Approaches Used to Evaluate the Social Impacts of Protected Areas

Emiel de Lange1, Emily Woodhouse2,3, & E.J. Milner-Gulland1,4

1 Department of Life Sciences, Imperial College London, Silwood Park Campus, Ascot SL5 7PY, UK
2 Department of Anthropology, University College London, 14 Taviton Street, London WC1H 0BW, UK
3 Wildlife Conservation Society Europe, Charles Darwin House, 12 Roger Street, London WC1N 2JU, UK
4 Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

Keywords
Social impact; evaluation; wellbeing.

Abstract
Protected areas are a key strategy in conserving biodiversity, and there is a pressing need to evaluate their social impacts. Though the social impacts of development interventions are widely assessed, the conservation literature is limited and methodological guidance is lacking. Using a systematic literature search, which found 95 relevant studies, we assessed the methods used to evaluate the social impacts of protected areas. Mixed methods were used by more than half of the studies. Almost all studies reported material aspects of wellbeing, particularly income; other aspects were included in around half of studies. The majority of studies provided a snapshot, with only one employing a before-after-control-intervention design. Half of studies reported respondent perceptions of impacts, while impact was attributed from researcher inference in 1/3 of cases. Although the number of such studies is increasing rapidly, there has been little change in the approaches used over the last 15 years, or in the authorship of studies, which is predominantly academics. Recent improvements in understanding of best practice in social impact evaluation need to be translated into practice if a true picture of the effects of conservation on local people is to be obtained.

Introduction
Conservation interventions have wide-ranging social impacts—both positive and negative. For example, protected areas can alter resource use-rights and displace communities (West et al. 2006), but can also secure ecosystem services, and generate employment and income (Pullin et al. 2013). Conservationists are increasingly recognizing that their interventions should benefit people and improve human wellbeing (Campagna & Fernandez 2007), and this principle is enshrined in the Convention on Biological Diversity (Secretariat of the Convention on Biological Diversity, 1992). Robust and comprehensive monitoring and evaluation of social impacts is therefore essential to ensure greater transparency and accountability, improve learning, and support effective allocation of conservation resources (Grantham et al. 2009). Demonstrating positive social outcomes could also improve support among and cooperation with local people (Milner-Gulland et al. 2014).

Evaluations of the impacts of interventions on people are common in development (Baker 2000), but less so in conservation (PCLG, TILCEPA, UNEP-WCMC & WCPA/CEESP taskforce, 2007; Schreckenberg et al. 2010). An impact evaluation has three main components (Gertler et al. 2011): First, relevant indicators are needed to assess changes in human wellbeing caused by an intervention. Secondly, evaluations need to be designed so that wellbeing outcomes are linked to the intervention being studied rather than to other factors. Thirdly, data need to be collected in an appropriate way, both in terms of the methods used and the overall sampling strategy.

Wellbeing is a broad term with multiple meanings (Leisher et al. 2013), but there is increasing agreement in international policy circles that it encompasses objective material components, relational aspects, and subjective...
experiences (Stiglitz, Sen & Fitoussi 2009). Empirical research has shown that there are broadly five aspects which are held in common: material assets, health, social relations, security, and freedom of choice and action (Narayan et al. 2000; Millennium Ecosystem Assessment, 2005). A broad set of indicators that reflects each aspect of wellbeing in both objective and subjective dimensions allows for more accurate and valid assessments of the impacts of interventions than indicators which focus on specific components (King et al. 2014). Local relevance can be ensured through participatory research with local stakeholders (Abunge et al. 2013).

A full experimental design is rarely possible in conservation, but robust attribution of outcomes to interventions is feasible using other designs (McConnachie et al. 2015). In quasi-experimental designs the researcher selects control groups in order to estimate the counterfactual—the case in the absence of the intervention. A before-after-control-intervention design combines controls with baseline data to further control for initial conditions (Ferraro & Pattanayak 2006; Ferraro 2009). Collecting quantitative data in such designs allows attribution of impacts whilst reducing bias from confounding factors by using statistical techniques such as matching (Gertler et al. 2011). However, nonstatistical methods can also be used to make causal inferences (Stern et al. 2012). Participatory methods attribute change based on the perceptions of those impacted—a “reflexive counterfactual” (Franks et al. 2014). Researchers may also make inferences by comparing evidence to predictions from theory (He et al. 2008). Choice of study design and method of attribution ultimately depend on the requirements of the researcher (Mascia et al. 2014). Quasi-experimental statistical methods, when used appropriately, can answer the “what” question and estimate the magnitude of impacts, whereas alternative methods are better suited to explaining “why” and “how” impacts have occurred (Stern et al. 2012).

The choice of data collection and sampling methods largely depends on the question to be answered and the form of causal inference required. Different methods are better suited to collecting qualitative and quantitative or objective and subjective data types (Wongbusarakum et al. 2014). Furthermore, different types of people will be impacted in different ways, and sampling across relevant subgroups (e.g., livelihoods, genders, and ethnicities) will ensure heterogeneity is captured (Daw et al. 2011; King et al. 2014).

A number of conservation organizations are developing guidelines for assessing the social impacts of their interventions (Schreckenberg et al. 2010; Wongbusarakum et al. 2014), and understanding of best practice is improving (Ferraro & Pattanayak 2006; Roe et al. 2013; Woodhouse et al. 2015). However there is limited evidence on how impact evaluations have been conducted to date; specifically the way in which they have approached the three components of the process. A formal review of practice could reveal shortcomings or strengths, and whether changes in understanding of best practice have translated into real-world implementation. Effort spent improving methods should be justified with reference to evidence of past failure, not just theoretical future ideals.

We conducted a systematic literature search in order to provide an overview of the methods used for evaluating the social impacts of protected areas to date. We structured our review around the three components of evaluation (selection of indicators; research design; data collection) in relation to current understanding of best practice, and investigated whether the methods used have changed over time as understanding of best practice has improved.

Methods

Systematic literature search

A systematic literature search was carried out based on published guidelines (Pullin & Stewart 2006) with search terms adapted from a recent systematic review of the impacts of protected areas (Pullin et al. 2013). We searched the academic literature through online databases and the grey literature on the websites of 19 relevant organizations. The search terms used were chosen in order to capture both the different types of protected area interventions, and the full range of terminology used to describe the social dimensions of impacts. Terms were adjusted to the search capabilities and requirements of each database. Relevance screening was done by title and then by abstract. Publications were selected if they reported an assessment of the wellbeing impacts of a protected area on a local human population. This criterion allows for studies where a wellbeing evaluation is not the principal objective, but has nevertheless been undertaken. However, this criterion excludes studies such as economic valuations of protected areas at a national level, which do not attempt to assess the wellbeing impacts on a specific population. Only English-language publications were retained. Full details of the search are given in supplementary material file S1.

Data extraction

Key information on the methods used by each study was extracted and codified (the protocol is given in supplementary material file S2). The unit of analysis was the
study, each of which made up one entry, including those assessing multiple protected areas. Background data were collected from all studies, while detailed information on methodology was collected where possible. The database search returned a total of 8,679 results, of which 75 were selected in addition to 15 from specialist websites. Of total 90 publications were retained for data extraction as detailed in supplementary material file S3. In these, 95 studies were reported. However 5 did not report their methods adequately and were excluded from further analysis.

We used a predefined framework to categorize wellbeing into 6 aspects (Woodhouse et al. 2015). We split income from other material aspects due to its preponderance in studies and importance as an indicator at the national and international levels. We classified study design and the method used to link impacts to the protected area drawing on existing typologies such as Stern et al. 2012. Data collection methods were categorized based on Wongbusarakum, Madeira & Hartanto 2014. Data were classified as quantitative (numeric) or qualitative (text-based), or both, as well as objective (externally verifiable e.g. material assets) or subjective (feelings or perceptions), or both. Full details on categorizations used are provided in supplementary material file S4.

**Statistical analysis**

Statistical tests were carried out in R (R Core Team, 2014). As the data are nonnormal, nonparametric tests were chosen, including Chi-square and Spearman’s rank correlation. When Chi-square tests were used to investigate changes over time, years were divided into three roughly equal periods: 1999–2005, 2006–2010, and 2011–2015.

**Results**

On average, 2.2 (SD = 0.79) relevant studies were carried out annually between 1999 and 2006. From 2007 there was a linear increase in the number of studies, with 14 being carried out in 2014 (Figure 1). Academic authors were involved in most studies (88%), with 67% of studies having only academic authors. The majority of the remaining studies were carried out by NGOs (21% of the total).

**Wellbeing outcomes assessed**

Material aspects of wellbeing (including income) were assessed in 89 of 90 studies. Income was assessed in 68% of studies while other material aspects were assessed in 87% of studies. Only 51% of studies assessed nonmaterial aspects of wellbeing, such as health, social relations, security and freedom (Table 1). A majority of studies (76%) examined multiple aspects of wellbeing, however if material aspects (income and other material) are combined this is reduced to 49%. Only one study examined all aspects of wellbeing, but 12 studies looked at 4 or more of the 6 aspects. No increase in the number of aspects assessed was detected over time (Spearman’s rank, \( \rho = 0.08, P = 0.47 \))

**Research design**

The snapshot (with no control or baseline) was the most common study design (66%). No significant change in prevalence of the snapshot design was detected over time (chi-square, \( \chi^2 = 0.65, df = 2, P = 0.72 \)). Twenty-three percent of studies had a control (Table 1), but the before-after-control-intervention design was only employed in one study (Gurney et al. 2014). ‘Other’ study designs included gradients of proximity to the protected area. Fourteen studies used a combination of study designs; five used a control-intervention design combined with measuring post-intervention change over time in the intervention site, while four combined it with a snapshot study in the intervention site. For example, one study carried out quantitative interviews in an impacted and a control village, as well as participatory group discussions on perceived changes in the impacted village (Bashar 2013).

The most common method of attributing impacts to the protected area was through the perceptions of the people being studied (53%; Table 1). Other methods include inference by the researcher (36%), comparison with a control (23%) and the use of correlational or statistical relationships (12%). No change in the use of perceptions was found over time (chi-square, \( \chi^2 = 2.24, df = 2, P = 0.33 \)).

**Methods used**

The most common tool was the semi-structured interview survey, used in 76% of studies (Table 2). Other common tools included key informant interviews (38%), focus group discussions (31%), and self-complete questionnaires and open-ended interview surveys (both 13%). ‘Other’ methods (11%) included the use of secondary data, such as government records or censuses, or direct measurement of physical variables such as fish landings or market goods. 73% of studies used more than one method. Thirty-eight different method sets were reported in total, with the largest method-set comprising 4 methods. The most common included pairing semi-structured interview surveys with either
key informant interviews (11 studies) or focus group discussions (10 studies). There has been no significant change over time in the number of methods used in a study (Spearman’s rank, $\rho = 0.17$, $P = 0.11$).

Both quantitative and qualitative data were collected in 51% of studies, while 36% of studies only collected quantitative data. Similarly, 67% of studies collected both objective and subjective data while objective data alone were collected in 21% of studies. No significant changes in use of objective data (chi-square, $\chi^2 = 3.40$, df = 2, $P = 0.18$) or mixed objective/subjective data sets (chi-square, $\chi^2 = 2.20$, df = 2, $P = 0.33$) were found over time.

Most studies (89%) sampled at the household level while smaller numbers sampled at the individual (28%) or village (13%) levels. Only 21 studies (23%) reported sampling specific subgroups; most commonly livelihood groups (16 studies), followed by gender (7 studies), wealth (5), social status (2), age (3), and level of education (1).

Table 1 | The frequency of use of various aspects of wellbeing, study designs and methods of attribution

| Aspects of wellbeing | Studies | Study design | Studies | Attribution | Studies |
|----------------------|---------|--------------|---------|--------------|---------|
| Material - income    | 61      | Control–Intervention | 20      | Correlational | 11      |
| Material – other     | 78      | Before–After  | 4       | Matched      | 21      |
| Health               | 14      | BACI         | 1       | Perception   | 48      |
| Social Relations     | 25      | Snapshot     | 59      | Researcher inference | 32      |
| Security             | 8       | Change over time (postintervention) | 17      |              |         |
| Freedom of choice and action | 23 | Other | 3       |              |         |

Table 2 | Frequency of use of various methods of data-collection

| Methods                                                      | Papers |
|--------------------------------------------------------------|--------|
| Participatory observation                                   | 3      |
| Participatory rural appraisal                               | 7      |
| Interview surveys: Structured/semi structured/scales        | 68     |
| Interview surveys: open-ended questions                     | 12     |
| Focus group                                                 | 28     |
| Key informant interviews                                    | 34     |
| Self-complete questionnaires                                | 12     |
| Other                                                        | 10     |

Discussion

The literature search carried out here was comprehensive; only a small fraction (<0.02%) of the returned publications were relevant, suggesting that the search terms were sufficiently broad to capture most relevant
publications (Pullin & Stewart 2006). As a result, the data set is adequately representative of practice in assessing the social impacts of protected areas to date. The surge in published evaluations since 2007 could be a reflection of a general increase in publication within the field of conservation science. Alternatively, it could reflect growing recognition of the need to evaluate the social impacts of conservation (Ferraro & Pattanayak 2006; Cowling & Wilhelm-Rechmann 2007).

Overall, the view taken of human wellbeing by the studies was limited. Only one study examined the full breadth of aspects (Silva 2006), while the vast majority examined only a small number. A narrow view of wellbeing is unlikely to reflect reality as improvements in measured aspects could be offset by undetected declines in others. For example, one study found that, compared to controls, households in a national park showed improved health indicators but lower income and less trust in their neighbors (Foerster et al. 2011). Measuring only income could have led to the conclusion that proximity to the park decreased wellbeing, although the reality was much more complex. It is worrying then that material aspects of wellbeing are overwhelmingly dominant, with almost half of studies examining nothing else. As our conception of wellbeing is refined, evaluators should broaden the range of information they collect accordingly.

Different evaluation designs are appropriate for different research questions, audiences, types of intervention, and capacities (Stern et al. 2012). The strength of quasi-experimental designs, with a counterfactual, lies in attribution and estimation of impact magnitudes—information often desired by donors and policy makers building an evidence base. However, most studies favored nonstatistical methods of causal inference and study designs without controls or baselines, instead using snapshot designs and local perceptions or inference by the researcher. These patterns may have more power to explain and contextualize impacts, and so be more useful than quasi-experimental approaches for improving protected area management at the site level. However, the emphasis on nonstatistical attributions by academic authors is unexpected as their reliability is hard to ascertain, and these methods can be prone to bias and manipulation when not done systematically and with care regarding equity of participation (Catley et al. 2008; Ferraro 2009; Gertler et al. 2011). Overall, further analysis would be required in order to draw reliable conclusions on the quality and appropriateness of all the evaluation approaches employed to the specific circumstances of each case study. On another note, little is currently understood about the trajectory of change in impacts over time (Woolcock 2009) and the prevalence of snapshot designs means that these are rarely captured fully (Gurney et al. 2014; Woodhouse et al. 2015).

The final component of impact evaluation is data collection. A large majority of studies used a combination of data-collection methods. The most commonly used—the semi-structured interview survey—was frequently combined with key informant interviews or focus group discussions. These combinations are particularly useful for collecting mixed data, allowing structured or quantitative data from the interview survey to be supported by more in-depth qualitative data, and both objective and subjective measures of change in wellbeing to be captured. A large majority of studies sampled households as these are the basic unit around which economic activity is organized. This is consistent with the focus on material aspects of wellbeing such as income and assets. However, as different people conceive of wellbeing and are impacted in different ways, heterogeneity may exist within a household, for example across gender and age groups. It is important that this heterogeneity is captured through sampling individuals, not just taking household averages. This would enable distributional dimensions of equity to be captured (Daw et al. 2011). Similarly, impacts are manifested heterogeneously within the community, and this is poorly captured by most studies. Only a few ensured that specific subgroups were included in the sample, and of these only one or two subgroups were sampled. The fact that the most common subgroups were livelihoods-based reflects the material view of wellbeing adopted.

No changes in the approach to evaluating impacts over time were found; evaluators have not broadened their view of wellbeing, and remain largely reliant on snapshot studies capturing the perceptions of local people. Using perceptions data suggests a positive engagement with subjective aspects of wellbeing and with local people, whose support is vital for successful conservation. However, this should be complemented with evaluations providing robust evidence of causal linkages, ensuring wider legitimacy. The absence of change in the methodology suggests that the discussion underway in the academic conservation literature is not yet being translated into evaluation practice. As calls for impact evaluation in conservation were made relatively recently, with one of the earliest being Ferraro & Pattanayak (2006), it may still be too soon for adoption of new evaluation methods to be reflected in the literature. However as we covered both academic and grey literature, one might have expected some indication of new approaches being adopted if the grey literature is more rapidly published. Also, the sharp increase in the volume of literature after 2007 suggests that social impact evaluations are becoming more common,
even as the methods used remain unchanged. There may therefore also be barriers preventing implementation of new methods, such as budgets, time or technical capacity.

Guidance applicable to different scenarios of capacity, budget and objectives is beginning to emerge (IIED, 2014; Wongbusarakum et al. 2014; Woodhouse et al. 2015); and in time this will improve the quality of evaluation, provided that practitioners are given the right support. Direct collaboration between academic researchers and practitioners is a particularly powerful way to enable the translation of new thinking in academia into practice, and to ensure that academic research is grounded in the needs of practitioners (Gossa et al. 2015). Evidence that this collaboration is actively occurring would include co-authorship of reports and papers by people from both academic and practitioner (NGO, government) institutions. Our study’s finding that two-thirds of publications in this field are by academics alone is concerning, in the light of this need for collaborative learning. There is a pressing need for more and better evaluations of the social impacts of protected areas, and conservation in general, in order to improve the sustainability and local acceptability of conservation interventions. By highlighting the current state of practice, we hope to have contributed towards this aim.

Acknowledgments
This paper is a product of the ESRC/DFID funded project “Measuring complex outcomes of environment and development interventions.” We also gratefully acknowledge funding from an ESRC Impact Acceleration award to Imperial College London.

Supporting Information
Additional Supporting Information may be found in the online version of this article at the publisher’s web site:

Supplementary material file S1: Full details of the systematic literature search carried out.

Supplementary material file S2: Coding framework used to extract data from selected publications.

Supplementary material file S3: Publications Investigated.

Supplementary material file S4: Categorisations used in classifying methodologies.

References
Abunge, C., Coulthard, S. & Daw, T.M. (2013). Connecting marine ecosystem services to human well-being: insights from participatory well-being assessment in Kenya. *Ambio*, **42**(8), 1010-1021.

Baker, J.L. (2000). *Evaluating the impact of development projects on poverty: a handbook for practitioners*. World Bank Publications. Washington D.C., USA.

Bashar, S.A. (2013). Impact of MACH project activities on socio-economic and environmental conditions in Keuta Beel of Sherpur District. Pages 62-80 in J. Fox, M.G. Mustafa, Bryan R. Bushley, Stephen M. Brennan, and Laurie, editors. Connecting Communities and Conservation: Co-Management Initiatives Implemented by IPAC in Wetlands and Forests of Bangladesh. Dhaka: USAID; Integrated Protected Area Co-Management Project.

Campagna, C. & Fernandez, T. (2007). A comparative analysis of the vision and mission statements of international environmental organisations. *Environ. Values.*, **16**(3), 369-398.

Catley, A., Burns, J., Abebe, D. & Suji, O. (2008). *Participatory Impact Assessment: A Guide for Practitioners*. Feinstein International Centre, Tufts University, Somerville, MA, USA.

Cowling, R.M. & Wilhelm-Rechmann, A. (2007). Social assessment as a key to conservation success. *Oryx.*, **41**(2), 135.

Daw, T., Brown, K., Rosendo, S. & Pomeroy, R. (2011). Applying the ecosystem services concept to poverty alleviation: the need to disaggregate human well-being. *Environ. Conserv.*, **38**(04), 370-379.

Ferraro, P.J. (2009) Counterfactual thinking and impact evaluation in environmental policy. Pages 75-84 in M. Birnbaum & P. Mickwitz, editors. *Environmental program and policy evaluation. New directions for evaluation.* 122nd edition. pp.

Ferraro, P.J. & Pattanayak, S.K. (2006) Money for nothing? A call for empirical evaluation of biodiversity conservation investments. *PLoS Biol.*, **4**(4), e105.

Foerster, S., Wilkie, D.S., Morelli, G.A., et al. (2011). Human livelihoods and protected areas in Gabon: a cross-sectional comparison of welfare and consumption patterns. *Oryx.*, **45**(03), 347-356.

Franks, P., Roe, D., Small, R. & Schneider, H. (2014). *Social assessment of protected areas: early experience and results of a participatory, rapid approach*. International Institute for Environment and Development, London.

Gertler, P.J., Martinez, S., Premand, P., Rawlings, L.B. & Vermeersch, C.M. (2011). *Impact evaluation in practice.*, World Bank Publications. Washington D.C., USA.

Gossa, C., Fisher, M. & Milner-Gulland, E. (2015). The research-implementation gap: how practitioners and researchers from developing countries perceive the role of peer-reviewed literature in conservation science. *Oryx.*, **49**(01), 80-87.

Grantham, H.S., Bode, M., McDonald-Madden, E., Game, E.T., Knight, A.T. & Possingham, H.P. (2009). Effective conservation planning requires learning and adaptation. *Front. Ecol. Environ.*, **8**(8), 431-437.
Gurney, G.G., Cinner, J., Ban, N.C., et al. (2014). Poverty and protected areas: an evaluation of a marine integrated conservation and development project in Indonesia. Global Environ. Change, 26, 98-107.

He, G., Chen, X., Liu, W., et al. (2008). Distribution of economic benefits from ecotourism: a case study of Wolong Nature Reserve for Giant Pandas in China. Environ. Manage., 42(6), 1017-1025.

IIED. (2014) Social assessment of protected areas: a rapid approach to social assessment for conservation practitioners. International Institute for Environment & Development.

King, M.F., Renó, V.F. & Novo, E.M. (2014). The concept, dimensions and methods of assessment of human well-being within a socioecological context: a literature review. Social Indicat. Res, 116(3), 681-698.

Leisher, C., Samberg, L.H., Van Buekering, P. & Sanjayan, M. (2013). Focal areas for measuring the human well-being impacts of a conservation initiative. Sustainability, 5(3), 997-1010.

Mascia, M.B., Pailler, S., Thieme, M.L., et al. (2014). Commonalities and complementarities among approaches to conservation monitoring and evaluation. Biol. Conserv., 169, 258-267.

McConnachie, M.M., Romero, C., Ferraro, P.J. & van Wilgen, B.W. (2015). Improving credibility and transparency of conservation impact evaluations through the partial identification approach. Conserv. Biol. DOI:10.1111/cobi.12610.

Secretariat of the Convention on Biological Diversity. (1992). Convention on biological diversity. Montreal.

Silva, P. (2006). Exploring the linkages between poverty, marine protected area management, and the use of destructive fishing gear in Tanzania. World Bank Policy Research Working Paper 3831.

Stiglitz, J., Sen, A. & Fitoussi, J. (2009). Final report of the commission on the measurement of economic performance and social progress. Paris, France.

West, P., Igoe, J. & Brockington, D. (2006). Parks and peoples: The social impact of protected areas. Ann. Rev. Anthropol., 35, 251-277.

Wonghusarakum, S., Madeira, E.M. & Hartanto, H. (2014). Strengthening the social impacts of sustainable landscapes programs. Arlington, VA, The Nature Conservancy.

Woodhouse, E., Homewood, K.M., Beauchamp, E., Clements, T., McCabe, J.T., Wilkie, D. and Milner-Gulland, E.J. (2015). Guiding principles for evaluating the impacts of conservation interventions on human well-being. Phil. Trans. R. Soc. B. 370(1681), 20150103.

Woolcock, M. (2009). Toward a plurality of methods in project evaluation: a contextualised approach to understanding impact trajectories and efficacy. J. Dev. Effect., 1(1), 1-14.