 2008; Jantke and Schneider 2010). The large scope of plans may come at the expense of species-specific and site-specific actions (Hornaday and Bloom 2006). Rahn et al. (2006) found that multi-species habitat plans cover species for which occurrence has not been confirmed and did not feature specific management actions; it is simply assumed that generalized habitat conservation will adequately protect all species found in the habitat. This complexity is also reflected in poor consideration of monitoring activities within plans (Boersma et al. 2001; Campbell et al. 2002). In addition, empirical research has found a relatively poor performance of multi-species plans over time: for example, Lundquist et al. (2002) found that single-species plans have a higher degree of implementation and were implemented faster than multi-species plans. Boersma et al. (2001) found that species covered by multi-species recovery plans were four times less likely to present a status improvement than species covered by single-species plans. These results suggest that multi-species plans are actually less cost-effective management tools than single-species plans (Boersma et al. 2001; Clark and Harvey 2002; Cullen et al. 2005). Taylor et al. (2005) correlated species with recovery plans in operation with their conservation status and observed that the species within single-species plans appeared to fare better than those in multi-species plans. The reason for the reduced effectiveness of multi-species plans is unclear, but could be attributed to a lack of management focus and consequent lack of sufficient attention to the needs of each species (Clark and Harvey 2002), or to recovery planning being obscured by the complexity of addressing ecosystems that contain a high number of species and multiple threats. Other limitations are related to the difficulty of establishing partnerships and management agreements in conflict situations and among contrasting stakeholder interests, which may cause years of delay (Alagona and Pincetl 2008). Such caveats in the recovery plans may lead to poor implementation effectiveness and failure of recovery. As other authors have suggested, single- and multi-species approaches should be complementary rather than exclusive (Flather et al. 1998; Meffe et al. 2002; Ortega-Argueta et al. 2011).

Regarding hypothesis 2, no significant differences were observed between the plans of the old and new legislation periods in two of the three criteria: 'addressing gaps in the scientific information' and 'addressing identified threats' (disaggregated criteria analysis). Single-species plans under the EPBCA performed better only in relation to the criterion 'consistency of established plan objectives with performance evaluation criteria'. The overall consistency of plans (aggregated QI analysis) did improve under the new legislation period, with the combined effect of single-species planning. There may be two possible explanations for this. Better quality scores for plans under the new legislation could be a result of accumulated management knowledge, or may be due to improved legislation and introduction of planning guidelines that did not exist before. Both cases feature time as a common denominator with a positive effect on improving recovery planning. We recall that recovery has been a long-term endeavor of governments, with 40 years in the US and 20 in Australia.

Regarding the first likely explanation ('accumulated management knowledge'), previous studies state that time, rather than revision, was associated with improved trends of species status (Boersma et al. 2001). The explanation given is that recovery plan revisions are a response to, rather than a cause of, changes in species status (Boersma et al. 2001) that capitalizes on improved scientific and management information and understanding and uses adaptive management to promote more effective recovery planning. Time may also be associated with the availability of improved management capacity for implementing recovery plans. For example, Lundquist et al. (2002) found that plan revision, assignment of a recovery coordinator and establishment of a recovery database were associated with a higher number of tasks implemented. Revised plans, although a minimum number of cases in Australia, probably prescribe management actions more clearly than the original plans, and administrative strategies that include a coordinator and database probably serve to improve the effectiveness and efficiency of recovery efforts (Lundquist et al. 2002). Complementarily, revised and updated plans tend to present more scientific information than the corresponding original versions (Harvey et al. 2002).

Regarding the second likely explanation ('improved legislation and introduction of the planning guidelines'), analysis of the disaggregated criteria only detected significant differences for one criterion. Our evidence is therefore limited to validating our hypothesis 2 in terms of the effect of new legislation and planning guidelines in improving quality of plans. However, we did find evidence in the aggregated QI that the combination of the new EPBCA plans and the focus on single-species improved the consistency of planning. A previous study (Ortega-Argueta et al. 2011) that assessed the compliance of Australian plans against a list of mandatory planning attributes, found that plans showed a better coverage of planning topics in the new EPBCA period, which also included the new guidelines. Here, the guidelines may have been a useful tool for managers and authors involved in recovery planning. However, adequate compliance with the guidelines checklist has not necessarily been translated into improved plans in terms of the internal consistency measures assessed in our study. More detailed instructions may be necessary in the official guidelines in order to provide advice on the attributes of internal consistency. Planning methods such as the 'logic models' and 'theories of change' (Wyatt Knowlton and Phillips 2013) are useful for improving understanding regarding the construct of recovery plans and the assumptions that underline the consistency and causal links between management needs and the necessary prescription of interventions.

It is difficult to separate the two likely effects of accumulated scientific and management knowledge and improved legislation on the quality of plans. In our assessment design, we did not include measures to do so. We can assume that a larger number of scientific publications referenced in plans may indicate a better knowledge. In this way, we could measure some indication that recent plans with a larger number of referenced publications improved quality through better scientific knowledge. Recovery planning continues in Australia with incorporation of new plans every year. A study covering a longer time span of recovery planning could identify more complete historical patterns and generate better explanations regarding the quality of plans.

Conclusion

The analysis revealed that recovery plans are internally consistent in some areas, such as the prescription of research actions to address gaps in the scientific information. However, planning improvements are necessary for addressing threat processes, especially those of greater complexity. In contrast, due to the lack of quantitative/qualitative measures and thresholds, internal consistency was poor when it came to definition of performance evaluation criteria. Our study identified planning deficiencies regarding scientific and managerial aspects. Such deficiencies could be addressed by improving recovery planning guidelines and more carefully reviewing the drafting and adoption of new recovery plans. The consistency criteria and questions that we formulated for conducting our assessment could be incorporated into the planning guidelines. As the research has demonstrated, the compliance of plans with official planning requirements may not necessarily be reflected with adequate internal consistency. Careful definition of planning elements is essential for effective management. We emphasize the critical importance of establishing appropriate objectives, performance criteria and monitoring programs that strengthen the reliability of recovery plans, regardless of their scope. We do not advocate between single- or multi-species plans; we have raised the issues and limitations that recovery plans face under both approaches and that should be addressed. There is room for further improvement of the quality of design of plans for the recovery of threatened species worldwide. Research should be directed toward examining the effective incorporation of science into recovery planning, the decision-making process for prioritizing implementation of recovery actions and evaluation of effectiveness of recovery plans at multiple management scales. Our evaluation framework could assist in identifying scientific knowledge gaps, addressing threats more effectively and establishing quality recovery criteria. We believe that the criteria used in this study could be incorporated into official guidelines in order to strengthen the planning process. Given the universality of the evaluation framework, it can also be applied to improve the design and planning of a broad range of biodiversity conservation programs.

Supporting Information

The list of recovery plans evaluated in the study is shown in the Online Resource 1. The judgment criteria and the scoring system for assessing the quality of recovery plans are available in the Online Resource 2. The IUCN classification scheme of threat categories is shown in the Online Resource 3.

Acknowledgements Research was funded by the National Council of Science and Technology of Mexico, the University of Queensland (UQ) and the Wildlife Preservation Society of Australia. The database was designed with assistance of K. Mehltreter and S. Garza (INECOL) and B. Maher (UQ). Jose Villamizar assisted with data collection. Our grateful thanks go to D. Rao, C. López, M. Mendoza, K. MacMillan, M. Hernández and three anonymous reviewers who provided constructive comments that improved the manuscript.

References

Abbitt RJF, Scott JM (2001) Examining differences between recovered and declining endangered species. Conserv Biol 15:1274–1284

Alagona PS, Pincetl S (2008) The Coachella Valley Multiple Species Habitat Conservation Plan: a decade of delays. Environ Manage 41:1–11

Australian Government (2016a) Recovery plans. Department of the Environment, Commonwealth of Australia, Canberra, ACT. https://www.environment.gov.au/biodiversity/threatened/recovery-plans. Accessed May 2016

Australian Government (2016b) Recovery planning - Compliance checklist for legislative and process requirements to be provided with recovery plans for terrestrial threatened species and ecological communities, Department of the Environment, Commonwealth of Australia, Canberra, ACT. http://www.environment.gov.au/resource/recovery-planning-compliance-checklist-legislative-and-process-requirements-be-provided. Accessed May 2016

Australian Government (2016c) Approved threat abatement plans. Department of the Environment, Commonwealth of Australia, Canberra, ACT. https://www.environment.gov.au/biodiversity/threatened/threat-abatement-plans/approved. Accessed May 2016

Boersma PD, Kareiva P, Fagan WF, Clark JA, Hoekstra JM (2001) How good are endangered species recovery plans? Bioscience 51:643–649

Bottrill MC, Walsh JC, Watson JEM, Joseph LN, Ortega-Argueta A, Possingham HP (2011) Does recovery planning improve the status of threatened species? Biol Conserv 144:1545–1601

Boyer DM (2001) Single-species versus multiple-species recovery plans: a review of the Stock Island Tree Snail recovery plans. Endanger Species Update 18:223–226

Brigham CA, Power AG, Hunter A (2002) Evaluating the internal consistency of recovery plans for federally endangered species. Ecol Appl 12:648–654

Burbidge AA (1996) Essentials for a good recovery plan. In: Stephens S, Maxwell S (eds) Back from the brink: refining the threatened species recovery process. Surrey Beatty & Sons PTY Limited, Chipping Norton, pp 55–62

Burgman MA, Keith D, Hopper SD, Widyatmoko D, Drill C (2006) Threat syndromes and conservation of the Australian flora. Biol Conserv 134:73–82

Campbell SP, Clark JA, Crampton LH, Guerry AD, Hatch LT, Hosseini PR, Lawler JJ, O'Connor RJ (2002) An assessment of monitoring efforts in endangered species recovery plans. Ecol Appl 12:674–681

Carroll R, Augspurger C, Dobson A, Franklin J, Orians G, Reid W, Tracy R, Wilcove D, Wilson J (1996) Strengthening the use of science in achieving the goals of the endangered species act: an assessment by the ecological society of America. Ecol Appl 6:1–11

Clark JA, Harvey E (2002) Assessing multi-species recovery plans under the Endangered Species Act. Ecol Appl 12:655–662

Clark JA, Hoekstra JM, Boersma PD, Kareiva P (2002) Improving U.S. Endangered Species Act recovery plans: key findings and recommendations of the SCB Recovery Plan Project. Conserv Biol 16:1510–1519

Cork S, Sattler P, Alexandra J (2006) Biodiversity: theme commentary prepared for the 2006 Australian State of the Environment Committee. Department of the Environment and Heritage, Canberra

Cullen R, Moran E, Hughey KFD (2005) Measuring the success and cost effectiveness of New Zealand multiple-species projects to the conservation of threatened species. Ecol Econ 53:311–323
Dickman CR (1996) Incorporating science into recovery planning for threatened species. In: Stephens S, Maxwell S (eds) Back from the brink: Refining the threatened species recovery process, Surrey Beatty & Sons PTY Limited. Chipping Norton, NSW, pp 66–73
Doremus H, Pagel JE (2001) Why listing may be forever: perspectives on delisting under the U.S. Endangered Species Act. Conserv Biol 15:1258–1268
Flather CH, Knowles MS, Kendall IA (1998) Threatened and endangered species geography. Bioscience 48:365–375
Foin TC, Riley SPD, Pawley AL, Ayres DR, Carlsen TM, Hodum PJ, Switzer PV (1998) Improving recovery planning for threatened and endangered species. Bioscience 48:177–184
Gerber LR, Hatch LT (2002) Are we recovering? An evaluation of recovery criteria under the U.S. Endangered Species Act. Ecol Appl 12:668–673
Gibbs KE, Currie DJ (2012) Protecting endangered species: do the main legislative tools work? PLoS ONE 7:e35730. doi:10.1371/journal.pone.0035730
Hadlock TD, Beckwith JA (2002) Recommendations to improve recovery of endangered species in the United States. Hum Dimens Wildl 7:37–53
Harvey E, Hoekstra JM, O'Connor RJ, Fagan WF (2002) Recovery plan revisions: progress or due process? Ecol Appl 12:682–689
Hoekstra JM, Clark JA, Fagan WF, Boersma PD (2002) A comprehensive review of Endangered Species Act recovery plans. Ecol Appl 12:630–640
Hornaday K, Bloom V (2006) Multispecies recovery planning: benefits and challenges. Endangered Species Update 23:10–11
IUCN (2012) Threats Classification Scheme (Version 3.2). International Union for Conservation of Nature. http://intranet.iucn.org/webfiles/doc/SSC/RedList/AuthorityF/threats.rtf. Accessed February 2015
IUCN-SSC (2008) Strategic planning for species conservation: a handbook, Version 1.0. Gland, Switzerland: IUCN Species Survival Commission. http://www.iucn.org/backup_iucn/iucn.org/about/work/programmes/species/who_we_are/ssc_specialist_groups_and_red_list_authorities_directory/mammals/african_elephant/strategies_plans/index4133.html?uPubsID=3553. Accessed January 2015
Jantke K, Schneider UA (2010) Multiple-species conservation planning for European wetlands with different degrees of coordination. Biol Conserv 143:1812–1821
Jewell SD (2000) Multi-species recovery plans. Endangered Species. Bulletin 25:30–31
Joseph C, Gunton TI, Day JC (2008) Implementation of resource management plans: identifying keys to success. J Environ Manage 88:594–606
Kleiman DG, Reading RP, Miller BJ, Clark TW, Scott M, Robinson J, Wallace RL, Cabin RJ, Felleman F (2000) Improving the evaluation of conservation programs. Conserv Biol 14:156–365
Lawler JJ, Campbell SP, Guerry AD, Kolozsvary MB, O'Connor RJ, Seward LCN (2002) The scope and treatment of threats in endangered species recovery plans. Ecol Appl 12:663–667
Lundquist CJ, Diehl JM, Harvey E, Botsford LW (2002) Factors affecting implementation of recovery plans. Ecol Appl 12:713–718
Meffe GK, Nielsen LA, Knight RL, Schenborn DA (2002) Ecosystem management: adaptive, community-based conservation. Island Press, Washington, DC
Mertens DM, Wilson AT (2012) Program evaluation, theory and practice: a comprehensive guide. The Guilford Press, NY
Miller B, Reading R, Conway C, Jackson JA, Hutchins M, Snyder N, Forrest S, Frazier J, Derrickson S (1994) A model for improving endangered species recovery programs. Environ Manage 18:637–645
Morris LL, Fitz-Gibbon CT, Lindheim E (1987) How to measure performance and use tests. SAGE Publications, Inc., Newbury Park
National Recovery Working Group (2005) Recovery Handbook (ROMAN). 2005-2006 Edition, October 2005. Recovery of Nationally Endangered Wildlife, Ottawa, Ontario
NMFS-USFWS (2010) Interim Endangered and Threatened Species Recovery Planning Guidance, Version 1.3. National Marine Fisheries Service and U.S. Fish and Wildlife Service, Silver Spring, MD
Ortega-Argueta A, Contreras-Hernández A (2013) Proposal for a monitoring and evaluation framework for recovery programs of species at risk. Gest Polit Publica 22:457–496
Ortega-Argueta A, Baxter G, Hockings M (2011) Compliance of Australian recovery plans with threatened species legislation. J Environ Manage 92:2054–2060
Pulliam HR, Babbitt B (1997) Science and the protection of endangered species. Science 275:499–500
Pullin AS, Knight TM (2003) Support for decision making in conservation practice: an evidence-based approach. J Nat Conserv 11:83–90

R Development Core Team (2012) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
Rahn ME, Doremus H, Diffendorfer J (2006) Species coverage in multispecies habitat conservation plans: where's the science? Bioscience 56:613–619
Roberts JD, Hamann M (2016) Testing a recipe for effective recovery plan design: a marine turtle case study. Endanger Species Res 31:147–161
Sattler P, Creighton C (2002) Australian Terrestrial Biodiversity Assessment 2002. National Land and Water Resources Audit on behalf of the Commonwealth of Australia, Turner, ACT
Schultz CB, Gerber LR (2002) Are recovery plans improving with practice? Ecol Appl 12:641–647
Schwartz MW (2008) The Performance of the Endangered Species Act. Annu Rev Ecol Evol Syst 39:279–299
Smallwood KS, Beyea J, Morrison ML (1999) Using the best scientific data for endangered species conservation. Environ Manage 24:421–435
Taylor MFJ, Suckling KF, Rachlinski JJ (2005) The Effectiveness of the Endangered Species Act: a quantitative analysis. Bioscience 55:360–367
Tear TH, Scott JM, Hayward PH, Griffith B (1995) Recovery plans and the Endangered Species Act: are criticisms supported by data? Conserv Biol 9:182–195
Venables WN, Ripley BD (2002) Modern applied statistics with S, 4th edn. Springer, NY
Wallace RL (2003) Social influences on conservation: lessons from U.S. recovery programs for marine mammals. Conserv Biol 17:104–115
Wyatt Knowlton L, Phillips CC (2013) The logic model guidebook, better strategies for great results, 2nd edn. SAGE Publications, Inc., Thousand Oaks