Journal of Environmental Management 139 (2014) 164–171

Contents lists available at ScienceDirect

Journal of Environmental Management

journal homepage: www.elsevier.com/locate/jenvman

Measuring the accuracy of management effectiveness evaluations of protected areas

Carly N. Cook [a,b,*], R.W. (Bill) Carter [c], Marc Hockings [a]

[a] School of Geography, Planning and Environmental Management, University of Queensland, St Lucia, QLD 4072, Australia
[b] Quantitative and Applied Ecology Group, School of Botany, University of Melbourne, Parkville, VIC 3010, Australia
[c] Sustainability Research Centre, University of the Sunshine Coast, Maroochydore DC, QLD 4558, Australia

ARTICLE INFO

Article history:
Received 26 February 2013
Received in revised form
22 January 2014
Accepted 23 February 2014
Available online 1 April 2014

Keywords:
Expert opinion
Elicitation
Environmental management
Practitioner
Uncertainty

ABSTRACT

Evaluating the effectiveness of protected area management to help improve management outcomes is becoming an increasingly common practice. The evaluation tools developed and implemented in over 100 countries generally rely on the expert judgements of protected area managers. Despite the growing collection and use of management effectiveness evaluation data, there have been no previous attempts to measure the accuracy of these data. We measured the accuracy of managers' judgements about the conditions in their reserves by collecting independent field data. We also assessed how accurately the evaluation tool reflected managers' views by conducting semi-structured interviews with 23 protected area managers from New South Wales, Australia. We found that managers made highly accurate judgements of the extent of a common weed species, *Rubus fruticosus* (blackberry), but often misinterpreted the scope, scale and timeframe of the evaluation. These framing effects can lead to error being introduced into the evaluation dataset, affecting the precision of evaluations such that they cannot be reliably compared among reserves. We suggest that the wording of evaluation questions needs to be explicit about the assessment frame to minimize the influence of framing effects on management effectiveness evaluations.

1. Introduction

For protected areas to achieve their promise for biodiversity conservation they must be managed effectively (Chape et al., 2005; Redford and Taber, 2000). Clear management objectives and knowledge of whether management actions are achieving the desired conservation outcomes, or if changes are needed, are important for achieving conservation outcomes (Hockings et al., 2004). Efforts to measure the effectiveness of management in protected areas have yielded over 50 different tools, used in more than 100 countries, equating to approximately 5% of the world's protected areas being evaluated so far (Leverington et al., 2010). These evaluations have often been driven by pressure from governments and non-government organizations funding management activities, who want to know the conservation outcomes associated with their investment in protected area management. For example, organisations like the Global Environment Fund require management effectiveness evaluations (MEE) for sites where they fund conservation projects. There is also an increasing commitment from the international community to evaluate the effectiveness of protected areas with the United Nations adopting MEE of protected areas as an indicator of biodiversity conservation (Walpole et al., 2009). Likewise, the Convention on Biological Diversity (CBD) has now specified that meeting targets for the coverage of protected areas requires that they be "effectively managed" (CBD, 2010), and they have challenged signatory countries to evaluate 60% of their protected areas by 2015 (CBD, 2004).

In addition to demonstrating the conservation outcomes associated with protected area management, and identifying where management strategies might need to be adapted (Hockings, 2003), MEEs are used to prioritize management actions and guide the allocation of resources within and among protected areas (Cook and Hockings, 2011). Therefore, MEEs have the potential to significantly influence the conduct of management in a growing number of protected areas globally. A lack of empirical evidence about

* Corresponding author. Quantitative and Applied Ecology Group, School of Botany, University of Melbourne, Parkville, VIC 3010, Australia. Tel.: +61 3 9035 6118.
E-mail addresses: carly.cook@monash.edu (C.N. Cook), bcarter@usc.edu.au (R.W. (Bill) Carter), m.hockings@uq.edu.au (M. Hockings).

http://dx.doi.org/10.1016/j.jenvman.2014.02.023
0301-4797/

protected area management (Cook et al., 2010a) has resulted in most MEEs being based on judgements made by the managers (i.e., those responsible for the day-to-day decisions about management, and those responsible for strategic decisions about management priorities and resource allocation) of the protected area (Cook and Hockings, 2011). Expert judgement is commonly used as a substitute for quantitative data in disciplines such as ecology (Kuhnert et al., 2010), natural resource management (e.g., Baird and Flaherty, 2005) and conservation biology (Martin et al., 2010). Experienced biologists have a wealth of unpublished knowledge about biodiversity (Maddock and Samways, 2000) and while largely untested, it is reasonable to expect that this is also true for experienced protected area managers. However, personal judgements can be prone to a range of biases (Burgman, 2000). For this reason, the rigour of MEEs based on the judgements of protected area managers has been criticized (Cook and Hockings, 2011).

Those developing and implementing MEEs have responded to the criticisms made of evaluation tools by employing several measures to improve the rigour with which managers' judgements are elicited, such as using supporting information when it is available and facilitating discussion amongst evaluators by using workshops to elicit information (Cook and Hockings, 2011). While the improvements made to evaluation tools are grounded in theory, the accuracy of the data they generate is almost never assessed (Cook and Hockings, 2011). Studies that have examined whether the management capacity of protected areas (e.g., the adequacy of budgets, staffing levels and planning processes) is a good proxy for desired conservation outcomes have shown both positive and negative results. For example, fire frequency was found to be independent of how a protected area scored on the Management Effectiveness Tracking Tool (METT) (Nolte and Agrawal, 2013). However, an increase in METT scores was shown to be positively correlated with changes in key biodiversity populations (Zimsky et al., 2010). However, using the overall METT score to infer conservation outcomes is misleading, considering only one of the 20 questions actually address conservation outcomes (Zimsky et al., 2010) and the implementation of a conservation action can be a poor predictor of whether the project will achieve the desired effect (Kapos et al., 2009). Therefore, the overall score from a MEE cannot be used to infer the accuracy of the evaluation. Instead, observations of conservation outcomes must be directly compared with evaluations of management outcomes.

Management effectiveness evaluations often represent the only information available about the management of protected areas (Cook et al., 2010a). Yet, without an understanding of the accuracy of these data, it is difficult to know whether using MEEs to guide protected area management will lead to improved conservation outcomes. In this study, we measured the accuracy of State of the Parks MEEs (Hockings et al., 2009a), which are conducted by protected area managers in Australia and Korea (Cook and Hockings, 2011). This system-wide (i.e., all or most of the protected areas within the network) evaluation tool relies on managers' judgements of management effectiveness for up to 30 different management issues relating to biodiversity conservation, historic and cultural heritage management, and visitation (Hockings et al., 2009a).

The two major sources of error within MEEs based on managers' judgements include managers' perceptions of on-ground conditions in reserves being false, and error originating from the process of eliciting judgements. Therefore, to evaluate the rigour of MEEs we sought to assess: (i) the accuracy of managers' judgements relative to on-ground conditions; and (ii) whether the evaluation tool could accurately elicit the views of managers. By understanding the level of error associated with both these elements of the evaluation process, we could determine both the level of confidence to be placed in qualitative assessments of management effectiveness and identify opportunities to improve the rigour of future evaluations.

2. Materials and methods

2.1. State of the Parks management effectiveness evaluations

In this study, we focussed on the State of the Parks (SoP) evaluation tool used by the New South Wales Office of Environment and Heritage (NSW OEH), which asks managers to evaluate the effectiveness of management against a set of four assessment intervals (e.g., Table 1), representing an ordinal scale of management outcomes and acting as a set of management standards (Hockings et al., 2009a). The SoP tool involves some commonly used methods for oversight and standardization of evaluation results that aim to promote consistency within and between evaluations (Cook and Hockings, 2011). These include capturing multiple opinions by asking managers to complete evaluations as a group, training managers prior to evaluation, and providing them with written guidelines (Hockings et al., 2009a). This provides an opportunity to test the effectiveness of these commonly used approaches to reduce elicitation errors.

2.2. Sample selection

Evaluation questions that address the outcomes of management are likely to be difficult to address without empirical evidence available (Hockings et al., 2009b). Therefore, we selected a question within the SoP evaluation tool that addressed the effectiveness of invasive plant (weed) management, a common and actively managed threat to biodiversity in most Australian protected areas (ASEC, 2011). We selected protected areas where managers reported that weed management was a priority management issue for protecting important vegetation within the protected area (NSW OEH unpublished data). We chose to focus on the management of *Rubus fruticosus*, the second most commonly managed weed in NSW (NSW OEH unpublished data), and the one with the widest geographic distribution (Coutts-Smith and Downey, 2006). *R. fruticosus* is a Weed of National Significance in Australia (Thorp and Lynch, 2000), is highly visible, easily identified, and its impacts on biodiversity have been clearly defined (Coutts-Smith and Downey, 2006). This weed forms dense thickets, excluding native flora and providing harbour for introduced pests, and its impacts on biodiversity are proportional to the size of the infestation (Coutts-Smith and Downey, 2006). This allowed us to use the area of the infestation as indicative of the impact of *R. fruticosus*, making it an ideal species for this study. We used the 2007 State of the Parks evaluations to identify reserves where managers consider *R.*

Table 1
The State of the Parks question addressing the effectiveness of weed management.

Assessment intervals:	
☐	Negative impacts of weeds on reserve values are negligible
☐	Negative impacts of weeds on reserve values are diminishing
☐	Negative impacts of weeds on reserve values are unchanged
☐	Negative impacts of weeds on reserve values are increasing
Evidence used to make the evaluation:	
☐	Staff experience
☐	Specialists opinion
☐	Community opinion
☐	Research
☐	Planning documents
☐	Corporate database
☐	Monitoring
Explanation for evaluation given:	