Evaluating evaluation:
Exploring evaluation methods to assist WWF-UK programme management.

Robert Lindner

September 2012

A thesis submitted in partial fulfilment of the requirements for
the degree of Master of Science and the Diploma of Imperial College

DECLARATION OF OWN WORK

I declare that this thesis,

Evaluating evaluation: Exploring evaluation methods to assist WWF-UK programme management,

is entirely my own work and that where material could be construed as the work of others, it is fully cited and referenced, and/or with appropriate acknowledgement given.

Signature ……..

Name of student:

Robert Lindner

Name of Supervisors:

E.J. Milner-Gulland

Georgina Mace

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

CBAPO – Community Based Anti-poaching

CBD – Convention on Biological Diversity

CBO – Community Based Organisation

CCF – Cambridge Conservation Forum

CMP - Conservation Measures Partnership

COPY – Conservation Output Protection Year

CWT – Curbing Wildlife Trade in Nepal

IS – Impact Summary

KPI – Key Performance Indicators

PIPAL – People in Participatory Action for Life

ROI – Return on Investment

SCLL – Strengthening Conservation at the Landscape Level

SHL – Sacred Himalayan Landscape

TAL – Terai Arc Landscape

TPR –Technical Progress Report

TRA – Threat Reduction Assessment

TRN – Tigers and Rhinos in Nepal

WAZA – World Association of Zoos and Aquariums

WWF – World Wildlife Fund

ZMG – Zoo Measures Group

ABSTRACT

Conservation science must allocate its limited resources wisely. Even large international conservation organisations like WWF cannot conserve everything, they must decide what to conserve and how to do it. To conserve effectively and efficiently requires prioritisation, planning, implementation, monitoring, analysis and adaption, or a lot of luck. Despite the author's respect for luck, this thesis explores a suite of evaluation techniques for conservation projects; some can be used for prioritising and planning as well. The techniques are reviewed for strengths and challenges; applicability to WWF, and suggestions for use are provided. The Cambridge Conservation Forum Project Evaluation Tool is a lengthy and thorough, qualitative evaluation that challenges project managers to critically review all angles and stages of a project from multiple activity views, and can be helpful for planning. The Threat Reduction Analysis and Impact Summary are quicker reviews, based on threats and activities respectively, which provide convenient indices for comparison across projects. The Ranked Outcomes approach required collating outcomes and potential outcomes from a portfolio of projects, which are then ranked by a panel, and used for scoring each project. The approach maybe suited for comparing projects with different sites and objectives than similar projects like those in this thesis. Return on Investment was difficult to apply with limited investment data and limited measures of return, but the incorporation of the method promises objective evaluation of efficacy and efficiency, making it an excellent option for upper level management reviewing and prioritising for an entire organisation. The tools all present aspects of evaluation that could be beneficial to a conservation organisation, however the organisation has to decide what complements its current standards and what it can commit to.

ACKNOWLEDGEMENTS

Primary thanks go to my supervisors E.J. Milner-Gulland and Georgina Mace for providing excellent direction and support. I owe Will Beale and John Barker many thanks for taking time to introduce me to WWF and their standards, provide WWF programme information, and share feedback throughout the project. Thank you to Val Kapos and Caroline Howe for their guidance on the evaluation techniques they have developed. None of this would have been possible without my parents. And finally, thanks to everyone at Silwood for an excellent year.

1. INTRODUCTION

1.1 Conservation industry

Conservation receives much attention from science, media and public alike as it is widely recognised to be important for all species, including humans. However, other social, political and economic issues generally take priority on global scale. Fewer resources are therefore devoted to conservation compared to these other agendas. Nonetheless, the need for investment is great and many scientists agree it is growing greater (Cleary, 2006; James et al, 2001; Balmford et al, 2003). These scarce resources must be spent wisely, but there is often little agreement about how to do it. Previously, general beliefs held that any investment in a good cause was good, and more was better (Margoluis et al, 2009). However, success often requires more than "throwing money at the problem." Increasingly frequent questions from funders to conservation organisations about whether money is being used well led to a rise in attention applied to prioritisation and measures of effect, and many organisations have gone to great lengths to develop more sophisticated planning for their programmes (Halpern et al, 2006; Meyers et al, 2000). Unfortunately, less effort has been devoted to developing ways to assess if goals and objectives are being achieved (Halpern et al, 2006). Whether those priorities and resulting investments are actually affecting key factors (outcomes) and targets of conservation (impacts) remains a difficult question to answer, especially when fear of failure, harm to reputation and potentially reduced funding opportunities preclude sharing of less successful endeavours (Redford and Taber, 2000).

1.2 Evaluation paucity

Despite improvements in planning, the majority of attention is naturally still given to implementation of interventions, and evaluation of those systems remains the neglected step-child. This in not unexpected, as choosing what to do and doing it are exciting, and by a project's end the energetic and emotional investment leave little desire to critique the efforts. In addition, once the intervention has been made, it is easy to think the mission is accomplished and choose to spend the remaining precious time and funds on the next intervention, to look at the future instead of the past. Unfortunately, even in an altruistic endeavour like conservation,

good intentions do not guarantee good results. While monitoring and evaluation plans recognise the need to show results, even when they are not just an afterthought, they often fail to provide evidence of effective conservation (Saterson, 2004; Sutherland, 2004). Though efficacy and efficiency are important characteristics for success in any field, the fact that conservation resources are limited when problems are only growing larger and more complex strengthens the case for evaluation, as we cannot afford to spend more on activities that are not successful (Balmford, 2000; James, 1999; Saterson et al, 2004). Granted, if the dynamics of a situation are continually changing, retrospective evaluation may not be as telling for future conditions, as the future does not necessarily mimic the past (Ferraro and Pattanayak, 2006). However, if conservation science is going to live up to its name, projects should be evaluated in a scientific manner using robust methods and empirical evidence. To embrace developing industry standards, conservation organisations should ensure effective conservation through evaluation, which will additionally improve transparency and accountability, to ensure the conservation effort is doing the best it can (Salafsky et al, 2002; O'Neil, 2007; WWF, 2007).

1.3 Aims and objectives

This thesis developed from the desire of the WWF-UK Board of Trustees to explore new options for evaluating the success of the organisation's resource investments – exploring options for evaluating not only what is being conserved, but also how it is being conserved, either from quantitative data or expert opinion. In order to meet that aim, this thesis applies a suite of evaluation techniques to a WWF-UK programme, examining the strengths and weaknesses of current approaches in conservation science. Useful characteristics sought are as follows:

1) An analysis based on empirical data if possible, derived from observation or experimentation to reduce subjectivity.
2) It should require critical examination of planning, implementation, outcomes and impact; documentation of evidence for each of these; and linkage to reporting.
3) Costs should be included, expenditures should be analysed by activity, an estimate of what has been achieved for each investment is ideal, and identifying opportunity costs could be beneficial.

4) Impacts on biological resources, ecosystem function, and social components like attitude, behaviour, welfare and equity, should be examined using indicators derived from goals and outcomes explicitly defined during planning.
5) An evaluation should provide easily interpretable results. An index for comparing across projects and sites is helpful, but it must be meaningful.
6) Challenging as that is, the evaluation technique should balance those demands with usability and timeliness, using data easily obtained from reports.
7) Finally, the ideal evaluation should be robust, accurately representing reality in repeated testing and application.

While it is unlikely all of these characteristics will be found in a single evaluation, each method will likely contain a few. In the end, this thesis should provide examples of what makes an evaluation technique effective, as well as suggestions for refining monitoring and reporting in order to aid evaluation efforts.

1.4 Thesis structure

Following this introduction, more depth is provided on the background of conservation project evaluation as well as on WWF-UK and the Terai Arc Landscape, the programme location. The methods used in project selection and in each analysis are described as well as the results of each. Finally, the thesis finishes with a discussion of possibilities for WWF-UK reporting, the strengths and weaknesses of the analyses, and a brief summary of ideal characteristics of an evaluation method.

2. BACKGROUND

2.1 History of evaluation

Early natural resource conservation efforts developed from growing spiritual, aesthetic and economic appreciation for a target, which were then protected by the government, communities and cultural taboos (Jones et al, 2008). As conservation targets became increasingly imperilled, general recommendations for protection and a cautious approach were

made by default until more could be learned about the ecology of the target to better inform management (Walters, 1986). However, there was discord between research and management, as researchers were focused on finding the single best solution by gathering information and using it to make predictions, while managers wanted to know about the effects of their actions and were open to trying multiple approaches and learning from them (Ludwig, 1993; Walters, 1986). This divide gave birth to adaptive management and the idea that solutions may vary, but reviewing the work of others, experimenting with multiple actions and assessing their success would provide opportunities for learning through conservation action (Holling, 1987; Walters, 1986). Informing and adapting management with successes and failures relied on evaluation, however, evaluation itself was still not receiving much attention even in 1993, monitoring still stole the spotlight (Ludwig, 1993).

In 1990, Noss had published a seminal paper on the use and benefits of monitoring for biological indicators, fuelling the monitoring and learning movement. Others have since pointed out that using biological indicators can be challenging because there is often a lag between effects on the target and observable changes in biological indicators (Lee, 1999; Salafsky and Margoluis, 1999). Current recommendations are often to gather a mix of data, ecological and social, quantitative and qualitative (Brooks et al, 2006a; Ferraro and Pattanayak, 2006; Woodhill, 2000). Despite wide acceptance and support of adaptive management, review showed it to be more of a philosophical principle than an actual practice (Lee, 1999), and evaluation was no exception. While other professions like health care had integrated evaluation into their programmes' cores, conservation was generally still talking about it. All the talk had reduced the cohesion and increased the confusion around monitoring and evaluation, so organisations and collaboratives like the International Union for Conservation of Nature (IUCN) released guidelines to help unify concepts and terminology (Woodhill, 2000; Salafsky et al, 2009). They highlighted five components of projects that evaluation assesses: 1) Relevance, 2) Effectiveness, 3) Efficiency, 4) Impact, and 5) Sustainability (Woodhill, 2000). The Conservation Measures Partnership (CMP), a collaborative of conservation organisations and funders, joined the effort to develop agreed upon standards, including planning and evaluation

methods (O'Neil, 2007). As concepts, terminology and approaches became unified; conservation planning, monitoring and evaluation were reunited.

In 2005, Stem et al reviewed monitoring and evaluation and found that too many organisations were still attempting to develop their own systems without assessing lessons learned by others, they were still missing a key component of adaptive management. They also delineated four types of monitoring and evaluation: basic research, accounting and certification, status assessment and effectiveness management (Stem et al, 2005). Ferraro and Pattanayak (2006) pushed for the use of counter-factuals and natural experiments in evaluating effects of conservation activities, and while it should be encouraged, such opportunities are uncommon in conservation management. Even with highly sophisticated monitoring and evaluation designs, results may not be beneficial if objectives are not clearly defined (Stem, 2005; Yoccoz, 2001). Reporting often includes inputs meeting what had been budgeted, or outputs matching expectations, as measures of success, though they do not necessarily reflect actual change in the system (Ferraro and Pattanayak, 2006; Kapos et al, 2009; Saterson et al, 2004). Expected outcomes of activities, like observed change in indicators, developed from well-defined objectives are now being widely recognised as more appropriate indicators of project impact on the system in question (USEPA, 2006; Ferraro and Pattanayak, 2006; Kapos et al, 2009).

2.2 The power of money

Money is the ruler by which most things are measured. Even in conservation, where the focus is often on plants and animals, the influence of human currency can no longer be ignored. It is used by the companies that threaten targets and the governments that manage them, just as it is used by the organisations that try to conserve them and, with increasingly rare exceptions, the local communities that rely and identify with them. These stakeholders may not be able to agree on many things, but all would agree that money is valuable. Those with the most money can generally get what they want (Leech et al, 2007). If conservationists want to conserve in the modern global society, then they need to accept that money is a critical component of their effort. Conservation needs to learn from other industries that have embraced this common

denominator of society and have researched endlessly how to be successful, and how to use money to evaluate efficacy and efficiency. Let us stand on the shoulders of commercial and financial giants, take their lessons learned and apply them to our cause. This has begun to some degree: conservation organisations are some of the most elegant marketers and fund raisers around, but that is only the first step. Using those funds more effectively and efficiently, and being able to show evidence of it through measurable outcomes will bring more conservation, more confidence and satisfaction, and more funding, which will feed the cycle. Efficacy, efficiency, accountability and sustainability are characteristics that we could learn about from the business industry and strive for in conservation.

Pillars of the for-profit industry actually helped usher this approach into conservation. Funders with names like Moore, Hewlett and Packard applied the same principals to their investments in conservation as they did in their own lucrative companies (Christensen, 2003). Conservation organisations and researchers have taken heed, pounding fists about monitoring (Brooks et al, 2006a), covering the world with spots (Balmford et al, 2000; Brooks et al, 2006b), and even auditing their own projects (Kleiman et al, 2000; O'Neil, 2007). Costs, ever an issue, have slowly begun to be integrated into prioritisation and evaluation.

Economists have advocated for costs analyses to be used in conservation too (Hanley and Spash, 1993). The US Government began applying cost-benefit analysis (CBA) to natural resource management by the beginning of the 19th Century, and it has now become the standard for evaluating effects of policy on the environment (Hanley and Spash, 1993). Hughey et al (2003) provided excellent definitions and examples of the three primary approaches to cost analysis. In addition to CBA, they also illustrated the differences of cost-effectiveness analysis (CEA) and cost-utility analysis (CUA). One of the key differences highlighted are the analysis units: CBA evaluates returns in monetary units, CEA evaluates returns in nonmonetary units, and CUA evaluates returns of activities by the worth of the status change to the target (Cullen et al, 2001; Hughey et al, 2003). CBA is appealing because comparison of costs and

returns is possible regardless of context once everything is given a monetary value. However, valuing all returns in a project is a daunting task, and could prove impossible.

Upon initial examination, CUA is possibly the most appealing approach given healthcare's success with such a system. However, while healthcare's measure of increased life length and quality is easily captured by units of Quality Adjusted Life Years (QALY) and agreed upon by society, it is more challenging to develop a unit that is equally representative across species, like a fly and a tree, or even across entities: say a seahorse and a dry tropical forest. Cullen (2001) developed the Conservation Output Protection Year (COPY) as a comparable unit for CUA, however, it is dependent on the ratings of species status with and without management. Similar to Ferraro and Pattanayak's (2006) suggestion, scenarios that provide opportunity for direct comparison are not common in conservation, unlike healthcare. Ratings would generally depend on individual opinion, which could introduce bias. In addition COPY is limited to species management programmes, and does not account for economic and social effects.

Taking a different approach, CEA eliminates the challenge of developing a single metric and uses specific and meaningful metrics for each target. Interest in CEA received a huge boost from Australian researchers working primarily on prioritisation (Carwardine et al, 2012; Evans et al, 2011; Wilson et al, 2006, 2007). Laycock et al (2009; 2010) have been using CEA to evaluate efficiency and effectiveness, and provide excellent examples of how CEA can be used to evaluate programmes. Hockley (2010) makes a strong argument that CEA is not as easily understood by the rest of society, as measures of ecological outcomes are not directly comparable like money. However, the value of those outcomes can be decided by the stakeholders who have previously made those value judgements in conservation, but with CEA their decisions will be bolstered with investment data.

The UK Government has also begun to embrace valuing nature, evident in DFID's request for Value for Money reports from organisations it funds (DFID, 2011). While the ideas behind Value for Money are still being developed, organisations like the New Economics Foundation

(NEF)and WWF are leading the effort to explore the realities of the guidance and how to incorporate it into practice (Lawrence, 2012; NEF, 2010). These organisations have outlined four principles that reflect a recurring theme throughout conservation: the quality of inputs, the efficiency of converting inputs to outputs, how well intervention outputs are achieving desired outcomes, and how much impact interventions achieve (Lawrence, 2012; DFID, 2011). As Lawrence (2012) states, "Value for Money is not just about cutting costs. It is about being really clear what outputs, outcomes and impacts we can really expect from an intervention." Setting explicit and measurable goals is a common theme throughout much of the discourse on conservation evaluation and conservation success. Robust planning precipitates robust evaluation of success, whether it is ecological, economical, attitudinal, or behavioural (Brooks et al, 2006). Although it is not only about money, money is an issue.

2.3 Evaluation methods

2.3a **Return on Investment**

ROI is designed to link spending with objective outcomes, and result in efficient use of resources. A crossover from economics, it falls under CEA as defined by Hughey et al (2003). In conservation, it has mostly been used as a prioritisation strategy, but has been tested as an evaluation analysis a few times (). ROI requires two primary inputs: activity spending as the investment and ecological outcomes as the return. In this thesis, the activity spending of WWF projects Curbing Wildlife Trade (CWT) and Tigers and Rhinos Nepal (TRN) is analysed using rhino population and rhino poaching as returns.

2.3b **Threat Reduction Assessment**

The Foundations of Success developed the TRA approach as a quick and easy analysis for projects that could be used for comparison with projects of varying focus and context (Salafsky, 1999). Threat levels are used as the primary indicator of success. For this exploration, the TRA assesses the entire Terai Arc Landscape.

2.3c **Cambridge Conservation Forum Project Evaluation Tool**

The Cambridge Conservation Forum (CCF) Tool was designed on the premise that outcomes, instead of the often used inputs and outputs, are better indicators of project success and impact (Kapos et al, 2008). It was built upon the established techniques of results chains and conceptual frameworks. The tool is primarily a series of questionnaires: one on *General Info*, one on *Threats*, and then seven specific to categories of intervention (*Species Management*, *Site Management*, *Livelihoods*, *Policy & Legislation*, *Education & Awareness*, *Capacity Building* and *Research*), and finally one on *Overall Impact*. The CWT and TRN projects are evaluated with the CCF Tool for this thesis.

2.3d **World Association of Zoos and Aquariums (WAZA) Conservation Impact**

The WAZA Project Conservation Impact Summary Form is an adaptation of the Zoo Measures Group Impact SummaryAssessment analysis designed to provide a quick but robust and comparable method for zoos to evaluate their in-situ and ex-situ conservation work (Mace et al, 2007; WAZA, 2012). The analysis assesses success based on qualitative and quantitative outputs and outcomes of the project. It evaluates projects by six categories of intervention targets: *Education, Capacity Building, Communities, Research*, *Species* and *Habitat*, by three measures: *Importance, Volume* and *Effect*. CWT and TRN in TAL were evaluated.

2.3e **Ranked Outcomes**

Howe and Milner-Gulland (2011) developed the Ranked Outcomes approach as an alternative index for evaluating conservation projects by qualitative outcomes when quantitative information may not be sufficient for evaluation. Outcomes used are extracted from the TAL portfolio of projects - CWT, Strengthening Conservation at the Landscape Level (SCLL) and TRN.

2.4 WWF Standards

WWF standards have changed with time, generally following the expanding frontier of conservation. As one of the largest conservation organisations, it is required to continually reinvent and improve its methods, and it has consulted and collaborated with other organisations and experts. Reporting standards for such a large organisation can be

challenging. While WWF used to be a species-focused conservation organisation, it now takes a more holistic approach, incorporating social, economic and political targets into its work (WWF, 2011a). This diversity of programme targets and the respective myriad interventions makes devising one standard difficult. While documenting information can be important, not all information is equally valuable (Saterson et al, 2004). Because resources are limited, little should be wasted on monitoring and evaluation (Salver and Salafsky, 2006; Sheil, 2002). Therefore, an effective and selective monitoring system that allows for efficient evaluation, both during and after projects, and allows for adaptive management and shared lessons is crucial to a successful conservation programme (Margoluis et al, 2009). WWF-UK is making some major changes to this effect by 1) refining their standards, 2) developing a new organisation-wide database, and 3) exploring evaluation options through support of this thesis. In order to evaluate WWF needs most accurately, the techniques will be tested with the WWF-Nepal programme in the Terai Arc.

2.5 Terai Arc Landscape

The Terai Arc Landscape is located in the Eastern Himalayas of Nepal and India. It is best known for its savannas and other grasslands, but also contains moist deciduous, subtropical broadleaf and subtropical pine forests (WWF, 2011a). The varied landscape has provided a home to tigers, Asian elephants and one-horned rhinos; contains some of the highest densities; and is one of the last areas where they co-exist (WWF, 2011b). There are an estimated 99 species of mammals (39 of which are endangered), 550 species of birds (including 50 endangered species), 47 species of reptiles and amphibians, and 122 species of fish, many of which are endemic to the area (WWF, 2011a). An estimated 342 endemic plant species and 160 endemic animal species have been reported in Nepal, a country of only 147,181 km² (Nepal MFSC, 2009). This is not surprising when considering that it sits at the intersection of six floristic regions, ranging from 67 - 8,848 m, including the world's deepest gorge and its highest mountains (Nepal MFSC, 2009). The latest physiographic data shows that Nepal harbours 29% forest area, 10.6% shrubland and degraded forest, 12% grassland, 21% farmland, 2.6% water body, 7% uncultivated inclusions, and 17.8% other landscape types (Nepal MFSC, 2009). Nepal has

recognised its rich and unique biodiversity, and has established protected areas over more than 20% of the country (Nepal MFSC, 2009). However, the threats to biodiversity are also alarming. Recently pristine, TAL has changed much since the 1950's when the government began its campaign to eradicate malaria in the area (WWF, 2011a). The success of the campaign led to a immense human immigration which was followed by rapid deforestation and habitat destruction as people settled and developed (WWF, 2011a). It has since become one of the most productive agricultural areas in the region (WWF, 2011b). Nepal is a signer of the Convention on Biological Diversity (CBD), which was ratified in 1993 (Nepal MFSC, 2009), however it is also a politically unstable country with 26.6 million people representing more than 100 ethnic groups (UNDP, 2010). 25.4% of the Nepali people live in poverty and 85% of the rural population depends directly on natural resources (UNDP, 2010). Despite the many challenges, Nepal has been able to fulfil its obligations to the CBD (Nepal MFSC, 2009).

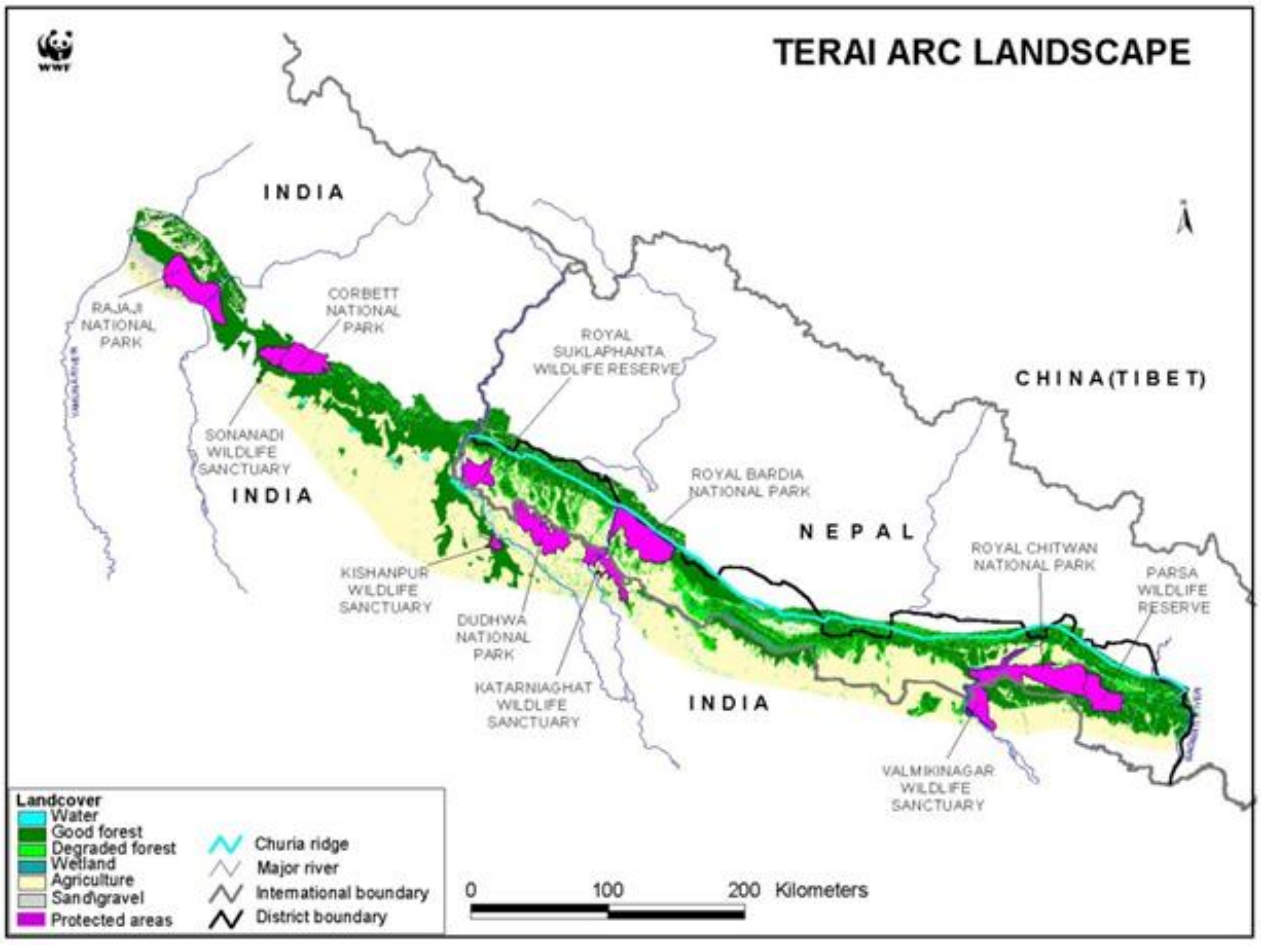

Figure 2.1- Map of Terai Arc Landscape (WWF).

3. METHODS

3.1 Project selection

Before evaluations could begin, a candidate programme had to be selected. Meetings with WWF-UK staff were held to share and discuss concepts and goal, and to outline a plan and project design. Each organisation is unique in how they function and the first task was to learn about WWF's standards including reporting. Visits to the WWF-UK Panda House were made for introductions to their organisational structure and standards, supporting documents were supplied for review, and thesis details were further developed. The discussion then turned to programme selection. The primary goal was to identify programmes that had data reported over 5-10 years, which included both ecological monitoring and activity spending. Ideally the programme area would have limited intervention by organisations outside the project, and access to managers and decision rationale would be possible. Programme selection and report acquisition required the help of WWF-UK staff members from the Design & Impact staff and Species Managers. Upon programme selection, reporting gaps were assessed and highlighted; then data was extracted and collated to facilitate analysis.

3.2 Return on Investment

In order to analyse ROI, both the investment and return must be quantified. WWF reports were searched for activity spending, generally found in the Workplans. After this data had been extracted, spending trends were identified and categories were developed for data collation. The reports were also examined for outcomes suitable for use as the return. Ideally this outcome would be a measure of change in a primary target assessed throughout the period the activity spending is available. The investment and return variables would then be plotted in a graph as the independent and dependent variables, respectively. The relationship between the two variables was also expressed in terms of return/investment units and investment/return units and explored for every year and over the entire period of investment. The investment and return were further examined by individual activities, again by plotting and

calculating ratios for each activity. Additionally, correlation between the return and activity investment was assessed through calculation of Spearman's Rank Correlation Coefficient.

3.3 Threat Reduction Assessment

The TRA approach requires one to describe the project site, identify threats to it, what 100% reduction in those threats would mean, and how close the project is to 100% reduction. The entire TAL programme could be assessed with this approach because the TRA approach does not require activity investment data. A TAL description and map were extracted from the reports as well as threats specific to the project targets in the landscape. Each threat was marked on a scale of 1-10 (10 highest) for the proportion of the landscape affect (*Area*), the severity (*Intensity*) and the immediacy (*Urgency*) of the threat (Margoluis and Salafsky, 2001); and this *Ranking* provides a weighting system for the threats. Explicit definitions of 100% reduction of each threat were created and then the actual amount each threat has actually been reduced was estimated. Like Anthony (2008), the TRA approach was modified to account for increases in threat by allowing a negative *Reduction Percentage*. The product of the sum of the *Rankings* (*Area, Intensity* and *Urgency*) for each threat and the corresponding *Reduction Percentage* provides a *Raw Score*. The sum of the *Raw Scores* is then divided by the *Rankings* total to provide a percentage referred to as the *TRA Index*, which signifies the amount of progress the project has made in reducing the threats to the targets.

3.4 Cambridge Conservation Forum Project Evaluation Tool

The entire TAL programme was also selected for evaluation with the CCF Tool. The order of the sheets in the CCF Tool was followed, skipping any activity category not included in the project, which in this case was *Species Management*. Beginning with the *General Info* and *Threats* sheets is particularly important to the assessment process. Most sheets begin with *Background* questions about the subject and follow with questions about *Implementation*, *Outcomes* and *Impact*. For many of the questions, a qualitative set of answer options are provided and project specific meaning and evidence are requested. After filling out the spreadsheets, they were shared with WWF-UK staff for verification of accuracy and feedback.

3.5 WAZA Conservation Impact Summary

This evaluation begins with questions of background information and follows with a request for objectives. Project objectives are then entered, and the evaluator is required to mark whether each objective has been achieved, partly achieved or not achieved; to quantify achievements; and to declare which type of conservation target the objective covers. Next, the categories of conservation targets included are scored based on the three measures, and the product of the scores provides a measure of impact. The average of the category impact scores represents the overall impact of the project. An opportunity for an external reviewer to evaluate is provided as well. The average of the overall impact from the two reviewers provides the final measure of the project's impact.

3.6 Ranked Outcomes

In this approach, a mix of outputs, outcomes and potential outcomes (all referred to as outcomes) were extracted from the WWF Eastern Himalayan reports based on an example provided by Howe (personal communication). These outcomes were grouped into four intervention categories (*Awareness & Education*, *Planning & Research*, *Infrastructure* and *Species & Habitat*), one *Legacy* category and one for *Negative Outcomes*. A panel of experts was asked to rank the outcomes within each category in the order of perceived contribution to conservation success. Agreement among the panel members was assessed by calculating the kappa coefficient. If an outcome did not receive a rank from a panel member, an average of the other panel members was given. Summary ranks were then used to score the projects giving the respective points for each outcome achieved by the programme. The total score for each project could then be used as a measure of success and for comparison with other projects analysed by the same set of outcomes.

4. RESULTS

4.1 Project reporting

Despite WWF-UK involvement in many programmes, the available reporting was limited; many programmes either lacked sufficient history or activity spending. Three general documents styles were selected from the general reporting, a *Workplan* that includes categorical spending data, an *R3* that includes budget and expenditure data, and an end of the year *Technical Progress Report* (TPR) that provides a summary of inputs, activities, outputs and outcomes. Latter year *TPRs* provided table summaries of the quantitative data in the *Key Performance Indicator* analysis. The WWF-Nepal programme was selected for analysis due to its 10 year programme history, the availability of reports and detailed ecological and spend data, as well as its expressed desire to improve its operation. Reporting was available for the three current projects: Curbing Wildlife Trade in Nepal (CWT), Strengthening Conservation at Landscape Level (SCLL), and Tigers and Rhinos in Nepal (TRN), which began in 2006, 2006 and 2010 respectively. The availability of the reports is documented in Appendix A. The reports generally followed the themes described above, however, some had different names and styles. Occasionally, attachments and appendices were mentioned but not included.

Curbing Wildlife Trade reports provided excellent coverage of desired data, especially the activity data in the consistently available Workplans. However the TPRs provided more information on inputs and outputs than outcomes. The most valuable ecological outcomes reported consistently were rhino population estimates (every three years) and rhino poaching estimates (every year). Workplans and the activity investment data were not available for SCLL, with the exception of the activity investment included in the 2006 TPR. This project had the largest budget and broadest range of targets, encompassing many of the activities covered by CWT and TRN, with additional investment in livelihoods, watersheds and climate change mitigation. Annual activity investment of the projects is available in Appendix B and summarised in Figure 4.1.

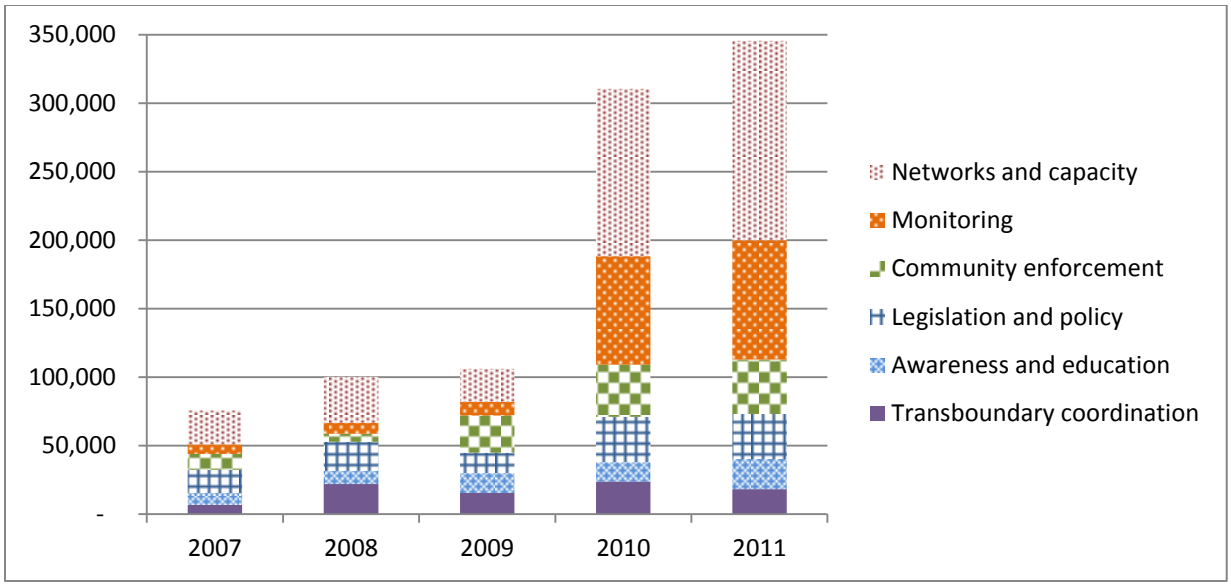

Figure 4.1 – Curbing Wildlife Trade (2007-2011) and Tigers and Rhinos Nepal (2010-2011) activity spending (USD) across years.

4.2 Rhino return

In the case of the Nepal Programme, activity spending was available for the wildlife trade focused projects CWT and TRN, and ecological data was available for Rhino population and poaching, which provided an opportunity for ROI. Investment was explored over the five year period which the data was available (2007-2011) by total and activity category and returns of rhino population change and poaching change were evaluated by each investment type respectively. Poaching data was available for each year and therefore was evaluated by total and categorical investment for each year.

Despite the limited ecological output data, the rhino population estimates and poaching numbers provided an opportunity to explore ROI and will provide more insight as data continues to be collected over time. A summary of the data used in ROI can be found in Appendix C. Only two population estimates were made during the period when investment data was available (Figure 4.2), so the insight ROI can provide at the moment is limited, but the example provides a clear demonstration of the key with the change in rhino numbers between the 2008 baseline and the 2011 re-count used as the return over the same period (Table 4.1).

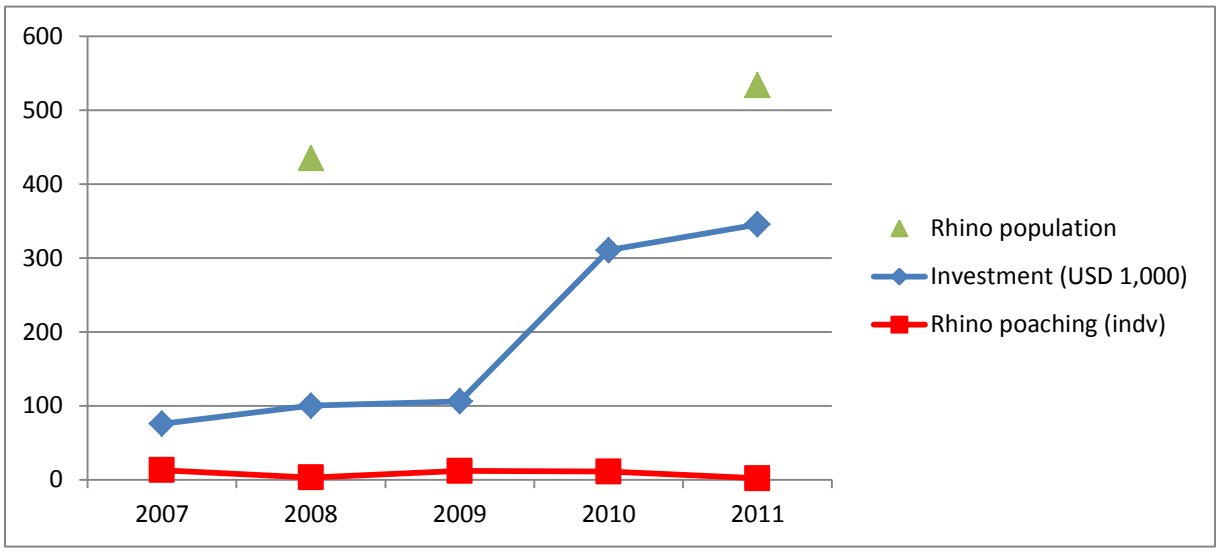

Figure 4.2 – Rhino and investment (CWT & TRN) trends.

Table 4.1 – Rhino population ROI (CWT & TRN).

Years	Investment (USD1,000)	ΔRhino population	ROI (rhino/USD1,000)	Unit cost (USD1,000/rhino)
2009-2011	762.058	99	0.13	7.698

Another ROI was performed with the more frequent poaching data to provide a more extensive example of the method, and more insight into the rhino situation (Table 4.2). Here the difference between the previous year's reported poaching and the current year's reported poaching have been evaluated by the current year's investment.

Table 4.2 – Rhino poaching ROI (CWT & TRN).

Year	2007	2008	2009	2010	2011	2007-2011
Cost (Total)	75.63	100.23	106.13	310.53	345.40	937.92
ΔRhino poaching	-12.00	-10.00	9.00	-1.00	-9.00	-80.00
ROI (Δpoach/USD1,000)	-0.16	-0.10	0.08	-0.01	-0.03	-0.09
Unit cost (USD1,000/Δpoach)	-6.30	-10.02	11.79	-310.53	-38.38	-11.72

ROI was established for individual investment categories based upon rhino population change and poaching change as well (Appendix D). However, the significance of these figures is questionable as investment data is incomplete and there is likely to be a lag between investment and return. So, trends were plotted (Figure 4.3) and Spearman's rank correlation coefficients were calculated to explore dependency between each investment category and poaching

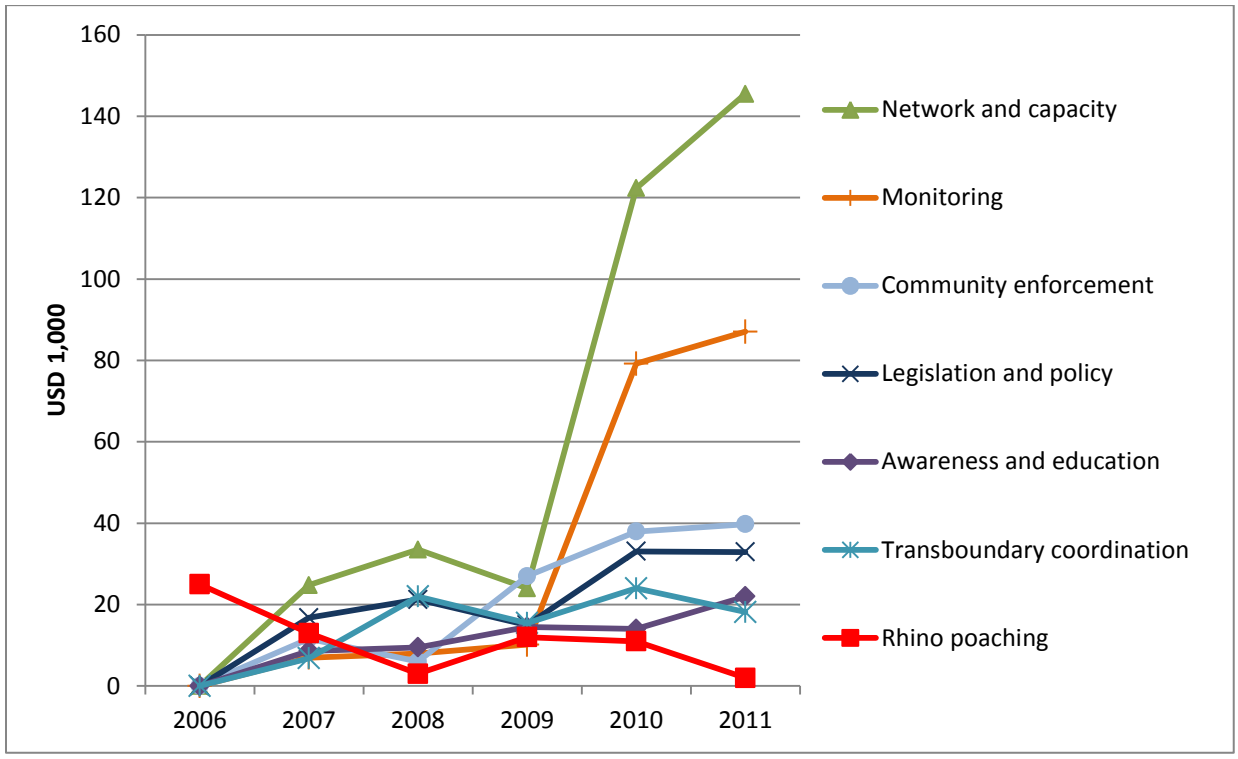

Figure 4.3 – Trends in rhino poaching and investment (CWT & TRN) by category.

The trends suggest some slight correlation between poaching and the categories. *Networks and Capacity* and *Community Enforcement* were expected to be the categories to have the most immediate effect on poaching numbers and therefore a negative correlation. More delayed effects were expected from investments in categories like *Transboundary Coordination* and *Legislation and Policy*. However, the Spearman's rank correlation suggested no significant correlation between poaching and the investment categories.

4.3 TAL threats assessed

The TRA approach began with a description of the Terai Arc Landscape (TAL) and the project period assessed, and then it quickly moves to the focus: threats (Table 4.3). Threats to the landscape were gathered from reports and then ranked based on the criteria (*Area*, *Intensity*, *Urgency*) resulting in *Conversion to agriculture* and *Unsustainable livestock populations & poor management* receiving the most severe rankings. Assessing how much a threat had changed since project implementation also required interpretation of report data, which is challenging if one is not practised in the task or familiar with the project. This translation of reported information to the TRA approach took more time and effort than it would if a project member was using the tool. Interestingly, *Wildlife killing/poaching* received a 90% reduction entry and *Unsustainable livestock populations & poor management* a -10% entry signifying an increase in the threat. The rankings and reductions combined for an overall threat reduction on the TAL of 18%.

Table 4.3 - The TRA approach results showing how *criteria rankings* are combined to provide threat *total rankings.* *Raw scores* are a reduction of the *total rankings*, and the *TRA Index* represents how much the site threats overall have been reduced..

	Criteria Rankings			Total Ranking	Threat Reduced	Raw Score
Threats	Area	Intensity	Urgency			
Conversion to agriculture	7	5	6	18	10%	1.80
Unsustainable livestock populations & poor management (grazing and fodder)	8	1	7	16	-10%	-1.60
Unsustainable timber management (market)	3	4	4	11	25%	2.75
Unsustainable extraction of forest resources (fuel)	5	2	8	15	25%	3.75
Wildlife killing/poaching	2	8	3	13	90%	11.70
Human wildlife conflict	1	7	2	10	0%	0.00
Expanding infrastructure development	4	6	5	15	10%	1.50
Forest fires (human induced)	6	3	1	10	0%	0.00
Total	**36**	**36**	**36**	**108**		**19.90**
TRA Index Formula	**Total Raw Score**			**Total Ranking**		**TRA Index**
	19.90			**108**		**18%**

4.4 TAL conservation outcomes examined

Apart from the combination and collation of investment activity necessary for ROI, the CCF Tool took the longest of evaluation methods. After general information had been documented, *a*

priori scoring of activity and project success is requested. *Policy and legislation* and *capacity building* received the highest estimates among the types, 4/5 each, while the projects received a 3/5. Similar to the TRA method, scoring was done with limited exposure to the technique or project. Threats had been outlined in the general info, but were broken down into a matrix based on mechanism, route and source. Examination of each activity type consumed the bulk of the effort with 30-55 questions listed for each type, many requiring evidence and project specific meaning in addition to answers. These questions examine the outputs, outcomes and impacts for each activity type and suggest that S*ite management* and *Capacity building* actually contributed the most to overall impact. In the case of WWF work in TAL, the tool highlighted these outcomes:

- target declines had slowed relative to expected without the projects and that was largely due to the projects;
- improvements are likely to persist with similar inputs;
- project successes have been documented and shared, but failures have not been actively shared; and
- human well-being has been positively affected.

4.5 TAL impact summarised

The WAZA Impact Summary (IS) has been developed from an early design by the Zoo Measures Group into an elegant Adobe document, into which data can be entered and a saved. The document adds pages as the evaluation progresses, forcing the evaluator to focus on one section at a time. General project information was required, and it is the first method to request information on focal species and their status; the CCF Tool asks for species only if they were being directly managed. The IS and CCF also bring investment into perspective, although, only the basic amount funded. The body of the technique is introduced, providing opportunity for six objectives. Once the objectives have been entered, a new section is added for analysing each in regard to completion, achievements, and activity type. Another section is then added to the document for scoring each of the listed activity types based on a range of options for

each criterion (*Influence*, *Number* and *Effect*). Finally, once the scoring has been completed by a project coordinator and a reviewer, the scores are automatically summarised and an *Conservation impact score* is provided (Table 4.6a). The *impact score* can be used for comparison across projects.

Table 4.5 - WAZA IS scoring comparing member and reviewer scores by for each activity. Final conservation impact score is calculated by averaging all activity scores.

Activity	Scorer	Importance	Volume	Effect	Activity Total
Education	Member				0
	Reviewer	3	4	4	48
Capacity	Member				0
building	Reviewer	3	4	4	48
Communities	Member				0
	Reviewer	3	3	4	36
Research	Member				0
	Reviewer	3	2	2	12
Species	Member				0
	Reviewer				0
Habitat	Member				0
	Reviewer				0
Project Total	Member				0
	Reviewer				144

Categories completed	
Member	
Reviewer	4
Conservation impact score	
Member	0
Reviewer	36
Final conservation impact score	
18	

4.6 TAL conservation outcomes ranked

There were 6 - 12 outcomes in each of the six categories, and 3 panel members. Overall agreement was low for each category. The panel members agreed the most on how *Legacy* outcomes contribute to success, but the overall agreement was only fair with a kappa value of 0.33. Due to the great variety in ranks, the overall rank used for scoring was decided by averaging the ranks each outcome received from panel members, and the final outcome rank

was determined by the order of the averages. If there were any ties, the ranks of the more senior panel members were weighted more heavily. Scoring the projects with the outcome rankings yielded practically the same result for each project, as they are all part of the same programme with similar objectives, so nearly all outcomes are attributed to each project.

5. DISCUSSION

5.1 WWF reporting

WWF's reporting standards have evolved throughout the years, which has inevitably led to inconsistencies. Changes in standards are an unavoidable necessity to continue improvement and adaptation of the organisation. Beyond changes in standards and protocol, inconsistencies in data can arise due to recorder bias as recorders for different projects may have different styles and interpretations of the standards. This can even occur within projects when they are longer. These inconsistencies primarily occurred in the presentation and type of data reported, although sometimes data was found to conflict amongst reports. Providing detailed instructions on, and examples of, ideal reporting could reduce the variation caused by recorder bias. Submitted reports should be reviewed and checked for accuracy and appropriateness to improve consistency and catch errors.

While the available reports were generally the *Workplan, R3* and *TPR*, sometimes reports of similar style but other names were substituted in a particular year. This should be avoided if possible, to streamline reporting and reduce confusion. The *Workplan* was integral to identifying how investments were made across activities. The *R3* lacked those details, though it did provide data on actual overall spending versus planned spending. The *TPR* contained the most qualitative data, and was the most subjective of the reports and therefore the most susceptible to recorder bias. Guidelines are likely provided, however the *TPR* is a free flowing word document, and the format and data reported varied across years and reporters for individual projects, which increased the difficulty of consolidating pertinent information. Nonetheless, the *TPR* is the report used for summarising outputs and outcomes and therefore can provide a wealth of information. It is important not to limit what recorders can report with

too strict of guidelines. However, more consistent structure across projects would be helpful, and providing detailed guidance and examples is recommended. Later reports began including tables of summarised quantitative data in the form of Key Performance Indicators, and this should be continued. Also, including logical frameworks, result chains, and/or goals and expected outcomes could be beneficial to remind project members each year what the project focus is, as well as provide an opportunity to update them as necessary in the spirit of adaptive management.

Accessibility of reporting data can also be challenging. To improve information sharing, WWF developed a network-wide database for searching reports. Unfortunately, some gaps in reporting persist. When exploring the Terai Arc Landscape programme, *Workplan*-like reports that contain activity investment were unavailable for the Strengthening Landscape Level Conservation (SLCC) project and therefore, ROI on rhino conservation cannot include the full investment by WWF-Nepal. In addition, reports sometimes referenced attachments and appendices that were not actually included. Placing the supplementary data in the document near the end could prevent this supporting information from being lost. Proposals provided helpful summaries of information, and would be beneficial if more widely available. WWF is developing a new database with improved usability, which will likely improve the amount of information successfully shared.

5.2 Return of ROI

Of the evaluation methods, ROI was the most quantitative. The majority of the time was invested in collating the spending data and extracting outcome indicators from reports. Once the data was assembled the calculations were simple. ROI shows great promise as it is based on quantitative, empirical data and provides clear and graphically presentable results. However, this example does not to show the analysis's full potential. ROI benefits from long-term data sets, which can be challenging when assessing ecological indicators that are slow to change like rhino population growth. Another rhino count will likely be made in 2014, and that

count will provide a better understanding between the relationship of rhino conservation investment and the rhino population.

Until another rhino count is available, one could use the current data to explore possible future trends. Assuming the 2011 investment level in CWT and the projected budget for TRN are maintained through 2014, the 2009-2011 ROI of 0.13 rhinos/USD1,000 could be extrapolated to estimate a 2014 rhino population of 706 individuals (Figure 6.1). It is unlikely the same return can be expected in the next three years due to project factors like investment distribution and ecological factors like changing inter- and intra-specific competition and long reproductive intervals. However, this is interesting if examined with the previous rhino recovery and project goals in mind. The population growth in 2009-2011 was nearly all in Chitwan National Park (CNP) (406 -> 503), and the programme goal is to maintain >500 rhinos in CNP and establish a population of >100 in Bardia National Park (BNP). This data suggests that even if numerous rhinos were translocated from CNP to BNP and Shuklaphanta National Park, projected investment levels should be more than sufficient to provide the same levels of conservation that allowed the CNP population to grow by nearly 100 individuals in the previous period. This should be interesting to project members and organisation managers.

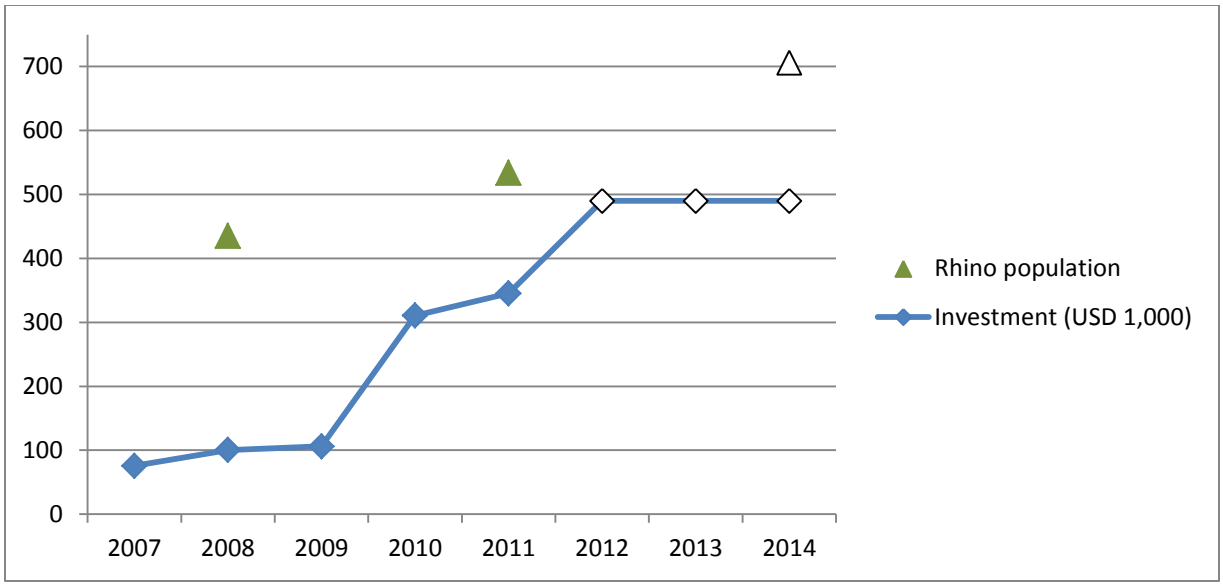

Figure 5.1 - Rhino and investment trends projected (projected estimates represented by hollow points). 2012-2014 investments estimated using CWT FY11 investment (USD 108,240) and TRN 2012-2014 budgeted investments (USD 381,600). Rhino population estimate product of 2009-2011 ROI (0.13) and estimated 2012-2014 total investment.

What may be most interesting is using ROI to analyse activity investment contribution to success. Conservation projects sometimes take a scattergun approach trying to impact targets from all aspects, in the words of Walters (1986), exploratory investment. While a holistic approach may be beneficial, not all activities are likely to contribute equally to success. Managing adaptively will facilitate informed reallocation of resources to increase impact, and ROI could assist the process by evaluating the contributions of activities to conservation success against each other. With monitoring data that allows evaluation of the intermediate outcomes of each activity, effectiveness of each investment could then be compared. This assessment of investment effectiveness would provide project managers with a tool for assessing how efforts are contributing to conservation priorities and organisation managers a tool for how efforts are contributing to conservation priorities.

5.3 Focusing on FOS's Threats

The Threat Reduction Assessment (TRA) functions as a site assessment more than a project assessment. It examines what changes have occurred in the severity of the threat to the site

and its conservation targets, and is less concerned whether those changes are a result of project outcomes and impact. Despite this limited perspective, the categories for assessing threat severity (*Area*, *Intensity* and *Urgency*) require critical review of the threat's characteristics for the analysis to be effective. This is beneficial to project members, because it challenges the group members to re-examine their beliefs about the threats, but an earnest assessment is required. Otherwise biased or inaccurate rankings and reductions will render the TRA Index useless and misleading. This tool could be used in planning a project as well, forcing project members to identify threats and to define and document what elimination of those threats means. In addition, it is a relatively quick assessment and it produces an index that can be used for comparison across projects and programmes regardless of context.

5.4 CCF summary

There is no doubt that the CCF Conservation Project Evaluation Tool is one of the most demanding evaluation methods tested. In an industry where time is limited, this can discourage adoption of the tool. Tool designers suggest it generally takes about a day to complete for each project, however, if project members are available, this time can be reduced (Kapos personal communication). In reality, one day is not very long, especially if the result is such a comprehensive evaluation like the CCF Tool. In this test, due to unfamiliarity with either the programme or the evaluation tool, the cumulative time required to find all of the necessary information and complete the evaluation took multiple days.

The CCF Tool does not produce an index for comparison, although one could be produced by repeating the a priori evaluation of project components again at the end of the evaluation and combining the component scores for a project or programme total similar to that of the WAZA Impact Summary (IS) (Appendix E). However, this score is derived more from the evaluator's impression of project success based on the tool, and less on quantitative data like the WAZA IS, so its significance might be more subjective. Another opportunity for an index lies in the attention the tool gives to threats. The TRA approach could be incorporated into the CCF Tool, complimenting and expanding the evaluation of threats. Nonetheless, the CCF Tool provides

much more than an index could. It provides a cohesive, guided format to identifying, documenting and evaluating the inputs, outputs, outcomes and impacts of a project by each activity type. This is where the tool's strength lies, in forcing project members to think critically about all aspects of the project. It builds upon earlier planning efforts, easily adapting the products of logical frameworks and result chains and providing a logical next step and additional value to those previous exercises. CCF provides excellent definitions that can be used for clarity and reference in an industry where jargon is frequently used frivolously. Key definitions include those of conservation, success, input, activity, output, and impact, as well as activity categories (Appendix F).

The CCF Tool provides a logical guide to examining projects, but could also provide assistance in planning by requiring identification and explicit, detailed documentation of targets, threats, goals, and desired outcomes. It also provides an excellent outline of considerations that should be reviewed before planning is completed and implementation begins. Use of the CCF tool could help development of a monitoring plan by highlighting what information is actually needed to evaluate success. It could also provide excellent, concise guidelines for what to include in reports like the *TPR*. Monitoring and reporting designed with a tool like this in mind would likely increase the ease of evaluation with the tool, and reduce the time needed to complete it.

5.5 Wazzup with WAZA?

The World Association of Zoos and Aquariums (WAZA) improved the appearance of the Impact Assessment suggested by the Zoo Measures Group (ZMG) by providing it in a user-friendly Adobe PDF form, to which information can be added and saved. This saves the user from having to build their own document in a spreadsheet, and provides a more easily understandable and elegant format for new users. However, WAZA did not heed the suggestion from ZMG to vary the scales of the measures to each project in order to make them more appropriate (Mace et al, 2007). For instance, a target group for *Education* may be a mix of the provided categories, and the new locked format makes it difficult to account for this.

Also, this new format limits the user's ability to customise other aspects. The user is limited to six objectives, and although more than six objectives for a project may be considered too complicated, the reality is conservation projects are becoming increasingly more complex as the threats become more complicated and development is increasingly incorporated into the work (Salafsky and Margoluis, 1999). The use of *Dissemination of research* as the volume measure may slightly bias the scoring to organisations that are more research focused than those doing research to primarily assist their own management. Offering an opportunity to weight categories could improve the accuracy of the WAZA Impact Summary (IS); surely not all activities receive the same amount of attention or carry the same import to a project. Finally, the WAZA IS does not provide an option for evaluating policy work as one of the *Project Types*, though policy work is a common and important component to many conservation projects.

However, the WAZA IS was developed to assist zoos evaluate their contributions to conservation *in situ* and *ex situ*, and the assessment has not yet been modified with all conservation projects in mind. It provides a quick and easy format for evaluation based on quantitative data and requires documentation of evidence to support scoring decisions. With some adjustments, such as those mentioned above, the WAZA IS could provide a quick evaluation and comparable index for use by all types of conservation organisations. Some changes like category scale and weighting would affect the comparability of the index, but use within an organisation could be guided to minimise those effects.

5.6 Ranking the rank

The Ranked Outcomes approach provides a new avenue for developing a comparable index based on qualitative data. This method could include all potential outcomes for conservation projects covered via general statements. If this were possible, the resulting score for an individual project could theoretically be compared to any other conservation project regardless of context. However, this would ignore what specific threats each project was actually facing and therefore the resulting score could be completely unrelated to project potential for success. One way to reduce this would be to rank them by region, but this would then reduce

the desired comparability and does not completely eliminate the first problem. The outcomes used for ranking are actually limited to those achieved and those that could potentially be achieved by the project or group of projects. This reduces the ability to use the index for comparison, now effectively limited to the group used for ranking, but if the evaluation is not open ended, that would make the issue moot.

Evaluators compiling outcomes must be careful to select only outcomes, and not the commonly confused outputs or inputs. Again, using the concepts and definitions as laid out by CCF or the Conservation Measures Partnership (CMP) as guidelines is advisable. Similarly, categorising the outcomes could be challenging. The categories originally used were modified and it may be beneficial to modify them again to mimic those used in other evaluations. One challenge observed in this example was that due to the complexity of conservation, some outcomes will be achieved by a combination of activities and therefore cannot be assigned to just one category. This could lead to individual outcomes being listed in multiple categories, similar to other potential complications of having many less significant outcomes that misrepresent the true overall contribution to success and coincidently reduce the value of more powerful outcomes in the scoring. A weighting system may be one option to address this.

Gathering a sufficiently representative population of evaluators presents another challenge. The preferences of those ranking the outcomes will affect their ranking of the outcomes, and that bias may ultimately affect the score of the projects. As pointed out by the authors, views on contribution to success vary, and larger sample sizes could reduce this effect (Howe and Milner-Gulland, 2012). The method does provide an opportunity to examine beliefs about success across evaluators, and to investigate them by graphical methods and the kappa statistic (Howe and Milner Gulland, 2012). In this iteration of the method, only three individuals were available for ranking outcomes, including the author. The panel members had varying degrees of experience in conservation and familiarity with the projects. This is probably the cause of the low level of agreement in this test, but with only three members, the odds for overlap are small to begin with. If the panel size were increased, the chances of agreement would improve,

unless distinct views are equally represented in the panel population. Members of the panel commented that some of the outcomes were unclear in their descriptions, which could be another factor in the limited agreement. This further highlights that employers of the Ranked Outcomes approach must be explicit and clear in their selection and presentation of outcomes.

This application of the Ranked Outcomes approach to projects sharing sites, objectives and activities did not provide much insight beyond highlighting outcomes and the opinions of a small panel. However, there is clearly potential for evaluating projects with distinct strategies. Evaluating programmes, projects or even activities with different focuses or locations would likely be more productive and more in line with the original application. There is potential for this technique at an organisational level and further development of this unique idea is certainly warranted.

5.7 Recommendations for monitoring and reporting

What comes first, the data or the evaluation? The data being collected through monitoring and reporting should assist the evaluation of success. In other words, the identification of outcome indicators and deciding how they will be evaluated should be done before deciding how to monitor - the evaluation is first in theory. However, this is generally not the reality. Evaluation is often an afterthought. Truthfully, all of these aspects of a project should not be viewed as individual tasks, but components of an integrated method, each building and relying upon each other to produce a cohesive project. After identifying the targets and threats, logical frameworks and result chains provide an excellent platform from which to then define explicit outcomes and the indicators and methods required to effectively evaluate them. These decisions then determine what should be monitored and reported.

WWF could benefit from emphasising the importance of linking these elements. Monitoring plans should be developed with evaluation in mind. If improved conservation education and raised awareness are outcomes selected in planning, then surveying for changes in attitude and behaviour would be an appropriate method for providing the necessary data to evaluate if they

implemented activities are successful. Similarly, reporting should contain and clearly present the data necessary for evaluation. Providing detailed annual summaries of poaching enforcement outcomes, like arrests and confiscations, could provide a more complete picture to evaluate anti-poaching success. Possibly the biggest gap in WWF reporting could be closed by including annual investment by activity for every project and thereby provide opportunity to evaluate projects and programmes on the effectiveness of investments. Reporting should also adhere to widely recognised definitions and concepts like those of CCF and CMP. Although those of CMP have been incorporated into WWF's standards, they are not necessarily reflected in all reports.

5.8 Concluding remarks

It is clear that no panacea exists for conservation evaluation. Evaluation techniques vary widely in their approach and their results, and none provide all the characteristics listed in the introduction. Nor should they, it is unrealistic to expect an evaluation technique to cover all of those components in a simple and easy manner while ensuring that it is applicable to the full spectrum of conservation projects. In addition, information necessary to make effective decisions at the project level may differ from those required at the organisation level. Therefore, no single evaluation is likely to provide information to meet all needs (Margoluis et al, 2009). The challenge for an organisation like WWF is to establish a balance between providing necessary information for adaptive management and effective allocation of the organisation's resources, while not creating more work and devoting more time to the effort than necessary (Salzer and Salafsky, 2006; Sheil, 2002)

WWF has invested much time and effort into improving its conservation effectiveness. As members of CMP they have adopted the partnership standards, have developed an audit programme to ensure programmes abide by those standards, and have used Miradi to develop conceptual models and results chains. The progress continues with an updated version of the WWF Standards nearly finished, a new network-wide database for programme reports now online, and a good practice self-assessment being tested within WWF-UK. A sign of a good set

of standards is the willingness to improve, and one aspect frequently highlighted in the audits is a need for improving evaluation (Beale, personal communication). WWF has outlined five criteria evaluations should asses: “quality and relevance of project design, efficiency of delivery of outputs, effectiveness of delivery of intermediate results, impact on ultimate conservation targets, sustainability of progress, benefits, and impact realised, and adaptive capacity of monitoring, evaluation, adaptation, and learning” (WWF, 2007). Among the attributes of the evaluation techniques reviewed, the most impressive was the shared recognition and effort to assess project impact. Each of the tested evaluation approaches deserves consideration by WWF, and all conservation organisations, as tools to choose from. However, one approach addresses a gap identified in this thesis and WWF - ROI assesses effectiveness and efficiency of implementation. For this reason, WWF should continue to explore how ROI can be used to achieve these goals in their programmes, developing a model analysis with hypothetical data to provide more examples of how the analysis can inform project and organisation managers alike.

REFERENCES

Anthony, B.P. (2008) Use of modified threat reduction assessments to estimate success of conservation measures within and adjacent to Kruger National Park, South Africa. *Conservation Biology,* 22 (6), 1497–1505.

Balmford, A., Gaston, K.J., Rodrigues, A.S.L. & James, A. (2000) Issues in international conservation Integrating Costs of Conservation into International Priority Setting. *Conservation Biology*, 14 (3), 597–605.

Balmford, A., Gaston, K. J., Blyth, S., James, A. & Kapos, V. (2003) Global variation in terrestrial conservation costs, conservation benefits, and unmet conservation needs. *Proceedings of the National Academy of Sciences of the United States of America*, 100 (3), 1046-1050.

Beale, Will. Design and Impact, WWF-UK. (Personal communication, 2nd, September 2012).

Brooks, J.S., Franzen, M. A., Holmes, C.M., Grote, M. N. &Borgerhoff Mulder, M. (2006) Testing hypotheses for the success of different conservation strategies. *Conservation Biology*20 (5), 1528–1538.

Brooks, T. M., Mittermeier, R. A., da Fonseca, G. A. B., Gerlach, J., Hoffman, M., Lamoreux, J. F., Mittermeier, C. G., Pilgrim, J. D. & Rodrigues, A. S. L. (2006) Global biodiversity conservation priorities. *Science,* 313 (5783), 58-61.

Carwardine, J., O'Connor, T., Legge, S., Mackey, B., Possingham, H. P. & Martin, T. G. (2012) Prioritizing threat management for biodiversity conservation. *Conservation Letters,* 00, 1-9.

Christensen, J. (2003) Auditing Conservation in an Age of Accountability. *Conservation in Practice*, 4 (3), 12-18.

Cleary, D. (2006) The Questionable Effectiveness of Science Spending by International Conservation Organizations in the Tropics. *Conservation Biology*. 20 (3), 733–738.

Cullen, R., Fairburn, G. A. & Hughey, K.F.D. (2001) Measuring the productivity of threatened-species programs. *Ecological Economics*. 39 (1), 53–66.

Cullen, R., Moran, E. & Hughey, K.F.D. (2005) Measuring the success and cost effectiveness of New Zealand multiple-species projects to the conservation of threatened species. *Ecological Economics*. 53 (3), 311–323.

DFID (2011) *DFID's Approach to Value for Money (VfM).* Department for International Development.

Evans, M.C., Possingham, H.P. & Wilson, K. A. (2011) What to do in the face of multiple threats? Incorporating dependencies within a return on investment framework for conservation. *Diversity and Distributions*. 17 (3), 437–450.

Ferraro, P.J. & Pattanayak, S.K. (2006) Money for Nothing? A Call for Empirical Investments. *PLoS Biology*. 4 (4), 482-488.

Halpern, B. S., Pyke, C. R., Fox, H. E., Haney, J. C., Schlaepfer, M. A. & Zaradic, P. (2006) Gaps and mismatches between global conservation priorities and spending. *Conservation Biology*, 20 (1), 56-64.

Hanley, N. D. & Spash, C. L. (1993) *Cost-Benefit Analysis and the Environment.* Cheltenham, Edward Elgar Publishing Ltd.

Hollings, C. S. (ed.) Adaptive environmental assessment and management. International Series on Applied Systems Analysis. New York, John Wiley & Sons.

Howe, C., & Milner-Gulland, E. J. (2012). Evaluating indices of conservation success: a comparative analysis of outcome and output-based indices. *Animal Conservation*, 15, 1469-1795.

Hughey, K. F. D., Cullen, R. & Moran, E. (2003) Integrating economics into priority setting and evaluation in conservation management. *Conservation Biology,* 17 (1), 93-103.

James, A. N., Green, M. J. B. & Paine, J. R. (1999) *Global review of protected area budgets and staff*. World Conservation Monitoring Centre. WCMC Biodiversity Series No.10.

James, A. N., Gaston, K. J. & Balmford, A. (2001) Can we afford to conserve biodiversity? *BioScience*, 51, 43-52.

Jones, J. P. G., Andriamarovolona, M. M. & Hockley, N. (2008) The Importance of Taboos and Social Norms to Conservation in Madagascar. *Conservation Biology*. 22 (4), 976-986.

Lee, K. N. 1999. Appraising adaptive management. *Conservation Ecology*, 3 (2), 3-16.

Ludwig, D., Hilborn, R. & Waters, C. (1993) Resource Exploitation, and Conservation: Lessons from History. *Science*, 260 (5104), 17&36.

Kapos, V., Balmford, A., Aveling, R., Bubb, P., Carey, P., Entwistle, A., Hopkins, J., Mulliken, T., Safford, R., Stattersfield, A., Walpole, M. & Manica, A. (2008) Calibrating conservation: new tools for measuring success. *Conservation Letters,* 1 (4), 155-164.

Kapos, V., Balmford, A., Aveling, R., Bubb, P., Carey, P., Entwistle, A., Hopkins, J., Mulliken, T., Safford, R., Stattersfield, A., Walpole, M. & Manica, A. (2009) Outcomes, not implementation, predict conservation success. *Oryx*, 43 (03), 336-342.

Kleiman, D. G., Reading, R. P., Miller, B. J., Clark, T. I., Scott, M., Robinson, J., Wallace, R. L., Cabin, R. J. & Felleman, F. (2000) Improving the Evaluation of Conservation Programs. *Conservation Biology*. 14 (2), 356–365.

Lawrence, K. (2012) *WWF Value for Money Briefing Document*. WWF-UK.

Laycock, H., Moran, D., Smart, J., Raffaelli, D. & White, P. (2009) Evaluating the cost-effectiveness of conservation: The UK Biodiversity Action Plan. *Biological Conservation*, 142 (12), 3120-3127.

Laycock, H. F., Moran, D., Smart, J. C. R., Raffaelli, D. G. & Piran, C. L. W. (2011) Evaluating the effectiveness and efficiency of biodiversity conservation spending. *Ecological Economics*. 70 (10), 1789–1796.

Leech, Beth L., Frank R. Baumgartner, Jeffrey M. Berry, Marie Hojnacki and David C. Kimball (2007) 'Does Money Buy Power? Interest Group Resources and Policy Outcomes', paper presented at the annual meeting of the Midwest Political Science Association, 12–15 April, Chicago.

Mace, G. M. & Balmford, A., Leader-Williams, N., Manica, A., Walter, O., West, C. & Zimmerman, A. (2007) Measuring conservation success: assessing zoos' contribution. In: Zimmermann, A., Hatchwell, M., Dickie, L. A., & West, C. (eds.). *Zoos in the 21st Century: Catalysts for Conservation?* Cambridge, Cambridge University Press. pp. 322–342.

Margoluis, R. & Salafsky, N. (2001) *Is our project succeeding? A guide to Threat Reduction Assessment for conservation*. Biodiversity Support Program.

Margoluis, R. Stem, C., Salafsky, N. & Brown, M. (2009) Design Alternatives for Evaluating the Impact of Conservation Projects. *New Directions for Evaluation*, 122, 85-96.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B. & Kunst, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, 403 (6772), 853-858.

Nepal MFSC. (2009) *Nepal Biodiversity Strategy*. MFSC.

New Economics Foundation. (2010) Value for Money (VfM) in International Development. NEF.

Noss, R. F. Indicators for Monitoring Biodiversity: A Hierarchical Approach. *Conservation Biology*, 4 (4), 355-364.

O'Neil, E. (2007) Conservation Audits: Auditing the Conservation Process Lessons Learned, 2003 - 2007. Conservation Measures Partnership.

Redford, K. H. & A. Taber. (2000) Writing the Wrongs: Developing a Safe-Fail. *Conservation Biology*, 14 (6), 1567-1568.

Salafsky, N. & R. Margoluis. (1999) Threat Reduction Assessment: A practical and cost-effective approach to evaluating conservation and development projects. *Conservation Biology*, 13, 830-841.

Salafsky, N., Margoluis, R., Redford, K. H. & Robison, J. G. (2002) Improving the Practice of Conservation: a Conceptual Framework and Research Agenda for Conservation Science. *Conservation Biology*, 16 (6), 1469-1479.

Salzer, D. & Salafsky, N. (2006) Allocating Resources Between Taking Action, Assessing Status, and Measuring Effectiveness of Conservation Action. *Natural Areas Journal,* 26 (3), 310–316.

Saterson, K. A., Christensen, N. L., Jackson, R. B., Kramer, R. A., Pimm, S. L., Smith, M. D. & Wiener, J. B. (2004) *Conservation Biology*, 18 (3), 597-599.

Sheil, D. (2002) Conservation and Biodiversity Monitoring in the Tropics: Realities, Priorities, and Distractions. *Conservation Biology*, 15 (4), 1179-1182.

Stem, C., Margoluis, R., Salafsky, N. & Brown, M. (2005) Monitoring and evaluation in conservation: A review of trends and approaches. *Conservation Biology,* 19 (2), 295-309.

Sutherland, W. J., Pullin, A. S., Dolman, P. M. & Knight, T. M. (2004) The need for evidence-based conservation. *Trends in ecology & evolution*. 19 (6), 305-308.

UNDP. (2010) *Millennium Development Goals Needs Assessment Report for Nepal 2010*. United Nations Development Programme.

USEPA. (2006) *Expanding the Use of Outcome Measurement for EPA's Office of Enforcement and Compliance Assurance*. U. S. Environmental Protection Agency.

Walters, C. J. (1986) *Adaptive Management of Renewable Resources.* New York, MacMillan Publishing Company.

Wilson, K. A, McBride, M.F., Bode, M. & Possingham, H.P. (2006) Prioritizing global conservation efforts. *Nature*. 440 (7082), 337–340.

Wilson, K. A, Underwood, E.C., Morrison, S. A., Klausmeyer, K. R., Murdoch, W. W., Reyers, B., Wardwell-Johnson, G., Marquet, P. A., Rundel, P. W., McBride, M. F., Pressey, R. L., Bode,

M., Hoekstra, J. M., Andelman, S., Looker, M., Rondinini, C., Kareiva, P., Shaw, M. R. & Possingham, H. P. (2007) Conserving biodiversity efficiently: What to do, where, and when. *PLoS Biology*. 5 (9), 1850-1861.

Woodhill, J. (2000) *Introduction to Key Concepts, Approaches and Terms*. IUCN. Global Monitoring and Evaluation Initiative.

WAZA. (2012) *Project Conservation Impact Summary Form*. [Online] Available from: http://www.waza.org/en/site/conservation/conservation-impact [Accessed 3rd September 2012].

WWF (2007) *WWF Standards of Conservation Project and Programme Management*. WWF.

WWF - Nepal (2011a) *Programme Proposal: People in Participatory Action for Life (PIPAL)* WWF. Project Reference Number: NP090301.

WWF – Nepal (2011b) *Programme Proposal: Conserving tigers and rhinos in Nepal*. WWF. Project Reference Number: 9Z140261.

Yoccoz, N. G., Nichols, J. D. & Boulinier, T. (2001) *Monitoring of biological diversity in space and time*. Trends in Ecology and Evolution, 16 (8), 446-453.

APPENDICES

Appendix A - Available reports for Nepal Programme

	Curbing Wildlife Trade			Strengthening Conservation			Tigers and Rhinos Nepal		
Year	**WP**	**R3**	**TPR**	**WP**	**R3**	**TPR**	**WP**	**R3**	**TPR**
FY06	-	X	-	-	-	X			
FY07	X	X	X	-	X	X			
FY08	X	X	X	-	X	X			
FY09	X	X	X	-	X	X			
FY10	X	X	X	-	X	X	X	X	-
FY11	X	X	X	-	X	X	X	X	X

Appendix B - Activity investment in the Eastern Himalayan Programme by project

Curbing Wildlife Trade in Nepal							
Activity Category		FY07	FY08	FY09	FY10	FY11	Total
Policy		16,774	21,228	15,000	8,000	7,000	**68,002**
Transboundary coordination		6,838	22,000	15,453	23,987	18,200	**86,478**
Informant and enforcement network		24,773	33,500	24,000	45,500	37,000	**164,773**
Community anti-poaching and anti-trafficking		11,612	6,000	27,000	25,000	17,000	**86,612**
Awareness		8,645	9,500	14,500	14,000	22,000	**68,645**
Monitoring wildlife trade		6,992	8,000	10,175	8,000	7,040	**40,207**
Total		**75,635**	**100,228**	**106,128**	**124,487**	**108,240**	**514,718**

Strengthening Conservation at the Landscape Level							
Activity Category	FY06	FY07	FY08	FY09	FY10	FY11	Total
Policy and advocacy	3,000	-	-	-	-	-	3,000
Institutional Strengthening and Coordination	175,895	-	-	-	-	-	175,895
Sustainable Forest Management	305,889	-	-	-	-	-	305,889
Sustainable Livelihoods	265,122	-	-	-	-	-	265,122
Species and Ecosystem Conservation	365,000	-	-	-	-	-	365,000
Churia watershed conservation	11,050	-	-	-	-	-	11,050
Conservation Education and Awareness	108,917	-	-	-	-	-	108,917
Total	**1,234,873**	**490,990**	**926,950**	**732,415**	**1,183,195**	**907,513**	**5,475,936**

Tigers and Rhinos in Nepal							
Activity Category					FY10	FY11	Total
Policy					25,100	25,943	**51,043**
Informant and enforcement network					39,960	45,000	**84,960**
Infrastructure					36,871	63,458	**100,329**
Community anti-poaching and anti-trafficking					11,686	13,090	**24,776**
Community anti-encroachment					1,255	9,665	**10,920**
Monitoring PA					42,283	50,000	**92,283**
Monitoring ecology					28,892	30,000	**58,892**
Total					**186,047**	**237,156**	**423,203**

Appendix C - Rhino and investment (CWT and TRN) data

Year		2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Investment (USD1,000)	1. Legislation and policy								16.774	21.228	15.000	33.100	32.943
	2 Networks and capacity								24.773	33.500	24.000	122.331	145.458
	3. Community anti-poaching								11.612	6.000	27.000	37.941	39.755
	4. Transboundary coordination								6.838	22.000	15.453	23.987	18.200
	5. Awareness								8.645	9.500	14.500	14.000	22.000
	6. Monitoring								6.992	8.000	10.175	79.175	87.040
	Total								75.635	100.228	106.128	310.534	345.396
Return	Rhino population	512					372			435			534
	Rhinos poached	16	5	35	31	11	19	25	13	3	12	11	2
	ΔRhino poaching					-20	8	6	-12	-10	9	-1	-9
					ROI (Δpoach/USD1,000)				-0.159	-0.100	0.085	-0.003	-0.026
					Unit cost (USD1,000/Δpoach)				-6.303	-10.023	11.792	-310.534	-38.377

Appendix D - Rhino population ROI across spending categories

Δ rhino 99			
Investment	USD1,000	rhinos/USD1,000	USD1,000/rhino
1 Legislation and policy	81.043	1.22	0.819
2 Networks and capacity	291.789	0.34	2.947
3. Community enforcement	104.696	0.95	1.058
4. Transboundary coordination	57.640	1.72	0.582
5. Awareness and education	50.500	1.96	0.510
6. Monitoring	176.390	0.56	1.782
Total	762.058	0.13	7.698

Appendix E - CCF Tool index possibility. This is the standard CCF Tool form, but the a priori scoring format can be re-used at the end of the evaluation, totalling the scores to create an index. In this case, the attempted components score 17 out of the possible 24. The ratio depicts how broad of an approach the project is taking and how successful those projects have been.

26. Which of the following components are included in the project? And how **successful** have they been to date? (Conservation is increasing the likelihood of persistence of native ecosystems, habitats, species and/or populations in the wild. Conservation success is the achievement of this (without adverse impacts on human well-being).)	For each component included in the project, please provide an a priori estimate of its success	
Species management (*Management of species and populations*) – actions directly involving species themselves, such as clutch management, captive breeding, etc.	Not included in project	
Site management (*Management of Sites, Habitats, Landscapes and Ecosystems*) – actions directly manipulating or managing a particular site (May include work on specific locations and in the wider landscape including: Protected areas (covering identification of new protected areas, establishment, management, and expansion); Maintenance/Conservation actions (e.g. fire regimes, forestry and farming practices, water management); Habitat restoration/creation (e.g. removal of invasive species, reafforestation, corridor establishment)).	2- made some contribution to increasing likelihood of persistence	Protected area buffer zones enlarged; management improved; CBAPO's formed; site restoration
Livelihood enhancements and alternatives (*Efforts to Enhance and/or Provide Alternative Livelihoods*) - actions to improve the well-being of people having impact on the species/habitats of conservation interest (including through sustainable resource management, income-generating activities, conservation enterprise, direct incentives)	3- achieved a significant amount that will contribute to increased persistence	Alternative energy schemes; carbon financing; micro-financing; alternative livelihoods
Policy and legislation (*Efforts to develop, adopt or implement policy or legislation*) - actions to establish frameworks within the processes of government, civil society or the private sector that make conservation goals official or facilitate their accomplishment. (May include development, implementation and/or enforcement of legislation, management plans, sectoral policies, trade regulations, among others)	4- highly successful, demonstrably increased likelihood of persistence	CITES implementation; formation of Species Action Plans; working towards policy changes relating to the effective management of protected areas and protection of wildlife in coordination with its like-minded partners

Capacity Building (*Training and capacity building*) - actions to enhance specific skills among those directly involved in conservation (such as such as wardens, land/site managers, NGOs, forestry departments, working groups, and researchers). Includes both building individual skills and improving the many components of organisational capacity.	4- highly successful, demonstrably increased likelihood of persistence	Good governance is practiced in almost all community-based institutions; minorities and women are included in committees and decision making
Education & awareness raising (*Education and awareness-raising) - actions* directed at improving understanding and influencing behaviour among people not necessarily directly involved in conservation action. (Covers all forms of communication, including campaigns, lobbying, educational and publicity/awareness programmes, and production of materials.)	2- made some contribution to increasing likelihood of persistence	Eco-clubs formed and strengthening; work carried out in schools; work with media
Research (*Research and conservation planning*) - actions aimed at improving the information base on which conservation decisions are made, including survey, inventory, remote sensing, mapping, development of new technologies	2- made some contribution to increasing likelihood of persistence	Regular scientific monitoring of tiger, rhino and their habitats

27. Please provide an estimate of the overall success of the project on the 1-4 scale given in question 26 above.	3- achieved a significant amount that will contribute to increased persistence	**Components scored 17/24**

Appendix F - CCF definitions and concepts

Conservation is increasing the likelihood of persistence of native ecosystems, habitats, species and/or populations in the wild. **Conservation success** is the achievement of this (without adverse impacts on human well-being). The challenge in measuring conservation success is to move beyond the commonly employed measures of **inputs** (money and time spent), **activity** (actions implemented) and **output** (concrete countable products), to measures that effectively assess the **outcomes** (the effects on a problem) and **conservation impacts** (effects on ecosystems/habitats/species/populations) of a project or initiative.

- **Management of Sites, Habitats, Landscapes and Ecosystems**- actions directly manipulating or managing a particular site
- **Management of species and populations** - actions directly involving species themselves, such as clutch management, captive breeding, etc.
- **Efforts to develop, adopt or implement policy or legislation**- actions to establish frameworks within the processes of government, civil society or the private sector that make conservation goals official or facilitate their accomplishment; may include development, implementation and/or enforcement of legislation, management plans, sectoral policies, trade regulations, among others.
- **Efforts to Enhance and/or Provide Alternative Livelihoods** - actions to improve the well-being of people having impact on the species/habitats of conservation interest, including through sustainable resource management, income-generating activities, conservation enterprise, direct incentives.
- **Training and capacity building** - actions to enhance specific skills among those directly involved in conservation, includes both building individual skills and improving the many components of organisational capacity.
- **Education and awareness-raising** - actions directed at improving understanding and influencing behaviour among people not necessarily directly involved in conservation action. Covers all forms of communication, including campaigns, lobbying, educational and publicity/awareness programmes, and production of materials.
- **Research and conservation planning** - actions aimed at improving the information base on which conservation decisions are made, including survey, inventory, remote sensing, mapping, development of new technologies.