

Using a theory of change to evaluate the impact of a conservation training programme: a practitioner's perspective

RACHAEL M. GERRIE, LIANNE CONCANNON, JAMIESON A. COPSEY
TIM WRIGHT and RICHARD P. YOUNG

Abstract Training plays a central role in the pursuit of conservation goals, and it is vital to know if it is having the desired effect. However, evaluating the difference it makes is notoriously challenging. Here, we present a practitioner's perspective on overcoming these challenges and developing a framework for ongoing evaluation of a conservation training programme. To do this, we first created a theory of change, describing the pathway of change we expect from training delivery to conservation impact. This provided the clarity and structure needed to identify indicators of change in the short, medium and long term. For data collection, we utilized both quantitative and qualitative methods to provide a more complete understanding of the change expected and capture any that might be unexpected. However, the more time that passes since a training event, the more difficult it becomes to attribute results; in response, we shifted predominantly to the use of qualitative methods to understand the long-term results achieved. After 3 years of implementation, this framework has enabled us to measure the difference our training makes to individuals and their work, and to provide evidence for the contribution it makes to achieving conservation impact. We believe that the lessons learnt can be used to improve the evaluation of training activities across the conservation sector and maximize the impact they achieve.

Keywords Capacity building, evaluation, impact, mixed-methods, qualitative, theory of change, training

Supplementary material for this article is available at doi.org/10.1017/S0030605321001551

Introduction

There is a continuous need for training within the conservation sector. It equips conservation practitioners with the skills, knowledge and personal qualities they need

to overcome new challenges and achieve positive results. To ensure training is having this effect, it requires systematic monitoring and evaluation. This process creates vital opportunities for learning and adaptation, ensures resources are used to greatest effect, and provides the evidence needed for reporting to donors, supporters and other stakeholders (Stem et al., 2005; Jones et al., 2013). However, evaluating the impact of training is also notoriously challenging and, in reality, there is a need to balance the ideal monitoring and evaluation framework (i.e. one that provides robust evidence of the results achieved), with what it is practically possible to measure, especially given limited conservation resources (Jones et al., 2013).

Training providers face a number of challenges when it comes to evaluating their efforts. Firstly, the results they achieve are dependent on the circumstances of each individual who takes a training course, making it practically impossible to predict what success might look like in every eventuality (Roche, 1999; Ortiz & Taylor, 2009). In some cases, the most important results can also be difficult to quantify, such as potentially pivotal changes in an individual's self-confidence or motivation (James, 2009; Vallejo & When, 2016). Furthermore, the desired conservation impact of training could take decades to achieve, and in this time results will inevitably be shaped by a combination of factors (Roche, 1999; James, 2001, 2009; Ortiz & Taylor, 2009; Simister & Smith, 2010). This makes the attribution of results to a single training event a significant challenge, especially given that experimental evaluation approaches, commonly used to determine causality, are typically not feasible because of the ethical implications of randomly assigning control groups and the practicalities of constructing a credible counterfactual (Roche, 1999; James, 2001).

In many cases, the evaluation of conservation training is limited to the quantity and quality of delivery (Bruyere et al., 2020), and although success stories are commonly used for communication purposes (Conservation Leadership Programme, 2021; Tropical Biology Association, 2021), the methods used by training providers to evaluate the effectiveness of their efforts are not readily available. In the wider education sector, one of the most well-known methods for training evaluation is Kirkpatrick's four-level model (Kirkpatrick & Kirkpatrick, 2006). This offers a structured approach to evaluation by guiding practitioners through four-levels of criteria: reaction (i.e. were trainees satisfied

RACHAEL M. GERRIE (Corresponding author, orcid.org/0000-0002-8197-9753, rachael.gerrie@durrell.org), LIANNE CONCANNON, TIM WRIGHT and RICHARD P. YOUNG Durrell Wildlife Conservation Trust, Les Augres Manor, La Profonde Rue, Trinity, Jersey, JE3 5BP, Channel Islands

JAMIESON A. COPSEY IUCN Species Survival Commission, Conservation Planning Specialist Group, Apple Valley, USA

Received 2 March 2021. Revision requested 27 May 2021.

Accepted 26 October 2021. First published online 6 April 2022.

with a training event?), learning (i.e. did they learn anything?), behaviour (i.e. did they do anything differently?) and results (i.e. did training affect the wider system in which they operate?). Similarly, the use of a theory of change is also widely recommended for providing a systematic approach to evaluation in a range of contexts, including training (James, 2001; Stem et al., 2005; Kapos et al., 2008; Kapos et al., 2009; Ortiz & Taylor, 2009; Simister & Smith, 2010). This approach requires training providers to articulate how they expect their intervention to achieve the desired impact through explicit causal pathways and, in doing so, makes it easier to identify measures of success in the short, medium and long term.

However, an evaluation framework that relies solely on measuring pre-defined indicators of success can result in missed opportunities to learn from any unexpected results (Roche, 1999; Watson, 2010; Vallejo & When, 2016), particularly in the context of training, where results can be difficult to predict. To provide a more complete understanding of the results achieved, it is widely recommended to utilize both qualitative and quantitative data collection methods (James, 2001, 2009; Stem et al., 2005; Ortiz & Taylor, 2009; Simister & Smith, 2010; Watson, 2010; Vallejo & When, 2016). A mixed-methods approach can help in understanding the extent to which an intervention is achieving pre-defined indicators of success as well as capture any difficult to define or unexpected results. The use of qualitative methods is also useful for capturing the long-term results of training, where direct measurement and attribution is not possible (James, 2001, 2009; Simister & Smith, 2010).

Durrell Wildlife Conservation Trust (Durrell) is a non-profit wildlife conservation organization, whose mission is saving species from extinction. For over 40 years, a core part of achieving this mission has been to provide training for conservation practitioners, typically from the organization's priority regions, and for aspiring conservationists, including university graduates and early-career individuals. Until recently, attempts to evaluate the difference this training makes have been opportunistic and focused predominantly on the collection of case studies. However, in an increasingly evidence-led sector (Sutherland et al., 2004), we wanted to develop a more systematic approach that would enable us to evaluate and maximize our impact. To do this, we decided to use a theory of change and mixed-methods approach as this gave us an opportunity to clarify the assumed links between training and our conservation mission, better understand the complexities of training outcomes, and develop a framework that we could review and adapt over time.

Given Durrell's long history in conservation training, we believe this is a useful case study for exploring the practicalities of using a theory of change and mixed-methods approach for the evaluation of training impact. Here, we describe the approach taken by Durrell to develop a simple

theory of change and associated monitoring plan for its conservation training programme. In addition, we present a subset of preliminary quantitative and qualitative results to share the key lessons learnt in the first 3 years of implementation.

Methods

Developing the evaluation framework

In November 2016, we began the development process by reviewing findings from three retrospective evaluations of Durrell's training programme: an internal evaluation of Durrell's flagship 12-week Endangered Species Management course (Payne, 2015), followed by two Imperial College London MSc projects evaluating 1–12 week courses across Durrell's training programme (Ruzowitzky, 2015; Sawrey, 2015). Collectively, these captured the experiences of 514 individuals who attended a course during 1979–2015. Payne (2015) used a mixed-methods approach to identify the outcomes perceived by past participants and the external factors that helped or hindered their progress. Ruzowitzky (2015) developed an evaluation framework based on Kirkpatrick's four-level model (Kirkpatrick & Kirkpatrick, 2006) to measure the extent to which training influenced their professional development and conservation actions, and Sawrey (2015) developed a theory of change from the perspective of trainers and validated it against the outcomes perceived by trainees. The key results from each (Table 1), provided us with an information base from which to develop a theory of change.

To do this, we began by explicitly describing the desired conservation impact of training (to achieve our organizational mission of saving species from extinction) and the key change we expect training to affect for this to be realized (more effective conservation action). From here, we used if–then statements to determine the results required in the short and medium term for these long-term goals to be achieved, and the potential causal links that existed between them. In particular, an increase in self-confidence, motivation and self-efficacy, collectively described as perception of control, proved to be one of the greatest outcomes of training (Payne, 2015; Sawrey, 2015) and an important precursor for increased effectiveness at work (Sawrey, 2015). However, in some cases, unsupportive organizations and a lack of opportunities prevented past participants from applying new skills and progressing in their careers. We knew therefore that an enabling work environment would be an important condition for success. To ensure we considered different perspectives in the development process, we shared a preliminary draft of the theory of change with colleagues across Durrell, including training and non-training staff working in a range of countries, and therefore socio-

TABLE 1 A summary of key results from three retrospective evaluations of Durrell's conservation training programme.

Result	Description	Reference
Skills & knowledge	Individuals gained new & improved subject knowledge, & practical, interpersonal & professional skills	Payne (2015), Sawrey (2015), Sawrey et al. (2019)
Perception of control	Individuals felt more confident in their abilities, more motivated to succeed & more capable of achieving their goals	Payne (2015), Sawrey (2015), Sawrey et al. (2019)
Support network	Individuals benefitted from the network gained, using it to seek advice, create partnerships & co-develop conservation projects	Payne (2015), Ruzowitzky (2015), Sawrey (2015), Sawrey et al. (2019)
Personal effectiveness	Individuals used the skills & knowledge gained to improve their job performance	Payne (2015), Ruzowitzky (2015), Sawrey (2015), Sawrey et al. (2019)
Professional development	Individuals progressed within their conservation careers	Payne (2015), Ruzowitzky (2015), Sawrey (2015), Sawrey et al. (2019)
Conservation action	Individuals implemented more effective conservation action	Payne (2015), Ruzowitzky (2015)
Work environment	An individual's work environment can significantly affect their ability to apply new skills & progress in their careers	Payne (2015), Sawrey (2015), Sawrey et al. (2019)

cultural contexts, before inviting three external experts from the conservation training field to challenge our logic and comprehension of results. During this review process, the role of the support network in achieving impact was repeatedly called into question, with the general consensus being that it can sometimes be influential in the success an individual achieves but is not always required. This debate was also echoed in results from the retrospective evaluations (Payne, 2015; Ruzowitzky, 2015; Sawrey, 2015). In response, we decided to explicitly represent this as a non-essential link in the theory of change.

We used the resulting theory of change (Fig. 1) to guide the development of the monitoring plan by defining at what point in time we expected each result to be achieved and what we needed to know to verify if this had happened or not. To streamline data collection and minimize the risk of survey fatigue, we grouped these information needs into four key intervals: those requiring verification immediately post-training, followed by those requiring verification at 1-, 5- and 10-years post-training. Based on our organizational experience, we considered 10 years to be the minimum amount of time in which we could reasonably expect training to improve the effectiveness of conservation action and therefore benefit species recovery. However, when evaluated retrospectively, response rates dropped significantly among participants who had completed a course > 10 years ago (Ruzowitzky, 2015), making it the most suitable time frame for evaluation. Because of this long time frame, we opted for a pre-test post-test non-experimental evaluation design and used online questionnaires, designed using Smart Survey (2020), to collect the information required at each interval and to establish a pre-training baseline. Online questionnaires proved to be a valuable tool for capturing useful information in all three retrospective evaluation projects (Payne, 2015; Ruzowitzky, 2015; Sawrey, 2015) and in comparison to alternative and complementary tools, such as interviews and

focus groups, require minimal resources to implement and analyse, making them the most feasible choice for the scale required.

For results requiring verification immediately post-training, we developed multi-item Likert scales. For perception of control, this included a series of three multi-item Likert scales, one each for motivation, self-confidence and self-efficacy. To measure change in skills and knowledge, we created a list of seven competencies tailored to the learning objectives of each course written as short 'How to...' statements. This number gave us sufficient scope to evaluate the main learning objectives of each course without overloading participants with a long list of questions and risking so-called straight lining, in which respondents lose interest and select the same response for each question. To avoid misinterpretation and improve the accuracy of responses, we also kept the wording of statements as clear and concise as possible, for example 'How to write a grant proposal'. For each, we developed two Likert scales, one to measure level of knowledge and one to measure level of confidence to apply knowledge. We also expected participants to gain a support network immediately post-training, but because of its non-essential role in the theory of change, and in the interest of minimizing the number of questions, we decided it was not a priority for evaluation at this stage in the development process and did not design an associated indicator.

For results requiring verification at 1-, 5- and 10-years post-training, we increasingly combined quantitative and qualitative methods. For example, to measure personal effectiveness at 1-year post-training we transformed the same list of competencies used to measure skills and knowledge into a multiple response, multiple choice question, asking participants to select which, if any, they had applied in their work and to provide an example if relevant. Similarly, to monitor conservation actions, at 1-, 5- and 10-years post-training we used a multiple response, multiple choice question based on the *Conservation Action Classification 2.0* (Conservation

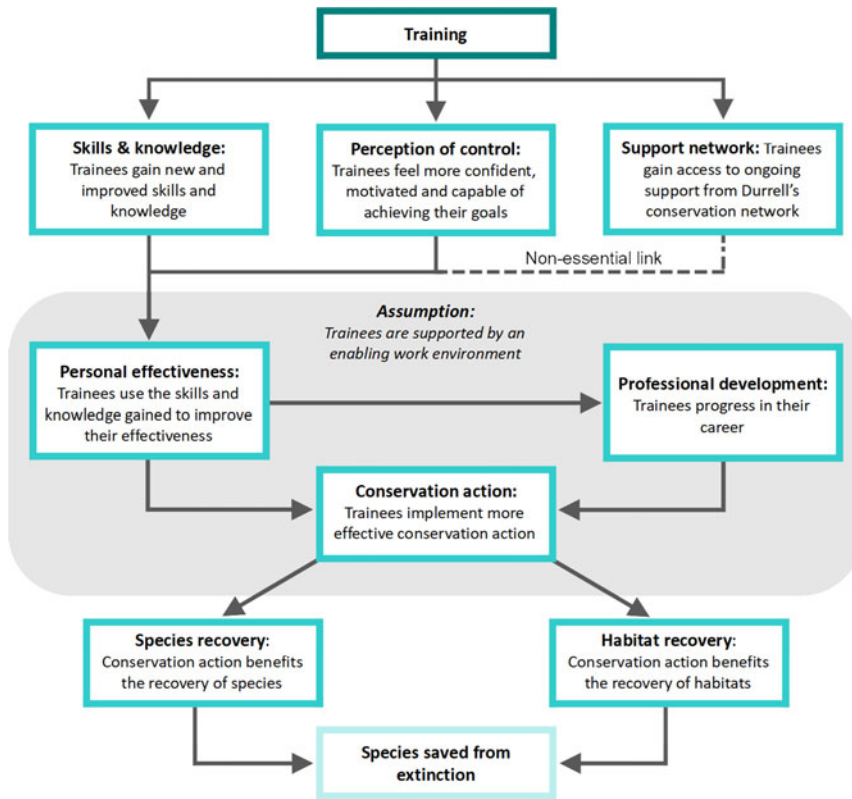


FIG. 1 The theory of change for Durrell's training programme: after completing a training course, an individual gains new and improved skills and knowledge, increased perception of control and access to ongoing support from Durrell's conservation network. An individual then returns to work and uses the skills and knowledge gained to improve their personal effectiveness. A more effective individual then progresses within their career and/or implements more effective conservation action. To achieve these results, we make the assumption that an individual returns to an enabling work environment; i.e. one that provides them with the opportunities and support needed to succeed (indicated by the grey box surrounding these results). The more effective conservation action taken benefits the recovery of species and/or habitats and ultimately achieves Durrell's conservation mission, to save species from extinction.

Measures Partnership, 2021), to understand the types of actions participants had taken in the previous 12 months followed by an open-ended request for an example. To further verify results against the theory of change as well as capture unexpected and long-term results, we used a simplified version of the most significant change method (Davies & Dart, 2005; Steadman, 2021), asking the open-ended question 'What is the most significant change you experienced as a result of this course?' in the immediate post-training questionnaire and 'What is the most significant change you experienced in the past 12 months?' in subsequent post-training questionnaires. To better understand the relationship between qualitative results and training, we included a rating scale for participants to report the extent to which training with Durrell influenced the result described, as per the approach used by Payne (2015), Sawrey (2015) and Ruzowitzky (2015). At 1-, 5- and 10-years post-training, we also wanted to understand the extent to which participants experienced an enabling work environment. To do this, we transformed a list of the seven most common barriers to the effective implementation of species recovery plans (Richardson, 2016) into a multi-item Likert scale, asking participants to rate the extent to which each affected their work, followed again by a request for a qualitative example if applicable. Finally, before implementing a full-scale trial, we pre-tested pre- and post-training questionnaires in person with a cohort of international course participants to ensure the language was accessible and completion time reasonable.

Preliminary implementation

During 2017–2019, we trialled pre-, post- and 1-year post-training questionnaires (Supplementary Materials 1, 2 & 3) across five face-to-face Durrell courses (Table 2). We e-mailed pre- and post-training questionnaires to all 88 course participants before and after their course and 1-year post-training questionnaires to the same 88 individuals 12 months later. For survey completion, we assigned each course participant a unique identification code, saved alongside their personal details in a secure database, to be entered at the start of each questionnaire. This allowed questionnaires to be completed and managed anonymously but still matched for analysis. To maximize response rate, we personalized e-mail correspondence, included a clear statement of purpose and sent a reminder one week after the initial request. We received 88 completed pre- and post-training questionnaires (100% response rate) and 58 completed 1-year post-training questionnaires (66% response rate). We summarized results from each training event on an ongoing basis before combining results from all five courses into one dataset to produce an appropriate sample size for statistical analysis and to draw general conclusions about the effectiveness of the training programme as a whole. We used the Wilcoxon signed-rank test in R 3.6.3 (R Core Team, 2020) to test for differences in indicators measured pre- and post-training, assuming statistical significance at $P < 0.05$. For the analysis of qualitative data collected

TABLE 2 An overview of the five Durrell training courses included in preliminary implementation of the evaluation framework.

Training course	Description
Endangered Species Management	A 12-week course for conservation practitioners to develop a broad range of skills including facilitation, leadership, animal husbandry, field research, project planning & management
Facilitation & Communication Skills	A 1-week course for practising & aspiring conservationists to develop the skills needed to facilitate multi-stakeholder meetings
An Introduction to GIS using QGIS	A 1-week course for practising & aspiring conservationists to develop basic mapping & analytical skills
Endangered Species Recovery	A 2-week course for practising & aspiring conservationists to learn the principles & practice of species conservation
Avian Egg Incubation	A 1-week course for conservation practitioners to develop practical skills in avian egg management

using the adapted most significant change method (Davies & Dart, 2005), we used a thematic coding approach to identify key words and phrases from each response and group them into distinct themes. We then reapplied these themes to determine the number of responses related to each.

Results

The evaluation framework

The evaluation framework for Durrell's training programme is underpinned by a theory of change describing the pathway of change we expect an individual to take from

completing a training course to achieving conservation impact (Fig. 1). To measure progress against this theory of change, and capture any unexpected results, we use a combination of quantitative and qualitative indicators (Table 3).

Preliminary results

Immediate post-training Results showed that trainees did gain new and improved skills and knowledge immediately after a completing a course, with participants reporting higher levels of knowledge post-training ($P < 0.001$) and higher levels of confidence to apply knowledge ($P < 0.001$). When results from each course were summarized independently, we were able to identify which competencies participants were the most and least knowledgeable and confident about (Fig. 2). However, preliminary results did not show the expected increase in all three elements of perception of control, with participants reporting an increase in their level of self-confidence ($P < 0.001$) and self-efficacy ($P < 0.001$), but not in their level of motivation ($P = 0.37$). Instead, over 88% of participants reported feeling highly or extremely motivated both before and after the course (Fig. 3). Similarly, when asked to describe the most significant change experienced as a result of the course, 44% of responses described an increase in self-confidence, whereas only 11% described an increase in motivation. Unexpectedly, 40% of the responses recorded immediately post-training described gaining inspiration and/or a new perspective to be the most significant change experienced. For example, as one participant described, 'it has allowed me to understand that many are working for the betterment of the environment and that although I come from a small country facing many threats, we are not alone, our conservation action is not in vain', and another, 'it has helped me to see the world differently and that there is hope in conservation'.

TABLE 3 The monitoring plan for Durrell's training programme.

Result/Assumption	Indicator(s)	Data collection
Skills & knowledge	% of trainees reporting a good/high level of knowledge; % of trainees reporting a good/high level of confidence to apply knowledge	Pre- & immediate post-training questionnaire
Perception of control	% of trainees feeling highly/extremely motivated; % of trainees feeling highly/extremely confident; % of trainees feeling highly/extremely capable	Pre- & immediate post-training questionnaire
Personal effectiveness	% of trainees applying each course-specific competency; qualitative example of competency applied	1-year post-training questionnaire
Professional development	% of aspiring conservationists vs conservation practitioners; % of conservation practitioners in each career level	Pre-, 1-, 5- & 10-years post-training questionnaire
Conservation action	% of trainees implementing each IUCN conservation action type; % of conservation actions influenced by training at Durrell; qualitative example of conservation action taken	1-, 5- & 10-years post-training questionnaire
Species recovery/habitat recovery	Qualitative stories of change; % of stories influenced by training at Durrell	1-, 5- & 10-years post-training questionnaire
Enabling environment	% of trainees facing major/complete barriers in their work; qualitative example of major/complete barrier faced	1-, 5- & 10-years post-training questionnaire

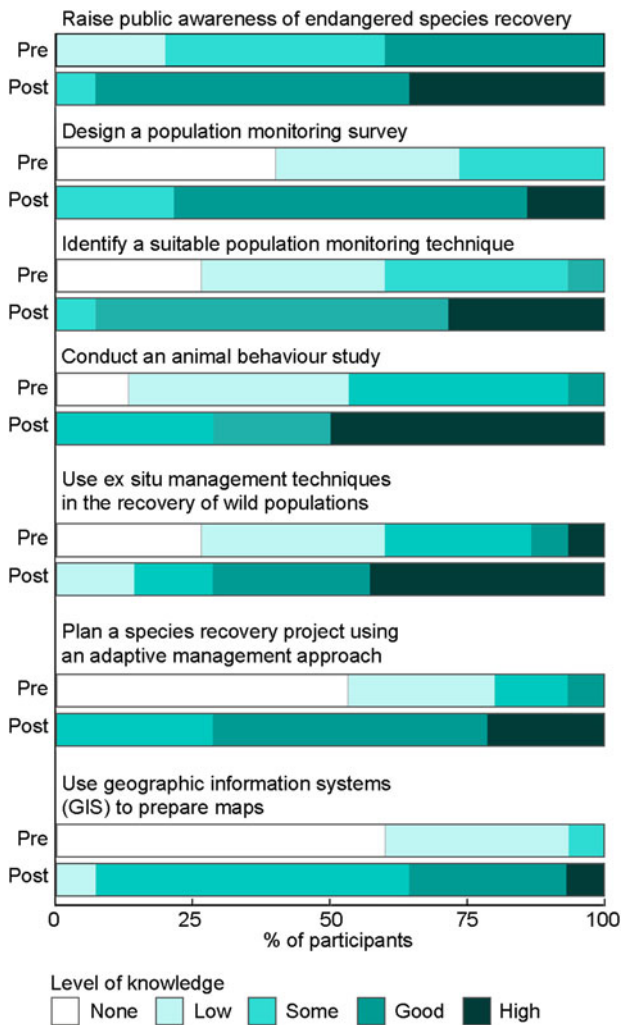


FIG. 2 Per cent of participants who reported each level of knowledge pre-training compared to post-training for each of the competencies assessed during the Endangered Species Recovery course in 2019.

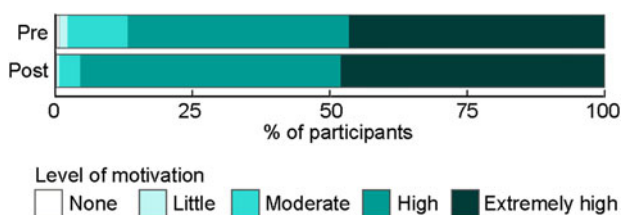


FIG. 3 Per cent of participants who reported each level of motivation pre-training compared to post-training across all of the courses included in this study.

One-year post-training One-year after completing a course, results showed that participants were using the competencies gained to improve their personal effectiveness and, following each course, we were able to identify which competencies were being applied by the greatest number of participants and which were being applied by the fewest

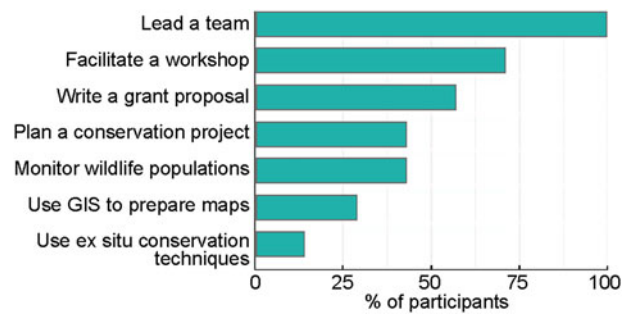


FIG. 4 Per cent of participants applying each competency in the 12 months after completing the Durrell Endangered Species Management course in 2018.

(Fig. 4). Qualitative examples provided further verification and illustrated the ways in which competencies had been applied, for example, in the 12 months since completing the Durrell Endangered Species Management course, one participant described using the population monitoring skills they gained as ‘I set up the monitoring of a threatened EDGE species’, and another described using the grant writing skills gained as ‘I wrote a proposal for a livelihoods group and they were successful in getting the funding’. Results also showed that training influenced the conservation actions taken by participants. For example, in the 12 months since completing the Facilitation and Communication Skills course, one participant described taking an action somewhat influenced by the training received as ‘I brought together an electric company, NGOs and a governmental institution to agree on the correction of a power line which is dangerous to raptors’, and following the Endangered Species Recovery course, one participant described an action significantly influenced by the training received as ‘I recently led my first offshore island invasive species eradication project’. In total, 58% of participants reported that their ability to take a given conservation action was significantly influenced by the training received, 33% reported it was somewhat influenced and 9% reported it was not at all influenced.

Discussion

Despite the challenges faced, we have shown that it is possible to develop a useful and practical framework for ongoing evaluation of a conservation training programme. Using a theory of change approach has helped us to clarify our assumptions about how training contributes to our conservation mission and identify appropriate measures of success in the short, medium and long term. The use of qualitative and quantitative methods to measure progress against the theory of change has provided us with a richer understanding of what training can achieve and enabled us to capture outcomes that were unexpected and difficult to define. The information gained throughout this process

will be used to refine the theory of change, improve the evaluation framework, and increase the effectiveness of Durrell's conservation training programme.

In line with the theory of change, preliminary results show that training is having the desired effect on the skills and knowledge of course participants and enabling them to be more effective at work. Immediately after completing a course, participants had more knowledge of the competencies taught and felt more confident to apply them in their work. Within 12 months, they were then using these competencies to improve their personal effectiveness; e.g. to write successful grant proposals and establish monitoring protocols. Within 12 months of completing a course, participants also reported that their ability to take specific conservation action was influenced by the training they received, including actions to remove threats and recover populations of threatened species. Although more difficult to assess, research has shown that measures of key outcomes such as these serve as powerful predictors of conservation success (Kapos et al., 2009) and have provided us with a reliable indication that training can achieve its goal of saving species from extinction. Quantifying the application of skills and knowledge has also enabled us to identify which competencies are the most useful to participants. For example, competencies relating to facilitation, leadership and project management were frequently reported to be the most applied in the 12 months after completing a course and we have begun expanding our training programme to meet these needs more fully. These findings support those of Barlow et al. (2016) and Englefield et al. (2019), who found project management and leadership skills to be in critical need across the conservation sector. In the future, we plan to broaden our definition and assessment of skills and knowledge to include more of the personal competencies we expect an individual to gain from training (Maggs et al., 2021), in particular for longer, more intensive courses such as the 12-week Durrell Endangered Species Management course.

In contradiction to the theory of change, preliminary results did not show a consistent improvement in all three perception of control elements. Participants reported higher levels of self-confidence and self-efficacy following training (i.e. they felt more confident in themselves and more capable of achieving their goals), but reported similar high levels of motivation both before and after training. In hindsight, this is perhaps not surprising, as we can expect individuals who seek out training to be highly motivated already. However, a combination of these perception of control elements previously proved to be an integral outcome of training (Sawrey, 2015; Sawrey et al., 2019) and warrant further investigation. We will continue to review the role they play within the theory of change and aim to improve the way they are defined and measured. In addition, we intend to explore in more detail the role inspiration plays in the theory of change as preliminary results suggest that feeling more

inspired and/or gaining a new perspective is one of the most significant outcomes of training and, as suggested in the wider education literature, this can increase the likelihood of an individual achieving their goals (Milyavskaya et al., 2012).

As with any evaluation, we needed to balance what we ideally wanted to know with what was practically possible to measure, and as a result we recognize a number of limitations with the method used. First is the reliance on self-reporting and its associated response biases, which can lead to inaccurate or false results. In particular, participants might report they feel more confident or more knowledgeable than they actually are because they think it is the correct answer. Second is the increasing risk of selection bias over time, where results become skewed because they only represent a subset of individuals who complete a questionnaire. In particular, participants who have had a negative experience or who did not continue to work in conservation might be less likely to complete a questionnaire and therefore be underrepresented in results. Although these biases cannot be fully overcome, their effect can be mitigated by following good survey design and we therefore consider questionnaires to be capable of generating useful information with minimal resources. Finally, the evaluation framework lacks a control group (i.e. a similar cohort of individuals in the conservation sector that do not receive the same training as course participants) against which we can compare rates of progress over time and truly understand the difference training at Durrell makes. This limits our ability to attribute long-term results and conservation impact to the training an individual receives. However, we have found a pre-test post-test non-experimental design to yield meaningful results and the use of qualitative data collection methods has enabled us to capture illustrative examples of success and determine the extent to which they are influenced by training. In addition, we found that these qualitative stories and the theory of change diagram itself to be valuable in communicating training impact to donors, supporters and prospective course participants.

We hope this work encourages others to adopt a systematic approach to evaluating the impact of training in conservation and we believe that the lessons learnt and the simple theory of change presented here can be used and adapted to facilitate this process. Finally, we hope this work will stimulate further discussion on the topic of impact evaluation in conservation training and encourage others to share their experiences.

Acknowledgements We thank The Balcombe Charitable Trust for funding the development of this work, Catherine Payne, Lucas Ruzowitzky and Brittany Sawrey for their preliminary research, Marianne Carter, Mark O'Connell and Rosie Trevelyan for providing feedback on the theory of change, and the two anonymous reviewers and Martin Fisher for their comments and support in preparing this article.

Author contributions Study design: all authors; data collection and analysis: RG; writing: all authors.

Conflicts of interest None.

Ethical standards This research abided by the *Oryx* guidelines on ethical standards, received ethical approval from Durrell Wildlife Conservation Trust and conforms to standards set out by the British Sociological Association. All participants provided informed consent prior to taking part and confidentiality and anonymity of research participants was maintained throughout.

References

- BARLOW, A., BARLOW, C.G., BODDAM-WHETHAM, L. & ROBINSON, B. (2016) A rapid assessment of the current status of project management skills in the conservation sector. *Journal of Nature Conservation*, 34, 126–132.
- BRUYERE, B., BYNUM, N., COPSEY, J., PORZECANSKI, A. & STERLING, E. (2020) *Conservation Leadership Capacity Building: A Landscape Study*. American Museum of Natural History, New York, USA.
- CONSERVATION LEADERSHIP PROGRAMME (2021) conservationleadershipprogramme.org/our-impact [accessed 26 September 2021].
- CONSERVATION MEASURES PARTNERSHIP (2021) ccnetglobal.com/resource/8i [accessed 26 September 2021].
- DAVIES, R. & DART, J.J. (2005) *The 'Most Significant Change' (MSC) Technique: A Guide to Its Use*. mande.co.uk/wp-content/uploads/2018/01/MSCGuide.pdf [accessed 26 February 2021].
- ENGLEFIELD, E., BLACK, S.A., COPSEY, J.A. & KNIGHT, A.T. (2019) Interpersonal competencies define effective conservation leadership. *Biological Conservation*, 235, 18–26.
- JAMES, R. (2001) *Practical Guidelines for the Monitoring and Evaluation of Capacity-Building: Experiences from Africa*. INTRAC, International NGO Training and Research Centre, Oxford, UK.
- JAMES, R. (2009) *Just Do It: Dealing with the Dilemmas in Monitoring and Evaluating Capacity Building*. International NGO Training and Research Centre, Oxford, UK.
- JONES, J.P.G., ASNER, G.P., BUTCHART, S.H.M. & KARANTH, K.U. (2013) The 'why', 'what' and 'how' of monitoring for conservation. In *Key Topics in Conservation Biology 2* (eds D.W. Macdonald & K.J. Willis), pp. 327–343. John Wiley & Sons, Hoboken, USA.
- KAPOS, V., BALMFORD, A., AVELING, R., BUBB, P., CAREY, P., ENTWISTLE, A. et al. (2008) Calibrating conservation: new tools for measuring success. *Conservation Letters*, 1, 155–164.
- KAPOS, V., BALMFORD, A., AVELING, R., BUBB, P., CAREY, P., ENTWISTLE, A. et al. (2009) Outcomes, not implementation, predict conservation success. *Oryx*, 43, 336–342.
- KIRKPATRICK, D.L. & KIRKPATRICK, J.D. (2006) *Evaluating Training Programmes: The Four Levels*. 3rd edition. Berrett-Koehler Publication, Oakland, USA.
- MAGGS, G., APPLETON, M.R., LONG, B. & YOUNG, R.P. (eds) (2021) *A Global Register of Competences for Threatened Species Recovery Practitioners: A Comprehensive List of Skills, Knowledge and Personal Attributes Required by Practitioners Working within Threatened Species Recovery*. IUCN, Gland, Switzerland.
- MILYAVSKAYA, M., IANAKIEVE, I., FOXEN-CRAFT, E., COLANTUONI, A. & KOESTNER, R. (2012) Inspired to get there: the effects of trait and goal inspiration on goal progress. *Personality and Individual Differences*, 52, 56–60.
- ORTIZ, A. & TAYLOR, P. (2009) *Learning Purposefully in Capacity Development: Why, What and When to Measure?* International Institute for Educational Planning, Paris, France.
- PAYNE, C. (2015) *DESMAN Capacity Building Survey Report*. Unpublished report. Durrell Wildlife Conservation Trust, Jersey, Channel Islands.
- R CORE TEAM (2020) *The R Project for Statistical Computing*. r-project.org [accessed 26 February 2021].
- RICHARDSON, H. (2016) *The global need for threatened species management programmes and the barriers that prevent effective implementation of recovery plans*. MSc thesis. Imperial College London, London, UK.
- ROCHE, C.J.R. (1999) *Impact Assessment for Development Agencies: Learning to Value Change*. Oxfam, Oxford, UK.
- RUZOWITZKY, L. (2015) *The conservation impact of training: the Durrell Conservation Academy as a case study*. MSc thesis. Imperial College London, London, UK.
- SAWREY, B. (2015) *Evaluating training impacts in conservation: a case study in Mauritius*. MSc thesis. Imperial College London, London, UK.
- SAWREY, B., COPSEY, J. & MILNER-GULLAND, E.J. (2019) Evaluating impacts of training in conservation: a case study in Mauritius. *Oryx*, 53, 117–125.
- SIMISTER, N. & SMITH, R. (2010) *Praxis Paper 23: Monitoring and Evaluating Capacity Building: Is it Really that Difficult?* International NGO Training and Research Centre, Oxford, UK.
- SMART SURVEY (2020) smartsurvey.co.uk [accessed 26 February 2021].
- STEADMAN, D. (2021) Towards ecological and social impact through collaborative governance of a seascape of marine protected areas in Honduras. *Oryx*, 55, 507–518.
- STEM, C., MARGOLUIS, R., SALAFSKY, N. & BROWN, M. (2005) Monitoring and evaluation in conservation: a review of trends and approaches. *Conservation Biology*, 19, 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, 305–308.
- TROPICAL BIOLOGY ASSOCIATION (2021) tropical-biology.org/impact [accessed 26 September 2021].
- VALLEJO, B. & WEHN, U. (2016) Capacity development evaluation: the challenge of the results agenda and measuring return on investment in the global south. *World Development*, 79, 1–13.
- WATSON, D. (2010) Measuring capacity development: Combining the “best of two worlds” in monitoring and evaluation of capacity development. In *Capacity Development in Practice* (eds J. Ubels, N. Acquaye-Baddoo & A. Fowler), pp. 251–263. Earthscan, London, UK.