Threatened Species Recovery Hub

National Environmental Science Programme

Valuing research: tools and approaches for assessing and enhancing the benefits from environmental research

Rachel A. Morgain, Ruth A. O'Connor, Timothy D. Johnson, Holly B. Vuong, Tyrone H. Lavery, Natasha M. Robinson, Christine A. Fenwick, Heather J. Christensen, Bradley Moggridge, Teagan Goolmeer, Oliver Costello, Nicholas A. Macgregor, Ben C. Scheele, Jennie Fluin, James A. Fitzsimons, Katherine E. Selwood, Rebecca Spindler, Brendan A. Wintle, David B. Lindenmayer

November 2021

Australian National University · The University of Melbourne · The University of Queensland Australia · University of Canberra · Government of South Australia Department for Environment and Water · The Nature Conservancy Australia · Zoos Victoria · Werribee · Melbourne · Healesville · Bush Heritage Australia

Cite this publication as: Morgain, R.A., O'Connor, R.A., Johnson, T.D., Vuong, H.B., Lavery, T.H., Robinson, N.M., Fenwick, C.A., Christensen, H.J., Moggridge, B., Goolmeer, T., Costello, O., Macgregor, N.A., Scheele, B.C., Fluin, J., Fitzsimons, J.A., Selwood, K.E., Spindler, R., Wintle, B.A., Lindenmayer, D.B. (2021) Valuing research: tools and approaches for assessing and enhancing the benefits from environmental research. NESP Threatened Species Recovery Hub Project 7.1 report, Brisbane.

Cover image: Quenda team at work in Craigie Bushland. Image: Nicolas Rakotopare

Figure i. Graphical Executive Summary

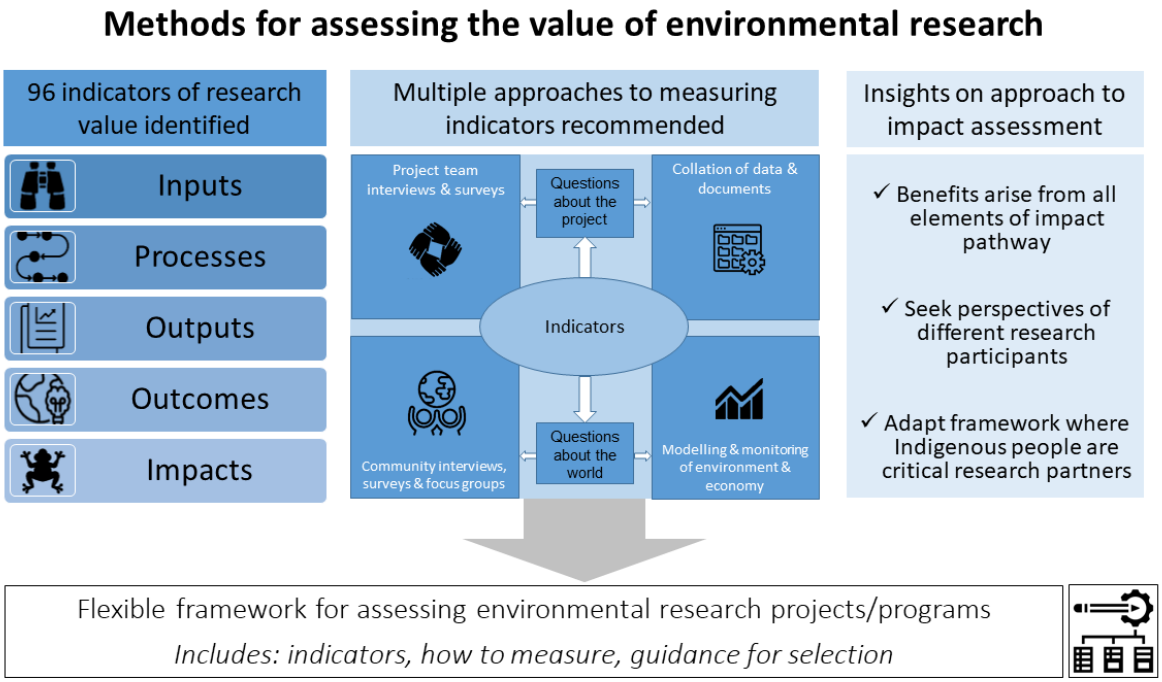
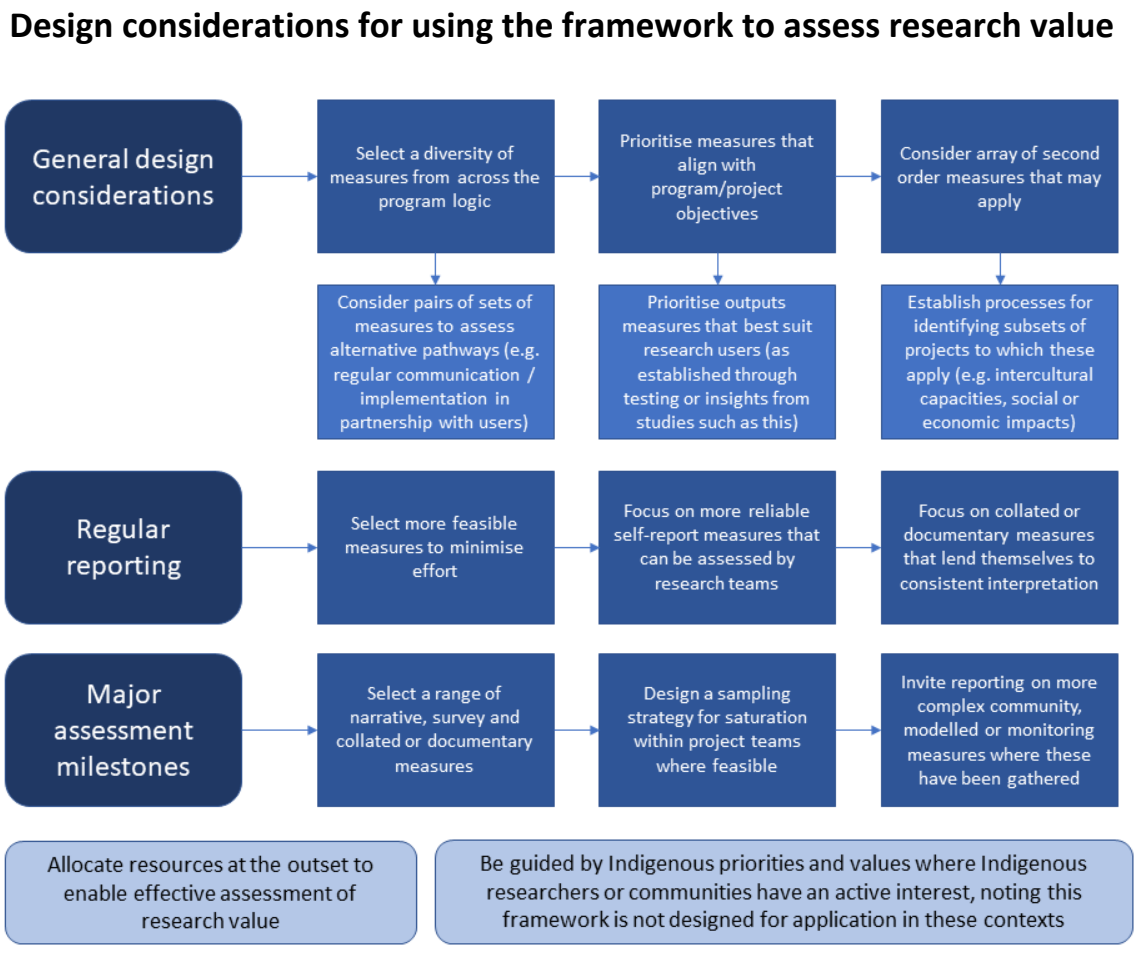


Figure ii. Decision principles for selecting and applying measures from the framework to assessing research value



Executive summary

I. This project has developed and tested a framework for assessing research value across the life cycle of conservation research projects and programs. This is the first comprehensive attempt of its kind to build and test a multimodal, integrated, qualitative and quantitative framework for achieving and assessing value in environmental research. Insights from this project have informed reporting within the Threatened Species Recovery Hub, and elements of the research are being used within NSW Department of Planning, Industry and Environment to better understand how they manage their knowledge outcomes.

II. The report focuses on research 'value', not just 'impact'. Achieving research value is strongly dependent on effective engagement and building appropriate 'pathways to impact'. Frequently, these pathways are circuitous and outcomes are achieved through unexpected avenues. Understanding and assessing the complex pathways through which research activities have flowed allows these diverse measures to be recognised as values in their own right. It also generates a more comprehensive and nuanced picture of research value in the face of this uncertainty and complexity than a narrow focus on achieving measurable environmental, social, cultural or economic impacts. Shifting the focus toward measures that help build a picture of trust, discoverability, participatory processes and respect for diverse points of view across project teams is arguably more telling and more inclusive for researchers at any career stage, institutional location and background.

III. The framework is designed as a flexible guide that can be drawn upon for a wide range of valuation purposes, and that can be applied (as appropriate) at any point through the life cycle of research projects or after completion. It is designed for practical application in agencies across a wide range of sectors, including research, research funding, government and land management agencies.
    A. It is important to note that this framework is not designed for application to situations where Traditional Owners, Indigenous researchers or Indigenous communities have an active interest in the projects being assessed. Assessing research value in these contexts should continue to be guided by the priorities and values of Indigenous communities and partners to the research.
    B. An Indigenous-led pathway, expert workshops and case studies to build an Indigenous-led approach to assessing research value is an important priority.

IV. The framework integrates survey, narrative, quantitative, documentary and independent qualitative measures against a subset of published indicators, which together are designed to provide a comprehensive, detailed and triangulated picture of the value delivered by a research project, suite of projects or program.
    A. In addition to its application as a valuation tool, the framework can be used to guide effective approaches to establishing and undertaking research;
    B. While the framework has been developed and tested for conservation research, many of the measures are potentially applicable, or adaptable to a wider range of environmental research disciplines and programs. This is something that could be explored with further work.

V. This report illustrates a range of ways the data collected using these approaches can be analysed to illustrate value, including:
    A. Aggregate analysis across a suite of projects
    B. Analysis by project
    C. Analysis by institution of respondents (project team members in research institutions versus those in other agencies)
    D. In depth analysis across particular matters of interest.

VI. In addition to providing an approach to assessing research value, the report sheds light on the processes and approaches that can help achieve value in research, including insights on:
    A. Tailoring research engagement, collaboration and co-design activities so these are fit-for-purpose
    B. Building effective relationships and generating trust
    C. How close relationships enable and generate research value
    D. The value of different kinds of research outputs

VII. There was strong agreement between non-academic and academic respondents on the value and importance of most measures. Non-academic respondents were more likely to value measures for:

Collaboration:
- Publications and other outputs produced in a collaborative manner
- Links between research team and key stakeholders improved

Different 'knowledge systems' and intercultural capacities:
- Methods drew on insights from different knowledge systems
- Different knowledge systems informed the outputs of the project
- Intercultural capacities of the research team have improved

Threatening processes:
- Reduction in threatening processes

Academic respondents were more likely to value measures for:

Monitoring and data management:
- Improved monitoring for species, ecosystems, heritage places, etc
- Improved data management, reporting and/or analysis about species, ecosystems, heritage places, etc

Tangible improvements for species, ecosystems and places:
- increase in population trajectory for threatened or significant species
- an improvement in condition of an ecosystem, wetland or marine environment

Public awareness:
- greater public awareness of the importance/ challenges of conservation

VIII. The major actions and recommendations arising from this research are summarised on the following page and discussed in detail at the end of the report.

IX. A table outlining what to consider for different assessment and valuation strategies follows after the actions and recommendations. The full framework is provided on pages 135-166 at the end of this report.

Summary of actions and recommendations

Assessing the value of research

1. Adopt a definition of research value that encompasses traditional aspects or research impact as well as the benefits arising from all stages of the impact pathway.
2. Approach research valuation critically and carefully when developing measures, valuation methods and approaches, drawing insights based on research such as this and/or independent tests of their importance, usefulness, fitness for purpose, feasibility for the circumstances, effectiveness in delivering meaningful insights, reliability and consistency, how they should be applied and any caveats surrounding their application.
3. Use caution in applying the framework to any circumstances in which Indigenous people are or should be considered critical partners in research projects

Applying this research valuation framework

4. Use the framework and measures to assess conservation research projects and programs, across their life cycle and across a diverse range of measures.
5. Draw on insights from the framework and tools such as the glossary to design a suite of measures suited to context, research phase and outcomes sought.
6. Give attention to context and baselines.
7. Where needed, draw on indicators from and beyond the wider Phase 1 suite as appropriate to reflect research value in a wide range of contexts.

Research practice and resourcing

8. Research funding bodies, researchers and research program managers: Invest time and resources in developing appropriate processes and relationships, including dedicated facilitation, brokering and valuation roles.
9. Researchers and research users: Co-develop research through early conversations and broad engagement of end-users and stakeholders.
10. Researchers, research users and research program managers: recognise and draw insights from the complexity and nuance behind what builds effective research relationships.

Further research

11. Undertake further research to improve research valuation methods across the research sector, including:
    a. The need for an Indigenous-led strategy for developing culturally appropriate and meaningful approaches to research valuation.
    b. Testing of quantitative, documentary and community insight measures and further testing of the survey measures with a wider range of projects to improve statistical assessment of reliability, stability and convergence of these measures.
    c. Specific research to test interpretations and meanings of important but potentially ambiguous concepts such as drawing on ‘different knowledge systems’, and to develop appropriate language and methods for asking these questions in research valuation contexts.
    d. Testing the extension of the framework to other environmental research contexts, and comparing with other ground-up methods for assessing research value in other disciplines, to identify whether overarching measures and approaches can be used as framing measures for research valuation across all disciplines.

Table i. Conceptual and practical considerations for applying the framework to different valuation circumstances

Conceptual considerations	Considerations for implementation
High level considerations: ● Select a diversity of measures from across the program logic ● Consider using pairs or sets of measures that provide information about alternative pathways for delivering outcomes; e.g. to assess the effectiveness of projects that are implemented in different ways, ask both about whether "research was implemented in partnership" and whether "frequent, quality communication" took place) First order measures: ● Prioritise measures that align with identified program/project outcomes (particularly for outcomes and impact measures, but may also apply to measures in other groups, e.g. data made available) ● Prioritise outputs measures that most align with the needs of research users (either through testing these directly as part of research co-conception, or by drawing on the general insights from studies such as this) Second order measures: ● Consider an array of second order measures that may apply to some projects, and establish processes for identifying subsets of projects to which these apply. (e.g. intercultural capacities, secondary (economic, social, community) outcomes/impacts or 'co-benefits') ● Consider additional outputs measures based on feasibility and value for program considerations (e.g. outputs that speak to research quality (academic metrics) or broad discoverability (dissemination metrics)	High level considerations: ● Select measures suited to the assessment strategies available at different points within the research process ● Allocate resources at the outset for data collection and analysis related to assessment of research value Regular reporting measures: ● Select more feasible measures to minimise effort ● Focus on more reliable self-report measures that can be assessed by research teams ● Focus on quantitative and documentary measures that lend themselves to consistent interpretations Major assessment milestones: ● Select a range of narrative, survey, quantitative and documentary measures ● Design a sampling strategy for saturation within project teams where feasible ● Invite reporting on more complex community, modelled or monitoring measures where these have been gathered as part of research design Implementation in practice: ● Use the glossary provided as a tool to help clarify unfamiliar terms in surveys and interviews, and to strengthen the facility of participants with research valuation concepts.

Glossary of key terms

Project teams: groups of people working together on research projects, including research users and collaborators as well as researchers from academic and research institutions.

Research cycle: the entire process of a research project, from ideas-forming and conception through to realisation of impacts; includes ongoing and final evaluations, feedback loops and adaptive strategies used throughout; likely to involve complex, non-linear relationships between different elements of the program logic.

Program logic: a means of conceptualising how change occurs through projects or programs by understanding different elements of inputs, processes, outputs, outcomes and impacts. These elements can occur at any point in a research cycle, and the relationships between these are likely to be complex and non-linear.

> **Inputs**: the resources available to a project.
>
> **Processes**: activities undertaken in the projects; implementation
>
> **Outputs**: what is produced directly from the project
>
> **Outcomes**: what has changed as a direct result of the use of the outputs produced; including immediate outcomes such as greater awareness, knowledge and skills and ultimate outcomes including changes to policy or management.
>
> **Impacts**: contribution of the project to wider and longer-term cultural, social, economic and environmental trends.

Value/research value: the overall benefit of research as understood with respect to all dimensions and aspects of the research cycle and across the program logic.

Indicator: a gauge of the value of research through different aspects of the research program logic.

Measure: means by which information is gathered to report against indicators (e.g. questions, data gathered).

> **Narrative measures:** are those that give long-form, descriptive and interpretive answers, usually in interviews or focus groups, but also potentially collected by other means (e.g. in surveys).
>
> **Survey measures:** are those that lend themselves to survey instruments, including multiple choice, likert-scale measures (e.g. multi-point scales measuring agreement) and short form text measures
>
> **Compiled measures:** measures related to project or program operations or benefits that can be compiled from documentary or quantitative information
>
> > **Quantitative measures**: are those compiled measures about a project or program that can be assigned a numerical value; they may potentially be used for aggregate or comparative purposes in some cases where care is taken to ensure consistency in how measures are assessed and quantified.
> >
> > **Documentary measures:** are those that can be sourced and demonstrated (such as a citation in a policy strategy, changes in social programs, new policies for inclusive decision-making, regulatory changes granting access to resources); such changes are not necessarily readily quantifiable but some could potentially (with some care) be aggregated and reported across large numbers of projects (e.g. number of policy changes across a program).
>
> **Community insight measures** may include quantifiable, semi-quantifiable (likert) measures and narrative approaches, but are designed to provide insights beyond the project team, suggesting a level of independent assessment beyond the project relationships and thus going beyond the survey and narrative measures discussed.
>
> **Modelled or monitored measures** are those measures that assess outcomes or impacts on environmental, economic or social values in a quantitative way.

Knowledge systems: different ways of understanding, knowing about or approaching questions, including Indigenous knowledge systems and western science, interdisciplinarity and other non-academic knowledge systems and sources of expertise, e.g. natural resource managers, farmers, etc.

Contents

Introduction

The idea for this project grew out of the challenge identified within the National Environmental Science Program (NESP) of developing effective measures for assessing the value of applied environmental research of the kind the program directly supports. Despite a global move across the research sector and funding bodies towards identifying and assessing research impact, the most common practice in assessing research outcomes across programs still rests in narrative case studies -- which have limitations in being selective rather than comprehensive -- and in established measures of academic success, such as citations and impact factors, which are arguably not a strong indicator of the value of research for non-academic purposes nor its potential or realised application in non-research contexts. Approaches to assessing impact within NESP draw on a similar pairing of narrative case studies with a somewhat wider breadth of quantitative measures (from Indigenous employment figures and numbers of students, to accessibility of research outputs). In the first phase of the NESP (2014-15 to 2020-21), this was complemented with an independent evaluation, which involved interviews and surveys used to collect qualitative insights and some quantitative data (for example, on frequency of contact between research teams and research users) (Charterpoint, 2018). While broader than the suite of measures promoted in other contexts such as the Australian Research Council (Australian Research Council, 2017), the reporting measures used in NESP do not fully reflect the breadth of a research program or the extent of the value it has realised.

Few approaches have been developed that focus on apprehending the array of complex and uncertain pathways that in practice often contribute to outcomes from environmental research, or on viable quantitative measures that could support aggregation of project-based assessments across projects and programs. Indeed, when the need for quantitative measures of conservation research value is discussed, many widely accepted approaches focus on measures of outcomes and impacts that are often not realised until well after research projects are complete, and are confounded by a wide array of factors beyond the research team's control. On the other hand, the strengths of narrative approaches in offering a detailed, nuanced and more holistic picture of research value and its realisation have not been fully explored in a way that would encourage their systematic use across various projects, including those that were not obvious cases of success.

The project set out to explicitly develop an approach to assessing and valuing environmental research that could address these challenges. The aim was to build and test a suite of indicators and measures across the entire value chain of research, and from the ground up. In order to be able to identify common measures that could apply across a wide range of projects, this complex challenge was broken down by focusing on conservation research. A suite of ready-to-use measures were generated from this place of common ground, that can be drawn on to assess conservation research projects and programs in a wide range of contexts. Using conservation research as a focus, the project aims to demonstrate proof of concept for developing an approach to assessing research value capable of dealing with many of the subtleties and complexities of achieving that value in practice. Our work is a building block in the wider endeavour to develop methods for assessing research impact that is taking place across a broad array of scholarly disciplines (e.g. Papageorgiou. et al, 2021). We also offer some indicative directions for extending the measures and approaches here to other research contexts, and in particular to research in other environmental disciplines.

The ambition of the project at the outset was to develop and test a framework and range of approaches to assessing research value that could:

- Provide a nuanced and timely picture through assessing research collaborations holistically, from establishment through process, to outputs and outcomes;
- Be applied (at least in part) at any stage throughout the life of a research project or program;
- Offer a suite of quantitative, qualitative, structured, likert-style and narrative measures and approaches for assessing the value of research;
- Suggest a range of options for assessing the value of research that can be readily applied in practice for a range of purposes;
- Provide some assessment of the uses and applicability of different measures and approaches, including limitations to their applicability and reliability;
- Recognise the complexity of how research becomes valuable in practice, including the non-linear pathways through which research value is often realised, the unexpected outcomes that can arise throughout the lifetime of research projects, and the many benefits ('co-benefits') of research pathway(s) when research collaborations and processes are undertaken effectively, including by strengthening research partnerships themselves.

While the primary purpose of this project has been to inform a conversation on how research programs and institutions measure and assess the value of research, the findings from this project can also be drawn on to give insights into what makes for effective research processes and partnerships. The framework and suite of indicators has been developed by drawing on insights from established good practice in research processes and partnerships. Thus, findings from the testing phase of this project deepen our understanding of what makes research processes and partnerships effective. The findings from the research can thus be used by those who regularly participate in setting up and undertaking research collaborations - researchers and research users alike - to guide more effective and productive research processes and partnerships.

Improving the value of research investment and how research is incorporated in policy and practice rests on understanding the complex pathways through which outcomes and impacts from research can be achieved. The importance of developing effective measures for valuing research goes well beyond ensuring greater accountability for public investment in research. Evaluation measures, when applied to research funding, can directly impact the processes and pathways of research, provide rewards and recognition for researchers who meet the measures, and may motivate particular kinds of research practice over others (Wilsdon et al. 2015). Use of the wrong measures, or a poor application of measures, has the potential to deliver perverse outcomes, presenting significant risks to the research sector, to funders, to researchers and research users, and ultimately to the benefits (social, cultural, environmental and economic) that research otherwise aims to deliver. On the other hand, increased focus on effective measures for delivering research value can ensure that real needs are prioritised for research attention, and that research is inclusive in design, communicated broadly, translated effectively, and has long-term benefits (e.g., Gibbons et al. 2008).

In this light, it is important to note a significant limitation of this research from the outset: that this framework has not been developed for use in research partnerships primarily involving collaborations with Indigenous communities and partners. In particular, due to COVID-related delays and the movement of key personnel, the original goal to begin work towards an Indigenous-guided approach

to assessing research value, likewise built from the ground up, and drawing together insights through case studies involving partnerships with Traditional Owner communities, could not in the end be realised. While contributions from Indigenous participants in the workshops and surveys have shaped the development of the framework, this work in its current form should not be considered culturally appropriate for assessing the value of Indigenous-led research and collaborations with significant Indigenous involvement. Significant work has been done elsewhere in this space, including on establishing effective research collaborations with Indigenous communities for research on Country and species (e.g. Moggridge 2020, Woodward et al. 2020) and on Indigenous-led frameworks for assessing the value of research (e.g. Tsey et al. 2016, 2019). Insights from that research, and from Indigenous co-authors, have been drawn on here to comment more fully on these limitations of the framework and give some insights into possible pathways forward. This is addressed in the next chapter in the section on "Indigenous approaches to evaluation and a caution on applying these measures to Indigenous partnerships and research", and returned to in discussion and conclusions.

Chapter 1. Context

Globally and in Australia, there is growing interest in tools and approaches for assessing and measuring the impact of research and improving the quality of research partnerships. Internationally, there have been a proliferation of tools and studies on research impact assessment methods in academic contexts and for use by research funding bodies (Kuruvilla et al. 2006, Ovseiko et al. 2012, Wilsdon et al. 2015, REF 2019, Mervis 2020). In Australia, research funding bodies, including the Australian Research Council, the National Health and Medical Research Council and the National Environmental Science Program have expanded their focus in recent years on both assessing research impact and on directing funding on the basis of proposed positive impact and assessment of plausible pathways to impact (Watt 2016, Deeming et al. 2017, Australian Research Council 2019a, b).

This focus on assessing, measuring and reporting on the value of research is motivated by both the need for accountability and a desire to focus research efforts more toward meeting societal needs, both in terms of contributing to the 'public good' and to achieve innovation and contribute to economic value. Among funding bodies and governments, there is a perceived need for research that is more closely targeted towards social and government priorities, and for growing accountability for public investment in research (e.g. Watt 2016). Among researchers, the motivation for an expanded focus on research impact reflects a strong desire to understand how research can best benefit humanity -- but also from increased competition for funding, mounting public pressure to justify research, a decline in research investment, and the short-term nature of funding cycles (Weiss 2007, Wilsdon et al. 2015, Hourihan and Parkes 2019).

These new approaches have attempted to move beyond traditional measures of academic impact, dominated by peer-reviewed publications, journal impact factors, article citation counts, and attracting (further) research funding (Radicchia et al. 2008, Scharnhorst and Garfield 2010, Vieira and Gomes 2010, Penfield et al. 2014, Australian Research Council 2019a), towards more holistic approaches to assessing both impact and research engagement. Such approaches increasingly seek to understand and assess how effectively research translates across diverse social, environmental, and economic contexts and institutions to generate tangible practical benefits (Penfield et al. 2014, Reale et al. 2017, REF 2019, Australian Research Council 2019a). These more holistic approaches at the moment rely heavily on the 'impact case study' model, a model which combines multiple streams of evidence into a story demonstrating how research findings and publications are being or could be used to inform domains outside of the research sector, such as policy or practice. Impact case studies often include narrative descriptions of impact pathways and outcomes from research, research-user testimony, evidence of citations in non-research outputs (such as policy documents), media and social media reach and other disparate sources.

Impact case studies are an incredibly valuable component in the repertoire of tools for assessing the value of research. They arguably provide a much more comprehensive and nuanced picture of the value of particular research projects for a much wider range of research users and purposes than traditional academic measures. However, impact case study reporting is generally limited by focusing exclusively on telling 'the best' stories, with little or no attention to comprehensively assessing the value of research across an entire program or institution. In particular, exemplary case studies do not generally attend to where (and to what extent) research across a program has had little positive

impact, or may have fallen short of realising its potential value. By 'cherry-picking' the best or most impressive stories, research impact case study processes can also often unconsciously reinforce existing institutional biases that favour more senior and better-represented demographics among researchers (Wilsdon et al. 2015). Furthermore, narrative approaches alone present limited scope for large-scale comparison across research projects, programs or disciplines.

Compounding the challenges of measuring and assessing the value of research, many of the tangible benefits that are most widely thought of when 'research impact' is discussed, such as changes in environmental, social, cultural or economic measures, are not likely to be felt for years after project completion. Even where research has a positive impact in one or more of these domains, many factors are likely to be present that confound measurement of the impact. In the case of threatened species research, for example, there may be time lags in species' responses to conservation actions (the phenomenon of 'extinction debt' - the likelihood of extinction of a species as a result of acts already taken in the past; Kuussaari et al. 2009 - and 'colonisation (or immigration, or species) credit', a delay in a species responding to a positive change; Hanski 2000; Jackson and Sax 2010). This may mean that species trajectories fail to improve, or even continue to decline, within the timeframes out from research projects likely to be measurable, regardless of any success a research project may have had in providing pathways to support species recovery (Watts et al. 2020). These intrinsic time lags are coupled with the more well-recognised confounding issues: that implementing meaningful changes is rarely fully in the control of the research team or even research partners; that multiple, interlinked projects can make it difficult to trace the impacts of any one project (Penfield et al. 2014, CSIRO 2020); and that changes often take significant time to implement, especially when these changes are complex (Morris et al. 2011, Pannell et al. 2018, see also Tsey et al. 2019), making it extremely difficult to pinpoint which contributing influences are most important (Barnett and Gregorowski 2013). This array of confounding issues makes tracing research value through impact measures alone particularly challenging, especially within the relatively short timeframes over which research impact is usually measured.

In response to this array of challenges, this project set out to create and test a set of approaches to valuing conservation research that could contribute to our conceptualisation of research value, and how it is measured and tested. The project builds on previous work that has begun to look more holistically at the value of research across the complex pathways and practice of research collaborations (e.g. Tsey et al. 2016, O'Connor et al. 2019, Pannell et al. 2018, Davila et al. 2016). This work provides an important foundation for conceptualising approaches to assessing research impact, including in environmental contexts. The more holistic approaches proposed in this literature have the potential to provide a more complete picture of the value of research collaborations, including insights into the potential or realised outcomes and impacts from research that are not visible when assessing impacts alone.

A narrative literature review undertaken in phase 1 of the project (Lavery et al. 2021) supported the development of the framework and identification of a preliminary suite of indicators to be measured and tested, drawn from other disciplines and from non-discipline-specific frameworks. It also explored more specific, tested frameworks and measures designed to evaluate conservation practice (e.g. Kapos et al. 2008) to learn how these might be applied to assessing the value of research. Twelve publications were found that included assessment frameworks or conceptual diagrams

tailored toward measuring impacts from conservation research, or conservation management projects. However, many of the impacts measured focused on explicit targets in conservation rather than taking a more holistic approach across inputs, processes, outputs, outcomes, and impacts. An additional eight publications had generic impact assessment frameworks tailored across all scientific disciplines, and another 12 publications had frameworks associated with related disciplines in environmental science and knowledge co-production (e.g. sustainability research, climate science, knowledge co-production, transdisciplinary research, forestry, agriculture). Key gaps among these existing frameworks and toolkits include that:

- many existing frameworks for measuring research impact are written at a high level of generality in order to be applicable to a wide range of research contexts and programs. Thus, they lack the specificity that research program managers and research partners likely require to understand the value of research, and/or they require further development to be applied in practice;
- few have been designed or tested with environmental or conservation research in mind; and
- those that were designed for environmental, conservation and land management research largely focus on measuring outcomes and impacts (although see Pannell et al. 2018 for an example that focuses on activities and processes for achieving policy outcomes).

A summary of previous frameworks reviewed, and an assessment of the aspects of the research process to which they apply, can be found at Appendix C.

1.1 Caution on applying these measures to Indigenous partnerships and research

One of the major limitations of this study is that it was not possible within the compressed timelines of the research caused by COVID-19 to develop the Indigenous-focused case studies that were part of the original conception of the research. While there were Indigenous participants in the expert workshops and surveys, this does not represent in any way an adequate process for gaining insights into the research from diverse Indigenous perspectives into what could be considered culturally-appropriate approaches to understanding and assessing research value. It is therefore not appropriate to apply the framework, measures and approaches presented here to contexts where Indigenous communities, organisations or individuals are directly involved in research as important research users, partners, participants and/or knowledge holders.

Importantly, there is also potential here for perverse incentives and structural disadvantages to be perpetuated in research spaces if culturally inappropriate measures continue to be applied to assessing or establishing research value. Given the long history of Indigenous concerns over the role of research in perpetuating colonial violence and structural disadvantage, there is an urgent need to address the inadequacies of current approaches to assessing research value from diverse Indigenous perspectives, and to create Indigenous-led approaches to evaluating research and to rethinking the incentive systems of research as it currently exists. This is a significant undertaking, and one which will take substantial care, time and resources. While addressing this fully was always likely to be beyond the scope of this project, some considerations of the implications of this for interpreting the framework presented here, and some possible steps towards this broader goal, are discussed further in chapter 6.6.

Chapter 2. Methods

The project was developed in two phases. The first phase involved expert workshops, a literature review and survey to scope conceptual frameworks for assessing the value of environmental research, and identify a suite of indicators that could be applied to measuring research inputs, processes, outputs, outcomes and impacts. The second involved developing and testing approaches and measures against these indicators, using NESP Threatened Species Recovery (TSR) Hub projects as case studies. NESP provides a valuable test-case for developing and refining environmental research change measures, since the program is focused on applied research to address environmental challenges for policy-makers, environmental managers or other research users. The program actively integrates many components found to be effective in achieving research adoption, such as co-design principles, a focus on stakeholder engagement, and dedicated knowledge brokering and communication resources.

2.1 Phase 1 methods: development and refinement of the framework and indicator suite

This project was first developed as part of the Early Career Researcher conference of the NESP TSR Hub in 2018. The parameters for this project and the broad framing principles for this research were identified as part of a workshop at the conference, attended by hub researchers, leaders and research users (see Box 1). Drawing on the conceptual framework piloted by van Kerkhoff and colleagues (van Kerkhoff, 2018), this workshop identified the need for an approach that assessed research value across the full spectrum of research activity, from (co-)conception to outputs, adoption and impact, and across multiple dimensions of measurement, from changes in narrative understandings that may have emerged through the project, to identifying and assessing the mechanisms developed for research co-production and delivery, through to an array fully quantifiable measures, from social media reach to research uptake into management or policy documents (see Table 2.1). The framework was conceptually refined and tested from participants' knowledge of research collaborations and partnerships, and an initial set of ideas for indicators identified that could be used to assess research projects against each of these dimensions.

Box 1. Initial framing questions from ECR workshop

1. What does TSR Hub research suggest are effective measures of success by which to evaluate the impact of research and improvements to monitoring and management for threatened species?

2. Can we identify intermediate measures that allow us to realistically assess environmental impacts of research in conservation monitoring and management (e.g. for threatened species) in an ongoing way?

3. Do researchers and research users agree on what are good measures of success, and what are successful projects?

This first workshop identified the importance of developing diversified and mixed-methods approaches to assessing the value of research combining narrative, quantitative, survey-based and semi-quantitative documentary measures, and that could draw together a full picture of research projects that could be used for comparing and assessing across disparate research projects. Several risks and concerns were also identified through this initial workshop,

including the risk that an undiscriminating application of a wide array of measures of potentially marginal value (e.g. social media statistics) risked implying to researchers and partners that all research projects must confirm to a fixed set of expectations to be considered 'impactful' (e.g. regarding social media usage in research projects). Drawing on insights from the workshop, the project team then developed and refined the scope, purpose and approach of the in-depth project in collaboration with key departmental research users.

Table 2.1. High-level framework conceptualisation from initial project workshop 2018

	Building Understanding	Shaping Options	Taking Actions	Pathway to Environmental Impact
Narratives	Reciprocal understanding, common framing and scope of research	Accessible findings and clarified options	Research findings adopted in monitoring management, regulation, etc.	New knowledge transforms understanding of other processes and/or enhances social connection
Mechanisms	Networks established and/or joint research undertaken	Research disseminated through relevant networks and/or options discussed in relevant forums	New mechanisms established for ongoing implementation, wider uptake, includes adaptive management	Wide and/or long-term mechanisms and processes adopted and resources, wide networks engaged/ established
Numbers/ Metrics	Quantifiable/ categorical evidence of shared understanding and reach of networks	Evidence of reach or of engagement with research findings	Evidence of changes to management, policy, actions; measures of species, and ecological and social values affected by these changes	Improvement in environmental (and social) measures

The first phase of the project concluded with a literature review and survey of practitioners working at the interface of conservation research and practice to refine the framework and identify a fuller suite of indicators for development and testing. The narrative literature review drew on peer-reviewed articles, books, and reports from government agencies and research institutions, with the aim of identifying and assessing tools, frameworks, and diagrams aiming to measure impacts in conservation research, and of extracting a list of potential qualitative and quantitative impact indicators. A logic model of inputs, processes, outputs, outcomes and impact was developed from the literature and organised into 'Domains' and 'Subdomains', corresponding to common dimensions and themes of research engagement and impact.

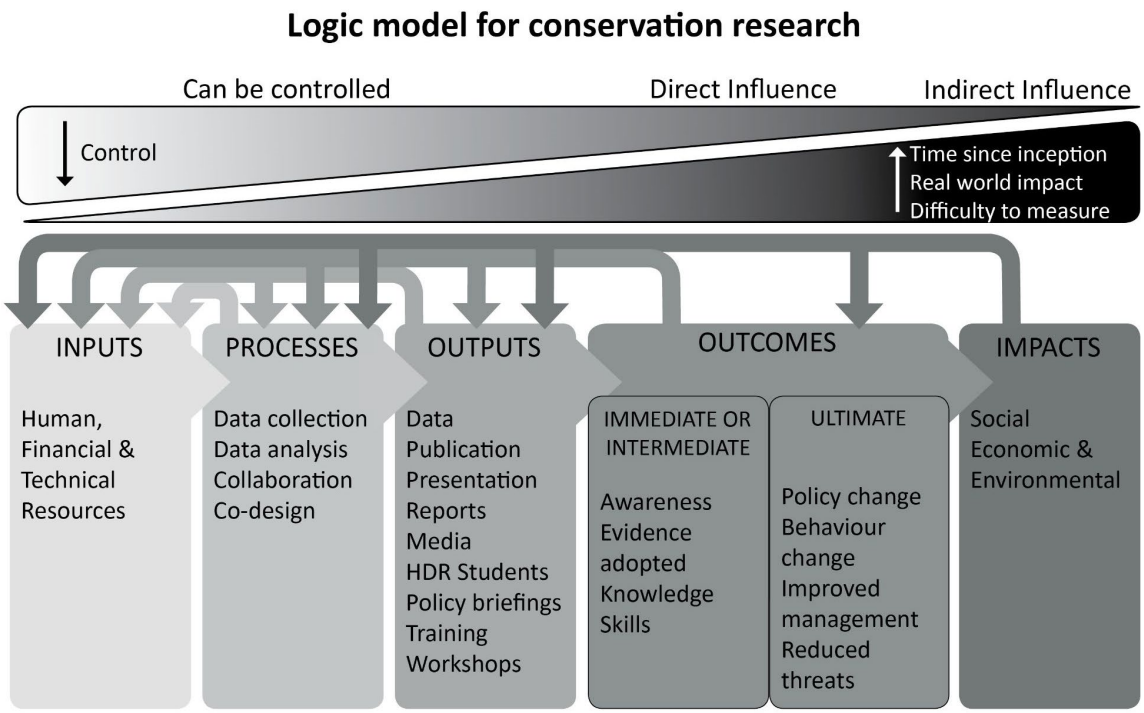


Table 2.2. Framework of domains and subdomains of indices developed in project Phase 1. Indicators can be both quantitative and qualitative thus the number of indicators within each subdomain may be smaller than the individual quantitative or qualitative indicators themselves.

Domains	Subdomains	# of indicators	# of quantitative indicators	# of qualitative indicators
Domain 1. Research Inputs	Subdomain 1.1 Research Direction and Resources (RDR)	13	6	13
Domain 2. Research Processes	Subdomain 2.1 Research Methods (RM)	3	0	3
	Subdomain 2.2. Research Management & Conduct (RMC)	12	5	11
Domain 3. Research Outputs	Subdomain 3.1 Academic Outputs (AO)	13	9	9
	Subdomain 3.2 Tool Development (TD)	3	3	3
	Subdomain 3.3. Research Dissemination (RD)	11	10	5
Domain 4. Research Outcomes	Subdomain 4.1 Research Uptake (RU)	3	3	1
	Subdomain 4.2 Increased Awareness & Responses (AR)	5	1	5
	Subdomain 4.3 Community & Stakeholder Engagement (CSE)	7	5	6
	Subdomain 4.4 On-ground Action (OGA)	6	5	6
	Subdomain 4.5 Public & Private Policy Development (PPP)	7	1	7
Domain 5. Research Environmental Impacts		13	6	12

The logic model is not intended to represent linear phases in a research cycle; rather, it is a means of conceptualising how change occurs throughout projects or programs. Any of these domains can in theory apply at almost any stage, often multiple times in different ways: outputs can be generated at the start leading to substantial on-ground changes; project stakeholder networks may coalesce late in a project, or re-form part way through. Projects often involve ongoing evaluations, adjustments, feedback loops and adaptive strategies, and thus the relationships between these elements are likely to be complex and non-linear.

A total of 65 conservation researchers and practitioners from research, non-governmental organisation, governmental, Indigenous and other organisations, who work closely at the interface of research and practice, assessed the relevance and comprehensiveness of our potential indicator list. A preliminary list of 76 potential research impact indicators were identified through the literature

review, and put to participants in a survey. Participants were asked to rank the candidate indicators for importance ('*How important is it to measure each of these items to understand the impacts of conservation research?*') using a four-measure Likert scale with no midpoint for a range of potential indicators (*1, irrelevant; 2, not important; 3, important; 4, very important*). Additionally, we asked them to respond to the open ended question of '*Are there any other items not included in the current list that are important for measuring the impact of conservation research?*' A further 20 recommended indicators were identified by survey participants, resulting in a final suite of 96 impact indicators.

A practitioner workshop of Australian-based applied research practitioners and conservation professionals then workshopped the indicators to identify potential options for and challenges involved in applying these indicators in practice. A subset of workshop participants also ranked these indicators for feasibility (n=12). This provided a broad indication of where the feasibility of measuring indicators may be a challenge (four-measure Likert scale). The feasibility measures suggested many areas where those indicators ranked as being of the highest importance (e.g. many of the impact measures) present substantial potential challenges for measurement, but also helped identify a subset of candidate indicators that could be both feasible to measure and assessed as relatively useful as a measure of research value. Both the feasibility and the wider discussion points raised at the workshop shaped the selection of indicators for testing through case studies in Phase 2 of the project.

2.2 Phase 2 methods: applying and testing the indicators in practice

The second phase of the project focussed on testing and refining approaches to assessing impact of environmental research in more depth. Specifically we sought to address the following research questions:

a) What are suitable ways to measure the indicators?

b) How well do the indicators capture the value of conservation research projects?

c) Do project participants from different types of institutions value similar aspects of conservation research?

Testing all of the indicators developed in phase 1 was beyond the scope and resources of the project. Instead we selected and tested a subsample. Both survey and interview data were collected under the human research ethics protocol approved for phase 1 (2019-895).

Case studies

Case study projects were drawn from a range of conservation research projects from within the NESP Threatened Species Recovery Hub, that had been conceived of, designed and implemented with a focus on research partnerships, uptake and impact. The nine case study projects were identified according to a suite of project characteristics including:

- Objectives. We included projects directed primarily at land management outcomes or other on-ground changes; projects designed primarily for policy outcomes; and projects that had more innovative or unusual approaches to achieving conservation outcomes.

- Diversity of research users and participants. We focused on three common "types" of research participants based on institutional affiliation: people based in research institutions (universities); people based in government agencies; and people based in non-government conservation organisations.
- Geographic spread. We attempted to cover several states and territories.

Project cases were not chosen based on preconceived ideas about whether or not they had been successful in achieving impact.

Indicator selection

A draft subset of indicators was selected based on several criteria. We gave preference to indicators ranked as important in phase 1, but also sought to cover a diversity of indicator types. In particular, we wanted to adequately capture the spectrum of activities and values along impact pathways that would make assessing research outcomes meaningful from different institutional and cultural perspectives. Several key themes repeated across indicator domains were either amalgamated, or one indicator was chosen to avoid duplication. Indicators were also amalgamated across several of the outcome and impact domains in order to organise the questions for greater comprehension (e.g. participants were provided a single frame of check-box questions addressing changes to monitoring, management, adaptive management, data management and threat reduction). Attention was given to particular themes that emerged repeatedly in the research impact literature, for example whether processes were collaborative and whether multiple knowledge systems were drawn upon.

This first subset of indicators was refined, and corresponding approaches and draft measures (questions) developed at a small workshop among the research team. The key mechanisms identified for testing the indicators included:

- Qualitative interviews and network-mapping;
- A semi-quantitative survey;
- Quantitative, documentary and community insight measures. These were not tested in the case studies. Insights from workshops in phase 1 were drawn together and developed further by identifying contemporary applications where similar measures have been applied. These were then assessed by three members of the project team with experience applying such measures in practice.

Table 2.3: Example of narrative, survey and documentary measures developed for an indicator

Indicator descriptor	Measures/questions addressing indicators	Type of measure	Indicator code*
Research needs, gaps, unanswered questions, new areas for research identified	How were research needs identified in the project?	Interview: narrative	RDR1
	Research needs or new areas of research were identified through the project	Survey: likert agreement and fit for purpose	
	Documentary evidence of research gaps that have been identified and prioritised, as well as those met through the project	Documentary (assessed, not tested)	

Forty six indicators were selected and tested through interviews and surveys and 53 developed as quantitative, documentary or community insight measures. The suite of indicators and measures tested or assessed is summarised in Appendix A. A summary of the full framework, along with

findings and considerations for the application of measures in practice is presented on pages 135-166 at the end of this report.

Developing the survey instrument

The survey instrument is reproduced in Appendix D.The team workshop identified several pre-cursor questions that could help streamline surveys and interviews e.g. asking what types of organisations participated in the project. The survey started with generic questions to provide some predictive variables (such as what type of institution the respondent worked for) and then sequentially went through the five domains of an impact pathway. We attempted to strike a balance between eliciting comprehensive and useful information, whilst keeping the survey short enough for people to answer every question. Questions weren't compulsory, except for the first question asking the respondent to identify which research project they would reference in the survey. This was in recognition of the fact that many participants in Hub research had been active in multiple projects. We asked respondents to refer only to their nominated project (suggested in the recruitment email). Institutional affiliation should also have been compulsory but fortunately all respondents answered this anyway. We also used skip logic to ensure respondents weren't asked irrelevant questions (e.g. how Indigenous consent was obtained if there were no Indigenous participants).

A combination of multiple choice, Likert-type questions and a text based open comment box for each of the five domains were used. We used a five point Likert scale with the options: Strongly Agree; Agree; Neither agree or disagree; Disagree; Strongly Disagree; and N/A or Don't know. We used N/A or don't know in recognition that not all research participants will be familiar with all aspects of a project and not all options would be applicable to all projects. Use of N/A in this circumstance will likely increase validity of the data by reducing selection of the midpoint when the question does not apply (Chyung, Roberts et al 2017).

The final question for each research domain (inputs, process, etc) related to how fit for purpose the measures (ie preceding questions) were: *"How fit for purpose do you think each of these [input] indicators are in terms of capturing the value of the research project? Note: an indicator may be important even if you think your project performed poorly against it."*

Given that many projects had not finished (although most were in their final stages), we expected that many outcomes and impacts may not yet have manifested. To accommodate this situation we phrased questions to allow respondents to speculate e.g. "*Please select which of the following outcomes related to on-ground practice your project has, or is likely to contribute*". For the impacts domain we also asked respondents to give a rough estimate of their confidence in their predictions. Pretesting of the survey was done by two project members outside the design team. Subsequent pilot testing with a subsample of the target population was not done because of restricted resources but is highly recommended (see Wardropper et al., 2021).

Recruitment delivering the survey instrument

The project leaders from each project identified people who had been involved in their projects that we could contact to recruit in the survey and or interview. The survey was delivered online via the Qualtrics platform with a link sent to 47 email addresses of TSR Hub research participants.

Development of the interview guide

Interview questions covered all 5 impact domains and were ordered according to an impact pathway (from inputs to impacts). We focussed on indicators that had been identified as likely best measured through a narrative response (e.g. thoughts about the role of trust in research) rather than more quantitative measures (e.g. how many outputs were produced). Interviews were semi-structured using an interview guide with 20 standard questions (Appendix E) that were altered depending on responses. Questions were piloted on two Hub researchers not in the case study projects and questions were slightly re-ordered and re-worded as a result. All interviews were conducted by one of the authors either via zoom/teams video (9) or phone (1) and ranged from 39 to 67 minutes duration.

All interviewees consented to their interview being recorded. Recorded interviews were transcribed automatically using the Otter.ai software package (Otter.ai, 2021). The rough transcripts were then edited and corrected by listening to the recordings while following the transcript.

Analysis of results

Interview transcripts were analysed using qualitative content analysis (Schreier, 2012) using QSR International's NVivo 12 Software for coding. All coding was done by the same author who conducted the interviews.The coding frame was based on the 20 interview questions given our primary interest was assessing the suitability of the impact measures. Further coding was not possible in the available time. Instead key themes in the responses to each question were summarised and aggregated. The percentage of transcript contributed by each answer and most common words were also calculated in Nvivo. Percentages are a relative guide given transcripts also included text from the interviewer.

Measures used in the survey were analysed in two primary ways in order to assess their viability and gain insights into their application in practice.

Fitness for purpose measures: respondents scored each measure as: highly fit for purpose (3); medium fit for purpose (2); or low fit for purpose (1) for their nominate project. They also had the choice of not applicable and these responses were omitted from analyses. Fitness for purpose measures were averaged across all academic (n=11) and non-academic respondents (n=7). Averages were then ranked into 4 equal categories from lowest (average fit for purpose scores 1 - 1.5) to highest (>2.5 – 3).

Average fit for purpose score	Ranking
>2.5 - 3	High
>2 – 2.5	Moderately high
>1.5 - 2	Moderately low
1-1.5	Low

Four projects had responses from both academic and non-academic respondents. Responses from this subset of four projects (n=13) were used to test **variability** and **within-project concordance.** The respondents' evaluation of their projects against different measures were as follows:

- Measures that tested likert agreement: Strongly Disagree=1; Disagree=2; Neither agree or disagree=3; Agree= 4; Strongly Agree =5; N/A or Don't know omitted/blank.
- Multiple choice measures that tested presence/absence: 1=present, 0=absent
- Likert likelihood: Highly unlikely=1; Unlikely=2; Possible=3; Likely=4; Almost certain/certain=5; N/A or Don't know omitted/blank
- Confidence measures: Limited=1; Medium=2; Robust=3

Variability was assessed by taking a within-project average for each of the measures, then assessing how often, how widely, and in what direction academic researchers and non-academic research users and partners departed from this average (individual "departure" results): e.g. if the project average was 3.4 and an individual project team member rated this measure 2, their individual departure result would be -1.4. The standard deviation of these "departure" results provides an indication of whether each research team was fairly consistent in the relative rating they gave projects, or whether their answers departed from project averages in markedly different ways. Variability was assessed proportional to the scale of the measure. For five-point likert-scale measures, variability was assessed as low if the standard deviation was below 0.75 (most respondents were well within half a point on the scale in either direction), high if it was above 1.25 (a significant number of respondents departed from the mean by at least half a point), and medium if it was between the two. Assessments for other measures were scaled proportionately.

Variability	Standard deviation of individual team member departure results from average project scores		
Score type	Low	Medium	High
5 point likert	<0.750	0.750 - 1.250	>1.250
3 point rating/confidence	<0.375	0.375 - 0.625	>0.625
Presence/absence	<0.300	0.300 - 0.500	>0.500

Within-project concordance was tested by averaging the ratings given by all academic participants within a project and all non-academic participants within a project, and comparing the two (group "departure" results). These group "departure" results allow us to see whether academic and non-academic participants rate their project similarly against each measure, or whether one or other group systematically rates the project higher or lower against the measure.

Concordance	Average difference between academic & non-academic ratings		
Score type	High	Moderate	Low
5 point likert	<0.2	0.2 - 0.4	>0.4 / beyond error range
3 point rating	<0.1	0.1 - 0.2	>0.2 / beyond error range
Presence/absence	<0.05	0.05 - 0.1	>0.1 / beyond error range

While the dataset is too small for definitive conclusions to be drawn from this comparative study (n=13 across four projects), it does provide some insights that are suggestive of both the options for and the caveats around using a survey instrument of this kind for assessing projects against the indicators, and the need to apply caution in using such an instrument to assess some of these indicators, particularly for outcomes and impact measures.

Quantitative, documentary and independent survey measures were assessed by three members of the research team with significant experience in applying these kinds of measures to assessing conservation research value. This assessment drew on insights provided at the scoping workshop and expert workshop, and from examples of recent application of such measures to assessing conservation research and conservation programs (particularly to draw in innovative examples of how quantitative and independent approaches have been applied to disentangling the complex factors involved in achieving research impact).

Overall assessments of narrative, survey and quantitative measures were made by drawing together insights from the analysis to provide an integrated 'rating' of the feasibility and usefulness of each of the measures tested and assessed, and to provide advice or recommendations on how to refine and apply the measures, and the circumstances in which they are likely to be relevant. The rubric used for these assessments is in Table 2.4. Assessments and insights derived from each of these kinds of measures were cross-checked against each other, in order to refine the interpretation and assessment of each measure, to test its feasibility and value relative to other, similar measures, and to identify significant gaps in the measures overall. Small adjustments were also made through the process to the wording of some indicators and their attribution within the program logic.

The data collected were also selectively analysed to illustrate the kinds of information yielded by a valuation exercise that uses the framework (Chapter 5).

Insights for research practice

The insights gathered through expert workshops, comments on the survey and in interviews provide further insights into what project team members and experts understand as good practice in research for generating research value. Thus the data gathered were also analysed for the light they shed on research practice and on contextualising and nuancing the 'common wisdom' and the wider research literature on what makes for effective research engagement and implementation.

A note on 'research impact' and 'research value'

The early phases of the project used the term 'research impact' to describe what we were assessing, reflecting colloquial use of this term. However, the technical meaning of 'impact' in evaluation systems usually refers to those changes that extend beyond the immediate control of research teams. Thus 'research impact' often lends itself to narrow interpretations (e.g. of measurable environmental benefit), rather than the complex pathways through which change can be measured and is usually achieved. This report uses 'research value' to refer to a holistic approach to valuation across the full research cycle. This was borne out by findings in phase 2 of the project, which showed that many of our case studies would have trouble demonstrating benefits at this stage against impact measures narrowly defined, but show considerable value across the broader suite of co-conception, shared processes, relationships, trust, co-created tools and products, and many other dimensions of value.

Table 2.4: Rubric of overall assessment of measures

Rating	Interview		Survey			Quantitative	
	Fitness	Concordance	Fitness	Variability	Concordance	Fitness	Feasibility
Good / Good: minor modification	Question could easily be answered; relevant responses	Similar responses from different members of project team	● High/moderately high ● High by non-academics	Low of low-moderate	Fair to good	Assessed as likely to be appropriate to a wide range of projects	Readily measured and/or benefit likely to outweigh measurement burden
Good: seek responses across team/ apply comparatively	Question could easily be answered; Relevant responses	Evidence of varying responses from different members of the project team	● High/moderately high ● High by non-academics	Low or low-moderate	Divergent	n/a	n/a
Good: preliminary filter to decide where relevant	n/a		Moderate	Low or low-moderate	Fair to good	Assessed as likely to be appropriate to many projects	Readily measured and/or benefit likely to outweigh measurement burden for relevant cases
						Assessed as likely to be appropriate to a wide range of projects	Potentially more challenging to measure
Possibly use, preliminary filter for relevance and value add	n/a		● Low, and/or ● A lot of non-committal answers, and/or ● Other approaches may be easier or more accurate	Low or low-moderate	Fair to good	Assessed as likely to be appropriate to select projects	Potentially more challenging to measure and/or other approaches may be easier or more accurate
Use with amendments / caution	Question could easily be answered; some relevant responses	Answers vary widely with individual perspectives. Can be addressed by significant changes to wording of measure/when it is asked /prompts	● High/moderately high ● High by non-academics	High/moderately high (qualifier: 'use with caution', 'seek response across project team' or 'apply comparatively')	Fair to good	Assessed as moderately relevant	Assessed by requiring care in application and interpretation
			● Moderate to low ● Low by non-academics (Add 'possibly' qualifier and preliminary filter)	Moderate (add qualifier: 'seek response across project team or 'apply comparatively')	Divergent (qualifier 'use with caution')		
Not recommended	Question difficult to answer; responses don't address indicators	Individuals poorly placed to answer meaningfully (other sources required)	● Moderate to low ● Low by non-academics	High	Divergent	Assessed as having low relevance	
Insufficiently tested			Low response size			n/a	

Chapter 3. Results from Phase 1 review and workshops

To support rigorous comparisons and build a richer picture of research value, our Phase 1 study focused on developing a list of impact indicators that included quantitative and qualitative measures. By incorporating both quantitative and qualitative metrics into our impact indicators list, we aimed to support reviews and assessments that can be more easily compared while still providing a rich narrative to showcase the research impact.

The full set of indicators, along with their importance and feasibility ratings, is provided in Attachment A. This provides a very broad suite of potential indicators across the many domains of research impact identified, from which anyone involved in considering and assessing research impact can select a subset in order to develop a relatively broad picture of the success of research collaborations and uptake.

Comparing the rankings given to these indicators by survey and workshop participants for importance and feasibility indicates the fundamental tensions at play in developing indicators of research value: the most important indicators in practice are also generally rated among the least feasible, particularly the subdomains of On-ground Action, Public and Private Policy Development and Research Environmental Impact (Table 3.1).

Table 3.1. Average ratings for importance and feasibility for the suite of indicators in each subdomain. Importance and feasibility scores are averages of responses by survey participants to questions "How important do you think this impact indicator is?" (1 = irrelevant, 2 = not important, 3 = important, 4 = very important); and "How feasible do you think it is to measure this impact indicator?" (1 = unfeasible, 2 = difficult, 3 = feasible, 4 = very feasible). Scores are shaded according to a scale from red (lowest average score returned for an individual impact indicator = 2.41; 1.50) to green (highest average score returned for an individual impact indicator = 3.82; 3.83).

Domain	Subdomain	Average Importance score (n=65)	Average feasibility score (n=12)
Inputs	Subdomain 1.1 Research Direction and Resources	3.26	3.21
Processes	Subdomain 2.1 Research Methods	3.33	2.69
	Subdomain 2.2 Research Management & Conduct	3.24	3.03
Outputs	Subdomain 3.1 Academic Outputs	2.92	3.51
	Subdomain 3.2 Tool Development	3.39	3.33
	Subdomain 3.3 Research Dissemination	3.28	3.54
Outcomes	Subdomain 4.1 Research Uptake (Citation)	2.97	3.39
	Subdomain 4.2 Increased Awareness & Responses	3.56	2.88
	Subdomain 4.3 Community & Stakeholder Engagement	3.28	2.86
	Subdomain 4.4 On-ground Action	3.59	2.65
	Subdomain 4.5 Public & Private Policy Development	3.50	2.11
Impacts	Subdomain 5. Research Environmental Impacts	3.53	2.06

The subdomains rated the most feasible were not surprisingly those most widely used to measure research success, but also those rated lowest in importance by our survey participants: academic outputs and research dissemination.

The distribution of responses for importance and feasibility for a series of example indicators is shown in Figure 3.1. Among those indicators ranked by survey participants in the top 50% for feasibility (48 indicators), only 18 were given an average score of 3 or higher for importance (important/very important). Examples of these are given in Table 3.2. The full suite of measures and their importance and feasibility scores are provided in Appendix B.

Table 3.2. Examples of indicators ranked in the top 50% for importance and feasibility. Importance and feasibility scores are averages of responses by questionnaire participants to the questions "How important do you think this impact indicator is?" (1 = irrelevant, 2 = not important, 3 = important, 4 = very important); and "How feasible do you think it is to measure this impact indicator?" (1 = unfeasible, 2 = difficult, 3 = feasible, 4 = very feasible). Scores are shaded according to a scale from red (lowest average score returned for an individual impact indicator= 2.41; 1.50) to green (highest average score returned for an individual impact indicator = 3.82; 3.83).		Importance (n = 65)	Feasibility (n = 12)
	Subdomain 2.2 Research Management & Conduct		
RMC 22	All participants and stakeholders contribute to designing or providing input to research questions	3.57	3.50
	Subdomain 3.3 Research Dissemination		
RD 51	Tailored reports, stories or summaries prepared on research and findings for specific audiences (e.g. land managers, Indigenous communities)	3.66	3.33
	Subdomain 4.4 On-ground Action		
OGA 71	Findings from research incorporated into real world experimental tests or trials	3.52	3.08

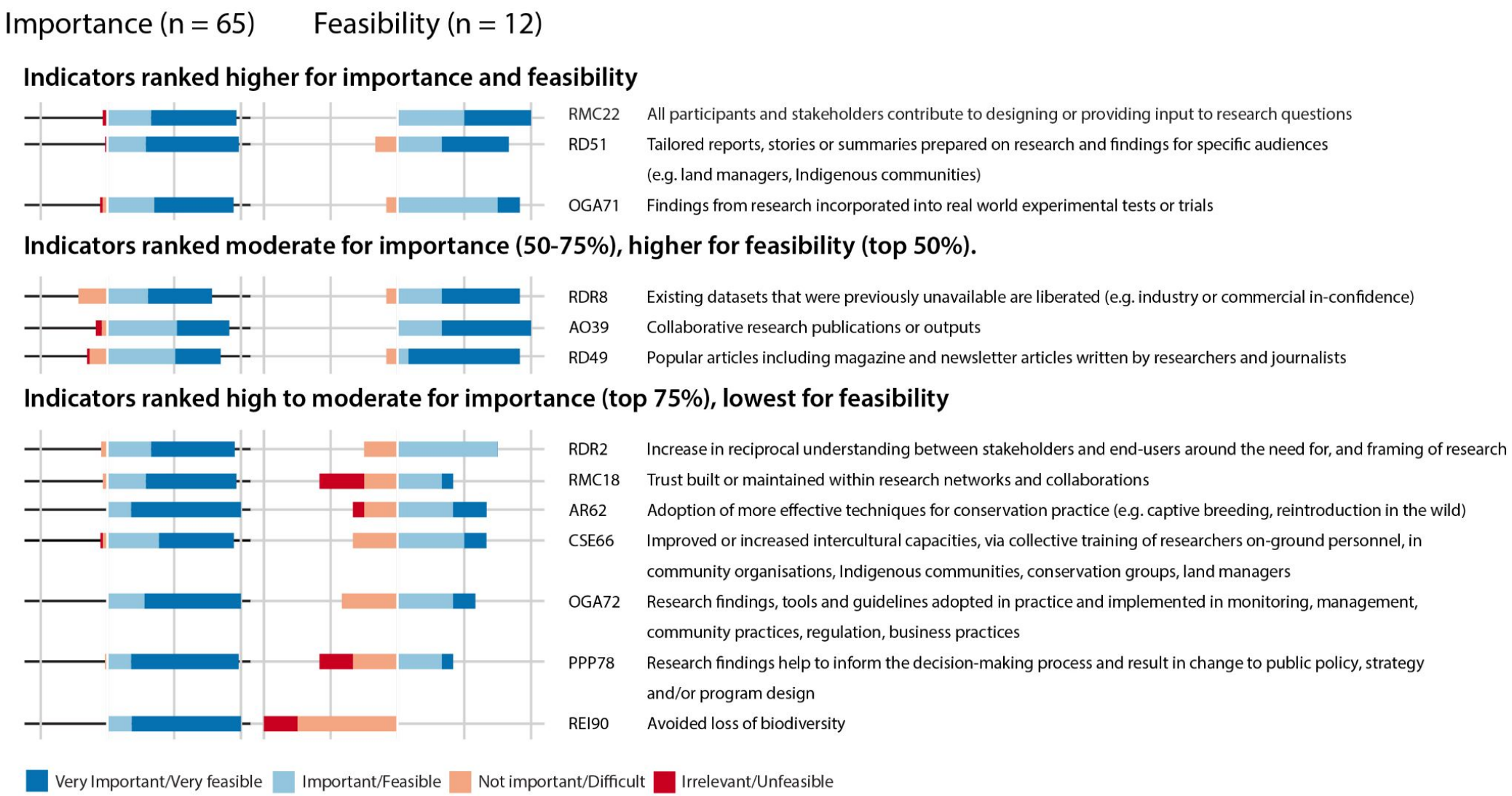
A further 12 indicators out of the 24 ranked in the lower 50-75th percentile for importance were ranked as relatively feasible (top 50%, see Table 3.3 and Appendix B).

Table 3.3. Examples of indicators ranked moderate for importance (50-75%), higher for feasibility (top 50%). Importance and feasibility scores are averages of responses by questionnaire participants to the questions "How important do you think this impact indicator is?" (1 = irrelevant, 2 = not important, 3 = important, 4 = very important); and "How feasible do you think it is to measure this impact indicator?" (1 = unfeasible, 2 = difficult, 3 = feasible, 4 = very feasible). Scores are shaded according to a scale from red (lowest average score returned for an individual impact indicator = 2.41; 1.50) to green (highest average score returned for an individual impact indicator = 3.82; 3.83).		Importance (65)	Feasibility (12)
	Subdomain 1.1 Research Direction and Resources		
RDR 8	Existing datasets that were previously unavailable are liberated (e.g. industry or commercial in-confidence)	3.26	3.50
	Subdomain 3.1 Academic Outputs		
AO 39	Collaborative research publications or outputs	3.26	3.67
	Subdomain 3.3 Research Dissemination		
RD 49	Popular articles including magazine and newsletter articles written by researchers and journalists	3.17	3.75

By contrast, a total of 43 of the 72 indicators were ranked by survey participants in the top 75% for importance and the lowest 50% for feasibility (see Table 3.4 and Appendix B). These lower feasibility indicators dominated most measures of outcomes and impact, including community and stakeholder engagement, on-ground action, public and private policy development, and research environmental impacts.

Table 3.4. Examples of indicators ranked high to moderate for importance (top 75%), lowest for feasibility (bottom 50%). Importance and feasibility scores are averages of responses by questionnaire participants to the questions "How important do you think this impact indicator is?" (1 = irrelevant, 2 = not important, 3 = important, 4 = very important); and "How feasible do you think it is to measure this impact indicator?" (1 = unfeasible, 2 = difficult, 3 = feasible, 4 = very feasible). Scores are shaded according to a scale from red (lowest average score returned for an individual impact indicator = 2.41; 1.50) to green (highest average score returned for an individual impact indicator = 3.82; 3.83).		Importance (65)	Feasibility (12)
	Subdomain 1.1 Research Direction and Resources		
RDR 2	Increase in reciprocal understanding between stakeholders and end-users around the need for, and framing of research	3.58	2.75
	Subdomain 2.2 Research Management & Conduct		
RMC 18	Trust built or maintained within research networks and collaborations	3.64	2.17
	Subdomain 4.2 Increased Awareness & Responses		
AR 62	Adoption of more effective techniques for conservation practice (e.g. captive breeding, reintroduction in the wild)	3.82	2.83
	Subdomain 4.3 Community & Stakeholder Engagement		
CSE 66	Improved or increased intercultural capacities, via collective training of researchers on-ground personnel, in community organisations, Indigenous communities, conservation groups, land managers	3.49	2.83
	Subdomain 4.4 On-ground Action		
OGA 72	Research findings, tools and guidelines adopted in practice and implemented in monitoring, management, community practices, regulation, business practices	3.72	2.75
	Subdomain 4.5 Public & Private Policy Development		
PPP 78	Research findings help to inform the decision-making process and result in change to public policy, strategy and/or program design	3.79	2.25
	Research Environmental Impacts		
REI 90	Avoided loss of biodiversity	3.82	1.75

Figure 3.1. Graphical representation of distributions of ratings from survey respondents for feasibility and importance for select indicators



Chapter 4. Results from Phase 2 applying and testing the indicators

Ten interviews were conducted and 18 people responded to the online survey (Table 4.1). Fewer than 28 people participated in total because several people did both the survey and interview. The ten people interviewed had participated in six of the projects. They ranged from project leaders in research institutions to government and NGO research partners. To answer our third research question: "Do project participants from different types of institutions value similar aspects of conservation research?", we clumped the government, NGO and Indigenous organisation survey respondents (those from outside of research institutions) into a single "non-academic" group.

Table 4.1: Breakdown of case study respondents according to anonymised TSR Hub projects

Project code	Number surveyed (academic/non-academic)	Number interviewed (academic/non-academic)	Total
PROJ01	1 (1/0)	2 (0/2)	**3**
PROJ02	3 (1/2)*	0	**3**
PROJ03	4 (3/1)*	0	**4**
PROJ04	2 (2/0)	3 (2/1)	**5**
PROJ05	0	1 (1/0)	**1**
PROJ06	2 (1/1)*	1 (0/1)	**3**
PROJ07	4 (2/2)*	1 (1/0)	**5**
PROJ08	1 (0/1)	2 (0/2)	**3**
PROJ09	1 (1/0)	0	**1**
total	**18**	**10**	

* Survey responses used to analyse variability and consistency of responses

4.1 The effectiveness of narrative approaches

In this section we reflect on how well individual interview questions were for eliciting useful information and the effectiveness of the sequence of questions.

Q2. Can you please describe the nature of your involvement in project X?

This question was designed to put interviewees at their ease and potentially provide some insights into the scope of the project, the interviewee's role, and what they self-identified as important aspects of the project e.g. *"I've sort of only been involved in the periphery"* and *"we are both an end user and a research partner"*. It didn't specifically relate to any of the indicators. The question was generally effective and everyone could provide an answer. The question also revealed that some people's roles were different than anticipated and some had changed during the course of the

project. E.g.*"So it's changed over time it evolved"* and *"what I really do is manage other contractors who are based in each of the states*". This is all useful contextual information and highlights any incorrect assumptions which is essential at the beginning of the interview.

Q3 Interactive partner mapping process

Many of our indicators are about collaboration in research. The mapping question was tested as a way of eliciting who actually took part in research so we didn't assume collaboration had taken place. We asked interviewees to identify participating organisations and also reflect on whether or not relationships had changed over the course of the project, in order to measure indicators related to improving professional networks and relationships. The mapping process often took up a significant portion of the interview and this response was the longest on average. Lengthy responses were generally not a problem because content was relevant. If a project involved more than around six organisations, we asked them to clump like organisations e.g. "regional NRM groups". Doing the mapping using an interactive and visual medium (we used the online zoom whiteboard function) is recommended.

The mapping was a good early question to get interviewees thinking about the fundamentals of the project. It elicited who was involved and sometimes what different groups did and how active they were. Where people from the same project were interviewed, the maps weren't the same, particularly in relation to minor contributors. The reason for variation appeared to be that people with different roles have a different awareness of who was involved in a given project. People participate in different ways in research, over varying time periods and have different values. This means individual first hand knowledge of a project will differ e.g. one would assume academic participants may know more about different expertise that was brought in, whereas NGOs might be more aware of who was collecting data. This illustrates variation in responses and results among participants within a project even with regards responses that seemingly involve little value judgement. This variability is also likely for surveys and demonstrates the value of asking the same questions of multiple project participants.

Q4. How did non-academic/your and other non-academic groups participate in the project?

All interviewees could answer this question and it had the second longest responses on average. The question provided good insights about how different individuals and groups participated (see section 5.2). In our cohort there were some differences in the types of observations from academics and non-academics. Specifically non-academics commented more on the drivers for them participating whereas academics focussed more on who did what and project roles (see section 5.2). One of the three projects where we had more than one interviewee gave different answers to this question. This example highlighted two issues. Firstly, as mentioned previously, individuals will know different things about a project depending on their role and length of involvement. Secondly there were different interpretations as to what constituted "participation" in a project. There can be a blurry line between engaging with non-academics (e.g. them participating in a webinar) and non-academics being research participants (e.g. targeted invitees to a workshop being asked for input to a project), particularly in the eyes of some academics. Network mapping, questions about production of outputs etc may help clarify and elicit this as well as prompts to clarify interviewee's interpretation of "research participation".

Q5. How were research needs identified in the project?

The issue of how research needs are identified sometimes came out during question 4 about non-academic participation but it is a good question to ask specifically. It also can yield other useful information such as research drivers. However, our results demonstrated that not all research participants will know the answer to this. This was because they weren't around at project inception or they have a more peripheral role. For those that did know, answers were comparable.

Q6. Do you feel trust was developed and/or improved among partners?

This was generally a very good question to encourage interviewees to reflect on interview processes and values of conservation research necessary to and beyond species/ ecosystem conservation (see sections 4.6 and 5.3). Trust is a complex concept with multiple components and so well suited to a narrative measure which can elicit valuable information beyond a level of agreement (e.g. prerequisites for trust, variations in trust, what trust looks like in practice). There was a common conflation of relationships and trust i.e. if you have good relationships there must be trust, if you don't have a relationship it might be through lack of trust (although lack of relationships may just reflect lack of opportunity). In terms of variability, the specifics of what people focused on varied somewhat, but any big issues seemed generally understood. A caution is that highly confident or senior project leaders may give a different perspective on this (and other partner-related) questions and should always be validated by interviewing others.

Q7. If Indigenous partners in the original map: How was informed consent obtained from Indigenous partners?

Only three people were asked this question because other interviewees did not identify Indigenous partners in question 3. None of the respondents were sure. This was either because it wasn't their responsibility or the processes for gaining consent were unclear *"it was implied that consent was given*" because a project was approved by traditional custodians but then " *there was no clear process that I understood of contacting them for their opinions or to consult with them"*. This seems to be information that may not be broadly understood. With the limited data available a tentative conclusion is that perspectives of Indigenous participants should be sought and this particular question reserved for project leaders.

Q8. Were there existing data sets/information necessary to do the research? If yes, were those made available? If not, why not & what was the consequence?

Asking about data is important. For example, some projects were specifically designed around available data. Information about data needs, accessibility, metadata and issues with data procured were elicited with this question (see section 5.2). The prompt about the implications related to limited access is also recommended because not being able to access data may be a key driver of non-delivery for a project (e.g. "*he's had to really scale back the complexity of some of the models*"). Academics may know more of the detail here but government agencies are often custodians of historical data and can also provide useful perspectives.

Two of the projects with multiple respondents gave similar results but in one project the 2 interviewees had completely different responses: "*It was not drawing upon existing data*" vs "*we definitely needed that existing material, but there was no one singular source, I think it was definitely an accumulation of many, many different bits of knowledge*". The latter quote being from the person responsible for acquiring that information. This illustrates the need to take into account individual roles (people may answer a question without the knowledge to do so accurately). It may also reflect different conceptualisations of data and information (the question needs to reinforce this, e.g. could be metadata or knowledge). Useful to ask or note whether sourcing the data was part of the interviewees responsibilities.

Q9. What is your view on the research methods used in the project? P1 Did it draw on different disciplines & knowledge systems

There was a tendency to simply describe the methods in response to this question so interviewees may need prompting about the qualities of the methods. It is also recommended to avoid using the term "knowledge systems" to describe different contributions as it is not broadly understood. The academics also tended to be more confident answering this question. However, where methods development was an expected project outcome, interviewees gave similar responses in relation to quality.

Q10. How would you characterise the knowledge produced by the project?

Having this as the first outputs question was suboptimal because it is a fairly complex question conceptually. This was reflected by quite varied responses. Knowledge was often equated with results and a key benchmark of quality was whether knowledge was useful/being applied (evidenced by the words "use/useful/using" being common). This also meant that answers sometimes veered into outcomes. This question related to Indicator AO30 "Legitimate, valuable, and rigorous knowledge developed according to the various knowledge systems involved in the research". Rigour was mentioned once or twice, but legitimacy was absent from interview responses. If these qualities are of particular interest, they may require a prompt. Asking what outputs people have seen or are aware of would be a better first question in this domain and can be followed by a prompt to think beyond papers if necessary. This knowledge question would be better as the last output question.

Q11. Can you describe how project outputs were produced, particularly who was involved? [process for production]

This question provided information on outputs produced and how this occurred and there appeared to be agreement among people from the same project. Where there were outputs, people seemed comfortable answering this question.

Q12. What do you think about how knowledge generated by the project has been packaged, presented and disseminated? [process of dissemination]

Only half of the interviewees were asked this question because a) some had indicated they had not developed outputs or b) they had previously described workshops etc. In retrospect, individual project participants may not be in a position to judge what appropriate dissemination is: *"I think that*

the fact that you have a project page and everything is assembled, and you can check everything on that page is really awesome...I just send that around to everybody...So I guess that's in terms of dissemination, well, I don't know how good is that?" Rather a communications professional may be a better judge. A more appropriate question might be asking people what outputs they know are being or will be used.

Q13. Do you think your project helped to develop the skillset of you and/or your partners to undertake this work? If yes, what types of skills were developed?

Skills and capacity as an outcome could be quite rich and identify important changes within the life of a project that may not be articulated in project plans (see section 5.5). People seemed comfortable answering this question but there was some variation within projects. This seemed to be because people had greatest awareness of their own skills and those of their colleagues. Learning was equated with skills and capacity by some and reflected in "knowledge" being one of the top ten unique words for the question.

Q14. Can you describe any changes to practice or policy the research may have contributed to?

There were several recurring issues for this and the following outcome and impact measures:

- For many projects it was too early (some didn't even have outputs) "*when I was doing the online survey, I feel like I said no to most of the outcomes. Because it was, did it change the trajectory of a threatened species? Or did it change how it's being managed? And I was like, well, actually, not yet. But it could*."
- Attribution - these types of change will rarely be attributable to a single research project and people felt reluctant to do this
- Terminology - the distinction between outcomes and impacts (and even outputs and outcomes) can be blurred for people. This has particular implications for designing survey instruments (i.e. may need to be quite specific about what an outcome is or what an impact is)
- When you ask about specific outcomes or impacts it's often irrelevant to the particular project because their scope is so variable

These issues meant that these responses were the shortest on average as there wasn't much they could say. In terms of within project variability for this question, big ticket items like a species listing or an overwhelming positive response from a target group of people are consistent. More subtle outcomes may not be. For example, in one project the project leader felt they had highlighted an issue with legislation but a respondent from the department was hesitant to attribute those outcomes. The government respondent was much more aware of how the project had an unintended outcome (contributing to the development of international standards). A third respondent from the project who was more on the periphery felt there were no outcomes at all.

For projects that are just finishing the best approach might be to ask one question about outcomes (with prompts as appropriate) beyond skills and capacity. It is recommended not to base any conclusions about outcomes or impacts on the feedback of academics alone.

Q15. Were there any other outcomes (expected or unexpected) connected at least broadly with the project? Eg investment decisions, program design, etc?

This question did elicit some new information and prompted people to think beyond outcomes they may have planned for. The data is too sparse to comment on within project variability. It is probably best used as a prompt to a generic outcomes question.

Q.16 We consider impacts to be broader than change in the immediate partners or organisations ie An effect on, change or benefit to the environment, society, culture, economy and/or research. Keeping this in mind, can you describe any environmental impacts your project has contributed to?

Q17. Can you describe any societal, cultural or economic impacts the project may have contributed to?

In three instances questions 16 and 17 were rolled into one, depending on previous answers. A lot of people struggled with this for similar reasons to the outcomes. Some gave outcomes rather than impacts and talked about potential outcomes. For projects that are just finishing, asking one question about impacts may be best (with prompts as appropriate).

4.2 Survey and likert-scale indicators and insights from the survey

Respondents to the survey were first asked to identify their own sector and to select the range of roles that they played within the project. Having a range of sectors and project roles represented in this test survey allowed us to cross-check the responses with projects between respondents based in research institutions ("academics") and those based in research-user and other partner agencies ("non-academics"). For any general survey-based valuation, ensuring representation from a range of sectors and roles allows for multiple perspectives to inform any assessment, providing a level of independence and robustness to the data collected. It also allows evaluators to analyse differences in responses offered by different sectors, and different project roles. Assessments of the measures by sector are provided below, while an example analysis of how the projects were assessed against outputs measures, by sector, is provided in chapter 5.4.

Team members who reported having a wider range of project roles also reported a wider range of sectors involved in the research, potentially reflecting incomplete knowledge among different project participants of how the project was rolled out, but also possibly reflecting differing views as to what constitutes "being directly involved" in a project. This variability in project sector mapping speaks to the importance of seeking insights from multiple project participants to build up a full picture of cross-sector research partnerships. It also suggests that there is potential value in asking further questions about the level of involvement or the nature of involvement of different sectors (e.g. did they provide incidental support, participate in a workshop, or provide data or expert advice, and/or was their involvement more active and ongoing in other ways?). A survey instrument could potentially ask some of these questions, but surveys need to balance eliciting comprehensive information and staying a manageable length. Fleshing out a fuller picture of project stakeholders was easier through the interview process, as discussed above.

Inputs

Inputs were tested in the survey through a 2-part yes/no multiple choice question seeking information on the diversity of sectors involved in the research, and five likert-scale questions

measuring level of agreement with positive statements about the research (5-point scale). Participants then rated the likert-scale questions for how fit for purpose each measure was for assessing the value of their research project. A summary of the measures tested and the results of this assessment is given in Table 4.2.

Fitness for purpose results

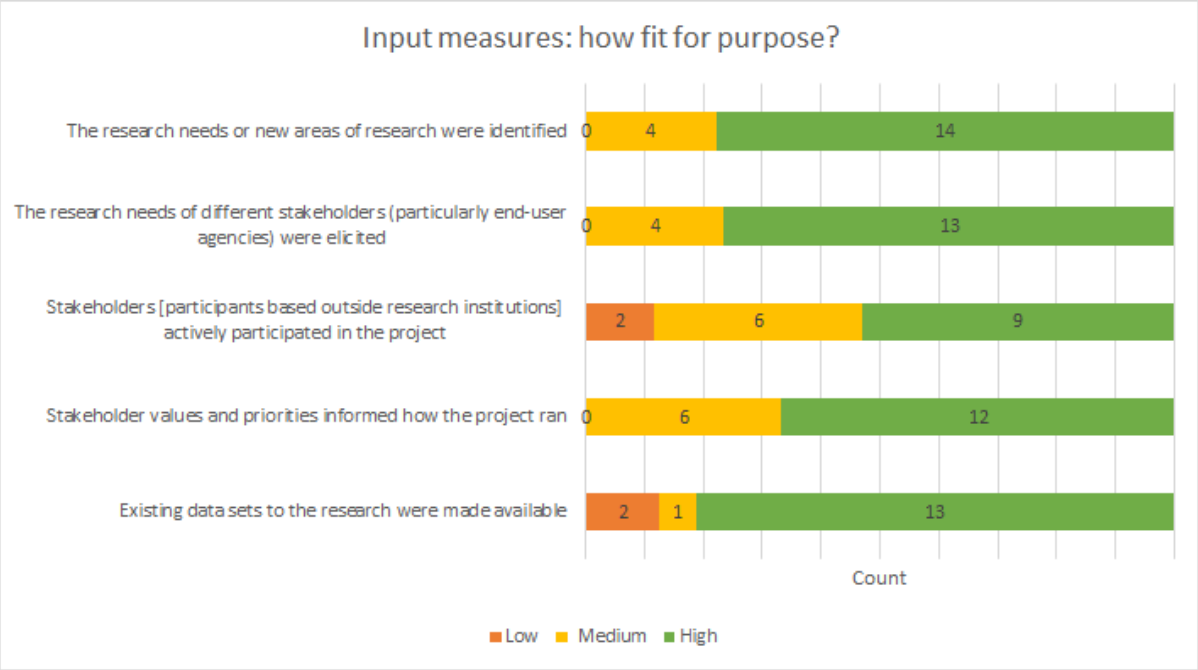
Participants on the whole responded very positively to the fitness of the Input indicators for capturing the value of their research project. Four of the five input indicators on average were rated highly fit for purpose, while one indicator, related to how actively stakeholders participated in the project, was rated moderate (Table 4.2). Responses were consistent among academics and those in non-academic institutions.

Table 4.2: Comparison of "Fit for purpose" survey responses from academic and non-academic organisations in relation to input indicators. Response options were: High (3); Medium (2); Low (1) & Not Applicable (-) (dark yellow: scores 1 - 1.5; pale yellow: >1.5 - 2; pale green: >2 - 2.5; dark green: >2.5 – 3).

Indicator	Average score (1-3 range)	
Inputs* **[overall average 2.67]**	Non-academic org avg (n=7)	Academic org avg (n=11)
The research needs or new areas of research were identified	2.86	2.73
The research needs of different stakeholders (particularly end-user agencies) were elicited	2.67	2.82
Stakeholders [participants based outside research institutions] actively participated in the project	2.43	2.40
Stakeholder values & priorities informed how the project ran	2.86	2.55
Existing data sets to the research were made available	2.67	2.70

The differences can be seen in the distribution of responses across the measure for "stakeholders actively participated" compared with other input measures (Figure 4.1). For other measures, a significant majority of respondents (12 or more out of 18) rated the measure as having a high fitness for purpose, while only just over half of respondents (9 out of 17, with one abstaining) rated the measure of active stakeholder participation as being highly suitable. This is somewhat surprising, given that active stakeholder participation in conservation projects is widely cited as an effective strategy for enhancing both learning and outcomes from the projects (Evely et al. 2011, Reed 2008). The relative concordance between academic and non-academic respondents on this point suggests that this is not just a result of researchers underestimating the importance of this step. It perhaps suggests that, while many participants do value this, achieving this active participation is not always as vital for research value as it is to have the other measures met, such as identifying stakeholder research needs and having stakeholder values inform research. This could be achieved in other ways, such as through regular communication between researchers and research users. This potentially suggests a need to adopt clusters of related measures that represent alternative pathways for achieving benefits from research (e.g. stakeholders actively participate OR there is regular communication across project teams).

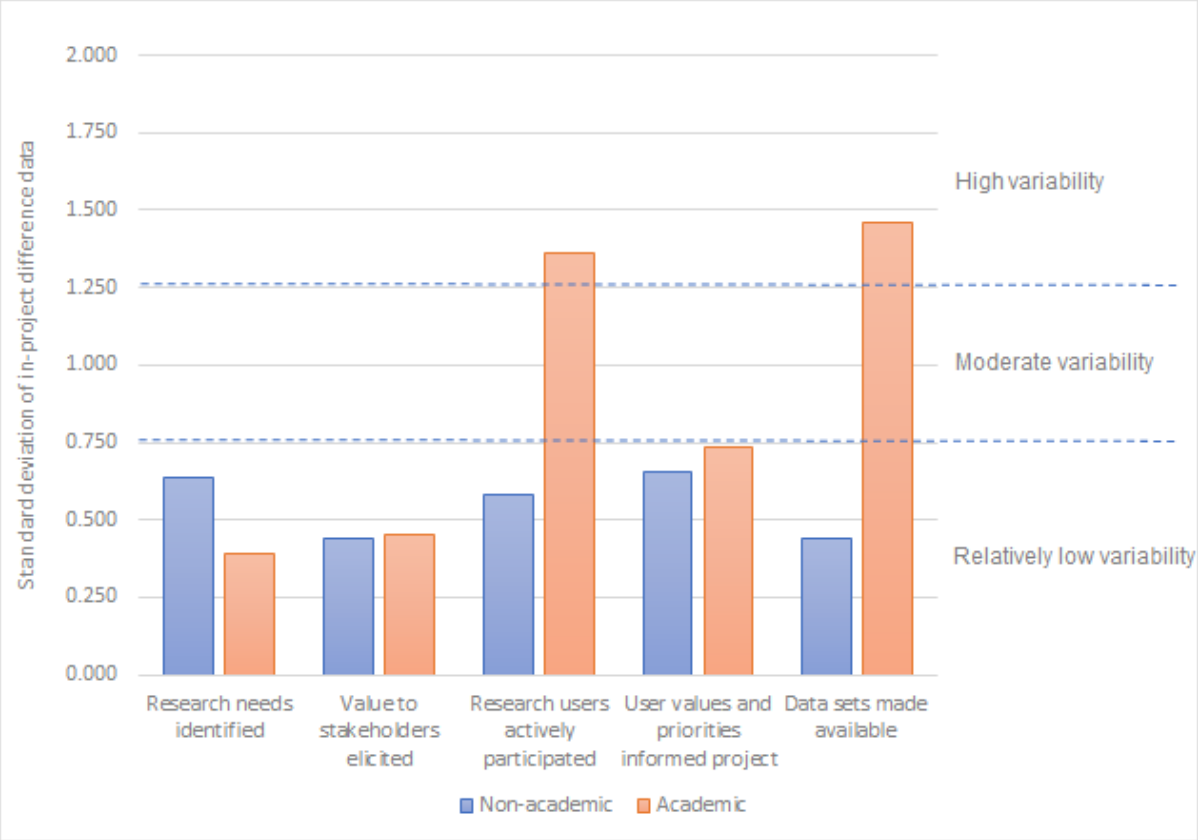
Figure 4.1. Summary of survey responses to fitness for purpose for Input measures.



Within-project concordance and variability results

Participant responses for each of the measures were also tested for the level of within-project variability and for any systematic differences between how project partners in and outside academic roles rated each of the indicators. Most of the input measures were rated in a consistent way by both academic and non-academic project team members (Figure 4.3). Responses from academics were highly variable on whether stakeholders actively participated and whether data sets were made available (Figure 4.2), but these results were strongly skewed by negative responses from a single respondent who, according to their project roles, was not as involved across the range of project activities as the others. Removing these responses, the variability of results on these two measures becomes low. So although the data set is small and therefore susceptible to these single outliers, this does perhaps also highlight that members of project teams can have variable levels of appreciation of different aspects of how the project engaged (such as stakeholder participation and data sets being made available), suggesting that care may need to be exercised in averaging results from across project teams. Where this does prove to be a problem, one approach to addressing this could be to weight the results to those academic team members more actively involved across a range of project activities, along with other potential weighting criteria (such as weighting the responses of research-users).

Figure 4.2. within-group variability for input measures, calculated as the standard deviation of data recording departure of individuals' ratings from project means, grouped by non-academic and academic respondents. Raw data in these fields is on a 5-point scale.



Looking at the level of within-project agreement for each of these measures between academic and non-academic respondents shows a strong level of agreement between these groups across four of the five measures. However, academic researchers consistently rated their assessment of whether "The research needs or new areas of research were identified" higher than non-academic respondents ($\bar{x}$=0.46 (5-point scale); p=0.004). This may be because researchers are focusing more closely on clarifying or eliciting new research questions, but it could also suggest that non-academic partners are not as aware of the research questions being developed as are the researchers.

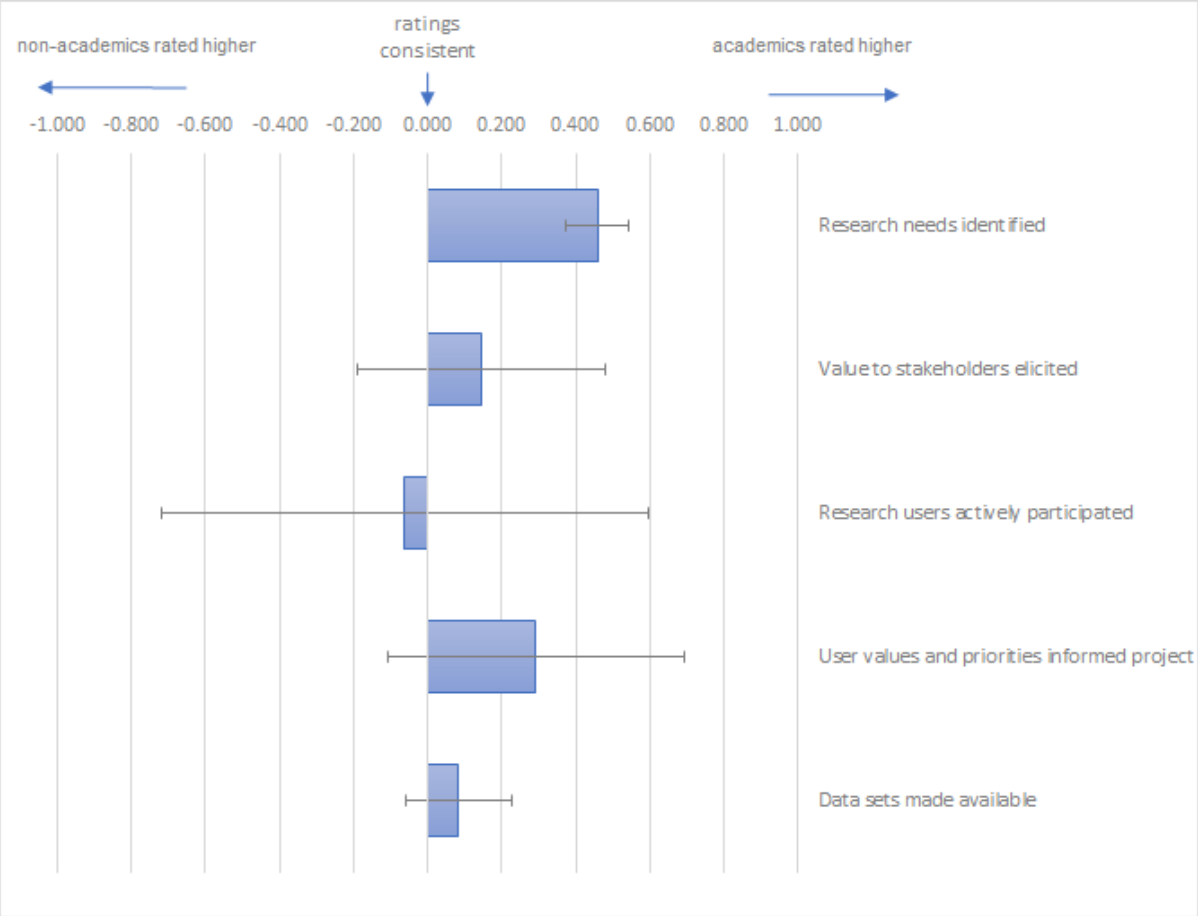
One of the reasons to test the consistency of these measures between academic and non-academic research partners is to assess whether researchers' self-report gives an accurate measure of research value. If, as part of a research valuation exercise, getting an accurate picture of whether research needs have been identified, and whether these needs are widely understood or shared among project partners, is important, it may be necessary to elicit the views of non-academic as well as academic respondents, or to find other means of gathering this information.

It is possible that academic respondents slightly over-estimate that "The values and priorities of research users informed how the project was run" compared with other project team members. Within-project differences on this measure are slightly larger than for most other input measures, though they are within error given high variability and low sample size ($\bar{x}$=0.29; p=0.14). However, this slightly higher weighting is consistent with comments made by an academic respondent, which did

not clearly make the distinction between eliciting values and priorities, and acting on them: “It is unclear how the first two [questions] are distinct.”

In general, assessing the consistency of researcher self-reporting against other project participants is useful, because self-report is used very often to garner a sense of how research projects and programs are developing and progressing, especially for interim reporting between ‘major’ evaluation exercises. Although the dataset is small, the relatively low variability of responses to these input measures and the consistency of within-project assessments between academic and non-academic respondents suggests that these inputs measures can offer a reliable set of measures that can be applied in a wide array of circumstances. In contexts where an assessment is reliant on researchers’ self-report (as opposed to garnering the views of whole project teams), other than the question of whether research needs were clarified through the project, asking academic researchers for insights on these inputs measures or using these measures within a survey instrument for academic project team members may be a relatively simple and reliable means of assessing the contribution of project inputs to research value.

Figure 4.3. within-project differences in input measures, calculated as the difference in project means between academic and non-academic assessments of projects against each measure. Raw data in these fields is on a 5-point scale.



Summary of findings for inputs measures tested in the survey

Table 4.2. Summary of inputs measures used in surveys along with considerations for applying each measure

Indicator descriptor	Measures/questions addressing indicators	Assess-ment	Comments and considerations for applying measure	Code
Research needs, gaps, unanswered questions, new areas for research identified	Research needs or new areas of research were identified through the project **(likert agreement 1-5)**	Good: survey across team	Rated highly fit-for purpose. Academics appear to assess higher; include partners outside research institutions to assess breadth of recognition of research needs identified.	RDR1
Need for research has been adequately established - and all perspectives gained on whether research is valuable (e.g. particular end-user agencies)	The value of the research to different stakeholders (particularly end-user agencies) was elicited **(likert agreement 1-5)**	Good	Rated high fit-for purpose. Academic self-report consistent with non-academic assessment.	RDRX
Active participation of stakeholders and end-users in research	· Were groups or organisations outside research institutions (like universities, CSIRO) directly involved in the project? **(yes/no)** · [if yes to above] What groups or organisations (other than research institutions) were directly involved in the project? **(multiple choice)**	Good	Academics may recognise slightly broader range of groups involved than non-academics, likely reflecting different perspectives on project operation. Collectively, responses could be aggregated to shed insights on breadth and sectoral diversity of project.	RDR4
	Research users and others based outside research institutions actively participated in the project **(likert agreement 1-5)**	Use with caution Survey across team	Rated moderate-high fit for purpose. Provisionally, results point to the possibility of diverging awareness of stakeholder participation among academic team members, thus need for care in averaging results across project teams. Could be useful to weight assessments toward project team members more actively involved across a range of project activities.	
New understanding of values and priorities captured (e.g. organizational agencies, citizen science, Indigenous priorities)	The values and priorities of research users informed how the project was run **(likert agreement 1-5)**	Good: survey across team	Rated highly fit for purpose. Possible that academic respondents report slightly higher than other project team members, suggesting value in surveying across project team.	RDR5
Existing datasets that were previously unavailable are liberated (e.g. industry or commercial in-confidence)	Additional data sets from outside of the research team were made available to support the research **(likert agreement 1-5)**	Use with caution Survey across team	Rated highly fit for purpose. Provisionally, variability of results point to the possibility of diverging awareness among academic team members. Thus, there needs to be care in averaging across project teams. Consider weighting assessment toward project team members who are more actively involved across a range of project activities.	RDR8

Processes

Processes measures were assessed through a combination of statements against which participants rated their level of agreement on a 5-point Likert-scale, with multiple choice questions on the standard of data collection, management and analysis. A summary of the measures tested and the results of this assessment is given in Table 4.4.

The majority of processes measures were rated highly fit for purpose and showed strong within-project concordance, suggesting that a survey approach is an effective way to assess the contribution of research processes to ensuring research value.

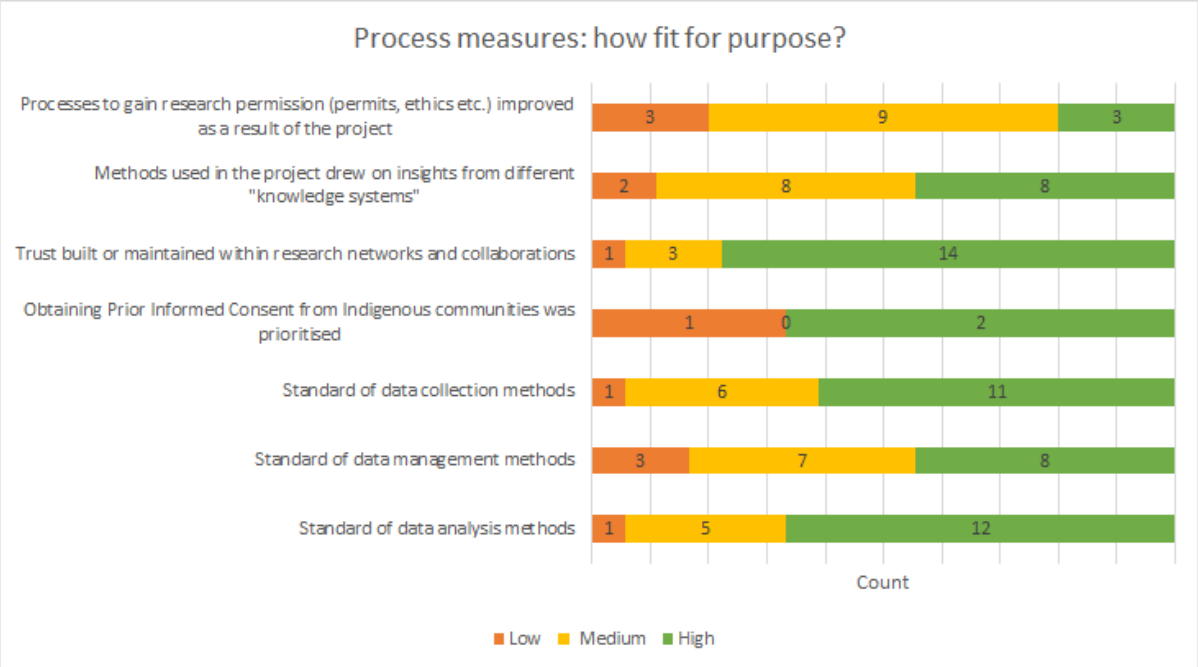
Fitness for purpose results

Most of the process indicators were rated as moderately or highly fit for purpose by survey participants (Table 4.3, Figure 4.4). The notable exception was the question of whether 'Processes to gain research permission (permits, ethics etc.) improved as a result of the project', which was rated by most participants as low-moderately fit for purpose. The question on data management methods were rated as moderately fit for purpose, as was the rating given by academics on whether the research methods drew on insights from different knowledge systems.

Table 4.3. Comparison of "Fit for purpose" survey responses from academic and non-academic organisations in relation to process indicators. Response options were: High (3); Medium (2); Low (1) & Not Applicable (-) (dark yellow: scores 1 - 1.5; pale yellow: >1.5 - 2; pale green: >2 - 2.5; dark green: >2.5 – 3).

Processes [overall average 2.42]	Average score (1-3 range)	
	Non- research org avg (n=7)	Research org avg (n=11)
Processes to gain research permission (permits, ethics etc.) improved as a result of the project	2.00	2.00
Methods drew on insights from different "knowledge systems"	2.72	2.09
Trust built or maintained within research networks & collaborations	2.57	2.82
Standard of data collection methods	2.57	2.55
Standard of data management methods	2.43	2.18
Standard of data analysis methods	2.57	2.64

Figure 4.4. Summary of survey responses to fitness for purpose for Process measures.



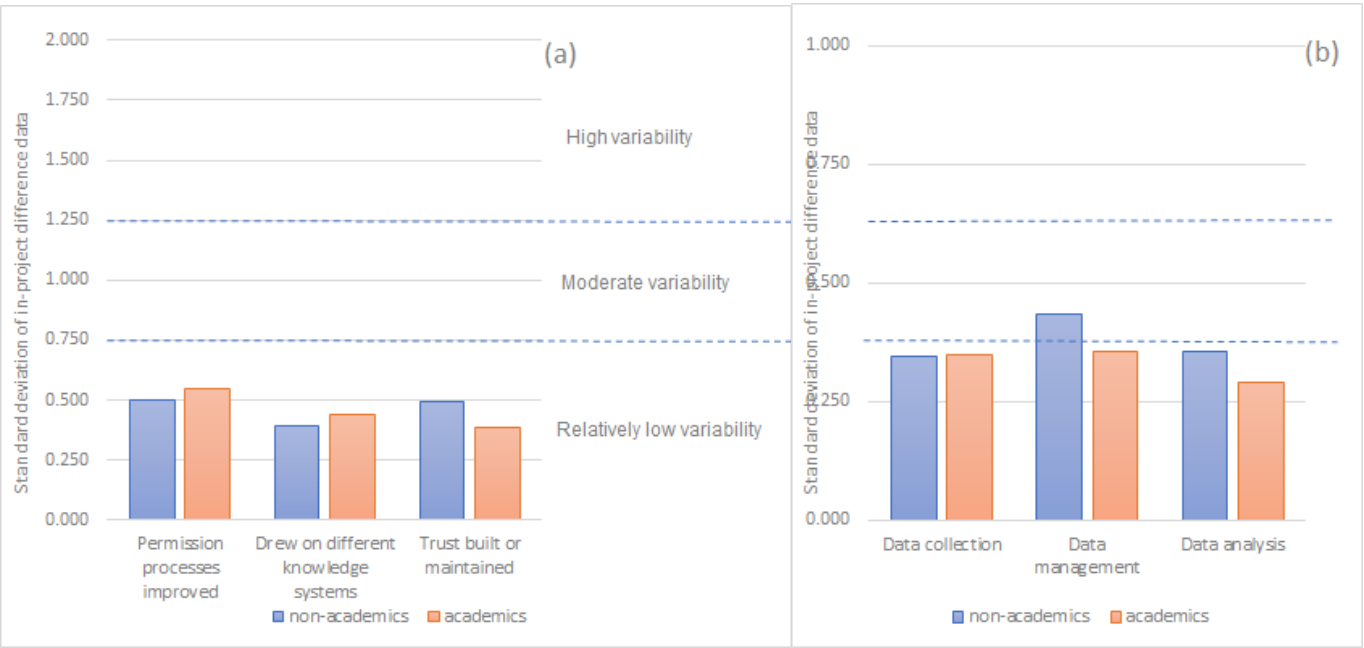
Within-project concordance and variability results

All processes measures demonstrated relatively low within-project variability, other than the standard of data management, for which non-academic responses were moderately variable (Figure 4.5). This degree of variability from research users is perhaps not surprising, if they had less visibility of data management methods, suggesting that a survey measure on this question may not be as useful as the questions on data collection and analysis for assessing research value. One of the survey comments highlights the potential ambiguity in these questions:

> "I think the data management and analysis question is really important, but I think answers might conflate whiz-bang with other measures of high standards, like that a research project effectively extracted all the best juice from the fruit, whether that be by a simple or complex analysis."

This points to some of the limitations of a survey-based approach, and the value of complementary narrative approaches in being able to effectively capture this kind of nuance. But it also suggests that some qualifying language might help sharpen this question, for example adding the statement 'e.g. the analysis gleaned the greatest value from the available data'.

Figure 4.5. Within-group variability for general processes measures (a), and data handling measures (b), calculated as the standard deviation of data recording departure of individuals' ratings from project means, grouped by non-academic and academic respondents. Raw data in these fields is on a 5-point scale (a) and on a 3-point scale (b).



The majority of measures showed good within-project concordance (Figure 4.6). The exception was the measure related to different knowledge systems ('Methods used in the project drew on insights from different knowledge systems'). Against the measure for different knowledge systems, academics appear to rate their projects somewhat lower than non-academic respondents, though this is not significant given the small sample size and variability in the data ($\bar{x}$=-0.31; p=0.19). This may be because respondents interpret the question differently, particularly for the measure on different knowledge systems, as the question itself seemed to cause some confusion. One respondent's

comment suggested they conflated this with the question on data management where these two measures had in fact been intended to understand quite different dimensions of research process:

> "I found these difficult to judge. Is a data management system or data set a knowledge system? They are only a partial match with the definition above to my reading."

This measure was intended to assess where different knowledge systems informed research, including Indigenous knowledge systems and western science, but also potentially encompassing interdisciplinarity and other non-academic knowledge systems. However, the confusion in the language used may be obfuscating other insights from this measure. The question would benefit from examples and from greater clarity as to which information is sought (e.g. drew on insights from diverse knowledge systems, e.g. natural resource managers, farmers, Indigenous knowledge, etc').

Figure 4.6. Within-project differences in general processes measures (a), and data handling measures (b), calculated as the difference in project means between academic and non-academic assessments of projects against each measure, for. Raw data in these fields is on a 5-point scale (a) and on a 3-point scale (b).

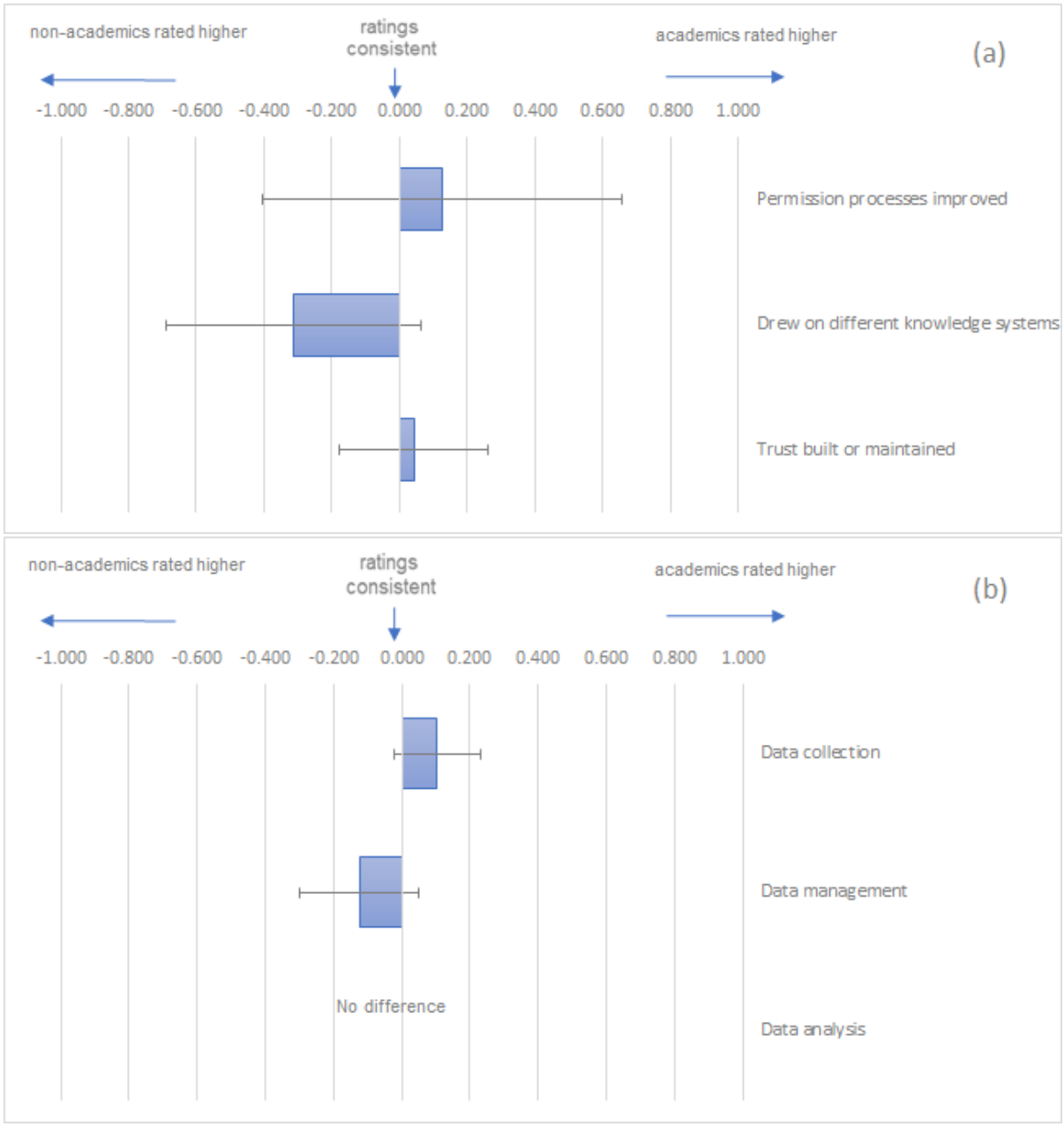


Table 4.4. Summary of processes measures and considerations for application of each measure from survey results

Indicator descriptor	Measures/questions addressing indicators	Assess-ment	Comments and considerations for applying measure	Code
Research management and conduct				
Trust built or maintained within research networks and collaborations	Trust was developed or improved among research partners **(likert agreement 1-5)**	Good	Rated high fit for purpose, low within-project variability, close concordance between academic and non-academic respondents.	RMC18
Obtaining Prior Informed Consent from Indigenous communities understood, developed, prioritised, excellent	Did the research need to go through permission processes (permits, ethics, etc)? **(yes/no branching question)**	Insuffic-iently tested	As a branching question, may want to word more specifically as research permission processes for Indigenous engagement.	RMC20
	[if indigenous partners were selected as a participating group] Obtaining Prior Informed Consent from Indigenous communities was prioritised in the project **(likert agreement 1-5)**	Insuffic-iently tested	Only applicable to 3 out of 18 respondents. Rated as highly fit for purpose by two out of three of these respondents.	
Processes to gain research permission (permits, ethics etc.) maintained, improved, made more meaningful and rigorous in relation to research content and processes, or streamlined where appropriate	Knowledge of and/or capacity to engage in ethical research by researchers and/or research users was improved as a result of the project **(likert agreement 1-5)**	Possibly use, prelim-inary filter for relev-ance, value add	Rated by most participants as only medium fitness for purpose. Reasonable within-project concordance. Many participants rated their projects 'n/a or don't know' or 'neither agree nor disagree' against this measure, suggesting low levels of conviction about this measure.	RMC28
Research methods				
Best methods for data collection, management and analysis developed or maintained across disciplines and knowledge systems	Methods used in the project drew on insights from different knowledge systems **(likert agreement 1-5)**	Good : survey across team with modific-ation	Rated moderate to high fitness for purpose. Academics may rate their project lower than non-academic partners against this, though within error. Question would benefit from clearer language specifying information sought. Ensure non-academic partners are surveyed where possible. See also survey measure for AO30 under Outputs.	RM16
	Data collection methods used in the project were: a) Below standard b) Standard c) Above standard/new d) Don't know **(multiple choice)**	Good	Rated high fit for purpose, low within-project variability, fair concordance between academic and non-academic respondents.	
	Data management methods used in the project were: a) Below standard b) Standard c) Above standard/new d) Don't know **(multiple choice)**	Good: prelim-inary filter for relev-ance	Rated moderate fitness for purpose. Moderate within-project variability, reasonable within-project concordance, though researchers may slightly under-report (within error).	
	Data analysis methods used in the project were: a) Below standard b) Standard c) Above standard/new d) Don't know **(multiple choice)**	Good: minor modific-ation	Highly fit for purpose, low within-project variability, strong within-project concordance. May be worth adding qualifier: "e.g. the analysis gleaned the greatest value from the available data".	

Outputs

Outputs measures were assessed through a range of statements on the value and accessibility of products, against which participants rated their level of agreement on a 5-point likert scale, and on their awareness of outputs of different types. A summary of the measures tested and the results of this assessment is given in Table 4.6.

The majority of outputs measures were rated highly fit for purpose and showed low within-project variability. Most of the likert-scale measures showed good within-project concordance, suggesting that a survey approach is an effective way to assess the contribution of research outputs to delivering research value.

Respondents from across sectors and project roles were overwhelmingly in agreement with the five positive statements about research outputs. On average the measure about knowledge outputs "Valuable knowledge was produced through the project" had the highest level of agreement (4.78 in possible range of 1-5). In contrast the measure related to different knowledge systems contributing to outputs had the lowest level of agreement (3.88). Both these measures related to Indicator AO30 "Legitimate, valuable, and rigorous knowledge developed according to the various knowledge systems involved in the research". This suggests it is right to unpack some of the complex research impact indicators that clump multiple concepts into separate measures for evaluation.

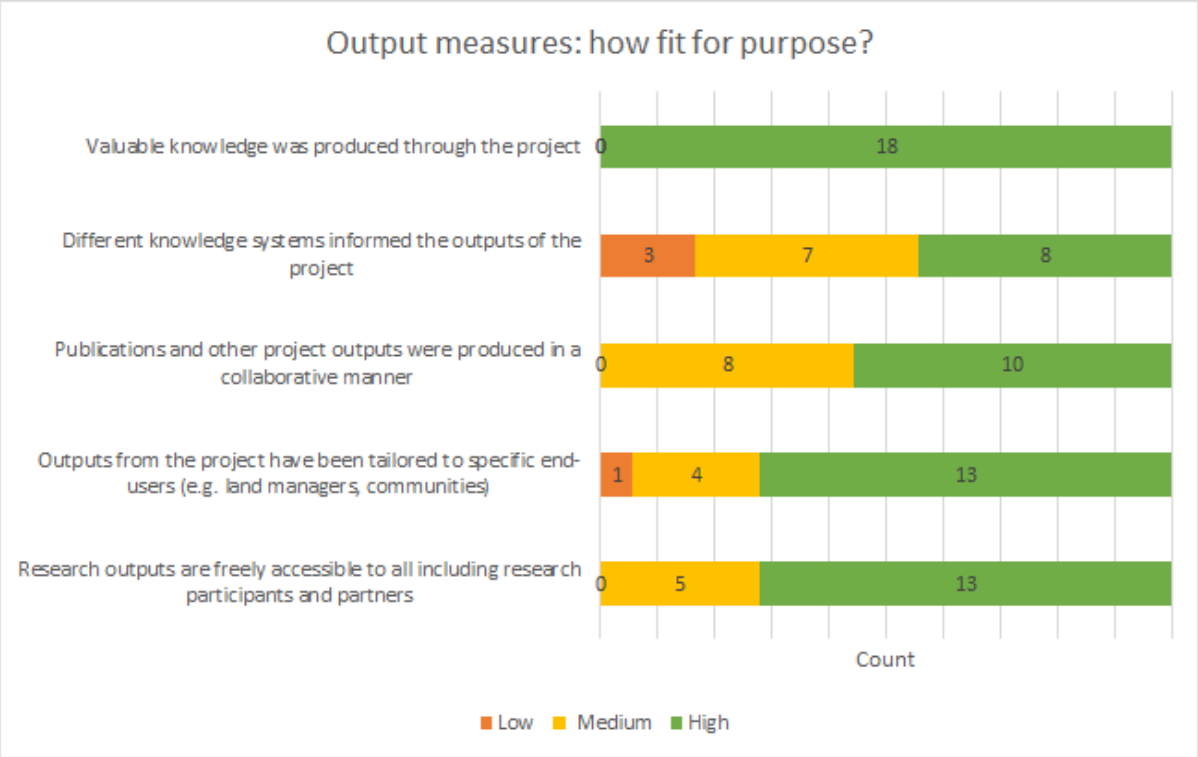
Fitness for purpose results

Most outputs measures were ranked by survey participants as highly fit for purpose. All project participants agreed that whether valuable knowledge was produced through the project was an important measure of the value of their project (Figure 4.7). Academics rated collaborative publication as somewhat lower (moderate-high), and whether different knowledge systems informed the outputs as moderate (Table 4.5).

Table 4.5. Comparison of "Fit for purpose" survey responses from academic and non-academic organisations in relation to output indicators. Response options were: High (3); Medium (2); Low (1) & Not Applicable (-) (dark yellow: scores 1 - 1.5; pale yellow: >1.5 - 2; pale green: >2 - 2.5; dark green: >2.5 – 3).

Indicator	Average score (1-3 range)	
Outputs [overall average 2.64]	Non- academic org avg (n=7)	Academic org avg (n=11)
Valuable knowledge was produced through the project	3.00	3.00
Different knowledge systems informed the outputs of the project	2.57	2.09
Publications & other outputs produced in a collaborative manner	2.71	2.45
Outputs from the project tailored to specific end-users	2.57	2.73
Research outputs freely accessible to all	2.57	2.82

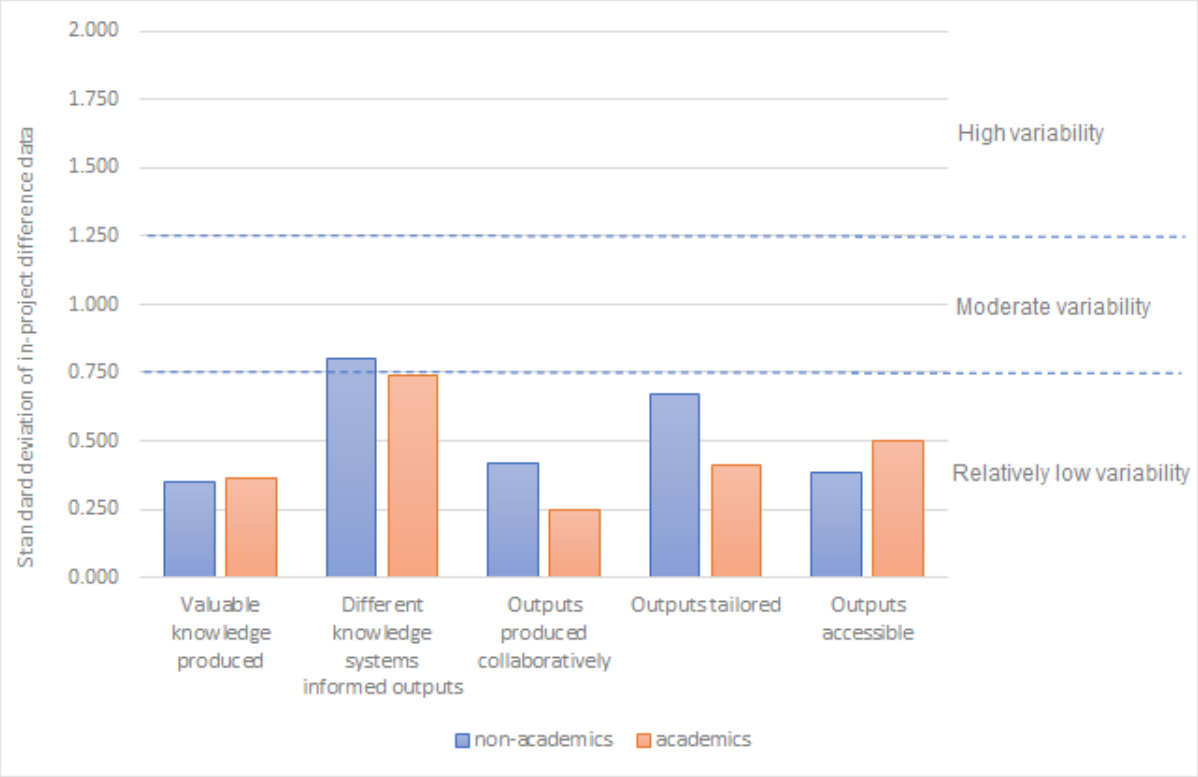
Figure 4.7. Summary of survey responses to fitness for purpose for Outputs measures



Within-project concordance and variability

Most of the outputs measures showed low to moderate within-project variability (Figures 4.8 and 4.10) and fair concordance (Figures 4.9 and 4.11). The measure of whether different "knowledge systems" informed the outputs of the project appeared to show some level of non-concordance between academic and non-academic project partners, with academic respondents appearing to rank their projects lower against this measure, though the high variability within this measure swamped any significance. The variability in part likely reflects different interpretations of the ambiguous phrasing "different knowledge systems", but the difference measure may also point to a greater appreciation among non-academic respondents of the influence of a range of knowledge systems (whether Indigenous or otherwise) at play in the research. As with the similar processes measure (RM16, see previous section on Processes), the ambiguity in language likely confounds this. As one respondent said, "the wording around the item 'different knowledge systems...' is impenetrable for me. I don't understand." This measure could usefully be reworded to clarify language, for example including both explicit reference to Indigenous knowledge systems (where relevant) and separate reference to other knowledge systems (where desired). There is also potential value in including an open-ended text option to describe the knowledge systems use, in order to seek greater clarification. Further research might draw out how different groups understand this concept of differing knowledge systems.

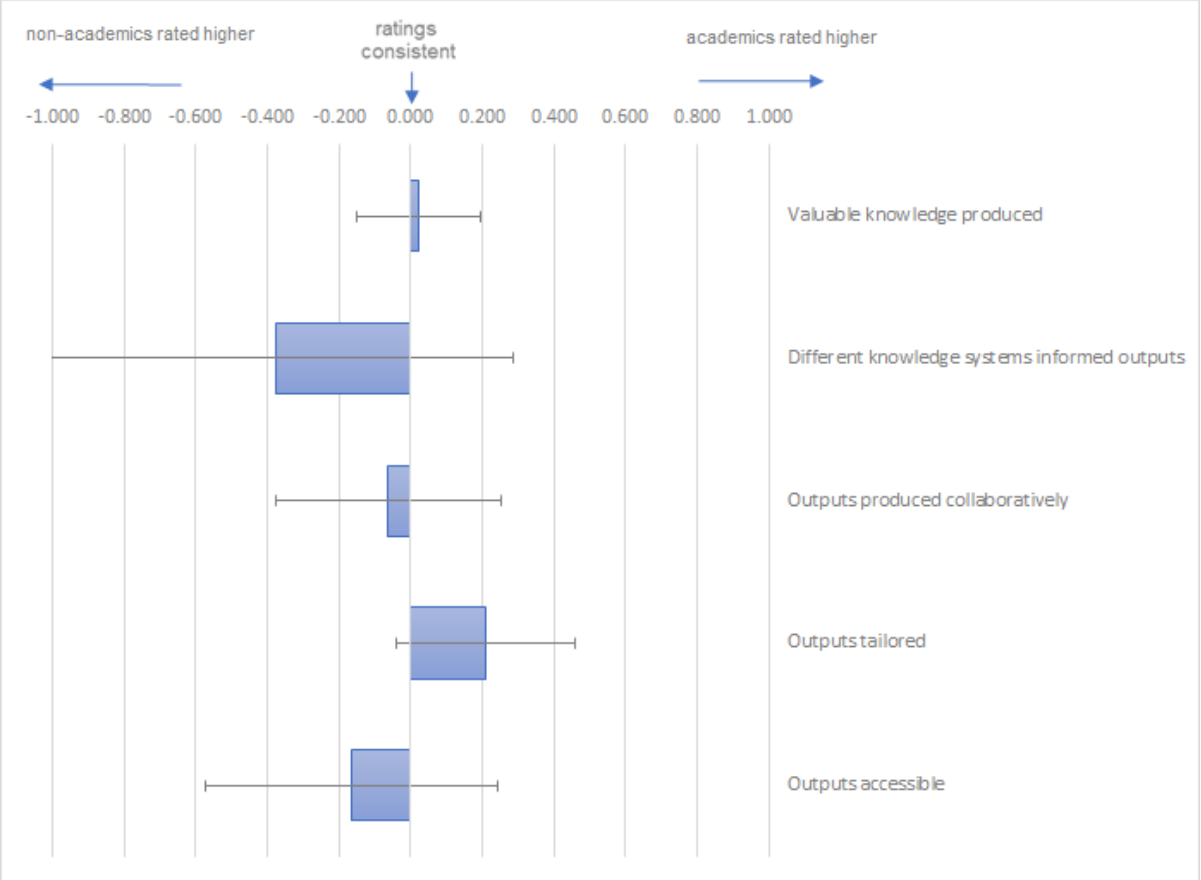
Figure 4.8. Within-group variability for outputs measures, calculated as the standard deviation of data recording departure of individuals' ratings from project means, grouped by non-academic and academic respondents. Raw data in these fields is on a 5-point scale.



One respondent suggested that the measure on whether research outputs are freely accessible may be better focused to emphasise broad accessibility: "I would reorder the indicator about accessibility to emphasise the all (as in public, non partners and participant orgs), in the current sentence. I feel the partners and participants are emphasised but that is the easier criteria to satisfy." This could also be addressed by separating the question into accessibility for research participants, research partners (especially community partners in the case of groups that may not access research on scholarly or research-based websites), and the broader public.

Non-academic respondents appeared to show generally less awareness of products from the project than their academic partners. This was the case across the board, with the exception of presentations (which all survey participants reported) and collaborative publications (where there was a high level of concordance). The differences were statistically significant adding up across all tangible output types ($\bar{x}$=0.88; p=0.001). The consistency of this pattern suggests that there may be real patterns of difference here. Survey questions on different types of outputs may be a valuable way to give an indication of breadth of awareness of products across project teams. This is particularly true if the survey methodology allows for comparison between project team members or with 'official' lists of project outputs, but aggregate levels of difference between numbers of products and research user awareness of products across programs could also be revealing (see Chapter 5 for a worked example). It is potentially worth modifying these to ask how many of each product type.

Figure 4.9. Within-project differences in outputs measures, calculated as the difference in project means between academic and non-academic assessments of projects against each measure. Raw data in these fields is on a 5-point scale.



It should be noted that not all research team members necessarily have an interest or need to know about every kind of product, and nor do all academic team members always show awareness of all of the products (particularly those who are mainly in an advisory capacity). So some level of judicious application of these comparative measures of awareness would be warranted. Nevertheless, if measures of this kind were to be asked of research users, it may be a useful overall gauge of the breadth of productivity of a research program in terms of the range of outputs of which research users are aware.

All respondents agreed that presentations had been given, suggesting this measure may not be so useful in its current form. Where there is interest is in the range of presentations given, and the diversity of audiences for these, there could be value in this question if it were to ask about presentations to particular groups of interest (e.g. academic audiences, research collaborators and partners, wider groups of potential research users, community or popular audiences, schools etc) or particular agencies (e.g. government departments).

Figure 4.10. Within-group variability for outputs measures, calculated as the standard deviation of data recording departure of individuals' ratings from project means, grouped by non-academic and academic respondents. Raw data in these fields is yes/no (binary scale).

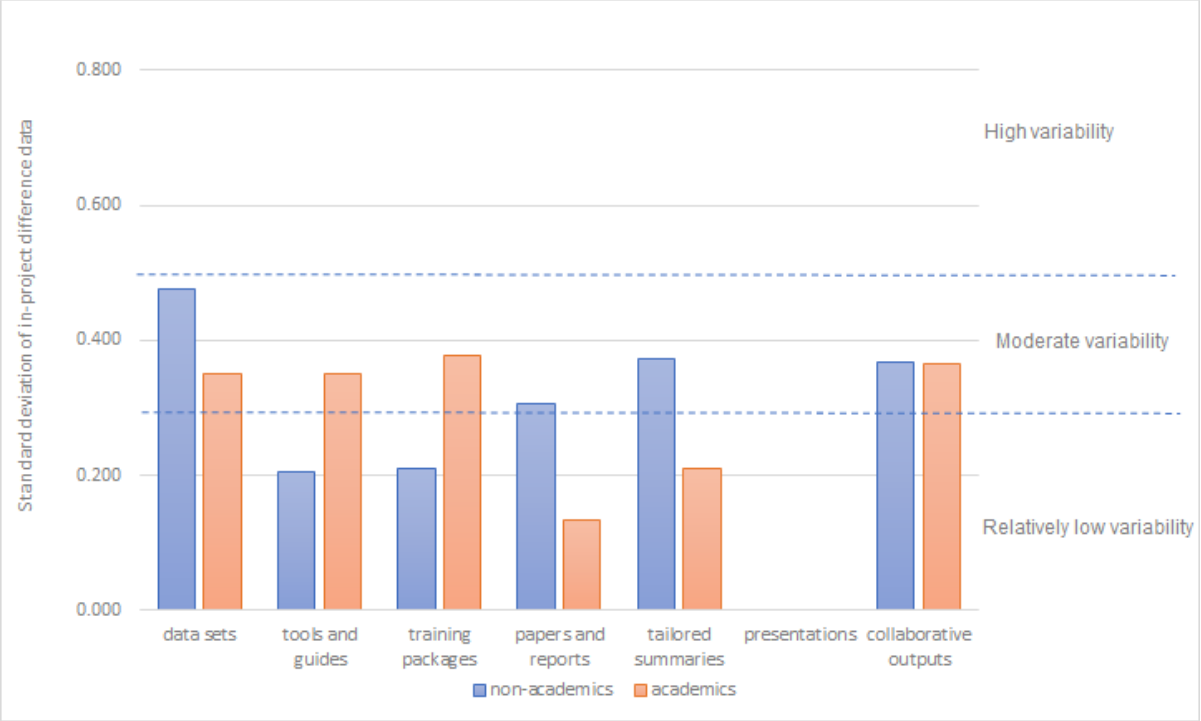
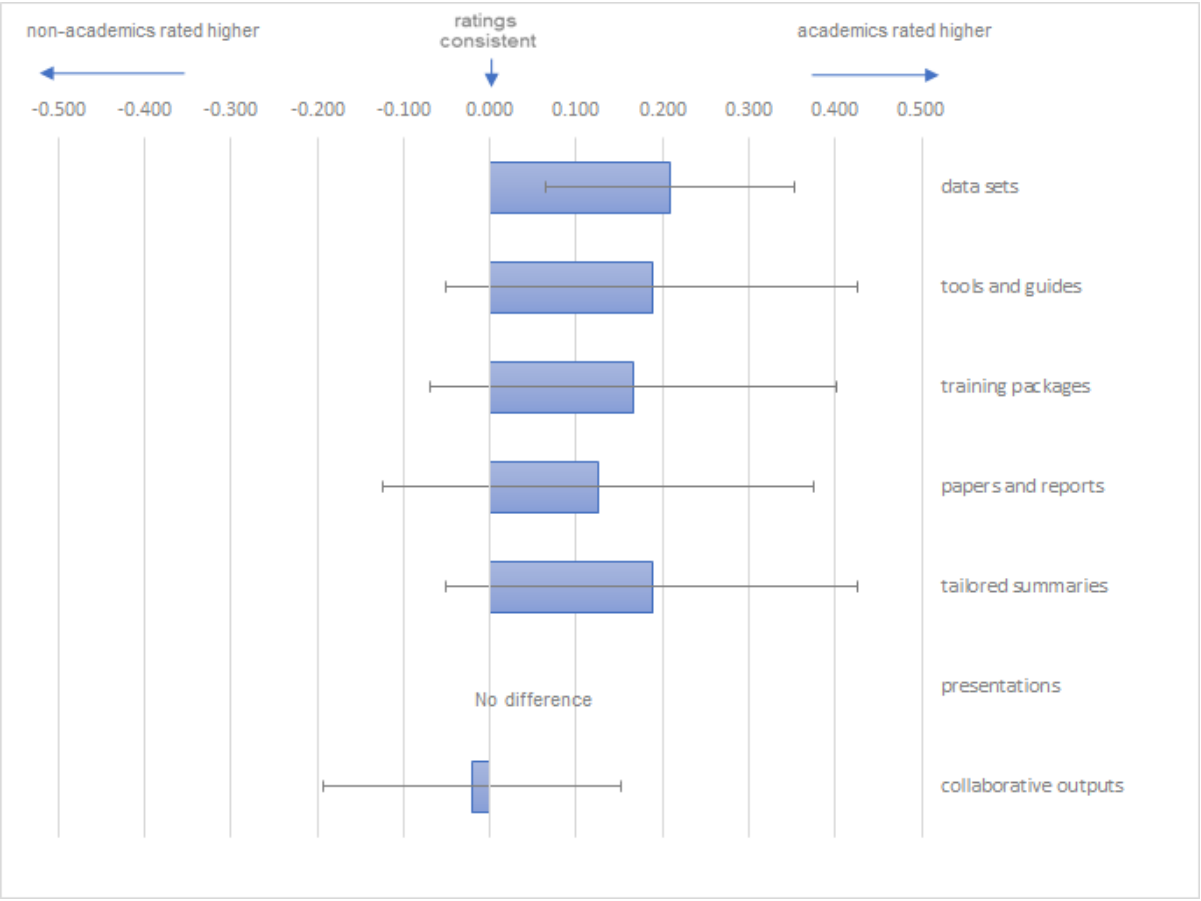


Figure 4.11. within-project differences in outputs measures, calculated as the difference in project means between academic and non-academic assessments of projects against each measure. Raw data in these fields is yes/no (binary scale).



Summary of findings for outputs measures tested in the survey

Table 4.6. Summary of outputs measures and considerations for application of each measure from survey results

Indicator descriptor	Measures/questions addressing indicators	Assess-ment	Comments and considerations for applying measure	Code
Production of quality outputs				
Legitimate, valuable, and rigorous knowledge developed according to the various knowledge systems involved in the research	Valuable knowledge was produced through the project **(likert agreement 1-5)**	Good	Rated as highly fit for purpose by all respondents. Low internal variability, high consistency between academic and non-academic respondents.	AO30
	Different "knowledge systems" informed the outputs of the project **(likert agreement 1-5)**	Use with amend-ments Apply compar-atively	Rated as highly fit for purpose by non-academic respondents, only moderately fit for purpose by academic respondents. Moderate within-project variability among both groups. Academics ranked their project lower than non-academic partners against this, though within error. This question would benefit from clearer language specifying information sought. Ensure non-academic partners are surveyed where possible. Consider analysing data separately for comparison across team roles. Potentially include write-in field for types of knowledge systems. See also measure for RM16 under Processes.	
Legitimate, valuable, and rigorous knowledge developed according to the various knowledge systems involved in the research Publications relating to the research authored or co-authored by researcher or team Research publications freely and openly accessible to all including research participants and partners	What type of project outputs have you seen or are aware of? · technical publications (reports, peer-reviewed papers) **(multiple choice yes/no)**	Good: apply compar-atively, minor note on implem-ent -ation	Question gives both an indication of the outputs produced, and how aware non-academic partners are of the range and types of products delivered. Relatively low within-project variability, but non-academic partners show less awareness of these products. Potentially modify to ask numbers as well as types.	AO30 AO33 RD46
Research tailoring and dissemination within networks				
Tailored reports, stories or summaries prepared on research and findings for specific audiences (e.g. land managers, Indigenous communities)	Outputs from the project have been tailored to specific end-users (i.e. land managers, communities) **(likert agreement 1-5)**	Good: survey across team	Rated highly fit for purpose. Relatively low within-project variability. Academic respondents may rate slightly higher. Survey both groups where possible. Consider analysing data separately for comparison across team roles.	RD51

Table 4.6 (cont). Summary of outputs measures and considerations for application of each measure from survey results

Indicator descriptor	Measures/questions addressing indicator	Assessment	Comments and considerations for applying measure	Code
Research tailoring and dissemination within networks (continued)				
Guidelines, guides, checklists, standards for dispersed but like-minded audiences (e.g. parties who do environmental monitoring, all parties who participate in species translocations) Tools made available to land managers, practitioners and citizens/communities for on-ground action (management plans, planning frameworks and processes) Training packages developed for practitioners, citizens, community members and stakeholder groups Data sets (including spatial) made available to other researchers, practitioners and citizens/communities Summaries for policy makers prepared (policy options papers, submissions to policy forums Contribution to public policy advisory committee(s)	What type of project outputs have you seen or are aware of? · tools for implementation (e.g. planning frameworks, guidelines, checklists) · training packages for practitioners, citizens, community members · data sets · summaries and other outputs tailored for specific audiences (e.g. factsheets) **(yes/no)**	Good: apply comparatively w mod-ification	Question gives both an indication of the outputs produced, and how aware non-academic partners are of the range and types of products delivered. Low to moderate within-project variability. Non-academic partners show less awareness of these products on the whole. Consider surveying across project team to gauge awareness of products. Potentially modify to ask numbers as well as types.	RD52 TD43 TD44 AO32 RD51 RD53 RD54
	What type of project outputs have you seen or are aware of? · presentations **(yes/no)**	Good: apply comparatively w mod-ification	Split question about presentation between those specifically designed for research users (relevant here), and more general awareness-raising and broadcast presentations.	
Presentations to schools, community groups and forums	What type of project outputs have you seen or are aware of? · presentations **(yes/no)**	Good: apply comparatively w mod-ification	Split between presentations specifically designed for research users, and more general awareness-raising and broadcast (relevant here). Likely of greater value to specify different groups of interest (e.g. academic audiences, potential research users, community).	RD55
Collaborative research publications or outputs	What type of project outputs have you seen or are aware of? · collaborative outputs (may be any combination of the outputs displayed in this section) **(yes/no)**	Possibly use, preliminary filter for relevance, value add	Moderate within-project variability among both groups, but good within-project concordance. Consider surveying the whole project team. Useful if information is sought on awareness of this across project team. Where information is sought on the extent to which collaborative outputs were produced, could be more effectively measured quantitatively.	AO39
	Publications and other project outputs were produced in a collaborative manner **(likert agreement 1-5)**	Good	Rated as highly fit for purpose by non-academic respondents, moderate-high by academic respondents. Low within-project variability among both groups and good within-project concordance.	
Research promoted widely				
Research publications freely and openly accessible to all including research participants and partners Data sets (including spatial) made available to other researchers, practitioners and citizens/communities	Research outputs are freely accessible to all including research participants and partners **(likert agreement 1-5)**	Good: minor modific-ation	Rated as highly fit for purpose by both groups. Low within-project variability, high concordance between respondents. May be worth modifying question to emphasise broad accessibility over sharing with partners and research participants.	RD46 AO32

Outcomes

Outcomes measures related to whether the research had or was likely to lead to different types of change (e.g. to on-ground action or decision-making in government, NGO or for-profit sectors) and how confident respondents were of this. For on-ground practice measures (Table 4.8), they were also asked to distinguish between change within their project region or partnership, and change extending beyond these immediate sets of relationships and sites of research. We did not make this scale question contingent on which types of outcomes an individual had selected. This was a mistake as four individuals did not nominate a scale for at least one of the outcomes they had selected. It is therefore unclear if they forgot which ones they had selected or deliberately skipped the scale question. This should be rectified in future applications.Two likert-scale questions sought participants' levels of agreements to strengthened relationships and intercultural capacities.

Many of these indicators displayed moderate to high degrees of within-project variability, or differences between the responses of academic and non-academic respondents that were significant or suggestive of the need for further investigation. When the degree of confidence was taken into account, the degree of variability increased (as would be expected), but the direction of these differences remained in most cases. Most measures that were more consistent were rated as being less fit for purpose, with the answers of most respondents eliciting less assent and less confidence altogether. The one outcomes measure that proved more reliable and consistent between project teams was the measure on the strengthening of relationships between researchers and stakeholders. For those measures assessing changes to policy or practice, and for that on intercultural capacities, it is likely going to be most useful to survey both academic research team and partners, and/or to ask for further, specific, clarifying details, in order to gain more confidence in the results.

Fitness for purpose results

Outcomes measures received a significantly more mixed assessment of their fitness for purpose than measures in other categories (Table 4.7). The outcomes measures rated moderately high to highly fit for purpose on average were improvements to on-ground monitoring, management and adaptive management, and changes to public policy. Other measures showed marked differences between academic and non-academic respondents. Of the measures for on-ground action, changes to threatening processes was assessed as a highly relevant question for a substantial minority of projects (7 respondents out of 17, Figure 4.12), but almost an equal number felt it to be of low fitness for purpose for applying to their projects (5) while the remainder rated it as moderate. However, looking at the differences between academic and non-academic respondents, those outside of research institutions rated this measure as moderately high on average, while academic respondents rated it as lower. Along with monitoring, management and adaptive management, measuring changes to threatening processes is therefore likely to be useful to assess for conservation projects of the kinds tested in this set of case studies, as this measure is valuable to on-ground partners.

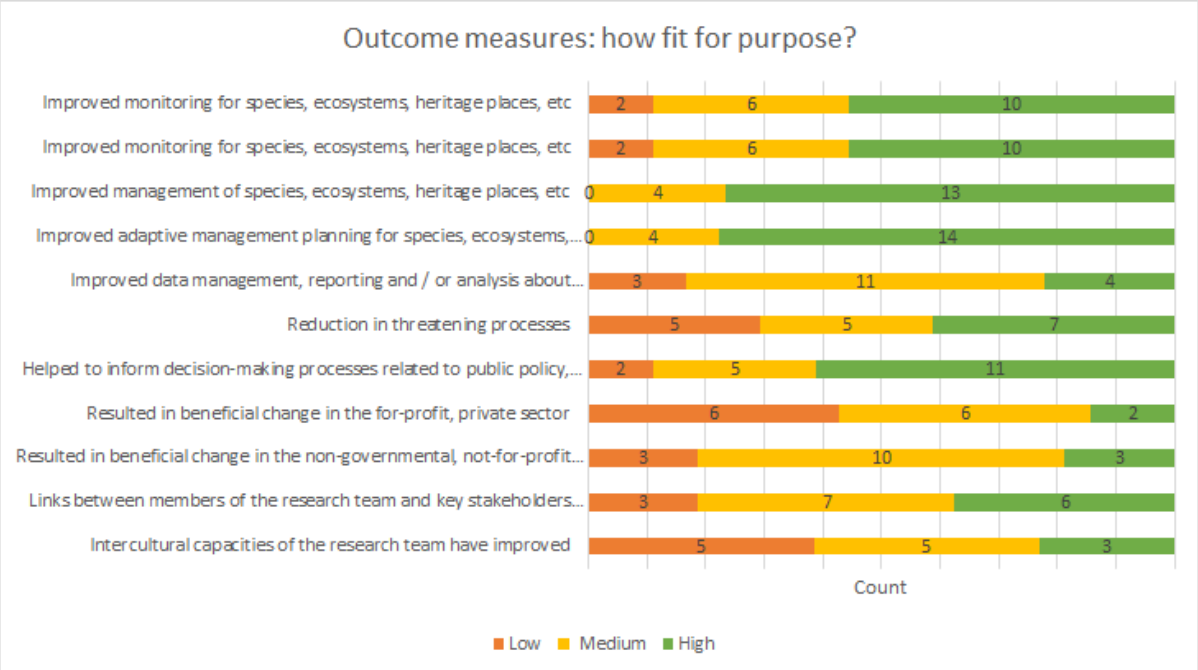
The likert scale measures of improved relationships and intercultural capacity also received a mixed reception. The most common rating for improved relationships was moderately fit for purpose (7), while five respondents rated improved intercultural capacities low and the same number moderate (Figure 4.12). However, again, these measures were rated substantially higher by project team

members outside of research institutions than by academic respondents, suggesting that these could usefully be retained, given they appear to be valuable to on-ground and policy practitioners

Table 4.7. Comparison of "Fit for purpose" survey responses from academic and non-academic organisations in relation to output indicators. Response options were: High (3); Medium (2); Low (1) & Not Applicable (-) (dark yellow: scores 1 - 1.5; pale yellow: >1.5 - 2; pale green: >2 - 2.5; dark green: >2.5 – 3).

Indicator	Average score (1-3 range)	
Outcomes [overall average 2.24]	Non- academic org avg (n=7)	Academic org avg (n=11)
Improved monitoring for spp, ecosystems, heritage places, etc	2.29	2.55
Improved management of spp, ecosystems, heritage places, etc	2.67	2.82
Improved adaptive mgmt planning for spp, ecosysts, heritage places	2.57	2.91
Improved data management, reporting &/or analysis about spp, ecosysts, heritage places, etc	1.86	2.18
Reduction in threatening processes	2.50	1.91
Helped inform decision-making: public policy, strategy +/or prog design	2.29	2.64
Resulted in beneficial change in the for-profit, private sector	1.80	1.67
Resulted in beneficial change in the NGO, not-for-profit policy sector	1.83	2.10
Links between research team & key stakeholders improved	2.50	2.00
Intercultural capacities of the research team have improved	2.50	1.56

Figure 4.12. Summary of survey responses to fitness for purpose for Outcomes measures



.

Questions on improved data management, and changes in decision making among NGOs elicited more positive responses from academics, but less so from non-academic partners. However, noting the low number of NGO respondents in the survey, there would be value in cross-checking this with a wider array of NGO partners and projects oriented toward working with NGO agencies. Changes to decision-making in the for-profit sector was rated as of lower importance by both groups, with almost half of respondents (6 out of 14) considering this to be of low fitness for purpose for their projects (Figure 4.12). All three of these indicators, but particularly improvements to data management and changes in the for-profit sector, could therefore be used more selectively and sparingly, where these matters are more relevant to projects or to the purposes of the assessment itself.

Within-project concordance and variability

Variability in responses to whether there had or would likely be changes within and beyond the project region or partnership was low to moderate across these measures (Figure 4.13 (a) and (b)). Some of the variability in responses may arise from a lack of clarity in the phrasing of 'Change within/beyond project region or partnership' suggesting there may be some value in including a clearer definition and explanation of what is meant by these terms. However, much of the variability is quite possibly attributable to inherent uncertainties and differences in perspectives on whether these changes have happened or are likely to occur.

Estimates of outcomes within the project region showed fairly consistent responses between academic and non-academic respondents within projects, except for the measure for improvements in monitoring. Here, academic respondents appeared somewhat more likely to attest to improvements in monitoring within the project area, though results were not significant given the degree of variability (Figure 4.14). By contrast, academic respondents were no more likely to attest to improvements in monitoring beyond the project region or partnership, but appear to be more likely to suggest that improvements in management had or would take place beyond the project region or partnership ($\bar{x}$=0.21 binary scale; p=0.06). There also appeared to be some divergence on adaptive management and data management beyond the project region or partnership (Figure 4.15). These results were not significant, again given the high variability and small sample size, but point to a possible need to be cautious in applying these measures without taking care to survey whole project teams and/or request more detailed information.

Adjusting for confidence rating generally increased within-project variability while not substantially altering these concordance trends. This would be worth testing over larger sample sizes, which could reveal whether moderating responses through confidence ratings reduces or removes some of the within-project differences (as it does within this small sample for the question of adaptive management beyond the project region or partnership).

Figure 4.13. Within-group variability for outcomes measures of change to on-ground action (a) within project boundary and (b) beyond project boundary. Calculated as the standard deviation of data recording departure of individuals' ratings from project means, grouped by non-academic and academic respondents. Raw data in these fields is yes/no (binary scale).

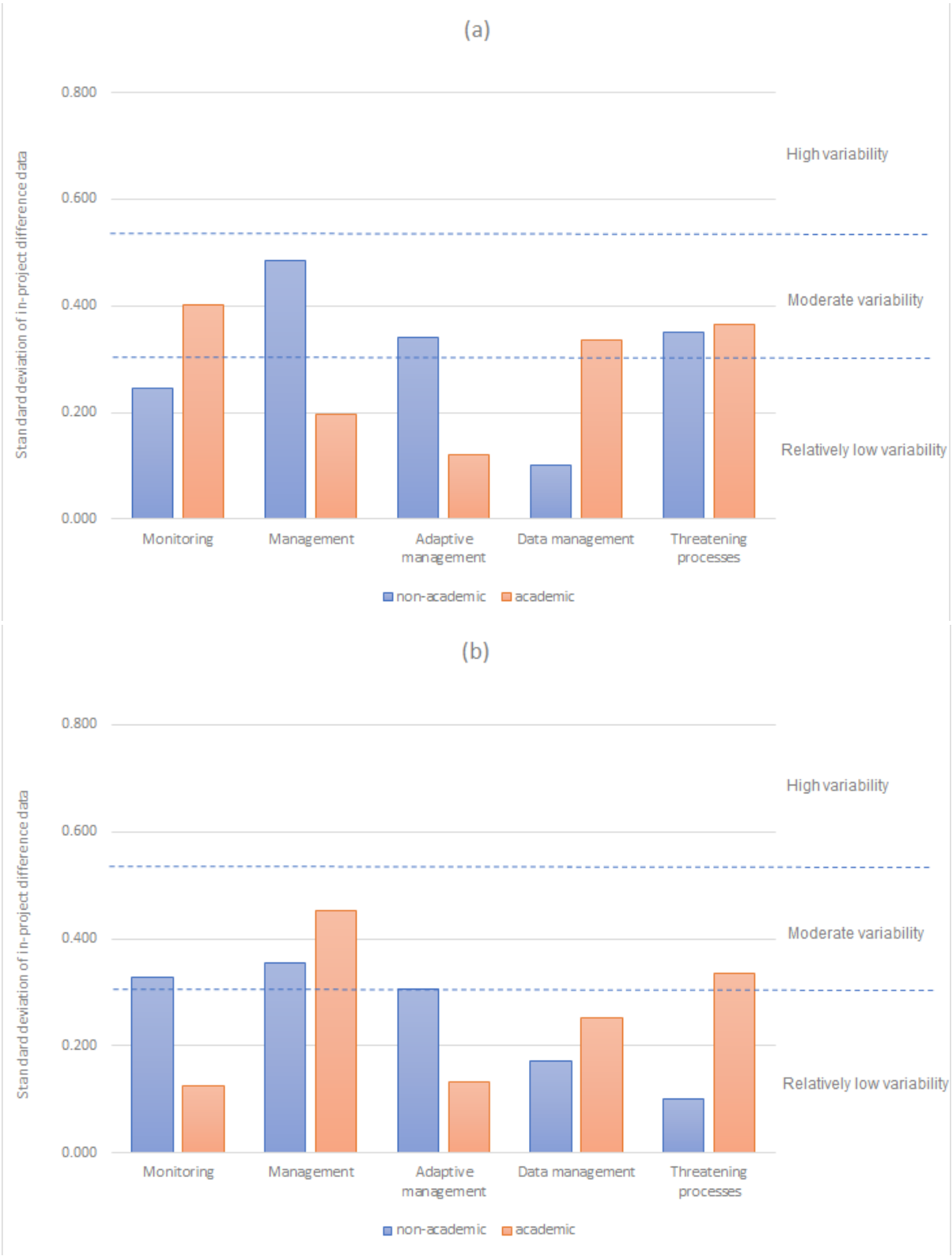
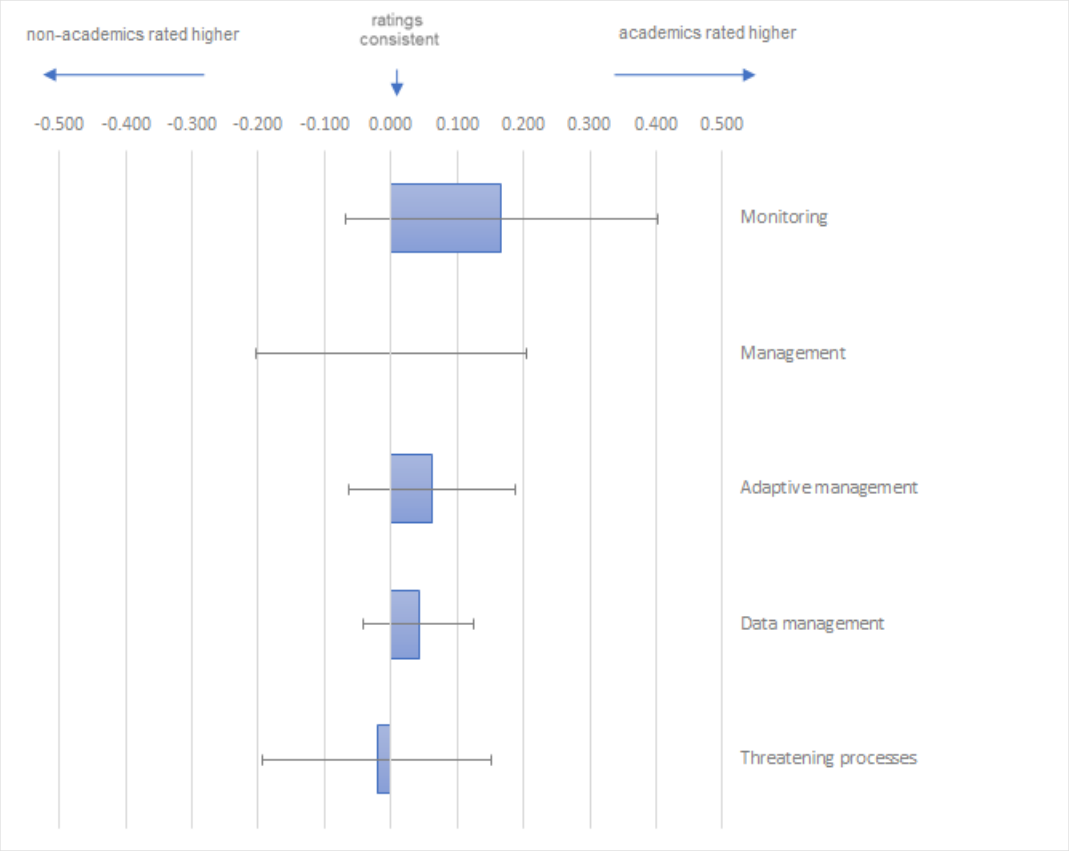


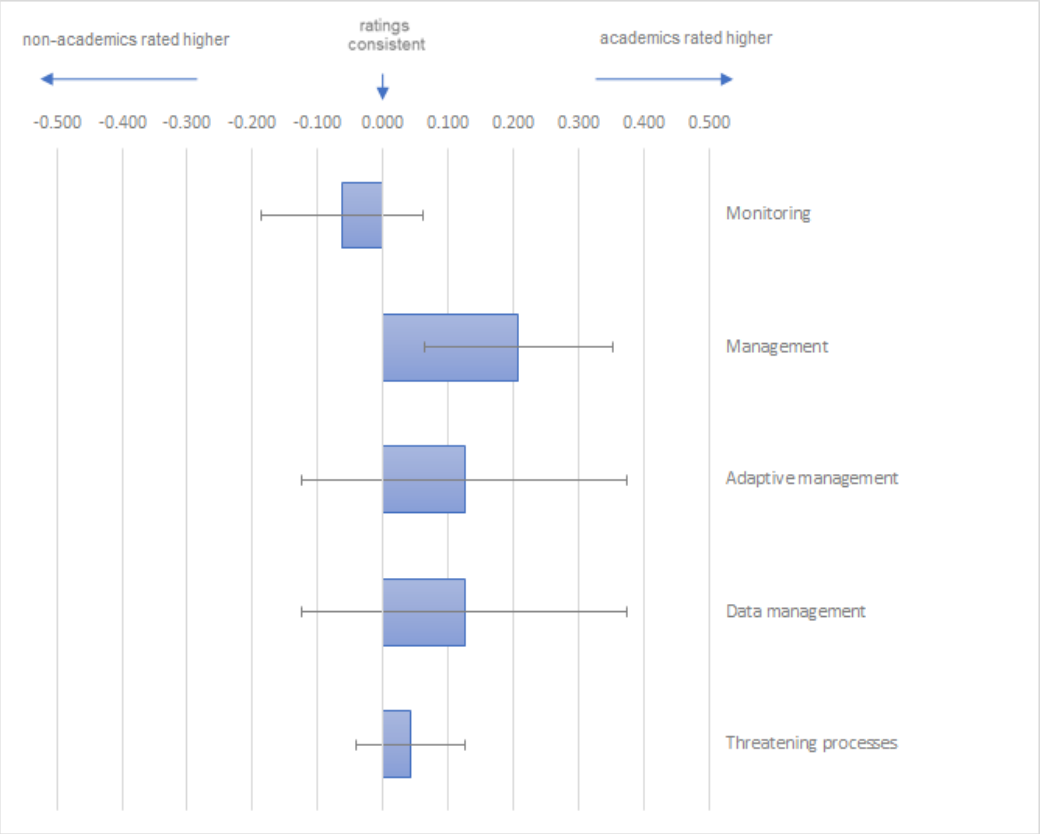
Figure 4.14. Within-project differences in outcomes measures for changes to on-ground action within the project partnership or region, calculated as the difference in project means between academic and non-academic assessments against each measure. Raw data in these fields is on a yes/no (binary) scale.



As noted above, some of the within-project variability and differences between project team members is likely related to the fact that this framework is designed to be capable of being applied during the life of research projects as well as after the discrete piece of work is formally completed. As one academic respondent commented: "The project is still ongoing so it is difficult to answer whether change has actually occurred for some indicators." It is natural to expect that inputs, processes and (to some degree) outputs measures are more likely to yield positive results at earlier stages in project cycles, or even (as with the projects tested here, in the later stages of projects but before there has been a broad opportunity for uptake).

Asking whether outcomes measures will occur or are likely to occur, and measuring the degree of confidence of respondents, allows room for projected and expected outcomes within the life of the project. If these measures were to be taken at multiple points during and shortly after a research cycle, assessments might be expected to reach a greater level of certainty and hopefully convergence at later points in the development of a project. Adding explanatory comments within the question framing that reassures respondents that projects underway may not have achieved all of these outcomes, and encouraging projections where possible, could help alleviate expectations that projects at earlier stages will be negatively judged for not yet having realised a significant body of changes in practice. Likewise moderation of the question with a confidence rating potentially helps give reassurance to participants and those collecting the data that the estimates are within reason, and provide indications of which outcomes are more certain between projects and/or across project teams. As one researcher states: "The Q22 confidence question is a really good addition to the evaluation of these outcome statements."

Figure 4.15. Within-project differences in outcomes measures for changes to on-ground action beyond the project partnership or region, calculated as the difference in project means between academic and non-academic assessments against each measure. Raw data in these fields is on a yes/no (binary) scale.

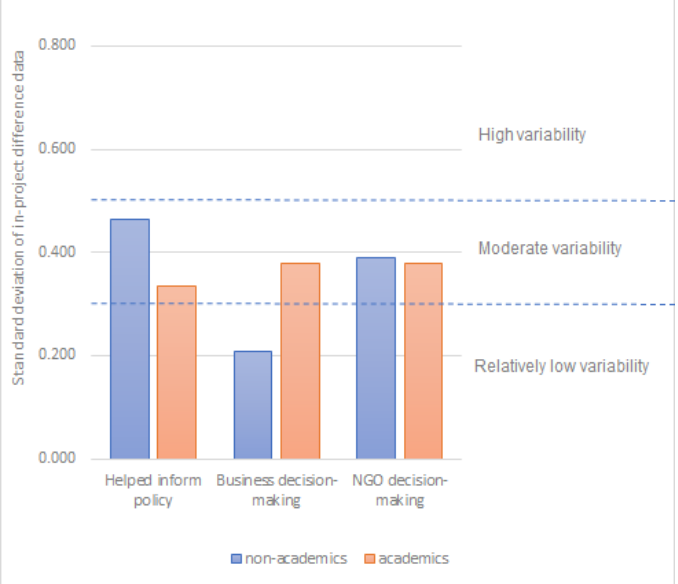


Within-project variability for responses to whether the project had helped inform decision-making in governments, for-profit and NGO sectors was generally moderate (Figure 4.16). Though responses were within error, it is notable that non-academic respondents, most of them from government, were more likely to attest to the project helping inform government decision-making (Figure 4.17). Academics were more likely to attest to the project helping NGO decision-makers, but in this case, adjusting for confidence rating removed these differences for this measure. It should be noted that there were a low number of NGO respondents among the non-academic responses, suggesting it could be potentially valuable to test these measures with NGO participants on projects with a particular focus on informing NGO partners, to understand if these differences are systematic and/or any patterns that might emerge here.

More robust assessments of on-ground action and adoption in decision-making, could likely be made by surveying both academic and non-academic project team members and seeking a higher degree of saturation in project teams than we were able to do. Seeking additional information is also likely to be of benefit, because of the inherent uncertainty in whether changes will likely take place, and because different project team members may have inherently different insights into the uptake and adoption of findings. Firstly, we suggest splitting the options between changes that have taken place and changes that are anticipated to take place in future. This could help remove variability arising from predicting change versus observing change. Secondly, including text-based responses inviting descriptions of changes, where feasible, could furnish concrete examples and contextual information. While this places an extra burden of time on respondents, it could significantly improve clarity by

resolving differences in perspectives between project team members, including those based in research institutions versus practitioner and policy agencies.

Figure 4.16. Within-group variability for outcomes measures for change to decision-making of government, for-profit and NGO sectors. Calculated as the standard deviation of data recording departure of individuals' ratings from project means, grouped by non-academic and academic respondents. Raw data in these fields is yes/no (binary scale).



Of the two questions testing agreement with outcomes measures on a likert scale, improvements in the links between researchers and stakeholders showed low within-project variability and high concordance, whereas the measure for improved intercultural capacity garnered moderate to high within-project variability (Figure 4.18). Academic respondents also appear to be less likely to attest to this latter measure than non-academic respondents (though the results are not significant with this small sample size and given the high degree of variability, Figure 4.19). Some ambiguity in the language of - or judgement of - intercultural capacities may be at play here in the variability, but it is possible that non-academic and academic respondents view this differently. This should be further investigated through an expanded survey sample, with greater clarification of the meaning of 'intercultural capacities'.

Figure 4.17. Within-project differences in outcomes measures for helping inform decision-making in government, for-profit or NGO sectors, calculated as the difference in project means between academic and non-academic assessments against each measure. Raw data in these fields is on a yes/no (binary) scale.

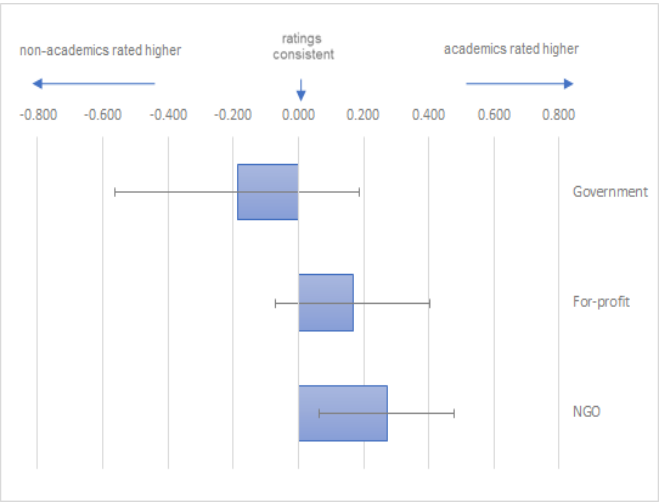
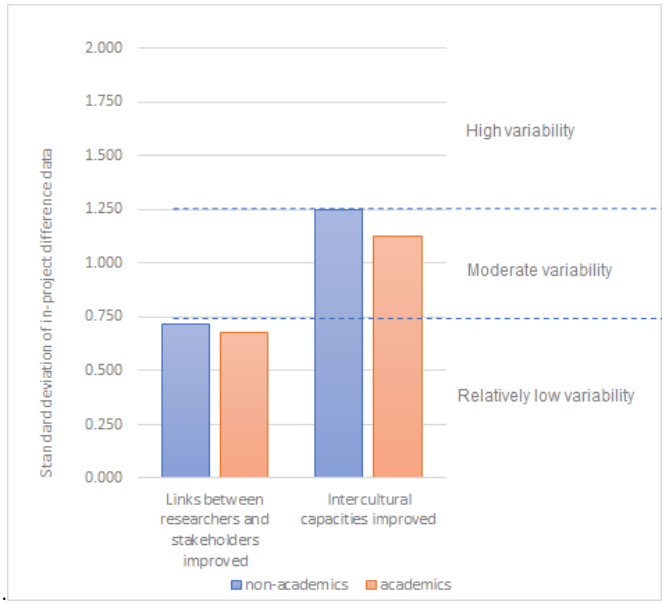


Figure 4.18. Within-group variability for change to capacities. Calculated as the standard deviation of data recording departure of individuals' ratings from project means, grouped by non-academic and academic respondents. Raw data in these fields is on a 5-point scale.



.

Figure 4.19. Within-project differences in outcomes measures for improved capacities, calculated as the difference in project means between academic and non-academic assessments against each measure. Raw data in these fields is on a 5-point scale.

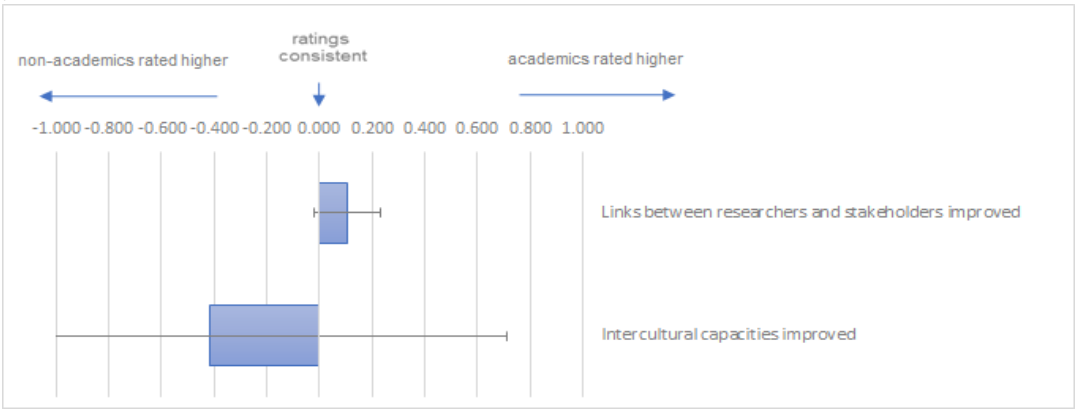


Table 4.8. Summary of outcomes measures and considerations for application of each measure from survey results

Indicator descriptor	Measures/questions addressing indicators	Assess-ment	Comments and considerations for applying measure	Code
Changes in-ground action (within-project region or partnership)				
Research findings, tools and guidelines adopted in practice and implemented in monitoring, management, community practices, regulation, business practices	· Please select which of the following outcomes related to on-ground practice your project has, or is likely to contribute: Improved monitoring for species, ecosystems, heritage places, etc · Please select which option(s) best applies to the scale of this outcome that your project has, or is likely to contribute: a) Change within project region or partnership **(yes/no and confidence)**	Use with amend-ments, use with caution, survey across team	Rated as moderately fit for purpose by non-academic respondents and high by academics. Moderate within-project variability among academic respondents. Academic respondents appear to rate higher (though within error). Requires further testing to understand whether these differences are systematic. Survey on-ground partners along with academic team. Split changes that have happened from likely changes, add free-text field for details, where feasible.	OGA72
Adoption of more effective techniques for conservation practice (e.g. captive breeding, reintroduction in the wild) Findings from research incorporated into real world experimental tests or trials	Improved management of species, ecosystems, heritage places, etc: a) Change within project region or partnership **(yes/no and confidence)**	Use with amend-ments, survey across team	Rated highly fit for purpose. Moderate within-project variability among non-academic respondents, but good within-project concordance between academics and non-academics. Survey whole project team where feasible. Split changes that have happened from likely changes, add free-text field for details, where feasible.	AR62 OGA71
Establishment of adaptive management trials, close coupled science-practice feedback loops	Improved adaptive management or management planning for species, ecosystems, heritage places, etc: a) Change within project region or partnership **(yes/no and confidence)**	Use with amend-ments, survey across team	Rated highly fit for purpose. Moderate within-project variability among non-academic respondents, but good within-project concordance between academics and non-academics. Survey whole project team where feasible. Split changes that have happened from likely changes, add free-text field for details, where feasible.	OGA73
Research findings, tools and guidelines adopted in practice and implemented in monitoring, management, community practices, regulation, business practices	Improved data management, reporting and/or analysis about species, ecosystems, heritage places, etc: a) Change within project region or partnership **(yes/no and confidence)**	Possibly use with amend-ments if relevant, survey across team	Rated low fitness for purpose by non-academics and moderate by academics. Moderate within-project variability among academic respondents, good within-project concordance. Consider applying in select circumstances where this information is a priority. Survey whole project team where feasible. Split changes that have happened from likely changes, add free-text field for details, where feasible.	OGA72
Reduction in threatening process(es)	Reduction in threatening processes a) Change within project region or partnership **(yes/no and confidence)**	Use with amend-ments, survey across team	Fitness for purpose rated moderately high by non-academic respondents but moderately low by academic respondents. Moderate within-project variability, but good within-project concordance between academics and non-academics. Survey whole project team where feasible. Split changes that have happened from likely changes, add free-text field for details, where feasible.	OGA76

Table 4.8 (cont). Summary of outcomes measures and considerations for application of each from survey results

Indicator descriptor	Measures/questions addressing indicator	Assess-ment	Comments and considerations for applying measure	Code
Changes in-ground action (beyond-project region or partnership)				
Research findings, tools and guidelines adopted in practice and implemented in monitoring, management, community practices, regulation, business practices	Improved monitoring for species, ecosystems, heritage places, etc: b) Change beyond project region or partnership **(yes/no with confidence)**	Use with amend-ments, survey across team	Rated as moderately fit for purpose by non-academic respondents and high by academics. Moderate within-project variability among non-academic respondents, but good within-project concordance between academics and non-academics. Survey whole project team where feasible. Split changes that have happened from likely changes, add free-text field for details, where feasible.	OGA72
Adoption of more effective techniques for conservation practice (e.g. captive breeding, reintroduction in the wild) Findings from research incorporated into real world experimental tests or trials	Improved management of species, ecosystems, heritage places, etc: b) Change beyond project region or partnership	Use with amend-ments, use with caution, survey across team	Rated highly fit for purpose. Moderate within-project variability. Academics appear to rate significantly higher than non-academics for same projects, which may be overestimation or because they are in position to observe changes beyond the project relationships. Survey on-ground partners along with academic team. Split changes that have happened from likely changes, add free-text field for details, where feasible.	AR62 OGA7
Establishment of adaptive management trials, close coupled science-practice feedback loops	Improved adaptive management or management planning for species, ecosystems, heritage places, etc: b) Change beyond project region or partnership	Use with amend-ments, survey across team	Rated highly fit for purpose. Low within-project variability. Academics may rate higher than non-academics for same projects, though within error. This may be overestimation or because they are in position to observe changes beyond the project relationships. Requires further testing to understand if these differences are systematic. Survey whole project team where feasible. Split changes that have happened from likely changes, add free-text field for details, where feasible.	OGA73
Research findings, tools and guidelines adopted in practice and implemented in monitoring, management, community practices, regulation, business practices	Improved data management, reporting and/or analysis about species, ecosystems, heritage places, etc: b) Change beyond project region or partnership	Possibly use with amend-ments if relevant, survey across team	Rated low fitness for purpose by non-academics and moderate by academics. Low within-project variability; low number of positive responses limits assessment. Academics appear to rate higher than non-academics for same projects, though within error. Requires further testing to understand if these differences are systematic. Apply in select circumstances where this information is a priority. Survey whole project team. Split changes that have happened from likely changes, add free-text field for details, where feasible.	OGA72
Reduction in threatening process(es)	Reduction in threatening processes: b) Change beyond project region or partnership	Use with amend-ments, survey across team	Fitness for purpose rated moderate by non-academic respondents and low by academic respondents. Moderate within-project variability among academic respondents. Low number of positive responses limits assessment, but provisional results suggest value in surveying whole project team for stronger convergence. Split changes that have happened from likely changes, add free-text field for details, where feasible.	OGA76

Table 4.8 (cont). Summary of outcomes measures and considerations for application of each from survey results

Indicator descriptor	Measures/questions addressing indicators	Assess-ment	Comments and considerations for applying measure	Code
Informed decision-making				
Research findings help to inform the decision-making process and result in change to public policy, strategy and/or program design Species or ecological community listed as threatened and / or protected by legislative instrument Landscape or threatened ecological community heritage listed Results trigger inscription of national park or conservation land tenure Increase in protected area size and/or quality	Has the project contributed to any of these other outcomes? · Helped to inform decision-making processes related to public policy, strategy and/or program design? **(yes/no and confidence)**	Use with amend-ments, use with caution, survey across team	Fitness for purpose rated as moderate (non-academics) to high (academics). Moderate within-project variability. Academic respondents appear to rate lower, but within error. Requires further testing to understand if these differences are systematic. Consider surveying whole project team, particularly government research users. Split changes that have happened from likely changes, add free-text field for details, where feasible.	PPP78 AR59 AR60 AR61 OGA75
Research processes and findings result in beneficial change to for-profit, business and private sector policy (e.g. private companies, financial sector consortium)	· Resulted in beneficial change in the for-profit private sector? **(yes/no and confidence)**	Possibly use with amend-ments if relevant, survey across team	Rated as low fitness for purpose by most respondents. Possibly apply in select cases where changes in business decisions are priority. Low number of positive responses limits assessment on variability and concordance; requires further testing. Split changes that have happened from likely changes, add free-text field for details, where feasible.	PP81
Research findings result in beneficial change to non-governmental and not-for-profit policy (e.g. International Union for the Conservation of Nature, Indigenous corporations)	· Resulted in beneficial change in the non-governmental, not-for-profit policy sector? **(yes/no and confidence)**	Possibly use with amend-ments if relevant, survey across team	Rated moderate fitness for purpose by academic respondents, low by non-academics; academics appear to rate higher (within error); possibly as a result of low number of NGO respondents. Adjusting for confidence removed differences. Requires further testing with NGO partners to understand if differences are systematic. Consider using for projects where relevant; ensure NGO partners are included. Consider splitting changes that have happened from likely changes, and adding free-text field for details, where feasible.	PP82
Improved capabilities				
Links formed/improved between members of the research team, community organisations, Indigenous communities, conservation groups, land managers, for on ground action	Links between members of the research team and key stakeholders (such as Indigenous communities, conservation groups, etc) have improved **(likert agreement 1-5)**	Good	Rated as moderate to highly fit for purpose. Low within-project variability, strong within-project concordance.	CSE65
Improved or increased intercultural capacities, via collective training of researchers on-ground personnel, in community organisations, Indigenous communities, conservation groups, land managers	Intercultural capacities of the research team have improved **(likert agreement 1-5)**	Use with amend-ments, use with caution, survey across team	Rated as moderate-highly fit for purpose by non-academic respondents and moderate-low by academic respondents. Moderate within-project variability. Researchers appear to rate lower (but within error). Requires further testing to assess whether differences are systematic. Consider language to clarify meaning of 'intercultural capacities', and surveying across project teams.	CSE66

Impacts

Survey respondents were asked to report on whether their research had or was likely to lead to a range of positive impacts for conservation, for local capacity and governance, and for social, cultural and economic areas (Table 4.10). They were also asked to assess their degree of confidence in their responses, and to rate how fit for purpose these measures were for their projects.

In general, measures assessing benefits for conservation and for local capacity were rated more highly fit for purpose than social, cultural and economic measures. However, most of these indicators displayed moderate to high degrees of within-project variability in their responses, while academic respondents rated many of these significantly lower than their non-academic partners, including when the degree of confidence was taken into account.

Fitness for purpose results

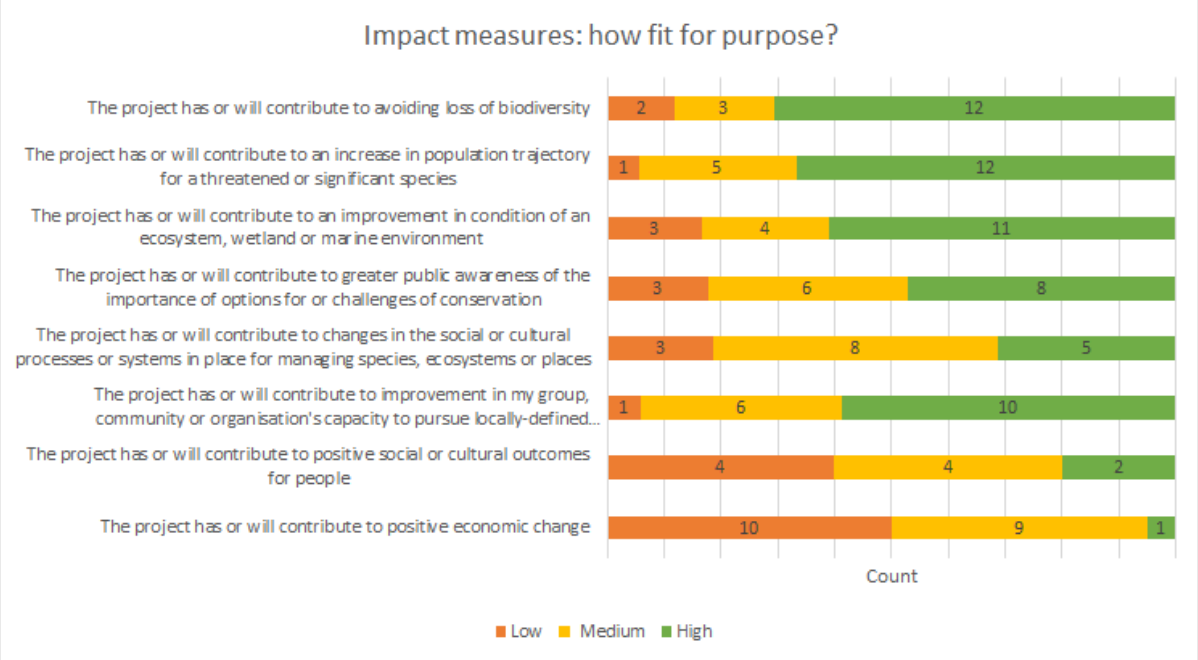
Table 4.9. Comparison of "Fit for purpose" survey responses from academic and non-academic organisations in relation to output indicators. Response options were: High (3); Medium (2); Low (1) & Not Applicable (-) (dark yellow: scores 1 - 1.5; pale yellow: >1.5 - 2; pale green: >2 - 2.5; dark green: >2.5 – 3).

Indicator	Average score (1-3 range)	
Impacts [overall average 2.24]	Non- academic org avg (n=7)	Academic org avg (n=11)
The project has or will contribute to:		
● avoiding loss of biodiversity	2.57	2.60
● increase in pop^{n} trajectory for threatened or significant sp	2.43	2.73
● an improvement in condition of an ecosystem, wetland or marine environment	2.29	2.54
● greater public awareness of the importance/ challenges of conservation	1.86	2.60
● changes in the social or cultural processes/ systems for managing spp, ecosystems or places	2.17	2.10
● improvement in my group's capacity to pursue locally-defined priorities, objectives, and goals	2.83	2.36
● positive social or cultural outcomes for people	2.00	1.78
● positive economic change	1.33	1.50

As with outcomes measures, fitness for purpose results of impact measures generated a mixed response from academic and non-academic respondents (Table 4.9). The only measure to be rated highly fit for purpose by both groups was that the project has (or will) contribute(d) to avoiding loss of biodiversity. Measures for increased population trajectory and improvement in condition of ecosystem, wetland or marine environment were rated highly fit for purpose by academic respondents and moderate-highly fit for purpose by non-academic respondents. The measure

assessing 'improvement in my group's capacity to pursue locally-defined priorities, objectives, and goals' was the only measure rated more highly fit for purpose by non-academic respondents. This is likely to reflect the wording asking about 'my group's capacity', as this form of wording is less likely to be as meaningful when applied to research institutions.

Figure 4.20. Summary of survey responses to fitness for purpose for Impacts indicators



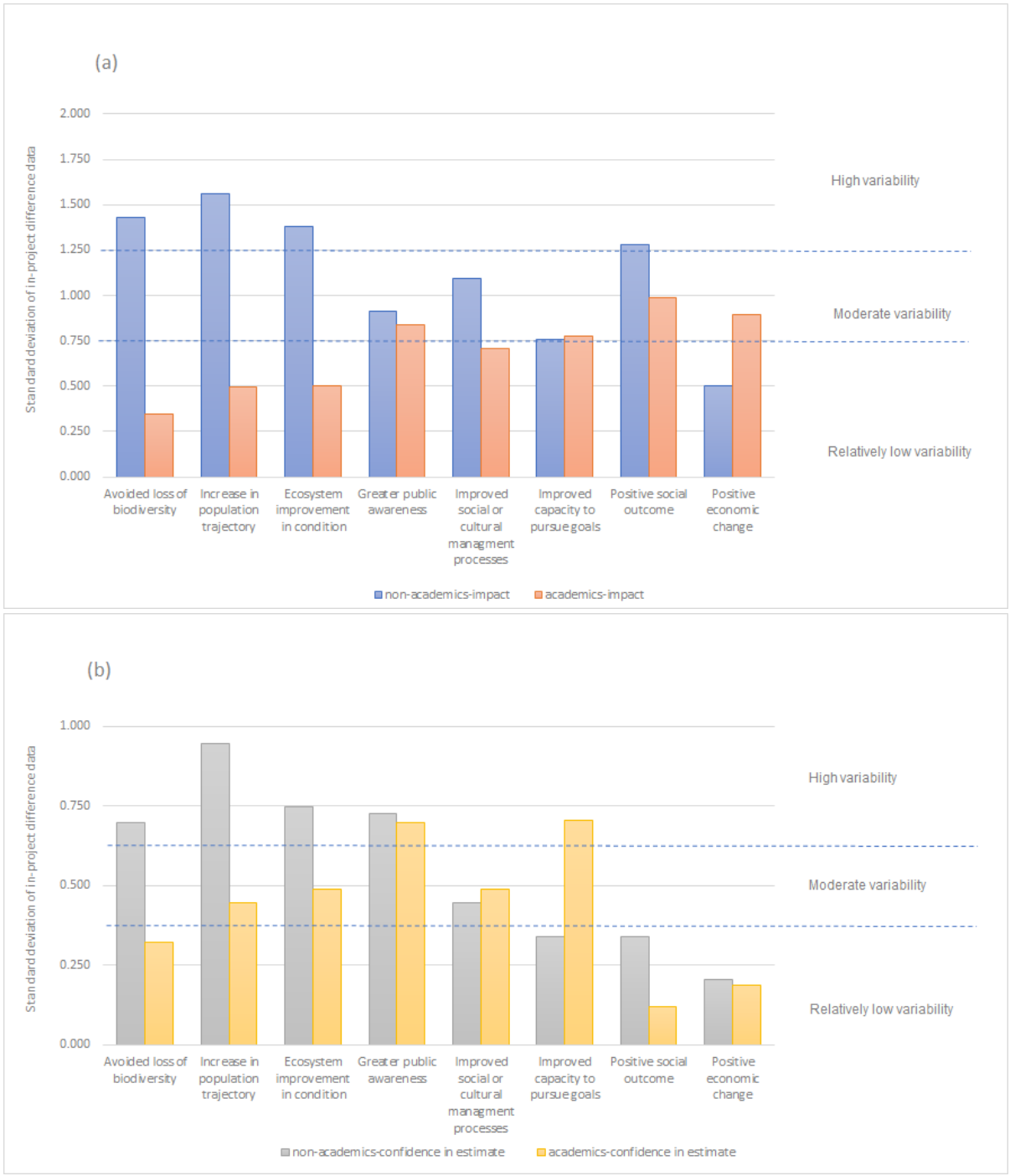
Academic respondents were much more likely to assess public awareness of conservation as being fit for purpose for their projects, while both groups of respondents assessed changes in the social or cultural systems for managing species, ecosystems or places as being moderately highly fit for purpose. Social and cultural outcomes were seen as of moderate to low suitability, while positive economic change was generally rated as not being fit for purpose for assessing the value of the projects tested here.

Within-project concordance and variability

Many of the impact indicators exhibited moderate to high variability in within-project responses (Figure 4.21), and a number of them showed substantial divergence between respondents inside and outside of research institutions (Figure 4.22).

While all three of the environmental impact measures tested - avoided loss of biodiversity, increase in population trajectory for threatened or significant species, and improvement in condition of an ecosystem, wetland or marine environment - were assessed as moderately or highly fit for purpose for these projects, only one of the three (increase in species population trajectory) generated good within-project concordance between academic and non-academic respondents (Figure 4.22), and all three indicators demonstrated a high degree of within-project variability among the assessments of non-academic respondents, relative to project averages (Figure 4.21).

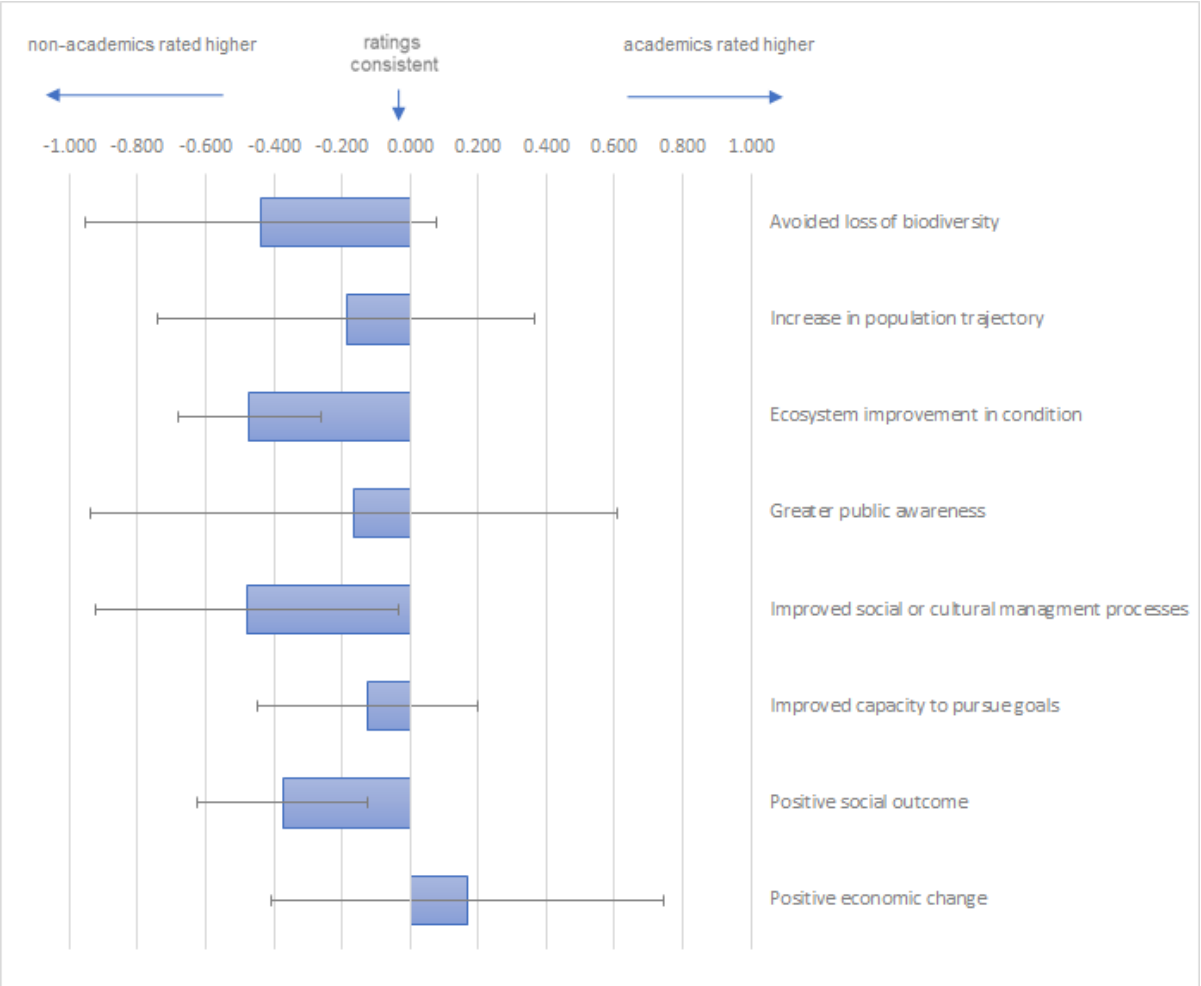
Figure 4.21. within-group variability for impact assessments on a 5-point scale (a) and confidence measures on a 3-point scale (b), calculated as the standard deviation of data recording departure of individuals' ratings from project means, grouped by non-academic and academic respondents.



This variability and divergence likely reflects both the inherent uncertainty of whether these impacts have been achieved or would likely be achieved, as well as variability in perspectives, insights and access to impact data among project partners. This pattern did not substantially change when degree of confidence was taken into account. This suggests that a survey approach may deliver inherently divergent responses depending on who is surveyed, particularly when project saturation is not reached, but also when averaged across all participants. Thus there is significant value in ensuring views are sought across project teams, and some care should be taken not to average or aggregate

these measures without ensuring clear distinctions between respondents with different roles and perspectives in the project. As with outcomes measures, separating out impacts that respondents assess to have happened from those likely to occur and including free text responses, would likely provide more clarity.

Figure 4.22. Within-project differences in impact measures, calculated as the difference in project means between academic and non-academic assessments of projects against each measure. Raw data in these fields is on a 5-point scale.



Measures for social and economic impacts, and for improved capacity, generated a more moderate degree of within-project variability (Figure 4.21). Two of the three measures in these categories assessed as being of moderate to high fitness for purpose (greater public awareness of the importance of, options for or challenges of conservation and improvement in my group's capacity to pursue locally-defined priorities, objectives, and goals) generated a low to moderate within-project variability, suggesting these are probably reasonably appropriate to apply in a range of circumstances. Within-project results for the latter would not be expected to reach convergence owing to the way the question was worded in this case study; this wording could be modified if the goal is to assess capacity for on-ground partners, to make this clearer to research team members in other institutions that they should answer in these terms.

The other indicator in these categories assessed as of moderate to high fitness for purpose (changes in the social or cultural processes and systems in place for managing species, ecosystems or places)

generated a moderate level of within-project variability and showed a substantial degree of divergence between academic and non-academic respondents (Figure 4.22). Academic respondents rated this lower than non-academic respondents, particularly once degree of confidence was taken into account ($\bar{x}$=-1.0 (rating*confidence 15-point scale); p=0.06), likely reflecting the different perspectives and insights of non-academic partners from academics on the social and cultural processes for management in place. This suggests the possibility that academic reports may under-represent the extent to which the project has supported positive change in social and cultural processes for management, and that a survey approach may deliver somewhat imprecise results when project saturation is not reached or when averaged across all participants. For this indicator in particular, surveying those directly involved in managing species, ecosystems or places may deliver more consistent results.

The remaining two indicators (positive social and cultural outcomes for people; and contribution to positive economic change) were assessed as having relatively less fitness for purpose for assessing these projects overall, although a majority of respondents still rated this measure as moderately fit for purpose for their projects. However, the survey question generated moderate to high degree of within-project variability for this measure, and respondents had a consistently low degree of confidence in their responses. There was some indication that academic respondents also appeared less likely to attribute positive social and cultural change ($\bar{x}$=-0.38 (5-point scale); p=0.06). Though this did not quite reach significance with the low sample size, this apparent divergence coupled with the high degree of within-project variability among non-academic respondents suggests that a likert-scale survey approach may not be the most useful measure for this question. Alternative approaches are discussed in the quantitative and narrative questions above.

The measure for positive economic change was not considered fit for purpose by most respondents, and most respondents rated their project as relatively low against this measure. Academic respondents seemed more likely to attest to a possibility of economic outcomes than non-academics, but the question generated a moderate degree of within-project variability among this group. This within-project variability and difference between academic and non-academic respondents reduced once the degree of confidence was taken into account, suggesting degree of confidence may be an important qualifier for this indicator, where it is used (for example, in select circumstances where this measure is considered important).

Summary of findings for impacts measures tested in the survey

Table 4.10. Summary of impacts measures and considerations for application of each measure from survey results

Indicator descriptor	Measures/questions addressing indicators	Assess-ment	Comments and considerations for applying measure	Code
Environmental benefits				
Species recovery or avoided loss, including: Population increase and/or avoided loss in threatened species Avoided loss of biodiversity	The project has or will contribute to avoiding loss of biodiversity **(likert likelihood 1-5 with confidence 1-3)**	Use with amend-ments, use with caution, apply comparat-ively	Rated highly fit for purpose, but generated high degree of within-project variability amongst non-academic respondents. Results suggest academics may assess lower than non-academic partners (though within large error range), including when degree of confidence is taken into account. Survey approach should seek responses across project teams, but may not deliver convergence. Apply with caution, avoid averaging across project teams. Split impacts that have happened from those projected. Request further information through free text where feasible.	REI87 REI90
	The project has or will contribute to an increase in population trajectory for a threatened or significant species **(likert likelihood 1-5 with confidence 1-3)**	Use with amend-ments, survey across team	Rated moderately to highly fit for purpose, but generated high degree of within-project variability amongst non-academic respondents. Academic and non-academic assessments within projects reasonably concordant, particularly once degree of confidence taken into account. Split impacts that have happened from those projected. Consider requesting further information through free text where feasible.	
Improved condition of places, including: Improvement in the maintenance and/or condition of an ecosystem, wetland, marine environment	The project has or will contribute to an improvement in condition of an ecosystem, wetland or marine environment **(likert likelihood 1-5 with confidence 1-3)**	Use with amend-ments, use with caution, apply comparat-ively	Rated moderately to highly fit for purpose, but generated high degree of within-project variability amongst non-academic respondents. Academics rate significantly lower than non-academic partners (divergence in the same direction but within error once degree of confidence taken into account). Survey approach should seek responses across project teams, but may not deliver convergence. Apply with caution, avoid averaging across project teams. Split impacts that have happened from those projected. Request further information through free text where feasible.	REI89
Enhanced governance and capacities for locally-defined priorities and inclusive decision-making				
Improvement in the abilities of parties involved in or affected by conservation research to pursue locally defined priorities, objectives and goals (e.g. Indigenous communities, community conservation groups, natural resource management authorities)	The project has or will contribute to improvement in my group, community or organisation's capacity to pursue locally-defined priorities, objectives, and goals **(likert likelihood 1-5 with confidence 1-3)**	Good: minor modific-ation	Rated highly fit for purpose by non-academic respondents. Low to moderate within-project variability. Would not expect within-project concordance due to wording of question regarding 'my group'. Academics rated significantly lower once degree of confidence was taken into account. Where interest is in local community capacity, question could be targeted primarily to respondents outside of research institutions and/or the wording of the question changed to ask about other group's capacities.	REI92
Improved capacities of management, governance and institutions to engage in effective, equitable and informed deliberation and decision-making around conservation issues	The project has or will contribute to changes in the social or cultural processes or systems in place for managing species, ecosystems or places (**likert likelihood 1-5 with confidence 1-3)**	Use with amend-ments, use with caution, apply comparat-ively	Rated moderately fit for purpose. Moderate degree of within-project variability amongst non-academic respondents. Academic respondents assessed substantially lower than non-academic respondents, more so once degree of confidence was taken into account, perhaps reflecting different perspectives and insights. Academics may underreport; surveying those directly involved in managing species, ecosystems or places may deliver more consistent results. Avoid averaging across project teams. Split impacts that have happened from those projected. Consider requesting further information through free text where feasible.	REI96

Table 4.10 (cont). Summary of impacts measures and considerations for application of each measure from survey results

Indicator descriptor	Measures/questions addressing indicators	Assess-ment	Comments and considerations for applying measure	Code
Social, cultural and economic benefit				
A better informed society with greater social licence for conservation Influence on public policy debate	The project has or will contribute to greater public awareness of the importance of, options for or challenges of conservation **(likelihood 1-5 with confidence 1-3)**	Possibly use where relevant with amend-ments, survey across team	Non-academic respondents rated fitness for purpose low-moderate. Academic and non-academic responses reasonably concordant, but survey generated a low-moderate degree of within-project variability. Split impacts that have happened from those projected. Consider requesting further information through free text where feasible.	REI94 PPP77
Improvement in human wellbeing derived from species and/or ecosystems, including access to resources, health, livelihoods, self-determination, social relations, and cultural and spiritual satisfaction: · Social and cultural measures of contribution to improvement in human wellbeing derived from species and/or ecosystems, including access to resources, health, livelihoods, self-determination, social relations, and cultural and spiritual satisfaction. · Economic contribution to improvement in human wellbeing derived from species and/or ecosystems, including access to resources and livelihoods.	The project has or will contribute to positive social or cultural outcomes for people **(likert likelihood 1-5 with confidence 1-3)**	Not recom-mended	Rated as moderately low fitness for purpose. Responses showed moderate to high degree of within-project variability. Respondents had a consistently low degree of confidence in their responses. Academics rate as significantly less likely than non-academic partners (including once degree of confidence taken into account). Suggests a likert-scale survey approach may not be the most appropriate, and alternative approaches to assess this should be sought. Where social and cultural outcomes are important to assessment and a survey approach is adopted, should seek responses across project teams, and may not deliver convergence. Apply with caution, avoid averaging across project teams. Split impacts that have happened from those projected and requesting further information through free text where feasible.	REI95
	The project has or will contribute to positive economic change **(likert likelihood 1-5 with confidence 1-3)**	Possibly use where relevant with amend-ments, use with caution, survey across team	Rated as low fitness for purpose by both academic and non-academic respondents. Many respondents rated this indicator relatively low, but academic respondents rated this slightly higher (though within error), with results showing a moderate degree of within-project variability among academic respondents. Within-project variability and differences reduced once degree of confidence was taken into account. Possibly apply, with degree of confidence qualifier, in select cases where economic change is a priority. Split impacts that have happened from those projected and consider requesting further information through free text.	Not in original study

4.3 Compiled, community insight and modelled/monitored measures

While testing quantitative, documentary, community and modelled/monitored measures was beyond the scope of this study, insights into the potential use and application of such indicators were considered by the initial project development workshop, the expert workshop and the research team. These insights have been drawn together and developed into a set of ideas for consideration. Examples are also given below of potentially applicable approaches developed in other contexts that could be applied to a range of these independent measures. These examples are not exhaustive, but are intended to provide an illustration of what could be done to develop quantitative assessment further.

These measures are broadly distinguished as follows:

- compiled measures represent information compiled from multiple sources that reflect on the implementation of the project and its uptake, and include both quantitative and documentary measures
- measures designed to assess changes beyond the project and its uptake, to environmental, economic, social and/or cultural variables, which include community insight measures and modelled or monitored measures

One observation at the expert workshop was that it is important to be "Careful about false comparisons in quantitative measures". This comment, made with respect to a particular quantitative indicator, is a caution usefully applied across the board of quantitative measures. It is tempting to assume that, because something has been given a number, it can be added, subtracted and compared. However, variability in how measures are understood and assessed, variability in intended project outcomes and reach, and other confounding factors - such as differing baselines at the commencement of research - can all contribute to significant differences in how quantitative measures are reported. This should not preclude all attempts at comparison, but suggests the need to apply significant care in any exercise of aggregation or comparison.

Quantitative, documentary and community insight inputs measures

Given the range of effective survey and narrative measures of research inputs, quantitative measures may not be so essential for research inputs as for other parts of the research process. Some options for quantitative, documentary and independent measures of research input are outlined in Table 4.11. Of these, not all are considered equally valuable or applicable. A measure of research gaps identified, through provision of documentary evidence of these gaps and prioritisation, was assessed by the expert workshop. Workshop participants felt that it should best be used selectively. Care should be taken to test whether the evidence provided gives an adequate assessment of the most important gaps and high priority research needs, and avoids the "shopping list" and "pet project" problems.

Likewise, independent community-based assessments of stakeholder involvement are not likely to be feasible in many cases, and may place an undue additional burden on research partner communities. They should be used with significant discretion. A range of documentary and quantitative measures of stakeholder involvement could be used as a partial proxy, though these do not give any indication on the quality of involvement, or whether this was appropriate to needs. Survey measures,

particularly of perspectives gained on the need for and value of research (RDRX), may be a more appropriate, ready-to-hand measure, in addition to or instead of these documentary and quantitative measures of involvement.

Table 4.11. Considerations in applying quantitative, documentary and community insight measures of research inputs

Indicator descriptor	Measures/questions addressing indicators	Assess -ment	Comments workshop and research team	Code
Research needs identified				
Research needs, gaps, unanswered questions, new areas for research identified	Documentary evidence of research gaps that have been identified and prioritised, as well as those that have been met through the project.	Use with caution	Best used very selectively if at all. Could possibly be used where active processes for collectively identifying and assessing research gaps have been undertaken as part of projects or programs. Difficult to assess concretely; critical assessment warranted of whether gaps analysis prioritises effectively and identifies the most important issues.	RDR1
Research users and stakeholders active setting of research directions				
Need for research has been adequately established - and all perspectives gained on whether research is valuable (e.g. particular end-user agencies)	**General inputs measures** · Number, diversity, geographic and sectoral coverage of stakeholder organisations involved in research planning, relative to project impact **Specific community-based assessments** Focus group, interviews, surveys, participatory evaluations to assess whether need and value of research has been established.	Possibly use, prelim-inary filter for relev-ance, value add	· General measures potentially more widely feasible, reported from project teams, but do not indicate how effectively research user values and perspectives were integrated. · Community-based assessments likely more feasible in single-project or targeted evaluations. Could complement or provide independent support to project team (academic and non-academic) survey measures, particularly in contexts where broader diversity of perspectives gained is essential. · Assessment should both assess and allow for churn among stakeholders where possible.	RDRX
Active participation of stakeholders and end-users in research	**General inputs measures** · Number of individuals from stakeholder community/ies actively engaged in research · Number, diversity, geographic and sectoral coverage of stakeholder organisations actively engaged in research, relative to project impact **Specific community-based assessments** Focus group, interviews, surveys, participatory evaluations to assess whether and extent of active engagement has been established.	Possibly use, prelim-inary filter for relev-ance, value add	· General measures potentially more widely feasible, reported from project teams, but do not indicate how effectively stakeholders were supported to take part in research and whether this was seen as desirable or appropriate . · Community-based assessments are likely more feasible in single-project or targeted evaluations. Could complement or provide independent support to project team (academic and non-academic) survey measures, particularly in contexts where broader diversity of perspectives gained is essential. · Assessment should both assess and allow for churn among stakeholders where possible.	RDR4

Counts of datasets made available through project networks and more widely are significantly more feasible and potentially valuable quantitative measures, complementing survey and narrative data. Information on the scope of these (e.g. numbers of species, geographic coverage relative to size of problem) may be more challenging to assess, but could in some circumstances provide valuable additional information for context, or with some care, for comparison and aggregation.

Table 4.12. Considerations in applying quantitative and documentary measures of research inputs - datasets

Indicator descriptor	Measures/questions addressing indicators	Assess -ment	Comments workshop and research team	Code
Previously unavailable datasets made available				
Existing datasets that were previously unavailable are liberated (e.g. industry or commercial in-confidence)	· Number and geographic scope and types of datasets newly accessible to researcher and research user networks · Number and geographic scope and data types of formerly inaccessible and/or resultant datasets now in public domain or otherwise accessible (with appropriate protections) · Documentary information and (where feasible) numerical information on data collected and geographic coverage of datasets made available, including differentiation between referents for data (e.g. species, ecosystem, geodata, remote sensing, models, genetic) and types of data (statistical vs qualitative data)	Good	Number of datasets newly available (e.g. to the project team or publicly) is potentially readily reportable by project teams later in or after the end of a project. Geographic scope and types of data collected would likely be documentary information that could provide additional information for context. It would take significantly more work and care to aggregate or compare information across programs on geographic coverage and content of data collected , but this could be done for like data (e.g. number of species for which time-series monitoring data has been made available).	RDR8

Quantitative and community insight processes measures

Table 4.13 provides a summary of the considerations of expert workshop members for process measures. A lot of this discussion focused on questions of how involved research users and partners should be in setting research directions, and on measuring and assessing indicators of trust. Workshop participants pointed out the importance of ensuring research design processes are fit for purpose, and engage stakeholders in the ways that they want to engage and that they are adequately resourced to do so. One participant pointed out the need “to fund co-designers if they are under-resourced (e.g. Traditional Owners)”, stating “if you don't do this, claims to co-design can be fraught.”

Although co-design is often expressed as the primary goal, as one participant pointed out, it can not always be assumed “that co-design is the only way or the best way”. Others pointed out that different approaches to research design are “going to be valuable or relevant to different processes” and that “Sometimes people want independent research - rather than co-design”. Another stated, “Feasibility differs greatly depending on what kind of conservation research project is underway” and warned against taking a “'tick-box' approach to 'yes we co-designed'”; rather, the focus needs to be on ensuring the outcomes wanted.

Designing appropriate measures to test the quality of involvement of diverse groups of people who have a potential stake in the outcomes of research therefore needs to be done with some care. Numbers of individuals invited and supported to be involved in research planning and design may be stronger indicators than the numbers who actually were involved, given that people may choose their appropriate level of involvement. Attention should also be given to appropriate representation from different organisations and communities in the process - that the right people are at the table who are authorised to represent communities and stakeholder organisations; and that the diversity and geographic extent of stakeholder communities is appropriate to the research scope and scale (see also comments on RDRX and RDR4 Input measures).

Table 4.13. Considerations in applying quantitative and community insight measures of research processes

Indicator descriptor	Measures/questions addressing indicators	Assess-ment	Comments workshop and research team	Code
Research management and conduct				
All participants and stakeholders contribute to designing or providing input to research questions	Extent of involvement of stakeholder community/ies engaged in research planning, including: · Number of individuals from stakeholder communities invited to design or provide input into research questions · Number of individuals from stakeholder communities adequately supported to design or provide input into research questions · Number of individuals from stakeholder communities designing or providing input into research questions Appropriateness of involvement of stakeholder community/ies engaged in research planning, including: · Proportion and number of stakeholder organisations represented appropriately · Representation of stakeholder group involvement relative to geographic and sectoral scope of research	Use with caution	Likely to be feasible and appropriate in many cases. Raw numbers of individuals involved in research design likely not as useful as representation, and numbers appropriately invited and resourced, but will also give an indication of trust and appropriate processes through demonstrating active involvement. It is important that these measures on providing input to research questions do not assume a one-size fits all approach to research design, and that the nature of the input of research-users is the best and most fit-for purpose process for the circumstances. Co-design is often promoted as the primary aim, but not all groups have the best range of tools or availability at their disposal to co-design research. A different range of tools may be warranted. Not everyone has the interest to actively participate in research design activities, though they may appreciate being asked. Sometimes research users appreciate independent research, rather than co-design, to guide the questions that can best be asked at a given point in time.	RMC22
	Community survey measures asking project stakeholders and community participants to assess the quality and appropriateness of the approaches taken to inviting and involving them in designing and providing input to research questions, e.g. through interviews and focus groups.	Possibly use, prelimin-ary filter for relev-ance, value add	Community interviews and focus groups are unlikely to be feasible except in individual project assessments and small, in-depth case studies. Likely to be more useful at mid-term evaluation stage.	
Trust built or maintained within research networks and collaborations	Before-and-after survey of project collaborators, partners, and community participants where appropriate, e.g.: · how confident would you be to work with these (researchers/project partners) again? · how likely would you be to recommend these (researchers/project partners) to others to work with?	Possibly use, prelimin-ary filter for relev-ance, value add	Could be used as a final evaluation measure in cases where a baseline has been established before projects progress far. Likely to be more feasible in individual project assessments, or small bodies of projects. Would need to allow for cases where trust is high at baseline, as well as where significant increases can be observed. Qualitative information on conditioning circumstances (e.g. external factors contributing to trust or erosion of trust) could be invited to contextualise assessments.	RMC18
Retention, continuity and evolution of research teams and partnerships achieved	· Number of students, postdoctoral researchers and other early career researchers supported to participate · Number of early career practitioners outside of research institutions supported to participate · Number of Indigenous rangers and early career researchers supported to participate · Number of citizen scientists enabled and supported to participate	Good	Likely to be feasible to measure, speaks to knowledge transfer and potentially growing capacity for quality, engaged research. Data on student numbers should differentiate between different kinds of students (e.g. honours, Masters, PhD).	RMC23

Workshop participants also expressed an interest in measures of trust and how to set a baseline, e.g. through before-and-after surveys of project partners and wider community. This is likely only to be

feasible in individual project assessments and/or small, in-depth case studies; in other circumstances, likert-scale survey measures and narrative approaches are likely to be more feasible and effective to implement. Interviews identified the importance of legacy (Section 4.4), including retaining and evolving the capacity of research teams. Thus, the corresponding indicators (RMC23 and AO31 under outputs) were reintroduced in the development of quantitative measures (Table 4.13).

Quantitative outputs measures

Outputs measures are among the most readily feasible to assess and report on quantitatively. Traditional measures of peer-reviewed academic outputs (e.g. number, reach and citations of academic publications) are widely used, but were rated by Phase 1 respondents as low importance. Formal training and development of early career researchers and qualifications gained were also rated relatively low. Expert workshop participants noted that they are needed for other purposes, for example in responding to university systems and reaching research networks. They were seen as being of less benefit for assessing research value, but some workshop and interview respondents commented that these measures can in some senses reflect the success of a project in developing rigorous academic research and growing research capacity. Thus, these readily countable measures have been retained in the framework to be used as needed (see QuantOutputs1). However, generic summaries, web blogs and other general products were rated of moderate to low importance.

Table 4.14. Considerations in applying quantitative measures of production and quality of research outputs

Indicator descriptor	Measures/questions addressing indicators	Assess-ment	Comments workshop and research team	Code
Production of high quality research and skilled researchers				
Legitimate, valuable, and rigorous knowledge developed according to the various knowledge systems involved in the research Publications relating to the research authored or co-authored by researcher or team	Number of peer reviewed academic articles	Good: prelim-inary filter to decide where relevant	Rated as among the least important indicators of research impact in Phase 1. However, may provide some information as a small component of a suite of measures of research impact, as peer review gives an indication of acceptance by research peers.	AO30 AO33
Postgraduate students completed, training and certificates completed by research participants, capacity of young researchers increased	Number of students and postdoctoral researchers trained through project; number of students and postdoctoral researchers having undertaken formal intercultural training.	Good	Likely feasible and relevant to building sectoral capacity	AO31
	Numbers of people outside of traditional research institutions trained or gaining skills through direct involvement in project, e.g.: · number of Indigenous rangers trained or growing in experience and capacity, through involvement in research · number of community members, landholders, etc trained and/or grown in capacity to undertake effective actions or further research · number of citizen scientists trained and/or grown in capacity to undertake effective actions or further research	Good: prelim-inary filter to decide where relevant	Likely somewhat feasible for use where relevant	

Beyond just counting products, workshop participants emphasised the need to ensure that products were of high quality, fit for purpose and targeted to appropriate audiences. The production of legitimate, valuable and rigorous knowledge was rated as important and relevant by both Phase 1

and Phase 2 survey participants. Quantitative measures do not give a strong indication of quality and targeting, though traditional measures of publication and citation of peer-reviewed academic outputs do give some indication of the legitimacy and quality of research (see Table 4.14). As one workshop participant pointed out, it is important "to link policy documents to peer-reviewed papers to showcase that academic outputs are key to research, but also that [peer-reviewed] evidence is needed for policy."

Tailored, specific summaries, guidelines, protocols, tools, and similar targeted outputs, were rated as the most valuable types of product deliverables. To have the greatest impact, these need to be co-produced, adapted, and socialised effectively, for example through tailoring and training workshops (Table 4.15). Collaboratively produced outputs were also rated highly, and can be readily used as a relatively simple proxy measure of the connection of research to collaborating authors (see QuantOutputs2).

Measures such as numbers and types of tailored products, measures of collaboration around research products, targeted sharing with potential research users, and wider reach figures for these products (Table 4.16), each point to different aspects of this system of tailoring and socialising research, demonstrating potential research value. While these are conceptually distinguished below, they can be assessed as a package, and should generally be readily feasible to measure, working with research and/or support teams.

Similarly, download and access statistics for tailored products (along with outcomes measures discussed in the following sections) provide some indication of the reception products are receiving (see Table 4.16). Nevertheless, narrative and qualitative survey methods to assess perceived quality, legitimacy and relevance of research findings and outputs are important for consolidating this picture, as complementary to these quantitative outputs measures.

Wider dissemination of research findings, e.g. through popular articles, social media, community presentations, and public accessibility of research, were considered of moderate importance, and are also discussed (see Table 4.16). Workshop participants emphasised not just accessibility, but discoverability. Research dissemination figures can provide some indication of this, indicating broad access to the outputs. The specificity of whether tools, guidelines and other tailored products are reaching key audiences can also be assessed through measures of targeted workshops and training to tailor products for adoption and facilitate uptake (see Table 4.15).

Table 4.15. Considerations in applying quantitative measures of research outputs - tailoring and dissemination of research through stakeholder networks

Indicator descriptor	Measures/questions addressing indicators	Assessment	Comments workshop and research team	Indicator code*
Research tailored for research users and disseminated through stakeholder networks				
Guidelines, guides, checklists, standards for dispersed but like-minded audiences (e.g. parties who do environmental monitoring, all parties who participate in species translocations) Tools made available to land managers, practitioners and citizens/communities for on-ground action (management plans, planning frameworks and processes) Training packages developed for practitioners, citizens, community members and stakeholder groups Data sets (including spatial) made available to other researchers, practitioners and citizens/communities Tailored reports, stories or summaries prepared on research and findings for specific audiences (e.g. land managers, Indigenous communities) Summaries for policy makers prepared (policy options papers, submissions to policy forums Contribution to public policy advisory committee(s)	Production and public availability of outputs tailored for use by specific groups of stakeholders/research users, e.g.: · number of guidelines, standards, protocols, etc · number of tools produced · number of training packages prepared · number of datasets published · number of tailored reports, stories, summaries · number of policy summaries, submissions, etc Active dissemination and promotion of specific, tailored outputs: proportion of products above actively promoted to relevant audiences of partners outside of research institutions, stakeholders and research users for each of the above	Good	Availability of specific, tailored products rated in phase 1 as consistently more important than academic outputs. Number of products of each kind, along with whether or not these have been actively disseminated to relevant audiences, provides a ready measure for aggregation across programs.	RD52 TD43 TD44 AO32 RD51 RD53 RD54
	Communication of research findings to audiences, including: Research actively shared with partners outside of research institutions, stakeholders and research-users: · Number and reach of workshops, discussions, forums specifically to tailor and enable adoption of research findings · Number and reach of training forums to support adoption of tools, new techniques developed through research · Number and reach of stakeholder-focused presentations given to support adoption · Number of participants and number and diversity of groups participating in forums Research actively shared with senior decision-makers: · Number of briefings given to Minister · Number of briefings given to senior decision-makers, Minister's advisers · Number and kind of appearances at inquiries, expert panels, advisory committees	Good: preliminary filter to decide where relevant	Number and reach of targeted briefings, workshops, training forums and presentations specifically to support adoption gives a strong measure of active research dissemination and potential uptake. May be difficult to capture across wide array of projects - could focus on small numbers of projects, or centrally-organised forums across larger programs.	
Collaborative research publications or outputs	· Numbers of each type of product co-authored with research partners outside of university/research contexts (including data, academic articles, popular articles, and tailored products) · Where relevant: Number of products co-authored with Indigenous partners · Number of presentations co-presented with research partners outside of university/research contexts · Where relevant: Number of presentations co-presented with Indigenous partners	Good	Rated among the more important academic outputs in Phase 1. Readily measured.	AO39

Table 4.16. Considerations in applying quantitative measures of wide promotion of research outputs

Indicator descriptor	Measures/questions addressing indicators	Assess-ment	Comments workshop and research team	Code
Research promoted widely				
Research publications freely and openly accessible {Specific indicators for tailored products being made available}	· Numbers and proportion of research products (of different types, as above) made freely, openly and publicly accessible, as appropriate. · Altmetric figures for publications. Distribution and download statistics for: · research publications · tailored products: guidelines, tools, training packages · tailored summaries (policy and management-accessible information) · data sets produced	Good	Relatively feasible and relevant measures.	RD46 RD52 TD43 TD44 AO32 RD51 RD53
Presentations to schools, community groups and forums Media and social media coverage for scientific publications authored or co-authored by researcher or team in the applicable time period Media and social media coverage for tools and other outputs (beyond research publications) by researcher or team in the applicable time period Websites, web pages, blogs produced Popular articles including magazine and newsletter articles written by researchers and journalists Generic reports or stories prepared on research and results (e.g. fact sheets, videos, animations)	**Higher priority:** · Number and audience reach of presentations to popular and general audiences, e.g. schools, community groups and public forums who are not primary users of research **Lower priority:** Number, download and reach statistics of articles and activities **developed by study authors** to popularise and share findings, including: · social media posts, videos, webinars and related outputs · websites, web pages and blogs · popular articles and opinion pieces	Good: prelim-inary filter to decide where relevant	Could use where important for assessment. Presentations to schools, community groups and forums both rated moderately high importance in phase 1. Popular articles, media and social media, generic summaries and web pages generated by research team were relatively lower in importance, but should be relatively feasible.	RD46 RD55 RD45 RD47 RD48 RD49 RD50

Quantitative and documentary outcomes measures

Quantitative, documentary and community insight outcomes measures may be a useful extension of the outputs measures outlined above, complementing other means of measuring outcomes, particularly considering the variability in some of the survey measures of outcomes. These approaches can be used to assess positive changes being implemented as a result of the research.

The first series of measures, general citations and mentions of the research in academic, popular and grey-literature contexts, for non-specific purposes (see Table 4.17). These measures were rated low to moderate importance in the Phase 1 study. While they are not a priority, they are relatively feasible to assess (particularly the first two), and could be included where there is particular interest or in the absence of feasible alternatives.

Table 4.17. Considerations in applying quantitative measures of research awareness and citation

Indicator descriptor	Measures/questions addressing indicators	Assessment	Comments workshop and research team	Code
Research awareness and citation				
Academic citations for scientific publications authored or co-authored by researcher or team in the applicable time period	Number of academic citations of academic publications linked to research	Good: prelim-inary filter to decide where relevant	Rated low in Phase 1 as measure of research value. Use if relevant to assess research quality and acceptance in the research sector. (See outputs RD46)	RU56
Research findings perpetuated via popular articles, newsletters, fact sheets, reports written by third parties	Number, download and reach statistics of articles and activities **developed by third parties** to popularise and share findings in general, including: · popular articles · social media posts, videos, webinars and related outputs · websites, web pages and blogs	Good: prelim-inary filter to decide where relevant	Rated moderate in Phase 1 as a measure. Readily measured.	RU58
Public and grey-literature citations for scientific publications authored or co-authored by researcher or team in the applicable time period	Citations and reach of grey literature citing findings and/or publications linked to research for general purposes, including general citations in policy or management documents (excluding e.g. citations specifically reporting changes in management, practice, policy or planning, e.g. citations in management guidelines)	Not recommended as a priority	Rated low in Phase 1 as a measure of research value. Currently can involve considerable work to assess, as traditional citation aggregation tools exclude grey literature.	RU57

Measures of change to on-ground practice and to policy were generally rated as more important than these generic citations. These have been drawn out in Table 4.18 and 4.19, and potential measures to assess these suggested. Many of these have been grouped, and a higher level of assessment applied, in response to suggestions from the expert workshop. Nevertheless, specific outcomes (e.g. adoption of more effective techniques for conservation practice, better management of protected areas, establishment of adaptive management trials, etc) can be assessed separately where this is useful.

Most of these indicators assessed here measure adoption of research findings or tools into practice, and are expected to be relatively feasible to measure, at least for small numbers of projects and/or where stakeholder groups likely to implement changes are relatively known or not too numerous. Measures of potential environmental value resulting from these changes to practice are likely to be more resource-intensive to assess, but could be applied relatively readily to specific projects, and/or as part of more significant assessments.

Table 4.18: Considerations for quantitative and documentary measures of research outcomes for on-ground action

Indicator descriptor	Measures/questions addressing indicators	Assessment	Comments workshop and research team	Code
Changes to on-ground action				
· Adoption of more effective techniques for conservation practice (e.g. captive breeding, reintroduction in the wild) Findings from research incorporated into real world experimental tests or trials Research findings, tools and guidelines adopted in practice and implemented in monitoring, management, community practices, regulation, business practices Establishment of adaptive management trials, close coupled science-practice feedback loops New skills and training developed through research implemented Increase in protected area size and/or quality Reduction in threatening processes	Evidence of incorporation of research findings into monitoring plans/process, management plans, policy, trigger points, recovery plans · citation for specific implementation in management plans, grey literature	Good	Should be relatively feasible to source over moderate number of projects	AR62 OGA72
	Immediate adoption to support new practice Documentary evidence of adoption of findings by partners e.g.: · plans resulting from research implemented by relevant stakeholders · evidence that research influences priorities, practices or policy (e.g. adoption in management operating procedures, planning guidelines, policy, regulation) Number of instances of changes to management, policy, actions adopting research findings, e.g.: · number and diversity of new sites established from research findings (e.g. new translocation sites, new management areas) · number and diversity of species included in new plans and sites **Broader adoption to support new practice** Measures of adoption of findings among wider stakeholders: · number and reach of organisations adopting research tools, guidelines, findings, techniques · number and reach of people trained in and applying new techniques, approaches **Environmental value resulting from changes to practices** Measures of environmental values within scope of changed management or practices e.g.: · number of threatened or culturally significant species under area affected by positive changes to action · coverage of threatened ecological community, wetland system or place affected by changed actions arising from research · coverage of protected area/IPA land positively influenced by research	Good	Measures of adoption of findings among project partners relatively readily assessed Numbers of sites positively affected moderately readily assessed for specific projects or small case studies Measures of broader adoption relatively feasible to assess for a set range of stakeholders Measures of environmental values within scope of changes implemented would require more work to estimate and aggregate, but are potentially relatively feasible where more information is required, e.g. for significant reporting milestones	AR62 OGA71 OGA72 OGA73 OGA74 OGA75 OGA76

Changes to policy can be measured in a similar way. Expert workshop participants suggested that additional factors in policy, such as the political environment, could deliver perverse incentives for the research team to deliver findings that are compromised to accommodate to perceived policy directions. These measures of policy change are thus potentially most useful when coupled with other measures, such as stakeholder value, research legitimation and trust.

Measures of policy change in for-profit or non-government contexts were generally rated as less important. However, they could be considered useful for particular contexts, and could be used as relevant to the projects and reporting needs.

Table 4.19. Considerations in applying quantitative and documentary measures of research outcomes for decision-making

Indicator descriptor	Measures/questions addressing indicators	Assess-ment	Comments workshop and research team	Code
Change in government policy				
Research findings help to inform the decision-making process and result in change to public policy, strategy and/or program design Species or ecological community listed as threatened and/or protected by legislative instrument Landscape or threatened ecological community heritage listed Results trigger inscription of national park or conservation land tenure Increase in protected area size and/or quality	**Changes to policy** Changes to implementation of policy instruments, including: · Listing, uplisting and/or downlisting of species and/or ecosystems as a result of research · Adoption of findings, guidelines or recommendations in policy instruments (recovery plans, conservation advices, assessments) · Integration of new spatial data (e.g. species distributions) into government planning and regulatory databases that underpin decision-making. Changes to operation of policy instruments, including: · Adoption of tools, findings, guidelines or recommendations into new policy instruments (planning processes, regulatory decisions) · Integration of new methods for assessing spatial data (e.g. species distributions) into government planning and regulatory databases that underpin decision-making. Research findings and tools adopted in major policy decisions including: · Adoption of findings into legislation or significant policy strategies · Inscription of protected area · New or targeted program or funding **Environmental value resulting from changes to policy** Measures of environmental values within scope of changed policy e.g.: · number of threatened or culturally significant species under area affected by positive changes to action · coverage of threatened ecological community, wetland system or place affected by changed actions arising from research · coverage of protected area/IPA land positively influenced by research	Good	Measures of adoption of findings among project partners relatively readily assessed Measures of environmental values within scope of changes implemented would require more work to estimate and aggregate, but are potentially relatively feasible where more information is required, e.g. for significant reporting milestones	PPP78 AR59 AR60 AR61 OGA75
Changes to decision-making in other sectors				
Research processes and findings result in beneficial change to for-profit, business and private sector policy (e.g. private companies, financial sector consortium)	Uptake and/or reach of economic tools and measures produced contributing to greater transparency in economic decision-making e.g.: · number of citations in industry ESG strategies, environmental accounting assessments, etc · number, reach and economic impact of businesses adopting research tools and/or findings	Good: prelim-inary filter to decide where relevant	Generally moderately low importance in Phase 1. Potentially readily sourced. May apply to specific studies with economic tools and products.	PP81
· Research findings result in beneficial change to non-governmental and not-for-profit policy (e.g. International Union for the Conservation of Nature, Indigenous corporations)	Evidence and number of cases of adoption in non-government planning or policy documents and frameworks	Good: prelim-inary filter to decide where relevant	Generally moderately low importance in Phase 1. Potentially readily sourced where applicable. May apply to specific studies with intended outcomes for NGOs.	PP82

A final set of measures relates to the potential for increased skills and capacities arising from research. Additional indicators from those shortlisted in the research workshop have been reintroduced here, where proposed measures also speak to these indicators (see Table 4.20). Some of these measures could be applied differently in different contexts, and care should be taken if aggregating. Nevertheless, where these are assessed, they potentially provide a useful source of insights into some of the range of outcomes a research collaboration could deliver beyond singular interpretations of environmental gain.

Table 4.20. Considerations in applying quantitative and documentary measures of research outcomes for improving capabilities

Indicator descriptor	Measures/questions addressing indicators	Assess-ment	Comments workshop and research team	Code
Improved capabilities				
Links formed/improved between members of the research team, community organisations, Indigenous communities, conservation groups, land managers, for on ground action	Evidence of new networks, partnerships and capacities built to implement effective practice guided by research, e.g.: · number and coverage of Indigenous ranger groups engaged and/or benefiting from changes arising from research · number and coverage of community or landcare groups engaged and/or benefiting from changes arising from research · number and coverage of citizen scientist groups and networks established and/or made more effective through implementing research	Use with caution	May be somewhat feasible, but may in places be difficult to define boundaries of engagement or to tie directly to research project. Caution should be applied in aggregating any examples as a result of variable assessments of what constitutes groups implementing changes.	CSE65
Increased support or legitimacy for community/citizen practices that help to achieve conservation goals Involved, engaged or affected stakeholders inspired/supported to engage further in independent research (e.g. citizen science programs, Indigenous knowledge projects)	Expanded capacity of people to undertake effective actions or further research, e.g.: · evidence of Indigenous ranger groups or communities developing further work to extend project · evidence of Indigenous ranger groups or communities asking to develop further research with research team · evidence of community members, landholders, etc growing networks to share or extend on research · evidence of citizen scientists undertaking effective independent actions or further research · evidence of extension of citizen scientist-based methods developed to new research projects or contexts	Use with caution	Documentary measures may be somewhat feasible, but may in places be difficult to tie to research project. Caution should be applied in aggregating any examples since assessments of what constitutes extension of practice.	AR63 CSE70
Improved or increased intercultural capacities, via collective training of researchers on-ground personnel, in community organisations, Indigenous communities, conservation groups, land managers	Number of Indigenous community members newly engaging in research Number of research team members and number of early career researchers and students newly engaging in research with Indigenous partners	Use with caution	Likely feasible to measure. Caution should be applied in aggregating as a result of different interpretations of this measure.	CSE66

Quantitative, documentary, community insight and modelled or monitored impact measures

Expert workshop participants were most active in discussing the challenges involved in measuring direct environmental impact indicators, including what to do in the absence of baselines, and the importance of robust frameworks for measurement. The importance of baseline data and robust

frameworks for measurement was discussed. The low likelihood of seeing measurable environmental changes also featured prominently.

A range of potential quantitative and documentary environmental impacts measures, and approaches to assessing these, is shown in Table 4.21. Some of these measures may require fairly substantial data (such as baseline and detailed population data) that may not be available in many cases. The application of a number of these measures is also likely to be time consuming and require specialist skills, but may be feasible and desirable in particular circumstances. Workshop participants also cited examples of frameworks and approaches to measurement in the absence of baselines, and other examples have been drawn from recent studies, including examples of where related work has been done to develop more feasible and cost-effective approaches to assessing environmental impacts which can be applied to the context of research impact assessment. This is intended as illustrative rather than comprehensive, but gives an indication of what may be possible in different circumstances and to suit different needs.

Table 4.21. Considerations in applying quantitative measures of environmental impact

Indicator descriptor	Measures/questions addressing indicators	Assess-ment	Comments workshop and research team	Code
Environmental benefits				
Species recovery or avoided loss, including: Population increase and/or avoided loss in threatened species Avoided loss of biodiversity	Modelled and/or monitored measures of project contribution to population increase in a threatened species; avoided declines in threatened species population or distribution; or avoided biodiversity loss. These could include: · Measurable improvement in population and/or condition metrics against baselines or from long-term monitoring trends · Likelihood of species persistence as assessed through Population Viability Analysis · Improvements in species or ecosystem trajectories, avoided decline or likelihood of avoided decline as assessed through expert elicitation	Possibly use, prelim-inary filter for relev-ance, value add	Quantitative measures of population may be resource intensive, and only feasible if measures were built into project design or developed for related purposes (e.g. listing). More feasible as in-depth measures for a specific project or small suite of projects, but aggregate approaches could be used in assessment, e.g., (Bayraktarov et al., 2021) for large-scale or intensive, locally focused research programs (such as action-research in one conservation reserve), where monitoring data are available. Where relevant, direct measures via monitoring (e.g. using before-after-impact-control design), may be desirable, but is dependent on baseline and presence of controls. May not likely see improvements during assessment timeframe except locally, where specific, targeted interventions are particularly intensive and successful, e.g., (Stojanovic et al., 2019). Projections e.g. via population viability analysis may be of benefit where data and analysis capacity allow. Avoided loss is harder to quantify against baseline, but more likely to be achieved within timeframes, and can usefully be assessed via expert elicitation (see Geyle, Garnett, Legge, & Woinarski, 2019, Maron & Evans 2018).	REI87
Improved condition of places, including: Improvement in the maintenance and/or condition of an ecosystem, wetland, marine environment	Modelled and/or monitored measures of project contribution to an ecosystem, wetland or marine environment, including measures and evidence of: · Improvement or avoided decline in function or condition of an ecosystem, wetland or marine environment · Improvement or avoided decline in extent of an ecosystem or threatened ecological community	Possibly use, prelimin-ary filter for rele-vance, value add	Similar considerations to population increase or avoided loss. Need a strong condition framework, indicators and specific metrics for measurement, e.g., (Sharp & Gould, 2014), but unlikely to see measurable change in time frames of projects. May need to consider feasibility of alternative approaches (e.g. expert elicitation of avoided decline) where data do not provide strong indication or are not available.	REI8 REI91

Measures on improved capacity and more effective, equitable and informed deliberation processes are harder to quantify than general environmental, social, cultural and economic impact measures. However, in some cases where these matters formed a significant focus or outcome of projects, documentary measures could add a valuable dimension to narrative and survey-based measures. This could include, for example, evidence of the application of research approaches or findings to wider processes (such as local planning or agency decision-making); evidence of the application of good practice processes from the research to other decisions or decision-making contexts;or higher levels of participation of research participants in wider deliberative processes. For more in-depth assessments, where appropriate, judicious use of community-based approaches such as focus groups, wider surveys or participatory evaluations may be warranted (Table 4.22). Again, these could be built into research design where these outcomes form an important, anticipated benefit of the research.

Table 4.22. Considerations in applying quantitative, documentary and community insight measures for enhanced governance and decision-making capacity

Indicator descriptor	Measures/questions addressing indicators	Assess-ment	Comments workshop and research team	Code
Enhanced governance and capacities for locally-defined priorities and inclusive decision-making				
Improvement in the abilities of parties involved in or affected by conservation research to pursue locally defined priorities, objectives and goals (e.g. Indigenous communities, community conservation groups, natural resource management authorities)	Project contributions to improvement in local capacity to pursue locally-defined priorities, objectives, and goals? Evidence of: · application of research approaches or findings to other local priorities; · local plans developed that draw from research; · higher levels of participation in decision-making bodies of research participants.	Good: preliminary filter to decide where relevant	Measures may not apply across many projects, but where present, documentary evidence would provide a strong independent indication of contribution of research to enhancing capacities to pursue locally-defined priorities, objectives and goals.	REI92
	Project contributions to improvement in local capacity to pursue locally-defined priorities, objectives, and goals? · Use of community surveys, focus groups or participatory evaluations, where relevant.	Possibly use, preliminary filter for relevance, value add	Community approaches require specialist expertise to administer and place an additional burden on stakeholder communities. However, in projects where these are important anticipated outcomes, these measures could be built into project design.	
Improved capacities of management, governance and institutions to engage in effective, equitable and informed deliberation and decision-making around conservation issues	Project contributions to improvements in capacities of agencies to engage in effective, equitable and informed decision-making: · Evidence of the application of good practice processes from the research to other decisions, or other decision-making contexts · Evidence of changes to policies, programs or processes improving decision-making for managing species, ecosystems or places as a result of research.	Good: preliminary filter to decide where relevant	Measures may not apply across many projects, but where present, documentary evidence would provide a strong independent indication of contribution of research to enhancing capacities for deliberation and decision-making.	REI96
	Project contributions to improvements in capacities of agencies to engage in effective, equitable and informed decision-making: · Use of community surveys, focus groups or participatory evaluations, where relevant.	Possibly use, preliminary filter for relevance, value add	Community approaches require specialist expertise to administer and place an additional burden on stakeholder communities. However, in projects where these are important anticipated outcomes, these measures could be built into project design.	

Contributions of research to health, wellbeing, education, connection to Country and sense of place were identified in the first scoping workshop as potentially important indicators of research value, alongside biodiversity-positive economic and social activities. The potential to apply these measures was given a very low feasibility rating in the subsequent expert survey. The indicator connected with improved human wellbeing derived from species and/or ecosystems was given the lowest feasibility rating of all the indicators assessed in Phase 1 (REI95, feasibility 1.50; see chapter 3).

Consideration is nevertheless given here to possible quantitative, documentary and independent measures (see Table 4.23), particularly since results from the broad social and cultural impact question tested in the survey ('The project has or will contribute to positive social or cultural outcomes for people') proved unreliable (see detailed analysis of survey measures). Part of the perceived feasibility constraints of assessing these indicators possibly derives from a comparative lack of familiarity with social research approaches by many of the environmental science researchers and managers who participated in our study. Thus examples are given in the table of where some of these approaches have been taken, within the context of environmental research projects.

Table 4.23. Considerations in applying quantitative, documentary and community insight measures for social, cultural and economic impact

Indicator descriptor	Measures/questions addressing indicators	Assess-ment	Comments workshop and research team	Code
Social, cultural and economic benefits				
A better informed society with greater social licence for conservation Influence on public policy debate	Project contribution to greater public awareness of the importance of, options for or challenges of conservation, including: Influence on public policy debate: · Documentary evidence of citation in reports from public inquiries Substantial direct engagement with findings by third parties, e.g.: · Documentary evidence of citation in opinion pieces and other public contributions to public debate that engages with research, written by people other than research partners · Significant syndication of media opinion pieces · Significant degree of likes and reshares of social media opinion pieces · Reach and engagement figures where available	Good: prelim-inary filter to decide where relev-ant	Could involve documentary and/or quantitative measures. Measures unlikely to apply across many projects, but where evidence is present, would provide a strong independent indication of contribution to wider public discussion. Examples and numbers of citations in public policy discussions, such as inquiry reports, and direct independent media and social media citations of the research for the purposes of drawing on it to inform public discourse could be ascertained across whole programs. Syndication, or likes and re-shares of social media pieces, could provide proxy measures in the absence of direct reach and engagement figures. Reach and engagement figures likely to be harder to assess, as these are not author-driven, but could data could be requested if needed for significant assessments in particular cases.	REI94 PPP77

The impacts discussed here likely apply very differently to different project contexts. A judicious use of many of these indicators is warranted. Some of these measures may be strongly relevant to only a small subset of conservation-focused projects, and may only represent a good investment of assessor, research partner and/or community time for this select number of highly relevant projects. As with environmental measures above, some of these measures require specialist skills to feasibly assess. This particularly applies to many of the independent survey and quantitative measures of social benefit and ecosystem services. However, in cases where these measures are particularly relevant, projects themselves could consider drawing on these approaches as part of project design.

Table 4.23 (cont): Considerations in applying quantitative measures for social, cultural and economic impact

Indicator descriptor	Measures/questions addressing indicators	Assess-ment	Comments workshop and research team	Code
Social, cultural and economic benefits				
Improvement in human wellbeing derived from species and/or ecosystems, including access to resources, health, livelihoods, self-determination, social relations, and cultural and spiritual satisfaction: · Social and cultural measures of contribution to improvement in human wellbeing derived from species and/or ecosystems, including access to resources, health, livelihoods, self-determination, social relations, and cultural and spiritual satisfaction. · Economic contribution to improvement in human wellbeing derived from species and/or ecosystems, including access to resources and livelihoods.	**Documentary evidence of changes to social or cultural programs** · Inclusion of project in cultural, educational, health, job-creation or wellbeing contexts **Contribution to economic measures or decision-making** · Uptake and/or reach of economic tools and measures (e.g. evidence of use of research findings in independent environmental accounting assessments) **Documentary evidence of economic and social value for local communities, e.g.:** · jobs created · programs created · enhanced access to resources, e.g. regulatory changes, new industries · economic value of new or enhanced industries **External community satisfaction or wellbeing surveys, interviews or focus groups** Could explore as relevant: · program satisfaction measures as a proxy for direct social measures · health and psychological wellbeing benefits · educational benefits · wellbeing benefits derived from greater access to resources and livelihoods · improved social relations · sense of self-determination · sense of place · increased cultural and spiritual satisfaction **Measures or proxies of current or potential changes to ecosystem services, e.g.:** · water quality and water provisioning services · carbon sequestration · pollination services	Possibly use, preliminary filter for relevance, value add	Documentary and quantitative evidence for improvements to human wellbeing derived from species and/or ecosystems is considerably harder to measure, attribute to research, or determine with any certainty. Direct evidence of uptake of inclusion of research in cultural and social programs or processes, or of uptake and reach of tools and measures into economic programs or processes, are unlikely to apply to many projects. However, where evidence is present, would be relatively feasible to gather, and provide a strong independent indication of contribution of research to social, cultural and/or economic well being. Evidence of jobs or economic programs created, changes to regulation or new industries are likely to be rare outcomes of conservation projects, but likewise, where present, would be relatively feasible to gather, and provide a strong independent indication of contribution of research to social, cultural and/or economic well being. Surveys of wider community to establish independent insights into social and cultural benefits and improvements in human wellbeing derived from species and/or ecosystems should only be used selectively, as they require specialist expertise to administer and place an additional burden on stakeholder communities. However, in projects were these are important anticipated outcomes, these measures could be built into project design, (see for example Ward et al., 2021). Measures of current or potential changes to ecosystem services require specialist skills and intensive resources to assess, and may not be seen within the life of research projects. However, where such outcomes are anticipated or desired as part of research (such as practice-change projects to achieve economic benefits), a processes for assessing these benefits could be build into research design. Proxies, such as water quality estimates in catchments with productive landscapes, could be used as a substitute where full environmental economics measures are not feasible. May be difficult to see benefit within assessment timeframe. Could be projected/modelled (e.g. Keith et al 2017).	REI95

4.4 Gaps in the measures tested via interview

The interview guide included questions asking the interviewees to reflect on the types of questions we had asked and how well they captured the value of their research (Appendix E). Answers to that question as well as new themes emerging in response to other questions are summarised in Table 4.24. Some of these have been picked up in the assessment of quantitative measures above (Chapter 4.3). However, survey and interview measures are also potentially valuable to address these. The value of these are discussed in Chapter 6, and corresponding measures have been added to the final framework (pages 142-173).

Table 4.24: Gaps in indicators and measures emerging from the interviews

Value domain	Gap identified or emerging from interviews	Existing corresponding indicator(s)
Inputs/ Outputs/ Outcomes/ Impacts	**1.** While we ask about how research questions were identified, there may be value in asking about the **objectives of a project** and (a) whether or not they were clearly articulated and whether research users were included in objective-setting (inputs measures); (b) whether or not they've been met and why (outputs, outcomes and impacts measures); and/or (c) whether the reasons for limitations, failures or challenges in the research have been clearly articulated.	(a) RDR2/RDR13 (b) OGA/REI unspecified (c) AO41
Inputs/ Processes/ Outcomes	**2.** The **quality of relationships** and their importance to successful research was repeatedly raised in the interviews. The network mapping process was useful for eliciting what relationships exist and how they function. Additional interview and survey questions to address the quality of relationships in the project could add value.	RDR3 RMC17
Processes	**3.** Within-project **communication processes**. This is different to output indicators which focus on communication of findings to those outside the team. Several interviewees pointed out that relationships are based on doing things together, having a reason to get together and this requires communication within teams.	RMC24 RMC27
Processes / Outputs/ Outcomes	**4. Legacy** of projects and how this has been planned for. It could be (a) how and where outputs are stored and how accessible and discoverable they are, (b) who has been trained in new skills and in the use of tools and knowledge, (c) funding to complete work, and planning for discontinued funding which may result in incomplete data, loss of key staff, etc	(a) RD46 (b) OGA74 (c) OGA / PPP / CSE unspecified
Outcomes	**5. Change in research focus** or approach as a result of non-academic contributions. The current focus is on change in research users not whether research users change researchers (although perhaps implied through co-design).	CSE64
Outcomes	**6. Learning** as an outcome. This should address both learning through the research and skills development in research engagement and achieving research value. There is also a legacy component (Gap 4), particularly with respect to creating a legacy of skilled researchers.	OGA74

4.5 Findings related to building effective research collaborations

While the focus of this research project is on developing, refining and testing measures of research value, we have also gained insights about what it takes to build effective research collaborations. These insights were drawn primarily from the interviews in Phase 2 of the project, and comments from the expert workshops.

Co-design - when and how should it be evaluated?

Academics and non-academics can collaborate in research at a range of levels including co-design. The concept of co-design and what it means for assessing research value was discussed in some depth at the expert workshop. Expert participants warned against a 'tick box' approach to undertaking co-design of projects or to encouraging such an approach in how co-design is assessed. Instead, they suggested the need to focus on ensuring that research collaborations adopt the most appropriate approaches to achieve the desired outcomes. In particular workshop participants pointed out:

- It is not appropriate to assume that co-design is appropriate for all research projects, or in other words "that co-design is the only way or the best way."
- Different co-design approaches are going to be valuable or relevant to different stages of the research process.
- The feasibility of undertaking co-design processes and of how to go about this differs greatly depending on what kind of conservation research project is underway.
- Research users such as land managers do not always know what it is they want in a research project, and do not always want or have time to participate in co-design processes. "Sometimes people want independent research - rather than co-design."
- The question should not be 'was co-design applied?', but rather "Are people using the best and most fit for purpose process?"
- Research users and partners "don't always have a good range of tools at their disposal to co-design research." Some work in this area is being developed testing a different range of tools to see where and how different dimensions of co-design or collaborative approaches can be applied appropriately in different contexts.
- It is important to fund co-designers if they are under-resourced (e.g. Traditional Owners). "If you don't do this, claims to co-design can be fraught."

One of the purported benefits of co-design and transdisciplinary research more broadly is that research is better targeted to user needs. This was apparent in at least one interview "*because we sort of crafted the questions really early on, and together the outputs were directly useful and relevant*".

Links between collaboration, relationships and trust and how these enable effective research practice

> *"it's all about building relationships and spending time in respectful, productive interactions"* [NGO interviewee]

Collaboration was an integral part of conservation research. Only one survey respondent said non-academics did not participate in research and all the interviewees mentioned at least three collaborating agencies. Collaboration ranged from a small cohort of researchers at a single institution working in consultation with one government agency, to complex networks of researchers, volunteers, government and non-government agencies and communities working together *"It's huge. We have several thousand volunteers, like three and a half thousand volunteers".* Given that collaboration is an integral part of conservation research, the question becomes how is effective collaboration enabled?

Collaboration in a time-constrained world was strongly connected to relationships. Interviewees observed that where there was trust and a relationship between individuals, interactions between those people happened more quickly and easily as they tended to be prioritised. So a research collaboration involving people who have a good relationship with each other may be more effective than where there is no relationship or a poor one. As one researcher noted of their relationship built with government staff: *"I'm definitely finding it easier to get a hold of people that I need to get a hold of now*". We specifically asked about trust in the interviews, and while it was often thought to be a key ingredient of good relationships and collaboration, there was recognition by some that it is a complex concept in research partnerships. For example, trust can be developed at different levels: trust in the quality of work done by technical experts; trust that you are acting in good faith (e.g. using data appropriately and usefully); trust in a method; as well as trust among individuals. Trust may also vary within an organisation depending on roles and responsibilities e.g. those on the operational level in government may trust the researchers, those higher up who are more risk averse may not. Lack of trust was equated in several instances to poor research uptake. On the other hand, at the expert workshop one person asked "What if policy-makers trust too much?"

Qualities of individuals was another component to trust developing during projects, which emerged from the interviews. Key individual attributes related to respect for others demonstrated by listening. One interviewee said of researchers that "*their inherent ability or willingness to turn up and listen"* was a key to success. Interviewees who reflected on effective relationships generally mentioned they had put deliberate effort into this: *"Part of my job was to encourage that relationship* [between two other groups]. *And that relationship, I think, is almost more important, because that will last, theoretically much longer. Even when* [Uni X] is *removed from the picture*." And as a government partner stated, "*you want to get something out of a relationship, you have to put the hard yards in*". The consequences of dysfunctional relationships were also noted: "*I now have a very good relationship with that team, the person who's leading it now. There's no problems. But, you know, there were difficulties at first... And essentially, the project wasn't being delivered, I mean, it's way overdue*".

Another ingredient in building effective research collaborations was having good communication. One researcher talked about giving informal updates to non-academic partners at their workplace. "*We have discussions, that's when we're fostering new ideas, they're able to ask really specific questions,... So that's when I felt most valuable to them*". Another NGO partner talked about annual results summaries for a long-term monitoring program. If managed well, research projects can provide an opportunity for academics and non-academics to have a shared purpose and reason to interact. "*It was actually quite a good opportunity to have a focal point to build relationships*".

As well as communicating about progress, results, etc. a couple of interviewees also talked about communicating findings that partners may not want to hear about "*So we had to manage that because we were pointing to a bad news story*". This is likely a common theme in conservation research: actions aren't having the hoped for results or policy instruments aren't delivering hoped for outcomes. Being able to have private one on one communication discussing these issues before they go out in the public domain can be critical for maintaining relationships.

Apart from issues between individuals, a key barrier to building effective research collaboration and relationships mentioned in the interviews was staff turnover. This could be researchers only employed for the duration of a research project or movement of staff within government agencies: "*The amount of change, the people who are cycling through...has meant that it doesn't have the buy in*". On a positive note, strong relationships may endure and maintain value where people stay working in the same field: *"The relationship is so strong between all these partners now that I feel, I would like to maintain those relationships going forward, even though I won't be employed at [X]"*

The benefits of robust relationships built on trust and respect within research teams go beyond being the foundation of research collaboration. The stakeholder mapping and interview question on trust elicited the following benefits:

- Researchers can become a trusted source for advice: "*They have to sign off on the threatened species side of things. And I provided them information so that they can make decisions on whether or not these things go ahead. That's a pretty big thing*." This translates into research adoption through expert advice.
- Several interviewees noted trusted relationships facilitate data sharing and more productive use of data: "*the project itself [provided] the opportunity to have that data analysed productively*". Most projects relied on existing information, and some were completely based on this. As well as access to data, a relationship with custodians facilitates understanding of what data represents and appropriate use of that data. The interviews also highlighted that data sharing and better data utilisation can be an extremely important component of transdisciplinary partnerships given non-academic partners are often data custodians. For example in relation to one state agency: "*the fact that the species have been monitored for many decades meant that there was a lot of existing data*".
- Relationships give you the ability to have open conversations about when something isn't working and what can be done about it. As a government partner noted: *"we're able to express concerns early and express them bluntly."*
- Good relationships were also noted as a key enabler for research teams to problem-solve effectively.

Chapter 5. Worked evaluation example: using the framework and questions for assessing the value of research

This project set out to develop and test approaches to assessing the value of environmental research using projects from the NESP TSR Hub as case studies. It did not set out to undertake an assessment of evaluation of these projects. Nevertheless, analysing the data obtained through these case studies can demonstrate some of the ways the framework and approaches developed here could be used for assessment and evaluation purposes. Here we look at the responses different project team members gave to the surveys and interviews, and we analyse these in different ways, to demonstrate how assessment of the data could be approached for diverse valuation purposes.

While the measures and framework can be applied to assessing the contribution, processes and value of individual projects, here we focus on their value for collective assessments across a suite of projects or a whole program. These program-wide and aggregate analyses could be used by research institutions, research programs and research funding bodies to look across a research portfolio to develop an overall picture of research value.

5.1 Understanding the contextualising information

Knowledge of who participated in research projects provides contextual material for assessments of the other measures, but it also provides an indication of the breadth of respondents and sectors involved enough in projects to warrant responding to case study surveys and interviews on research value.

Insights from the survey data

The Phase 2 case study survey had 11 respondents from research institutions and 7 from other institutions (government, non-government and Indigenous organisations). Respondents were able to select as many project roles as applied to their situation. They spanned all possible project roles, with the majority of research users sitting outside of research institutions and all of those involved in analysis being in research institutions. Those in research institutions were more likely to say they were responsible for conducting data analysis, communicating about the project and producing outputs (Figure 5.1). Those outside research institutions were more likely to provide advice to projects and be research users.

All but one of the respondents agreed that there were groups or organisations outside research institutions directly involved in the project. Participants were mostly likely to select that government agencies were involved (15), followed by non-government organisations (8), Indigenous organisations (3) and community organisations (2) (Figure 5.2).

Figure 5.1. Breakdown of roles academics and non-academics carry out in research projects

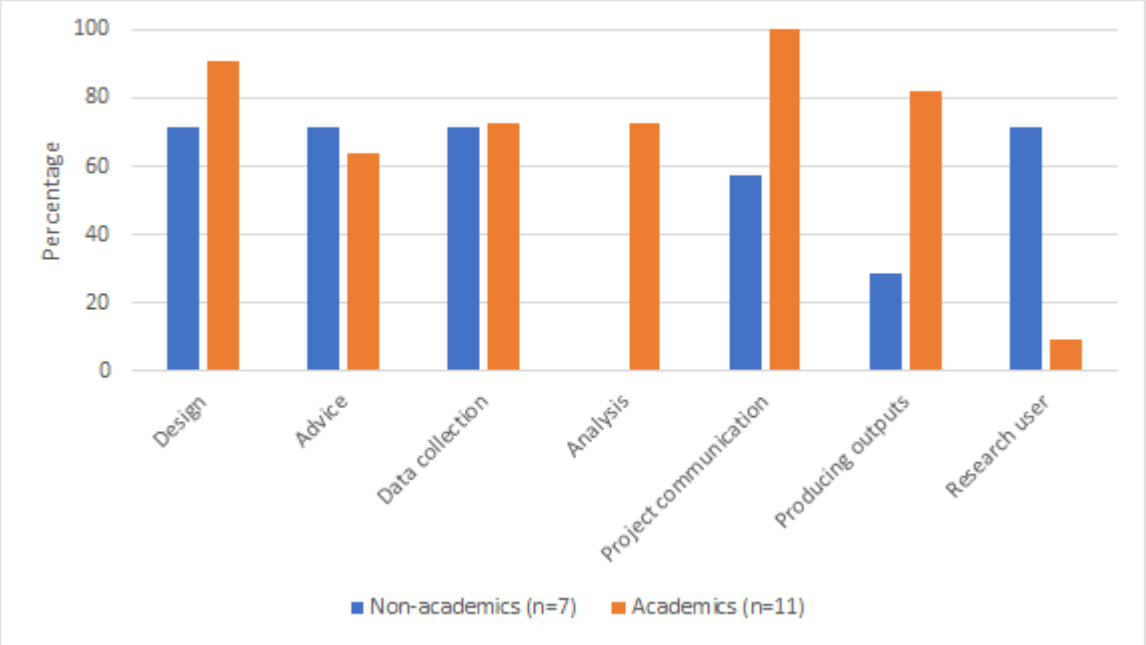
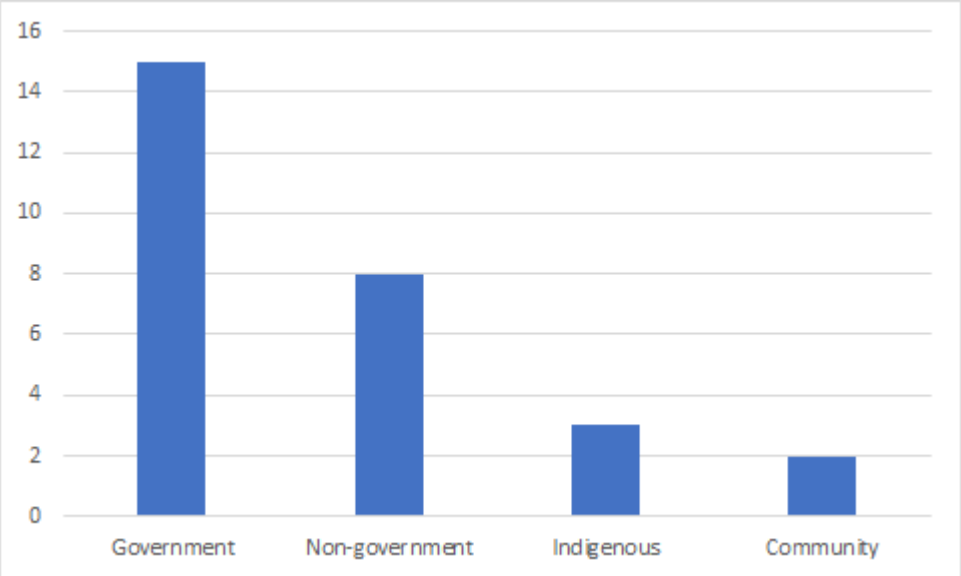


Figure 5.2. Number of respondents stating that organisations from particular sectors outside of research institutions were directly involved in the project



Insights from the interviews

In the interviews, Question 3 specifically prompted a mapping exercise of relationships that operated through the project. At a project level this approach elicited useful information such as the size and extent of relationships built; relationship connections between people (groups or individuals); practical examples of functional relationships and why they are important; whether relationships were brokered through individuals or entities; and how relationships developed. In addition, relationship mapping can also help contextualise the assessment of research value, through identifying issues such as individuals leaving and the nature of pre-existing positive or negative relationships/interactions. At a program level this could be used to situate the responses in the context of how many different types of organisations were identified as being connected with the program and the extent of pre-existing partnerships and relationships coming into the program.

Interviews also elicited important contextual information; for example, "*COVID really did knock it about so there was no field work possible after March last year*." It may be worth both seeking information on barriers to research and analysing narrative data for themes and frequency of these challenges being mentioned, so that contextualising information relevant to the capacity to achieve outcomes can be accounted for in assessing value (either on a project-by-project basis, or across a program where common impacts are widely observed).

5.2 Aggregate analyses: examples from inputs, processes and outcomes measures

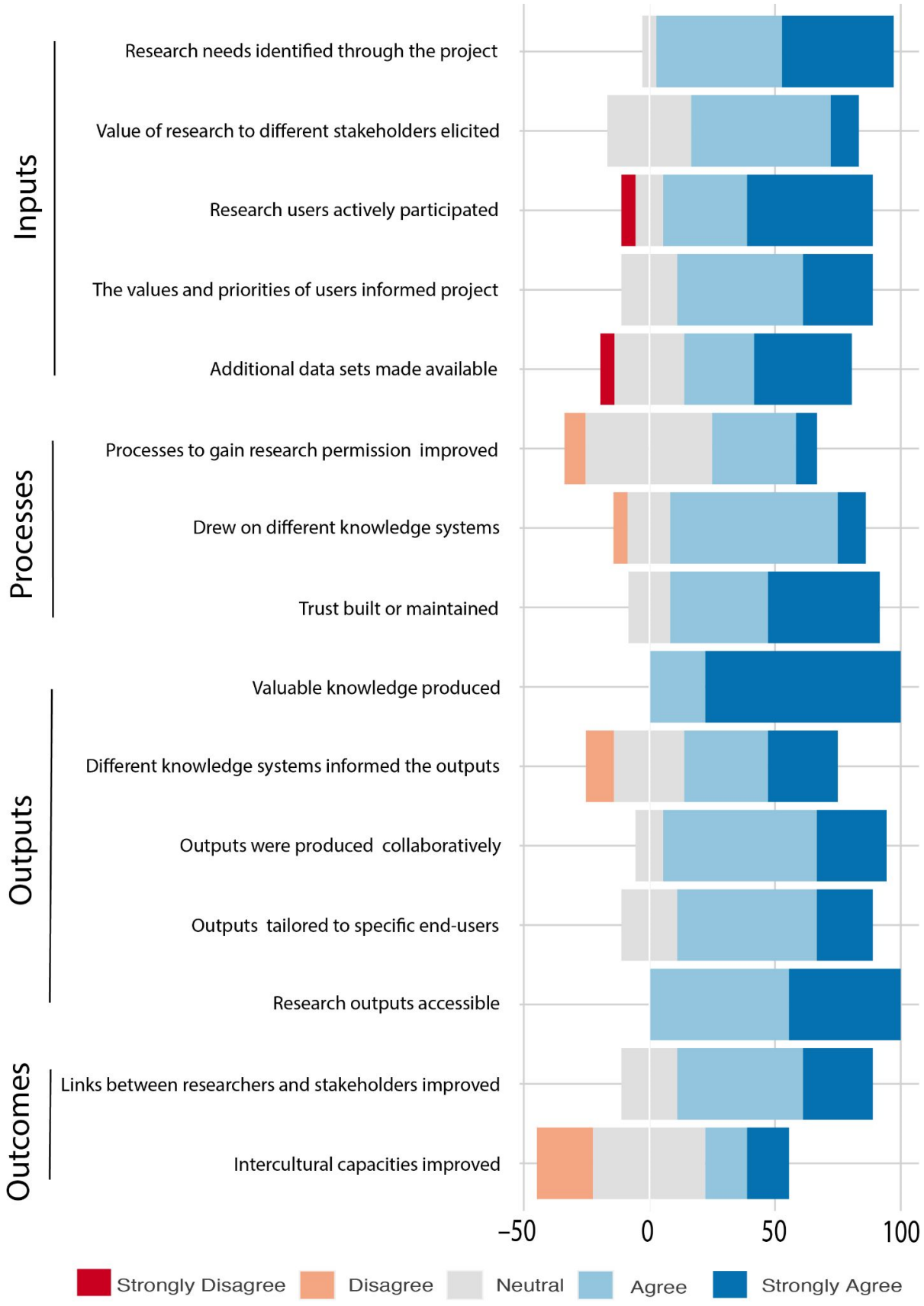
One of the simplest ways to provide an overall assessment of research value across a program or suite of projects is to provide aggregate analyses of survey, interview and quantitative measures across all respondents and all projects in the assessment (bearing in mind the pitfalls in aggregating some of the measures, identified in Chapter 4). This approach is applied here across measures in some of the elements of the logic model by way of illustration.

Insights from the survey data

Aggregate analysis of likert scale survey questions provides a simple, comprehensive picture of how a program or group of projects performed against the measures overall. For the 5-point likert scale measures across Inputs, Processes, Outputs and Outcomes, this suite of projects displayed generally very high levels of agreement across most of the measures tested (Figure 5.3). Some of the highest rated measures were "Research needs or new areas of research were identified through the project" (Inputs RDR1, average 4.39 in possible range of 1-5), "Trust was developed or improved among research partners" (Processes RMC18, 4.28) and "Links between researchers and stakeholders improved" (Outcomes CSE65, 4.06).

The statements "The value of the research to different stakeholders (particularly end-user agencies) was elicited" (Inputs RDRX, 3.88) and "Intercultural capacities improved" (Outcomes CSE66, 3.31) had the lowest level of agreement, but these measures nevertheless rated well above the midpoint on average (3 on a scale of 1-5). This suggests at the broadest level that the suite of projects could be assessed very positively from the survey measures.

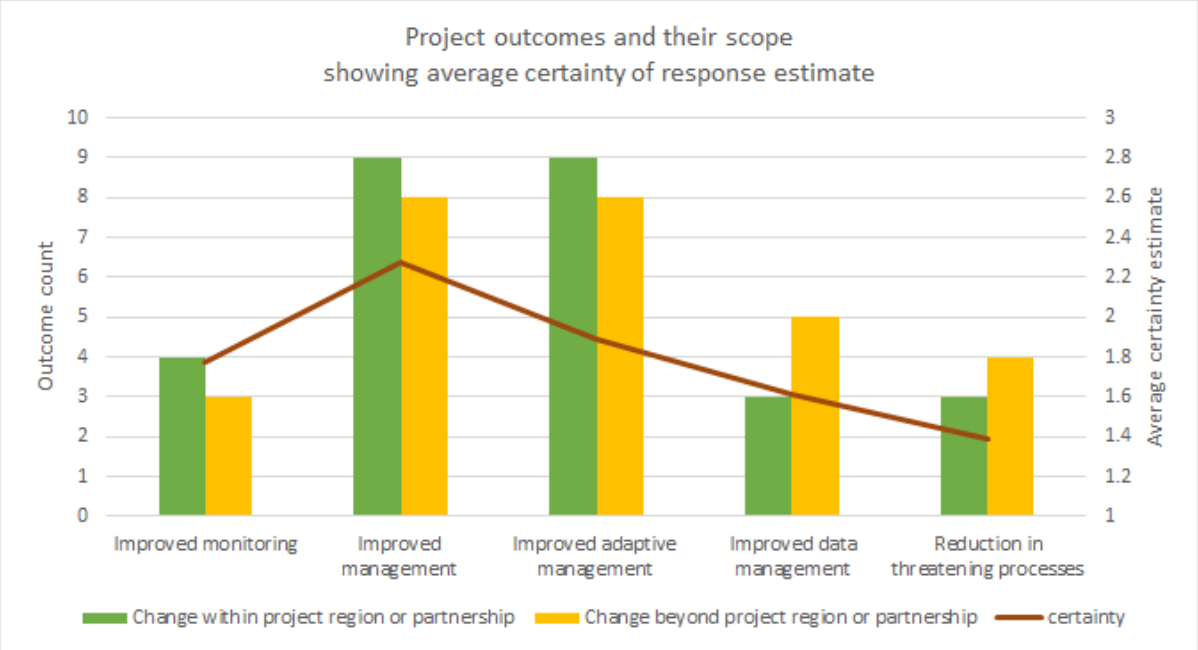
Figure 5.3. Likert 5 point scale data for the five research value domains transformed to percentages



The five outcomes measures related to on-ground practice were tested by asking about the scale where outcomes could be seen, whether within or beyond the boundaries of a project, and how certain respondents were of these outcomes having been or being likely to be achieved. The certainty estimates attached to each outcome were generally low, limited to medium on average, with the exception of the "Improved management" outcome (Figure 5.4). A robust level of confidence or certainty was defined as "an outcome has already occurred or a process is in train for it to occur (e.g. legislation being reviewed)". The fact that some projects were still in their final stages at the time of the survey meant there was limited time for outcomes to manifest, which is the likely reason for the lack of certainty.

Nevertheless, all research outcomes were nominated as occurring within and beyond the project region and partnership scale for at least some projects (Figure 5.4). Improvements in data management and reduction in threatening processes outcomes were more common at a broad than local scale whereas the opposite was true for the monitoring and ecosystem management measures (Figure 5.4). Only 17% of responses were that a specific outcome occurred at both scales. However, within and beyond counts were quite similar for a given outcome.

Figure 5.4. The scale of outcomes expected overlain with average level of certainty estimates that outcomes will occur (based on a three point scale: limited = 1; medium = 2; robust = 3)



Insights from the interviews

Analysis of interviews can draw out key themes for aggregate analysis across a program or suite of projects. This can be presented as summary highlights; for example, some of the key summary observations on outcomes across our suite of case study projects include:

- More examples were given of changes to practice than policy. This may reflect the suite of cases, but it may also be a reflection of a broader pattern that the impact pathway to changing practice is shorter and more direct e.g. if researchers are working with on-ground managers.

Table 5.1. Illustrative examples of how inputs and processes measures from the interviews show different aspects of research value

Category and measure/ question	Co-design and brokered partnerships	Researcher-led value	Challenges / limitations to collaboration
Inputs: Participation **Q4.** How did non-academic/your and other non-academic groups participate in the project?	**Brokering model of research:** "I was always responsible for working directly with the academic group to facilitate interest and design across the department" **Mutual understanding of drivers**: "the umbrella driver was we didn't really have a good understanding of the conservation status of the species in [X]" "then we sort of settled you know that's what we're trying to do"	**Consultative model of research:** "[University X] is the lead research organisation. So we develop the research questions and the monitoring programme, and write up the research, in consultation with all the partners"	**Limited/no collaboration:** "So the only people that participated in this project are the researchers"
Inputs: Question development **Q5**. How were research needs identified in the project?	**Direct response to need:** "[speciesX] had a translocation plan for [conservation reserve Y], and that gave the reason for the needs for the translocation" **Co-development:** "I didn't feel ever that it was a handball... it was, here's this idea, let's do this thing together" "they were developed, firstly, with [non-academic partner X] and then involved other partners as appropriate"		
Inputs: Data **Q8**. Were there existing data sets/information necessary to do the research?	**Close working relationships supporting data integration:** "the only reason we knew of all the work that had come before, is because [a member of the team] had been working in that country for 20 years" "context is really important in terms of how is it collected? What was the intent of its collection? So it's really important to have the relationship with the people who collected or know more about the data as well."	**Successful data integration:** "one of the most important and difficult tasks for us has been to actually get accurate [X] data."	**Accessibility and issues with available data, implications of no access:** "that wasn't as sort of straightforward as being handed this beautiful database" "we have no way of comparing our results to their results"
Processes: Methods **Q9**. What is your view on the research methods used in the project? Did it draw on different disciplines & knowledge systems?	**Two-way learning and integrating different knowledge systems:** "a really important part of the method for this was the delivery of the content..how are we going to deliver Indigenous knowledges and Western knowledges without Indigenous knowledge seeming tokenistic or stereotyped" **Improvements in methods successfully promulgated:** "our contribution towards this international negotiation of the standards was very important"	**Contributions and insights from non-academic partners:** "It was, I think, science led. Logistics and data. But then we got the input" "We [non academics] contributed the context specific natural history of the species and of the landscape the species was in". **Adaptability in methods:** "we tried this [X method] because we didn't have money at that stage to purchase radio transmitters"	

- Additional outcomes were identified along with those prompted: e.g. the researcher becoming a trusted expert, subsequent applications for funding to do work to protect the species, better understanding etc
- Respondents, when asked to reflect on skills developed through the project, included reflections on actual skills as well as learning and awareness (which may be thought of as capacity) e.g. you may go to a workshop and learn about a particular method which you then understand but may not be able to employ yourself (although you know when to ask someone else to do that for you) "*…knowledge is that sort of output. But that translates into things like monitoring methods or, targeted monitoring techniques. So that's a skill that comes out of that research.*"

Narrative data can also be presented across different aspects of project value, as in the example in Table 5.1 of themes from interview questions on inputs and processes. This could be undertaken illustratively, drawing out examples of differing responses to particular questions, as in the Table provided here, or as a more systematic thematic analysis that draws out key themes across the narrative dataset (possibly with a quantification of how often these themes occur). For example, a thematic analysis could be used to show the diverse ways in which positive value was delivered or demonstrated across a program, challenges or constraints in achieving research value, and examples of where and how programs have delivered less completely or fail to deliver against particular measures of research value.

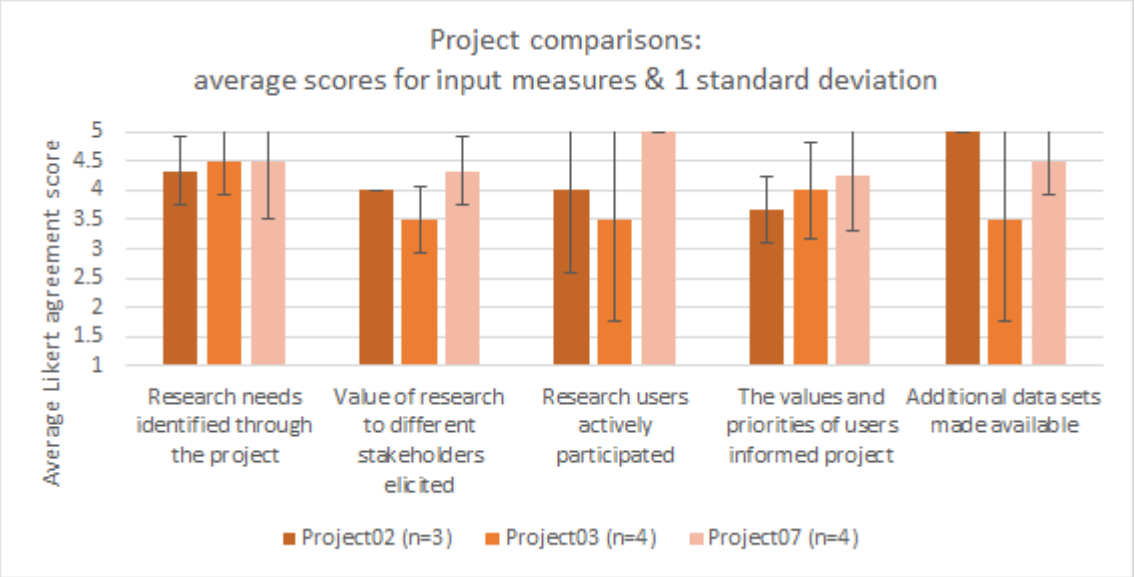
5.3 Between-project comparisons: examples from inputs and processes measures

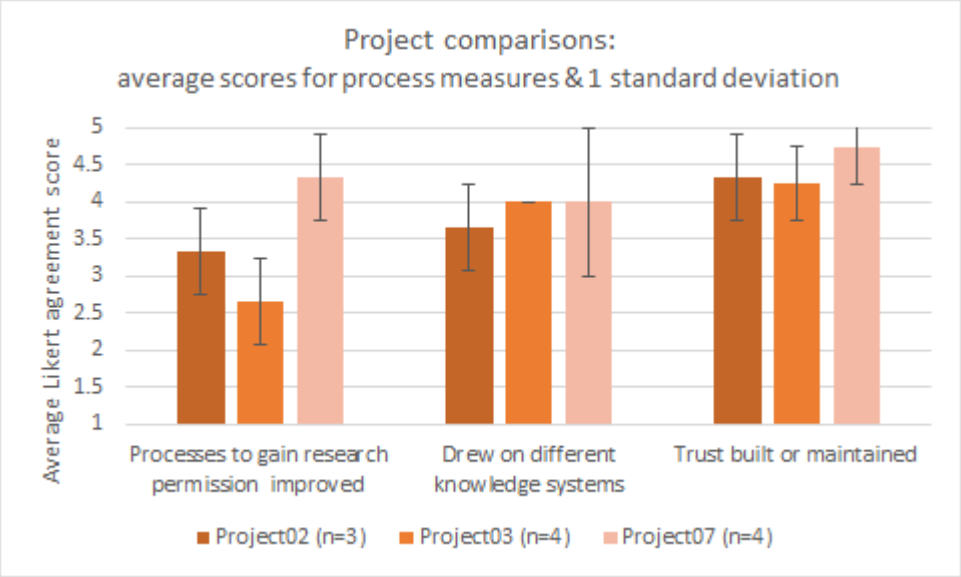
Survey, quantitative and narrative data can also be used for comparing between projects. This relies on an instrument that actively specifies which project the respondent is being asked to respond to, as well as the coordination of sampling strategies to elicit responses from multiple perspectives within individual projects. While project-level assessments necessarily have relatively few respondents who can adequately speak for projects, and therefore should always be treated with some caution, findings from these project-level assessments can support wider insights.

Insights from the surveys

The data from inputs measures in the survey highlight that there are some cases where there was a high degree of positive within-project agreement with some of the statements, e.g. that "research users actively participated" (Project 07) or that "additional data sets were made available" (Project 02); and cases where there was greater within-project variability and an overall less positive assessment against these measures (Project 03). The data from processes measures likewise highlights that there are specific measures where a high degree of variability exists between projects (e.g. the measure on 'processes to gain research permission improved'), and specific projects (Project 07) for which this measure rates particularly highly, along with the measure for whether trust has been built or maintained.

Figure 5.5. Average project ratings of inputs and processes measures, disaggregated by projects, for the three projects with n>=3 (based on a 5point scale: strongly disagree = 1; disagree = 2; neither agree nor disagree = 3; agree =4; strongly agree=5)





These specific project-level assessments demonstrate where research programs tend to deliver value across the board, and which measures elicit more variable responses. More importantly, larger evaluations could aggregate these data up into figures that can provide a fuller overview (e.g. proportion of projects sampled for which an indicator was rated above 3 (in this case 100% of projects sampled across all indicators).

Quantitative insights

Quantitative measures can help add concretely to the picture of within-project assessments by providing numbers for illustration and comparison. For example, overall responses to the survey question "Additional data sets from outside of the research team were made available to support the research" garnered fairly high levels of agreement on average across the respondents, especially among non-academic respondents. The cross-project comparison of this measure, however, showed

variability between the three comparison projects, with all respondents in project 02 rating this very highly, while the average response from project 03 was significantly lower.

Our recommendations note the need to take care in averaging likert responses across project teams for this measure. Adding documentary or quantitative evidence of examples, scope and geographic reach of the datasets made available to each of the projects would potentially be one way of ground-truthing these responses and shedding light on the differences between projects in assessment of this measure. For example, assessments for projects could be accompanied by documentary or quantitative measurement of:

- How many datasets were made available to researchers and research user networks for this project
- How many datasets will be available to the research team for future work (if different from above)
- How many formerly inaccessible datasets are now in public domain or otherwise accessible as a result of this research project (with appropriate protections)
- The scope and coverage of these datasets

This would help specify the value and accessibility of the projects in both bringing together datasets and making these available for use.

While care should be taken in aggregating or comparing figures across a program, options for aggregation or comparison could:

- Count datasets by extent of their availability, e.g. whether they are available to:
    - research team for future use
    - particular users (research, government) via a protected database
    - publicly available
    - published in a data repository with metadata and unique identifier

  Once this information has been provided for projects, aggregation would be a relatively simple exercise, but would provide only general information on value since there would be no accounting for type, scope or coverage of the datasets made available.
- Disaggregate into type, e.g.:
    - species monitoring and population datasets
    - ecosystem function and condition datasets
    - species distribution models
    - remote sensing data
    - spatial data - e.g. geospatial, geopolitical, environmental layers
    - ecological data
    - genetic data
    - image or video data
    - models and simulations
    - Radio tracking, cage trapping, weight monitoring, etc

  and count datasets grouped by their availability and type. This is a slightly more complicated exercise but would provide considerably more information on the value provided across a suite of projects or a program.
- Disaggregate by availability and type, and match scope and reach to each, e.g.

- Numbers of species covered by datasets
- Geographic scope covered by datasets

This would be a considerably more complicated exercise, but could be undertaken with appropriate resources.

Insights from the interviews

Project-level assessments of survey data can also guide where cross-referencing with narrative information might provide the most insights. For example, interview data from Project 07 can be used to provide more insights into the measures where this project rated particularly highly: that research users actively participated, processes to gain permission were improved, and trust was built or maintained. In a larger review, ideally multiple interview responses would be integrated to build this picture. In our smaller case study, we had one interview response from the researcher most actively involved in the day to day work of the project. The descriptions this researcher gave of both the effort put into building relationships and what that meant for fostering trust, facilitating and streamlining research processes, and delivering better results provide a strong illustration of why this project might have been rated so strongly against these measures:

> *[I]t's been a big emphasis that I've put on this project. ... I really wanted to develop a really collaborative and applied project…*

This involved substantial investment of time and effort, and the researcher describes often needing to actively initiate that in order to establish trust and value to the research partners. A wide range of strategies were involved in building that trust, centred around regular communication, updates, learning from their insights, field site visits, informal presentations and many other strategies. This was not always easy, nor readily recognised by the research institution's performance metrics:

> *That first year was really tough in terms of, I just spent a lot of time trying to build my relationships… unfortunately, that wasn't a metric that ... in ...my [academic] reporting after a year, like, what have you done? I was like, well, actually, I've just spent a lot of time talking with people and ..., you know, and spending time with people and trying to ... really understand their perspective and how they operate, and that ... sort of thing. .. But I definitely felt like it took me quite a while to really develop a strong relationship …*
>
> *Now, it's like this, it's a back and forth thing. ... I've got people asking me, did you have to get data at this site? What do you think of this? … but I think that took quite a bit of time to develop.*

The researcher's responses affirm that this investment of time, effort and trust on their part has been a significant factor in facilitating agreements and the processes around permission, and was ultimately a marker of the project's success:

> *So I think ... if you can show a little bit of trust on your end, and put some trust in them, then yeah, maybe that sort of fosters that relationship. But in terms of how I feel ... it's ... changing the outputs, or … how you measure the success or whatever of a project, I think it's been a huge part.*

In this example, the narrative insights gained in interview support and reinforce the high rating from across the research team members on these measures of research user participation and the trust built through the project. In other contexts, narrative material may provide a counterpoint to responses provided in the survey. But, regardless, altogether these responses can work to form a more complete picture of the value of projects across a suite of measures.

5.4 Analysis by sector: examples from outputs measures

In addition to aggregate analysis across the measures and analysis by project, analysis by sector (or alternatively, by project roles of respondents) can shed specific light on the success of research projects in achieving value, and can aid with the interpretation of particular findings. Examples are given here of how sector-by-sector analysis can help shed light on interpretation of outputs measures, in two ways. First, using an analysis directly comparing the differences in responses to some questions (in this case differences in awareness of outputs reported in the survey), we can draw direct provisional conclusions about the success of the project (e.g. in disseminating outputs). Second, by looking at the different themes emphasised by respondents from different sectors, and the likert-scale and narrative assessments provided by members of each group, we can gain a more nuanced synthesised picture of what is valued in the research program by each of these two groups, and some of the limitations or challenges each group perceives.

Quantitative outputs measures

Quantitative measures of different types of outputs give an indication of the productivity of research programs and the extent of tailoring and co-authorship within programs, as per the examples in Table 5.2.

Table 5.2. Aggregate figures for different types of outputs across the projects in this study

Product type	Number of outputs across all projects		Number co-authored	Awareness and reach data (e.g. altmetric / citations)	
	Total	# accessible		Altmetric (mean)	Citations
Academic articles	15	12	15	25	41
Guidelines, standards, protocols	1	1	1	n/a	n/a
Tools	1	0	0	n/a	n/a
Training packages	1	0	0	n/a	n/a
Tailored reports, stories, summaries for policy or management	14	13	8	n/a	n/a
Data sets	9	8	0	n/a	1
Videos	7	6	3	n/a	n/a

Where extra information is sought on outputs, evaluations could report on number, download and reach statistics on these products and on articles and activities developed by study authors to popularise and share findings, including:

- social media posts, videos, webinars and related outputs
- websites, web pages and blogs
- popular articles and opinion pieces

Further information on dissemination and training could be gathered from members of the project team, for example numbers of:

- tailored products shared with research users
- briefings to senior decision-makers, Minister's advisers, Minister
- appearances at inquiries, expert panels, advisory committees
- presentations co-presented with research partners outside of university/research contexts, and with Indigenous partners.

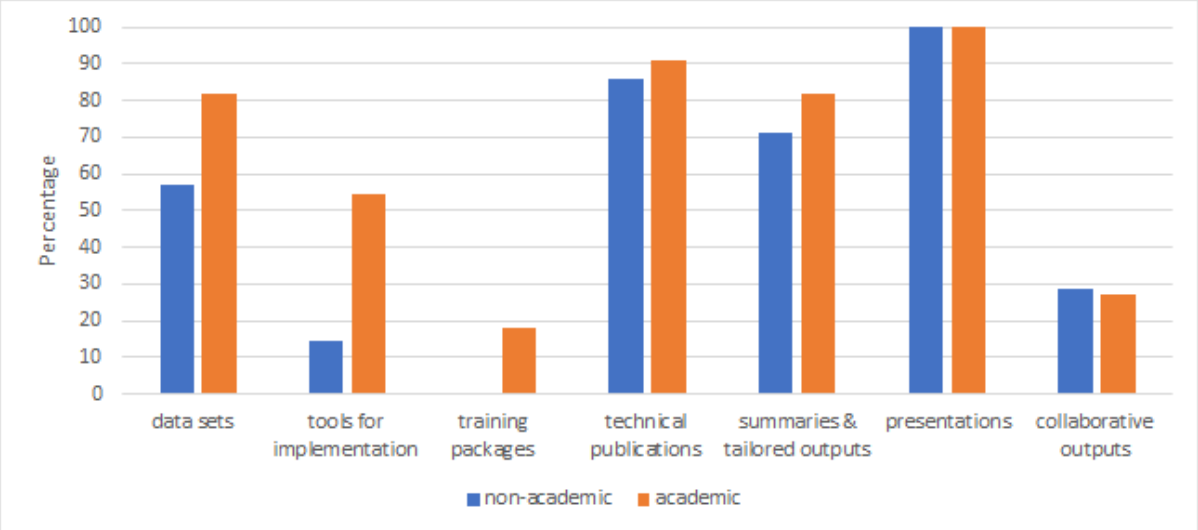
It could also look at audience and sectoral reach of:

- workshops, discussions, forums specifically to tailor and adapt findings
- training forums to support adoption of tools, and new techniques
- stakeholder-focused presentations given to support adoption and uptake
- presentations to popular audiences including schools, community groups and public forums

Insights from the surveys: comparing responses between academics and non-academics

Complementing these quantitative and documentary outputs measures, assessment of survey responses can tell us about what level of awareness there is of what kinds of outputs among different sectors, as opposed to quantitative measures of total research outputs.

Figure 5.6. Types of outputs from their project that survey respondents had seen or were aware of.



After adjusting for the fact that more academics than non-academics completed our survey, academics consistently showed a greater awareness of different types of outputs that were produced than non-academic participants except for presentations (all were aware of these, Figure 5.6). On the other hand, more non-academics reported awareness of collaborative outputs. This suggests that

there may be a gap between production of outputs and their dissemination, accessibility or the level of awareness of them among non-academic respondents.

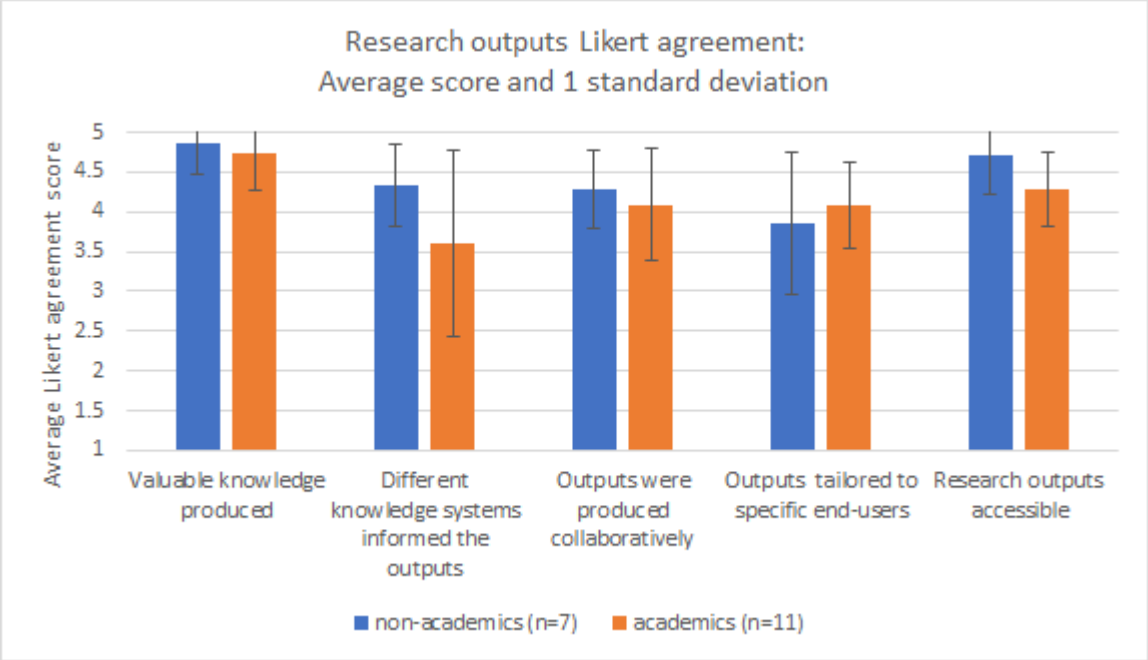
This is something that could be assessed at a point in time, or at repeated points, including later in a program cycle, especially as the level of dissemination and awareness might be expected to grow over the course. If the survey questions were adjusted to ask participants to estimate the number of different types of products they were aware of, as recommended in the analysis in Chapter 4, this would allow the comparison between the degree of awareness of products within different sectors to be quantified further.

Aggregate comparison of the likert-scale outputs measures between academic and non-academic respondents can also reveal how the value of research projects is assessed comparatively between sectors. Respondents from across sectors and project roles were overwhelmingly in agreement with the five positive statements about research outputs, with all responses being rated above average, and most significantly so (greater than 4 in a possible range of 1-5, Figure 5.7). On average the measure about knowledge generation itself - "Valuable knowledge was produced through the project" - had the highest level of agreement (4.78 in possible range of 1-5), with very little difference between academic and non-academic respondents, demonstrating wide agreement that the suite of projects as a whole has delivered significant value in terms of this measure of producing valuable knowledge.

Contrasts between responses from different sectors are also revealing. Non-academics on average assessed outputs to be more accessible than academics (4.71 versus 4.27) which indicates the research is contributing value in terms of perceived product accessibility among non-academics, despite an apparent lack of awareness of this among academics. This trust in product accessibility among non-academics is in contrast with their much lower awareness in general of the range of products generated by projects, discussed above. This points to the potential need to dig deeper into this apparently contradictory situation, for example through narrative questions that could explore the range of products known and being used and shared by research users or non-academic members of project teams, their accessibility, and any reasons for the disparities in reporting these outputs between members of project teams from different sectors.

On average the biggest difference in responses between people inside and outside academia was around the statement "*Different "knowledge systems" informed the outputs of the project*" which non-academics agreed with more than academics (4.33 versus 3.60; Figure 5.7; this difference is also observed in within-project comparisons, as noted in Chapter 4). This point is discussed further in Chapter 6.4, in terms of what it may mean for academic practices and building research relationships. However, in a comparative valuation exercise, these differences in responses between sectors could be interpreted to suggest that value is being added through different knowledge systems brought by and through non-academic partners, even where this is less apparent to research teams. Again, narrative measures can be used to shed light on these diverse views.

Figure 5.7. Comparison of responses from research participants inside and outside academia to research outputs measures.



Insights from the interviews

The variability in narrative responses to questions about outputs reflected a combination of differences in the types of research being done, the role of the interviewee in the project and their sector of employment. For example, in Question 10, non-academics tended to focus on how knowledge is being applied (or not) and its impacts while academics tended to focus more on a description of the actual outputs and what they were targeted at (Table 5.3). These differences potentially reflect both the values, and the components of the research process that people are most familiar with.

These differences in values and perspective were also evident when it came to reflecting on the processes to create outputs (Question 11). There was a tendency for academics to describe the outputs more than the process of producing them (Table 5.3). Academics mentioned some outputs where non-academics had contributed, but were more likely to say they had led the production of outputs, particularly when discussing academic papers. In contrast, the two non-academics that spoke about academic articles specifically mentioned co-authorship (these were from the same project, so this may not be a broader phenomenon). Non academics also focussed slightly more on outputs beyond peer-reviewed papers such as fact sheets and presentations.

Question12: Only five people were asked Question 12 (three academics and two non-academics) so responses were quite variable reflecting the different types of projects and stages of output preparation. As captured in the recommendations, the adequacy of dissemination is probably not the right question to ask in an interview. Academics possibly focus on papers a bit more (although they did mention other outputs).

Table 5.3: A selection of quotes from academic and non-academic interviewees highlighting their key topics of interest or value.

Question	Non-academic respondents	Academic respondents
Q10. How would you characterise the knowledge produced by the project?	"because we crafted the questions really early on, and together, the outputs were directly useful and relevant" "we did have confidence using it in conservation planning" "One of my concerns is that people might be taking some of these results as conclusive, but they're not"	"primarily, the research outputs were designed to assist with management decisions" "t's been a significant piece of work for measuring the success of the government's current policy"
Q11. Can you describe how project outputs were produced, particularly who was involved?	"what TSR has done is normalised some of that [working collaboratively]. I look at it now and go of course we helped write the research factsheets, because we're a research team" "I think everyone on the project team had some contribution, to the write up of the fact sheet summaries, the general public summaries and technical expertise or ownership of the data resulted in the scientific publication"	"their main contribution came in terms of editing, reviewing drafts, or approving the final product" "we've produced some threatened species hub materials like fact sheets and findings fact sheets and videos"
Q12. What do you think about how knowledge generated by the project has been packaged, presented and disseminated?	"it was a chance for everybody to come together and catch up on the project and hear about the results and meet everybody else...I think that's been really successful"	"it's the question we're dealing with right now. How to do it and who are we doing it for" "I guess they've sort of followed the traditional way of the researcher writing the paper."

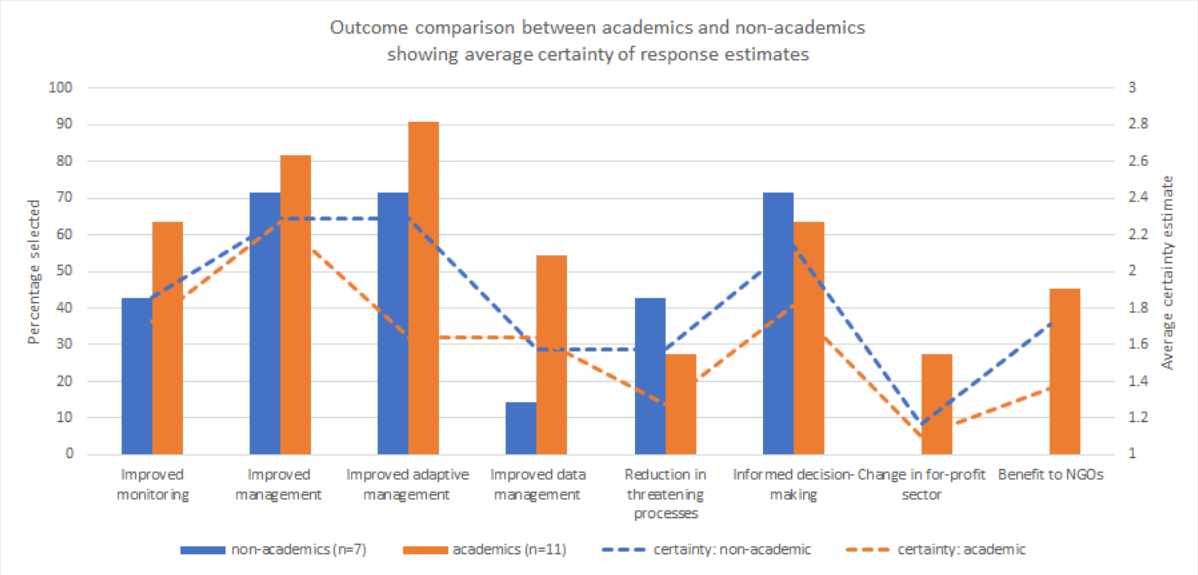
5.5 Deeper dive: examples from outcomes and impacts measures

A final approach to analysing the data available is through a deeper dive, either at the project level, or, as in the data presented here, at the program level. This can be particularly valuable to shed light on research value where aggregate data are highly variable or uncertain.

Insights from the surveys

Most respondents rated their confidence in most of the outcomes measures moderate to low (Figure 5.8). Respondents from different sectors tended to rate the likelihood of outcomes differently. Academics were more likely to feel research had contributed to outcomes than non-academics, with two exceptions: with regard to reduction in threatening processes and informed decision-making (Figure 5.8). Non-academics were on average more certain of their assessment of project outcomes than academics, especially improved adaptive management (which fewer saw as an outcome compared with academics) and informed decision-making (which more identified, Figure 5.8).

Figure 5.8: Comparison of project outcomes selected by research participants inside and outside academia overlain with average level of certainty estimates that outcomes will occur (based on a three point scale: limited = 1; medium = 2; robust = 3)



Impact measures were ranked lower than measures from the other four domains (Figure 5.9). With the exception of measures related to capacity to pursue locally-defined priorities, objectives, and goals (Indicator REI92) and greater public awareness about conservation options etc (Indicator REI94), impacts were more likely to be considered possible (the neutral option) to highly unlikely, than likely to highly likely. Positive economic change scored lowest overall, with none of the survey respondents rating this likely or highly likely. The certainty of respondents about impacts was also typically below the midpoint across most measures (Figure 5.10).

Figure 5.9. Survey scores for measures related to research impacts

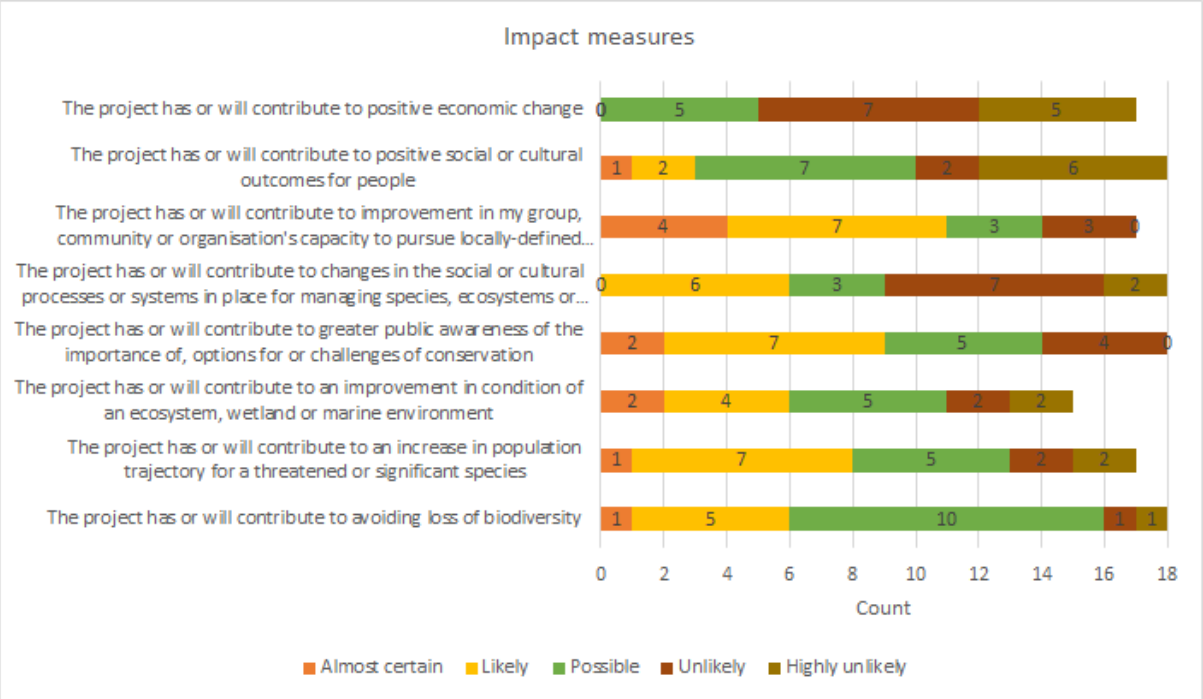
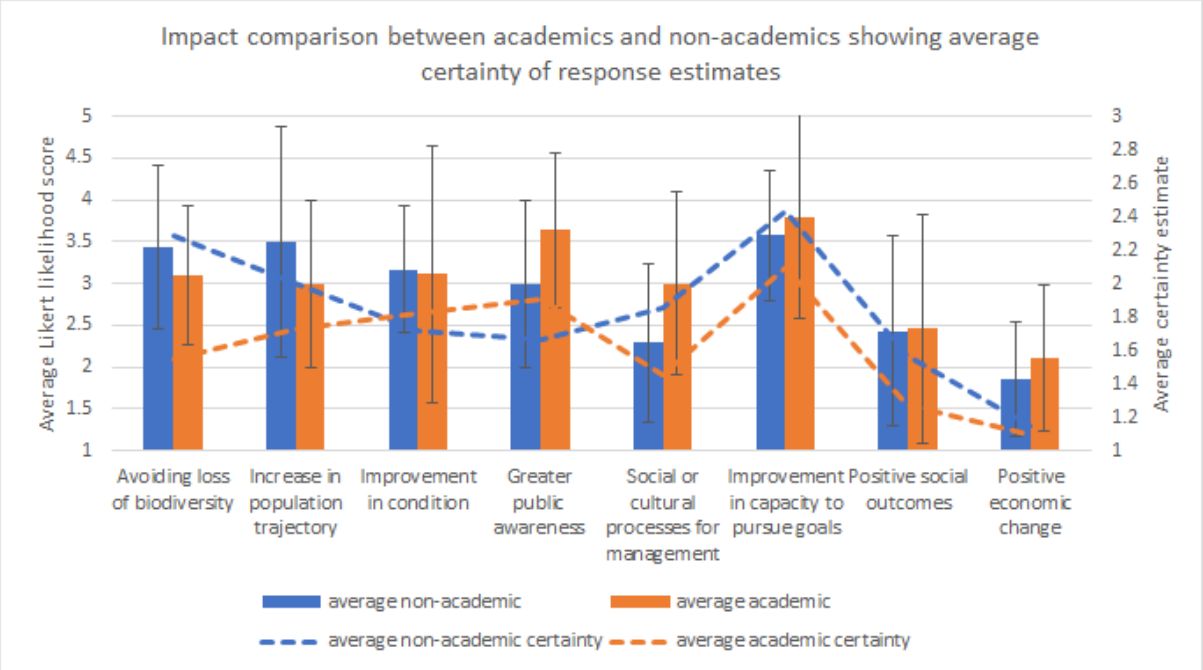


Figure 5.10. Comparison of project impact scores for research participants inside and outside academia overlain with average level of certainty estimates that impacts will occur (based on a three point scale: limited = 1; medium = 2; robust = 3)



One of the most significant reasons for this relatively low assessment of the likelihood of and confidence in outcomes and impacts appears to be because in many cases these have not had a chance to fully mature. As one respondent said in interview:

> *"even when I was doing the online survey. Just because of the wording, I feel like I said no to most of the outcomes, not the outputs, the outcomes. Because it was like, did it change the trajectory of threatened species? Or did it change how it's being managed? And I was like, well, actually, not yet. But it could."*

Nevertheless, for most of the impacts measures, some of the respondents assessed the chance of these impacts occurring as likely to highly likely. Likewise, the outcomes measures all garnered positive responses from at least some respondents.

For both of these categories, then, diving deeper into the outcomes and impacts that were achieved or seen as likely, both through narrative examples and through quantitative, documentary and community survey data (where available), can provide valuable, detailed insights into some of the greatest successes, as well as some of the challenges and constraints. The exemplary approaches used in this deeper dive may be similar to the kinds of information gathered for impact case studies. However, this approach also allows for assessment of key themes, reporting of limitations and challenges, and detailed exploration of particular themes across multiple projects (such as policy change, or changes in social and cultural perspectives).

Insights from the interviews

In the interviews from our case studies, examples of outcomes and impacts were provided spanning on ground outcomes for species, new funding, changes to policy, new partnerships, growing community connectedness, new skills, and greater definition of the problem. In a deeper dive, an

overall assessment of narrative themes or of insights from the survey data can be a starting point for more in-depth illustrative examples, which can draw out more of the detail and nuance of the value achieved and the breadth of areas in which this was achieved.

This can be done to draw together in-depth examples that illustrate common themes across a suite of projects or programs, or pointing towards a common set of goals. For example, in our case studies, project participants spoke of a wide range of ways by which species directly benefited from the research:

> *"we managed to educate and really have* [participants] *connect with the idea that these plants and ecological impacts are really, really important.. And these* [participants] *took it to heart. They planted a garden with their species, which I think meant that* [it is] *the largest population of* [that species] *in the state."*

> *"the research is principally about having a more robust current understanding of the status of the animal and it definitely ties into the nomination to uplift from vulnerable to endangered"*

> "*I know that it's changed management on the ground at* [ParkX] *in terms of management decisions around animals being released and managing risks"*

> "*I remember writing the words, with* [X] *saying, this will be the first safe haven, fenced safe haven for* [our species] *in Australia"*

Narrative data can also provide powerfully compelling insights into some of those areas of research outcomes and impact that are difficult to fully capture in likert-scale and quantitative measures, especially those that are less common across a suite of projects or program. For example, this can be in changes in the social or cultural processes or systems in place for managing a species, ecosystems or places:

> *"we've formed an informal recovery team for the species"*

> *"One of the biggest things to come out of this research was the fact that [X's] work demonstrated the species was in decline... That just made a whole bunch of people sit up and take notice"*

> *"we have a really good relationship with them. I would say it's almost as good as it could be for the short amount of time that this project has been happening. So as an example, we've also set up a community of practice, which is sort of their idea, but we both manage it and we meet every four to six weeks."*

> "*there's this body of young scientists out there now, and no matter where they end up, whether they end up in academia,or government or end up in*

NGOs, they've got this training in their DNA around how they collaborate with nonscientific organisations."

Or it might be an illustration of where programs have helped change community perceptions and values for conservation, such as this example of changing perceptions of the importance of conservation work and of species that are threatened and often undervalued:

"I would get stopped when I was out monitoring animals. A community member might stop by my car,and have a chat to me about their sightings or what they thought about the programme."

"To have all these [community members] saying, yes, this plant … you probably would just blink and you miss it not thinking much about it - it's a really important species and it has every right to exist, as I do…I thought that was really, really impactful… And it's one thing to say that on a ... questionnaire, but to hear those same thoughts kind of being spoken aloud when they're interacting with myself, I think that was a really surprising, in the best way possible, outcome... And [to] really have these [participants] connect with the idea that these plants and ecological impacts are really, really important. And they are going to take that with them. Because you don't, it's something that I think you can try and teach somebody, but if they actually take it to heart, that's a big difference… I think that's a huge impact. That's a mindset, that's not necessarily a one off thing, that is a change to the way that you think, that is a big, big difference."

At the same time, deep dive evaluations can also draw out examples across the program where some respondents may have observed particular barriers to achieving the desired research outcomes, such as the persistence in some cases of one-way or 'push' strategies of research communication:

"*In days gone by a challenge for the relationship between government and academics was, particularly some senior professors turning up and giving us all the answers before we've even given them the questions.*"

"*Whether or not we've been able to persuade* [researcherX], o*ne way or another on a particular topic, I can't think of an example*"

Quantitative, documentary and community insight outcomes and impacts measures

Quantitative, documentary and community insight outcomes and impacts measures drawn from the framework can provide another approach to 'deep dive', complementing, bolstering or contrasting with narrative information. While it is not possible to provide illustrative examples of quantitative, documentary or survey measures from these case study projects because of the need for confidentiality, quantitative data are available for example projects within this suite that include:

- Population viability analyses
- Ecosystem condition assessments
- Documentary evidence of policy change
- Documentary evidence of programs created

- Community surveys
- And many more

In some cases, these more difficult-to-measure quantitative and independent measures are available because they were part of the project methodology. These measures were built into the design of projects for the project's own purposes, but can then be adapted for assessing and reporting research value.

Chapter 6. Discussion and applications of the research

6.1 Conceptualising research value

Throughout this study as a whole, a core insight has emerged repeatedly: that **many elements need to fall into place for a conservation research project to achieve broader environmental, social or cultural impacts**. Even producing intended research outputs is fraught with uncertainty. While the Phase 1 study showed that indicators for Research Environmental Impacts (Domain 5) were consistently considered among the highest importance, these are difficult to measure and to disentangle from all other factors that contribute to achieving or not achieving these impacts. Understanding and assessing the complex pathways through which research activities have flowed, and identifying the gains made through research collaborations *even when they are not 'successful' in directly benefiting species or ecosystems or other values*, allows us to develop a more comprehensive and nuanced picture of the value (and shortfalls) of research projects. This is the thinking behind the focus on and language of research 'value' throughout this report, not just 'impact', and on the **use of a diversity of measures that span the research cycle to assess this value.**

The importance of using a diversity of measures is twofold. Firstly, as is increasingly recognised in the research impact and engagement literature, achieving environmental, social, cultural or economic impacts is not just a matter of generating high quality (or highly cited) research (Donovan 2011). It is strongly dependent on effective engagement and building appropriate 'pathways to impact'. Applying measures across the domains identified here allows an assessment of these pathways, how effectively they were navigated, and any limiting or conditioning factors that could not have been controlled, but which may have impinged upon the project's effectiveness.

Secondly, these measures represent values in themselves. They are not just steps towards a goal, but benefits in their own right. Building strong research partnerships that can take new work forward, training early career researchers in a different way of working, growing the interest and skills of citizen scientists, and producing discoverable outputs are all values that contribute to the overall societal capacity for conservation action. Frequently, outcomes occur through circuitous pathways and unexpected avenues. A management action is not taken in one place, but new relationships are built to enable the knowledge to be extended and tested elsewhere. A policy is not changed, but business investment is redirected, changing the conservation landscape. Measuring values across the spectrum of research activities allows an assessment that captures research value in the face of this uncertainty and complexity through which conservation gains are made.

This has implications for how research is established, practiced, and funded. Much of the debate in the research impact space, particularly in Australia, has focused on the limitations of academic measures (papers and citations), and on attempts to shift the focus to 'real world impacts' (e.g. Watt 2016, Australian Research Council 2017, but see Papageorgiou 2021). Even internationally, the challenge of achieving, measuring and disentangling 'real world impacts' has driven a focus on narrative case study assessments, but left the system in something of a stalemate in developing quantitative, survey-style, comparative or more standardised measures (e.g. Wilsdon et al. 2015). It has also tended to reinforce the privileging of more senior and advantaged researchers and institutions, with more access to direct influence and more resources to trace and assess this influence (Wilsdon et al. 2015). Shifting the focus toward measures that help build a picture of trust,

discoverability, participatory processes and respect of researchers for diverse points of view (including those of research users) creates a picture of value that researchers at any career stage and any background can potentially demonstrate. Having institutions and funding bodies both reward such approaches and invest resources in systems that support their assessment (and achievement) would encourage more engaged practice and iron out some of the perverse incentives that currently obtain in the research industry.

6.2 Considerations for applying these approaches for measuring and assessing research value

This research project emerged as a response to the need identified in the first phase of the National Environmental Science Program to develop appropriate measures to complement narrative case studies for use in annual reporting (self-report measures) and in independent evaluations. Recognising that traditional academic metrics of success are important measures of academic rigour, but are not sufficient to reflect the real value of engaged research projects and programs funded under NESP, the goal was to extend the range of metrics available to assess value in the program and the engagement practices it fosters. Here, value is about real world relevance as a complement to (not replacement of) academic rigour and excellence. It is a powerful additional criteria in research evaluation, one designed to increase and indicate how far and in what way a given piece of research has 'reached out into the world' to change it for the better (Rickards et al. 2020). In the course of this project, some of the approaches identified here have been applied to reporting within the TSR Hub. NSW DPIE have also drawn on the suite of indicators developed in this project (Lavery et al 2021) to generate discussion and draw on as a basis for developing future research and better understand how they manage their knowledge outcomes.

We anticipate that research funding bodies, research institutions, and research users will be able to use the framework developed here to plan, implement and evaluate research: to consider what they hope to achieve from research, determine the best foundation to try to deliver value, how they will plan research processes, and how they will assess success. The aim is not to provide a prescriptive set of indicators or measures suitable for every project or assessment rather a suite of indicators, measures and methods that can be tailored to suit the circumstances, feasibility and aims of a particular research context or valuation exercise. The framework also represents a way of thinking about valuation that stretches beyond traditional metrics or tangible measures of conservation impact.

While the focus of this project has been on indicators for conservation research, there are many potential parallels for applying these measures to other forms of environmental research, or research more broadly. More detailed considerations for applying the framework and measures to conservation research and to other disciplines are discussed in the following sections.

How well do the measures developed against these indicators capture the value of conservation research projects?

Overall, the measures tested and developed here paint a comprehensive picture of the value of conservation research projects. A strikingly wide range of the measures tested were rated as moderately or highly fit for purpose in the survey while in interviews research participants generally

responded well to our broader interpretation of research value. Many of the measures appear to elicit relatively consistent responses across project teams, or have differences that can be analysed to reveal research value, suggesting the measures are robust and applicable to use in a wide range of analyses.

These measures can be applied and analysed in a range of ways to assess the value of conservation research projects: in aggregate form, by project, by research institution of respondents (or similarly by research project role, separating out research users from research providers and other partners), and/or through drawing out a deeper analysis of a few projects to illustrate value, and the challenges or limitations in achieving value, more comprehensively. Integrated analysis across survey and narrative measures for the case study projects reveals both positive assessments of the suite of projects, and challenges, limitations and areas for improvement. Even where some projects were still in progress, and it was too early to demonstrate outcomes or impact, our approach showed that it was possible to provide some assessment of value by looking across the full impact pathway . Examples of quantitative, documentary and independent community insights in Chapter 5 indicate such measures could flesh out the picture painted by survey and narrative approaches and demonstrate the potential benefits of an integrated approach. Altogether, these measures appear to provide a broad and nuanced picture of the value of conservation research projects across the full program logic of research projects.

Applying the framework for measuring the value of research

The framework brings together, tests and assesses a range of measures that could be applied through a combination of narrative approaches (interviews or focus groups), surveys, and compilation of quantitative and documentary information (for example through inquiry to project team members and research support teams, if these are available). The aim was to suggest diverse types of measures - narrative, survey, quantitative, documentary, modelled, monitored and independent community responses - that play different roles in the research valuation ecosystem, but that can be drawn on as a suite to build an effective picture of research value across projects and programs, according to opportunity, feasibility, and purpose of a particular valuation exercise.

Our expert workshops suggested that modelled, monitored, quantitative, documentary, and independent survey measures are often the first (and sometimes the only) kinds of measures participants think of when asked to think about assessing research value and impact. They certainly have value in providing an additional point of reference for triangulating and testing these survey measures and upon which narrative measures can elaborate. In the case of some quantitative and documentary measures, these also lend themselves to aggregation and comparison, which can be useful to develop an overall picture across research programs and institutions. However, in all cases of aggregation and comparison, care should be taken to ensure that these measures are likely to have been assessed and reported against in consistent ways between the entities being compared. The temptation of quantitative measures is an assumption that they provide a greater level of 'truth' and that they can automatically be added and compared, but, as the narrative responses demonstrate, the nuance at play in processes of research collaborations, and in what they have been able to achieve, in reality often confound the data.

A number of the survey measures of research partnerships appear from this preliminary study to be relatively stable between research team members, and can be used relatively readily on a wide scale to develop a broad assessment of projects. Against some of these measures, noted in section 4.2, the rating of projects appears to be fairly consistent between academic and non-academic members of research teams, suggesting that, in these cases, researcher self-report may be relatively reliable as a measure of research value if this is the only feasible approach in a particular circumstance. Many of these measures are likely to be more reliable when tested across different members and diverse roles within research partnerships (academic and non-academic, researchers based in research institutions and research users). In some cases, it might be appropriate to give greater weighting to those who are closer to certain processes (e.g. those who need to have access to and use the data assessing the quality of data management) or to use the results from the survey to reveal important within-project comparisons (such as the levels of awareness of different types of outputs among different members of project collaborations).

Narrative measures provide depth, detail and context to an assessment, providing more confidence to survey and quantitative measures, but also giving 'flesh to the bones' of how change is or is not achieved. This is critical, not just for understanding and confirming whether and how specific outcomes or impacts are connected to research, but for lending nuance to the complexities of these assessments, the additional factors that were in play, and how these changes have (or have not) come about. Narrative approaches are also important for ground-truthing any assumptions about research value, identifying unintended impacts, and being guided by more holistic understandings. This may be of particular importance in Indigenous-led research (e.g. Tsey et al 2016).

These different types of measures are designed to be complementary, and this complementarity is a critical feature of the framework. It enables flexibility in the application of the framework, while increasing the potential for cross-verification across measures, adding reliability to any valuation exercise as a whole. For example, quantitative measures may be relatively less important for addressing the aspects of research where survey measures tend to be stable and effective (e.g. inputs and processes), but narrative measures could add value here for fleshing out detail and context. On the other hand, where likert-style survey measures appear to be less stable or consistent between different members of research teams, as with many of the impact and outcomes measures, quantitative, documentary and independent community insight measures may prove to be of greater relative importance. These measures can anchor and lend independent weight to less clear survey assessments and to insights shared in narrative form. While quantitative assessments of outcomes and particularly of impacts can be very challenging to measure, and will always be subject to some degree of uncertainty due to confounding factors, the examples cited in the analysis of where such approaches have been used in recent conservation research and policy contexts give an indication of where and how such approaches could potentially be applied, in contexts where this is deemed to be important and where resources are available for these more challenging analyses.

Directly attributing research to impacts can be difficult due to time lags associated with project completion. Desired outcomes may take years, decades, or centuries to eventuate. For example, improving the recovery of hollow-dependent mammals such as Leadbeater's possum (*Gymnobelideus leadbeateri*) is reliant on increases in hollow bearing trees that may take 120 - 150 years to develop (Lindenmayer et al. 2015). Additionally, interlinking projects with similar aims, or confounding

processes such as external policy, community or economic pressures, can make it difficult to attribute impacts to particular projects and to particular objectives. Applied research of the kind assessed here, particularly when done well, can have more direct pathways to being taken up and used in practice or decision-making, but fundamental research (for example, into the ecology of species and systems) are necessary starting points to conservation research. These initial studies may take years to gather information necessary to begin to understand the natural history of wildlife, plants and ecosystems in order to identify critical outcomes needed for species recovery.

This framework has the advantage of not being entirely reliant on disentangling these confounding factors, and on being applicable at almost any point in a research cycle. Which measures are likely to be effective will naturally vary with the stage the research is at. This is not only true of published products, outcomes and impacts measures, which are more likely to be measurable later in research projects, or in the months and years to follow. Measures showing the immediate outcomes from inputs and processes, from skills development and training, are conversely often best collected while a research project is underway, or not too far from the end, as the recollection of these is likely to be more immediate and accurate.

Setting a baseline at the start of a research project, program or cycle is useful but not necessary for most of the measures, most of which primarily rely on reflective reports from across project teams and triangulation of measures (quantitative, survey and narrative) to ground-truth responses. However, a number of the quantitative measures - ranging from before-and-after measures of trust, to species population trajectories - do rely on a baseline. The limits of available monitoring data for species and condition, and the feasibility of collecting such data, are likely to impact on the potential for effective baseline-setting, especially for measures of biodiversity impacts. On the other hand, because what is being tested here is the value of research, in some of these cases, where the ultimate goal is to achieve positive results for species or ecosystems, the most meaningful research measures may be set against a 'zero' baseline. For example, where the research is dedicated to improving monitoring techniques for a species that has been undetected for many years, the measure of 'number of detections' becomes a relevant independent, quantitative measure of research success, even where the impact on the species is not directly traceable.

Setting of baselines can also help in disentangling the contributions of particular projects from wider collaborations and processes, recognising that the boundaries around projects are often arbitrary, and/or that individual projects are frequently elements of broader relationships and research partnerships that can extend over many years, over different pieces of research funding and cycles, sometimes over whole careers. In this context, a baseline could help understand whether already-strong responses to measures (such as high levels of trust in relationships) function as a sound input into the research process, and/or to assess whether there are improvements in these measures, e.g. whether linkages are further strengthened, through the particular project or program phase in question. In the absence of a baseline, due to the evolving nature of stakeholders and research partners, or because this was simply not done at the time, contextualising questions on the length of collaborations and partnerships and/or previous work undertaken together could be used to frame and interpret the responses of the project-based valuation exercise.

On the other hand, the strength and durability of collaborations and the cumulative nature of relationship outcomes over these project cycles can also be assessed as a significant contribution to growing research value. As one research user observed in an interview *"thinking back to what it was, like working with researchers during CERF[1] and what it's like working with these researchers now, I think, wow there's been some really transformational shifts occur during that time*" [TSR01]. Many of these measures tested here can be applied across funding cycles and multiple projects, as well as within individual projects or cycles. Baseline setting and repeat measures can also be undertaken periodically within and between project cycles to assess cumulative change over time, as well as to adjust for unavoidable changes (such as changes in personnel).

Adjustments to the measures and their application

The application of measures in the case studies also highlighted some adjustments, gaps and areas requiring further clarification.

Allowing for inherent uncertainty in outcomes and impacts: The nature of assessing outcomes and impacts is inherently challenging and fraught in many cases. For many of the Phase 2 case studies tested here, processes for finalising the research and achieving outcomes were still underway. This is to be expected, and is part of the design of the framework, but it means that, for many outcome and impact measures we are asking people to look into the future and speculate as to what might happen. It is hardly surprising then that there are more variable responses to these types of questions than there typically are for more concrete assessments of past events. For this reason, we recommend separating out when projects *have already contributed* to certain outcomes and impacts, from their speculations as to the likelihood that they *could* contribute to outcomes and impacts in the future.

There is also an important question about the knowledge base different individuals have upon which to base their judgment. Thus collecting more detail on the certainty of these assessments, as well as on the sources of information on which the assessment is based, is potentially very useful. Retaining the confidence measures used in these surveys, but also adding free-text space for extra descriptive information on the outcomes and impacts, and the evidence upon which these assessments were based, would allow the likert-scale and multiple-choice ratings data from these questions to be interpreted more readily, and potentially lend insights to the different perspectives provided by members of project teams with different roles. This would strengthen the usefulness of survey-form questions for determining value. Triangulating these survey measures with quantitative, documentary and independent measures, and with the richness of narrative information, would lend further weight to these assessments.

Clarifying language and information sought on diverse knowledge systems: The Phase 2 study highlighted some confusing language in the survey and in the narrative measures. As discussed above, the language of 'different knowledge systems' was almost certainly confusing and confounded responses. One survey respondent commented on the process measures, "I found these difficult to judge. Is a data management system or data set a knowledge system? They are only a partial match

[1] The Commonwealth Environment Research Facilities program, an early predecessor to the National Environmental Science Program, which funds the research projects in this case study.

with the definition above to my reading", suggesting that the language around 'knowledge systems' was conflated with the questions on 'data systems'.

The language of 'different knowledge systems' is often in practice used as code for Indigenous knowledge. However, it could also mean practitioner knowledge, specialist knowledge such as that of farmers or educators, local community knowledge or even interdisciplinary academic knowledge. The outputs question (but not the related processes question) in the survey provided some clarifying text, as follows: "Different "knowledge systems" informed the outputs of the project (agents, practices, and institutions that organize the production, transfer and use of knowledge i.e. policy, management, societal, Indigenous)". Having clearer language explaining and giving examples of different knowledge systems, and potentially separating the question out to ask about specific knowledge systems where this is important (e.g. separating out questions on Indigenous knowledge systems from those focused on other forms of diverse knowledge) would help greatly clarify this measure and achieve more useful results.

Clarifying language on methods and data: The narrative measures related to the methods used in the interviews (*What is your view on the research methods used in the project? P1 Did it draw on different disciplines & knowledge systems?)* did not quite connect effectively to drawing out insights for assessing the value of research. Researchers in particular tended simply to describe what they did, rather than commenting on how what they did might reflect research value. Rephrasing this question to ask whether the choice of methods for data collection and analysis contributed to the success of the project, would better assess value. Follow up prompts could explore:

- whether the methods allowed for the most useful range of insights to be drawn from what was available (resource-effectiveness of the methods),
- whether the methods used helped build or strengthen networks, or
- whether new research methods were developed that would or could now be applied to ongoing work in this area.

Avoiding 'tick box' approaches to sensitive questions: It is important to take care in using appropriate methods to assess particular aspects of research value. For example, as discussed above, participants at the expert workshop warned against using a 'tick box' approach to assessing research co-design. A comment on the survey seeking greater nuance in the question on whether prior informed consent was sought suggested a similar hesitancy about 'ticking a box' in a survey on whether or not these happened. Surveys naturally lend themselves to a tick box approach - indeed, this is often one of their strengths. Thus, if survey measures are the only available instrument in a particular context, and these issues are important to the assessment of research value, it is important to take care around the wording of these questions to assess the quality of the application of these approaches and not just whether or not these happened (e.g. through using likert scales and free text). It is also important to ensure wherever possible that partners who would be expected to participate in any research design or informed consent processes, and whose views are thus most material to whether these processes were appropriately applied, are sought in the survey.

Collect contextualising information and branching questions: Some important contextualising information, such as the institutional position and project role of respondents, was collected in the surveys and interviews and used to inform analysis and interpretation of results. However,

conservation projects are extremely broad in scope, as even our small sample demonstrates. The most efficient valuation method would likely have a fuller suite of situating questions, and use a more sophisticated branching logic to tailor what questions follow to suit the context and purpose of projects.

In addition to the questions asked here, of who were the research participants and research users involved in the project, potential situating questions could usefully include:

- Geographic scope of the project at its conception and in its application, including whether or not an expansion of scope was made possible because of the value or success of the project
- The nature of data sought, and dependence on third party data
- The timing and context for how the individual project relates to a wider body of research (e.g. whether this is the first significant research in this area, or a continuation of 10 years previous work)
- The timing and context for how the individual project relates to a wider set of relationships (e.g. whether this is the first research for this group of research users in this partnership, or this is part of an ongoing program of work between these research teams and research users in this region or field).

Potential situating and branching questions could relate to:

- The primary purpose or aims and application of the project (e.g. whether it aimed to inform policy, management, planning, community participation, education, or other) and the intended domains of impact (e.g. social, cultural, economic or environmental).
- Research team roles of the respondent with respect to the project - this question was asked in the surveys and interviews, but could be applied as a branching question to determine which further questions are put - for example, only asking those close to the data analysis about the benefits of the methods used. Providing additional project role options, including 'building partnerships and collaboration', would allow more sophisticated branching or conditioning analysis, particularly around questions such as the extent to which research users were involved in the research.
- How long the respondent had been involved with the project or during what phases in the research cycle the respondent was involved (conception, identification of the question, planning, design, implementation (including as a partner or stakeholder), publication, dissemination, uptake). This should be used as a measure to condition interpretation of the responses, and could also function as a branching question for all of the later questions.

Questions on the primary purpose and aims of the project could particularly be used to determine the range and order of outcomes and impacts questions put to respondents. However, the branching should not preclude any unexpected benefits that might arise from the project. Thus a question logic (in a survey context), could foreground the likely most relevant questions (e.g. around on-ground management), then add a further branching question to ask 'were there any additional benefits from the project?' prompting with those areas not flagged as the primary purpose of the project (e.g. policy, planning, community, business). These could then branch into further questions on any of these domains of outcomes or impacts that the respondent has checked as relevant.

Contextualising factors during the research can also condition whether outcomes were achieved. The expert workshop particularly identified the "challenge of churn" in stakeholder relationships, that could impact what value the research is able to achieve, through no fault of the research partnership. As the workshop noted, "It is not always possible to control this, but may be important to develop measures to capture the extent to which this has impacted on research progression." Interview questions on the stakeholder map could ask for the duration of involvement of different individuals. Survey questions could ask whether (1) research users and (2) members of the research team were added, lost or replaced during the course of a particular project, and could also test for agreement/disagreement with the statements 'changes among research users and stakeholders impacted on research outcomes' and 'changes among the people undertaking the research impacted on research outcomes.'

Add new questions to address gaps identified through comparison of interview and survey approaches: as noted in Chapter 4, additional questions to round out the framework could address objective-setting, within-project communication processes, the quality of relationships, legacy planning, learning as an outcome, and governance and management of projects, including whether or not the project teams have assessment of whether objectives have been met and why. These measures have been added to the full framework provided at the end of this report (pages 135-166).

Questions addressing the extent to which objectives were clearly and inclusively articulated and met would shed light on the project's achievements and the governance and management of projects (e.g. how well project team members understand and recognise their objectives and monitor whether they achieve these objectives). This would ideally include measures as to whether joint understandings on the objectives of the research were achieved (inputs measures), whether or not these objectives were met and why (academic outputs and outcomes measures), as well as measures on whether any limitations, constraints, challenges or failures have been clearly communicated (outputs measures). While some of the original suite of indicators touched on aspects of these inputs and outputs measures, this overarching assessment of the governance principles within research projects would help guide research practices to articulate objectives more clearly and strengthen a focus on research governance.

The question of strong relationships is a theme throughout this report, but a focus on assessing the quality of relationships as well as gathering information on the types of partnership or collaboration established would help highlight this as a critical if less obviously tangible resource for delivering research value.

Survey and interview questions on the quality and regularity of communication processes within projects (processes), are an important dimension of effective research designed to meet the needs of research users, and can have the added benefit of being a valuable means of strengthening relationships. As our case studies showed, active participation of research users in research processes can be an incredibly effective way to ensure research value, but is not useful, feasible or desirable in all cases. Questions on the quality and frequency of communication within projects, including on whether communication was one-way or appropriately reciprocal, would help flesh out the quality of research processes. These questions should aim to determine whether the nature and frequency of

communication was sufficient and fit for purpose, noting that more frequent communication is not always better.

Questions could be added to ask explicitly whether there were any changes in research focus or processes as a result of research user input. While the absence of this is not inherently an indication of poor process or value, its presence would help deepen the assessment of how strongly projects were established with adaptive, feedback loops embedded, and how well the research drew on and responded to the insights and perspectives of research users.

The question of what arrangements have been made within projects for legacy could be added to both the survey instrument and the interviews. The quantitative measures on discoverability and reach partially inform the legacy picture, but specific questions to project teams on whether outputs are stored in 'perpetuity', with an eye to duration beyond the project life, planning for discontinued funding and most importantly the effective emergence of a new generation skilled, knowledgeable and experienced researchers, could help inform assessments of whether these outputs are likely to be accessible in the long-term and to whom, as well as helping ascertain how conscious, deliberate and durable legacy planning has been.

Finally, the learning achieved through projects is an outcome in itself. This includes learning about the topic, the wider policy and management challenges and context to which the project relates and the development of skills in effective approaches to research engagement and delivering research value. Additional questions could ask about what has been learnt in the project as a whole, whether there has been learning by individual project team members across each of these domains (self-report and a general question on learning by project team members), and specific questions on how well expertise and skills have been developed among early career researchers and students.

Considerations for aggregate evaluation strategies, e.g. across programs

In piloting the framework to apply to test-case projects, we have consciously taken a project-based approach, asking participants to respond with respect to a particular project that we proposed. As we have seen, taking a project-based approach to assessing research value across a program (e.g. aggregating data from surveys and interviews that are project-focused) allows for comparison across projects. For larger evaluation processes, it would also allow different types of projects to be aggregated and assessed collectively - for example, disaggregating data from projects with an on-ground management focus from those with a policy or decision-making focus. In addition, the advantage of a project-based method of assessment is that it allows respondents to consider responses that are concrete within the context of a particular project, and potentially grounded in a greater level of evidence (such as quantitative evidence or narrative descriptions of project links to outcomes), rather than responding on generalities that risk becoming significantly more heuristic.

For the purposes of this study, the suite of projects we chose to request that participants focus on in their responses was based on a series of pragmatic considerations and factors relevant to the pilot nature of this study. We were constrained by accessibility and interest of research partners to give their time to a pilot study of this kind. In order to test the indicators and their relevance, we also aimed to achieve some diversity among the representation of project stakeholders, and among different project purposes and pathways to impact, across land management, policy and community-

based outcomes. If these methods were applied in a context of program evaluation, more careful consideration would need to be given to sampling strategies, in order to do this kind of project-based assessment meaningfully and in a way that would allow data to be appropriately aggregated.

However, in practical terms, in terms of program-level evaluation, a project-by-project approach to evaluation alone runs the risk of requiring research users who are active across many projects to respond multiple times, particularly if a more comprehensive dataset is sought. Across coherent, network-driven research programs, such as NESP or Cooperative Research Centres, project-level approaches alone also risk missing key emergent properties that arise from the scale and higher-level collaborations that can be enabled through large-scale network-driven research programs. In applying an approach such as this to wider assessments (e.g. evaluations across institutions or research programs), consideration could also therefore be given to:
(i) Ensuring coordinated sampling strategies (e.g. using a cluster-based sampling strategy by research theme to disaggregate research users into coherent groups), to reduce burdensome resampling of individuals across multiple projects and allow for a more streamlined and coordinated approach to requesting participation from repeat 'super-users'.
(ii) Providing opportunities for a targeted group of research 'super-users' to respond to a series of program-level or other higher-level questions, to complement project-level responses.
(iii) Ensuring that open-ended questions on project-level assessments (e.g. in interviews) provide ample opportunities and invitations to draw out connections to wider processes, e.g. impacts that arise from the combined effect of multiple projects, further initiatives that have been made possible through the connections made within projects, or higher level strategies that emerge from the combined effects of coordinated research networks. This potentially has the added advantage of allowing program evaluations to draw connections between the operation of project-level research collaborations and wider processes of positive change.

The measures should, wherever possible, be selected at the outset of a project or program of research. This allows goals to be clarified and lays a stronger foundation of shared expectations. However, the measures selected in advance should not be rigidly enforced, but instead flexible and responsive to changes in the research program so that the suite of appropriate measures is suited to a changing context. While there are no prescriptive rules to selecting a suite of measures, Figure ii and Table i provide some general guidance.

Extending the framework to other research contexts

The approach to designing this framework was deliberately bottom up, built from a perspective of conservation research, and tested with projects with a deliberately contained range of approaches and purposes. The original conception was to focus exclusively on projects with a strong on-ground management component, but in the course of developing and testing the indicators and planning the case studies, it became clear that it was both feasible and more time-effective to expand to include projects that also have a primary focus on policy or community or other intended domains of impact. This relative coherence across the case study projects allowed for a deeper dive into the range of indicators and measures tested. Ideally, the extension of this approach to other research disciplines in environment (e.g. climate change or waste research) and to other disciplines (e.g. political science, anthropology, business management) would involve building bottom-up research valuation

frameworks within those disciplines, and then applying higher-level comparative analysis to draw together a synthesised suite of indicators and measures that could be applied across contexts. Such a process would represent an ambitious, multi-year program of research and development, but given the scale and influence of research valuation processes for motivating, funding and rewarding research, such a program is surely of great potential benefit.

In the absence of this more comprehensive approach, some general principles can be used to extend the application of this framework, at least to other environmental research contexts (and likely beyond). The logic model establishing the suite of indicators gives a place to start. Firstly, assuming that the same broad principles apply in these other fields around the importance and value of research user involvement in research, many of the inputs, processes and outputs indicators and measures would apply equally in this broader array of contexts. The outcomes and impacts measures used here were more tailored to conservation projects, but parallels could be readily sought in other domains. For example, while conservation research might focus on 'Adoption of more effective techniques for conservation practice (e.g. captive breeding, reintroduction into the wild)' in Subdomain 4.2 - Increased Awareness & Responses (AR 62), 'Reduction in threatening processes' under Subdomain 4.4 - On-ground Action (OGA 76), or 'A better informed society with greater social licence for conservation' under Domain 5 - Research Environmental Impacts (REI94), climate change research may want to adapt these indicators to focus on adopting effective techniques for climate modelling (Awareness and Response), improved adaptation management (On Ground Application), or a greater social licence for climate mitigation or adaptation action (Research Environmental Impact). Alternatively, for the same indicators, researchers in urban development might focus on adopting effective tools for assessing air quality (AR), or applying research to improve environmental benefits (greening, biodiversity) to a new development (OGA), or a greater social licence for integrated environmental urban planning (REI) as potential indicators of research value.

These indicators could then be adapted to measures relevant to the research field relatively readily. The measures for the domains where outcomes are expected were broadened in the survey context to be relatively generic (e.g. changes to management, policy, business decisions), and would likely change only slightly to reflect the language of the environmental research disciplines in question: 'adaptive management' may become 'adaptive planning'. 'reduction in threatening processes' may become 'reduction in polluting processes'. Likewise, the measures of impacts could be adapted to measure appropriate impacts from other domains, taking the higher-level aggregating principles of 'environmental benefits', 'improved governance and capacity building' and 'social, cultural and economic benefits' as a starting point. The most relevant domains of action and impact could be drawn from insights from research leaders and managers in each of these research areas.

6.3 Do project participants from different types of institutions value similar aspects of conservation research?

By and large, there was striking consistency between the assessments made by academics and non-academics on the fitness-for-purpose of the measures. Given this, the observable differences between these groups' assessments of fitness for purpose of the survey questions are striking.

Across the survey measures, non-academic respondents were substantially more likely to rate seven of the measures as more strongly fit for purpose than academics. With the exception of

“improvement in my group’s capacity to pursue locally-defined priorities, objectives, and goals”, which is inherently more relevant and tailored for research user agencies, these followed several recognisable patterns (ratings against these measures in brackets represent average responses of non-academics and academics, respectively).

Collaboration:

- Publications & other outputs produced in a collaborative manner (2.71 / 2.45)
- Links between research team & key stakeholders improved (2.50 / 2.00)

Although ratings on the suitability of measures for whether research needs of different stakeholders was elicited, non-academic participation in projects and whether stakeholder values and priorities informed how the project ran were comparable, non-academic project team members placed a higher premium on collaborative publication and the strengthening of links between research team members and key stakeholders. Despite a research context where skills and capacity for stakeholder-engaged research is highly valued and well-developed, this difference points to the persistence of some key cultural differences in expectations and priorities between academics and non-academics on the extent of collaboration and the importance of relationships.

Different knowledge systems and intercultural capacities:

- Methods drew on insights from different "knowledge systems" (2.72 / 2.09)
- Different knowledge systems informed the outputs of the project (2.57 / 2.09)
- Intercultural capacities of the research team have improved (2.50 / 1.56)

Not only were academics less likely to rate their projects strongly against these measures, they also consistently rated these as less fit for purpose for assessing their projects. This suggests a real divergence in the perception of the value of different knowledge systems and intercultural capacities for improving the effectiveness and value of research. Again, this is likely to be a reflection of institutional differences in priorities and expectations. Views on the importance of diverse knowledge systems and intercultural capacities for research are still evolving, but these differences suggest the evolution of these perceptions towards greater inclusivity and acceptance of intercultural work and different knowledge systems is perhaps less developed in academic contexts.

Threatening processes:

- Reduction in threatening processes (2.50 / 1.91)

It is unclear why this would be valued higher by non-academics than by academics, but it suggests that perhaps policy and management-focused practitioners, being closer to delivering on these outcomes, may be more likely to take a holistic view of managing species, ecosystems, special places and other values, and to understand reduction of threatening processes as key element of this. The interviews also suggest that some academics may be nervous about having their research judged against a yardstick that is largely beyond their direct control (e.g. implementing research findings to reduce threatening processes) and prone to time delays.

Academic respondents rated six of the survey measures as more strongly fit for purpose than non-academics. While one of these measures, that the project “Resulted in beneficial change in the NGO, not-for-profit policy sector (1.83 / 2.10)”, was less directly relevant to the largely government group of non-academic respondents, the focus of the remaining measures likewise reflects some of the priorities - and frustrations - of academic researchers.

Monitoring and data management:

- Improved monitoring for spp, ecosystems, heritage places, etc (2.29 / 2.55)
- Improved data management, reporting &/or analysis about spp, ecosysts, heritage places, etc (1.86 / 2.18)

Where academic respondents tended to rate other outcomes measures (such as adaptive management) as similar to or less important than non-academics, they were more likely to value improvements in monitoring and data management as relevant measures of research value. Being often closer to delivering on these outcomes, this suggests that academic respondents may be more likely to see their importance within the overall delivery of outcomes for species, ecosystems, and places.

Tangible improvements for species, ecosystems and places:

- increase in population trajectory for threatened or significant species (2.43 / 2.73)
- an improvement in condition of an ecosystem, wetland or marine environment (2.29 / 2.54)

It is unclear why academics should value these particular measures as relevant to their projects more highly than non-academic respondents. It may be that, for non-academic respondents, the focus on 'increased trajectory' and 'improved condition' is rated less fit for purpose as it is less realistic an impact than 'avoided loss of biodiversity', the other comparable impact measure. Academic respondents, being further removed from the site where actions are implemented, may be relatively less concerned about realism of a measure, and more focused on the ideal of improved environmental measures as an ultimate goal. This difference could be tested by testing a version of these two measures rephrased to address avoided loss rather than gain, e.g. 'avoided decrease in population trajectory' or 'avoided decline in improvement in condition of an ecosystem...'.

Public awareness:

- greater public awareness of the importance/challenges of conservation (1.86 / 2.60)

The idea of a need to raise awareness for conservation is something that pervades ecological research, reflecting a widespread perception that a significant factor behind the challenges of achieving greater action for conservation is that the public are not aware of the importance of conservation. It seems intuitive that public awareness of the challenges and importance of conservation should translate into changes in policy and investment, and it is thus striking that a measure focused on growing public awareness is considered substantially less fit for purpose for measuring research value by non-academics.

6.4 Insights from the research for building and supporting effective research collaborations

Our holistic framework for assessing research value throughout the complex cycle of research projects drew on insights from research and practice attesting to the importance of collaborative processes for optimising research value (e.g. Reed et al. 2014, Wyborn 2015). It was this fundamental insight that informed the logic model developed in Phase 1 of this project. In turn, the workshopping of these measures by experts, and their application and testing in real-world case studies, has allowed further insights on what it takes to build effective research partnerships to emerge.

Co-design

Co-design is a label commonly applied to research developed (and sometimes implemented) in collaboration with research users. Co-design is widely promoted as an effective principle for enhancing the value of research and its potential for uptake (van Kerkhoff and Lebel. 2015, Wyborn 2015, Nel et al. 2015, Parsons, Fisher and Nalal 2016, Muashekele et al. 2019, Jarvis et al. 2020). However, its meaning, application and importance vary widely with context and with the perspectives of those undertaking the work. In conservation research, the term is used loosely to apply to a wide range of activities, from processes that invite research users to set research questions to participatory research processes that draw together the expertise and guidance of research users and researchers throughout the life cycle of research (e.g. Miller and Wyborn 2020).

Participants at our expert workshop showed some reluctance around the idea that co-design of research was always appropriate and that benchmarks of good co-design practice were universally applicable. Rather, they felt evaluative benchmarks related to co-design should take into account whether design processes were "fit for purpose" for achieving the desired outcomes. Our Phase 2 results did highlight the complexity of applying a single set of measures to conservation research. However, the challenge of the "fit for purpose" idea is who decides what is "fit" and whose "purpose" is being served. Studies of deliberative processes involving academics and non-academics have shown the scientific voice can dominate (Davies 2013; Hansson and Polk 2018). This means that one of the key challenges for co-design is whether appropriate consideration and weight is being given to non-academic knowledge and values.

We tested measures of co-design principles in survey and research questions around drawing on different knowledge systems throughout research project stages, in particular, formulation of research questions and needs. At a higher level, if we want to evaluate decisions about what types of research a group like the TSR hub prioritises and how well that reflects the needs of all stakeholders, we may need to look at measures related to research governance arrangements.

Workshop participants also noted that a challenge for co-design is that stakeholders do not always know what they want or have the time to participate. These issues need to be decoupled in designing and evaluating conservation research. O'Connor et al (2019) suggest supporting communicative competence or the capacity of non-academics to make and question claims. So instead of asking Traditional Owners or park managers what they want from research, a broader conversation may be more appropriate.This could be about the management needs, cultural imperatives, associated knowledge gaps and the ways research may be able to fill these. In terms of time, O'Connor et al (2019) suggest support for participation be evaluated "according to how well it enables broad participation across stakeholder organisations and its flexibility to respond to participant needs" i.e. projects should be evaluated according to how they were designed to support collaboration not whether that support was actually utilised.

Each step of the research cycle can create value

Research is currently valued primarily based on its "impact". The rating by so many of the Phase 2 participants of most of the Inputs and Processes measures as being highly fit for purpose for assessing the value of their research projects supports the validity of our holistic evaluation framework. The value of conservation research is not only in whether it delivers measurable impacts or even observable changes to on-ground management or policy, it also lies in what emerges as a

result of how the research was undertaken. For example, the expression of stakeholder values and priorities that informed how the project ran, the collaborative and inclusive way research questions are developed, the building of trust, the instilling of skills, the production of really tangible, usable products, and many other matters. And, despite the relatively lower rating granted to academic publications in the Phase 1 study, research value was certainly seen also to lie in the generation of research findings that are legitimate and of high quality, and in such things as effective data collection and analysis. And as one survey respondent noted, this assessment of the quality of data analysis is not always about "whiz-bang" approaches, but "*that a research project effectively extracted all the best juice from the fruit, whether that be by a simple or complex analysis*."

The language of and concepts behind the logic-model type of approach is not always familiar to those undertaking effective research collaborations. Phase 2 results demonstrate that terms like "knowledge systems" can create confusion and so more self-explanatory would likely be more effective in gaining meaningful responses e.g. "Did people with different backgrounds and from different types of organisations contribute to research outputs?"

Thus this framework, particularly the indicators and measures rated more highly by participants across both surveys and workshops, can be used to support how projects are established, implemented, communicated and applied, so that greater value from the research might be achieved, as well as measured. This can augment other established guidelines that have been developed to facilitate effective collaborative work, such as the Conservation Measures Partnership (Conservation Measures Partnership 2018).

Understanding and reflecting upon the importance of relationships and trust in effective research collaboration

Collaborative research processes are generally accepted as an effective way to foster research uptake (see e.g. Reed et al 2014, van Kerkhoff and Lebel 2015, van der Hel 2016). While "relationships" between collaborators have been recognised as an element in the success of transdisciplinary research, our Phase 2 results illustrate the very tangible benefits of respectful relationships in research teams. Good relationships based on trust mean people work together more effectively, exchanging information, giving frank feedback, and being responsive. However trust needs to be matched with critical analysis and feedback among partners. "While trust can reduce information and processing costs between parties, trust can also lead to 'blind faith' or a lack of vigilance between parties, potentially causing cognitive lock-in, favouritism, uncritical commitment to a suboptimal course of action, or limitation of the integration of diverse ideas" (Lacey et al. 2018, 24)

Diverse knowledge systems and intercultural capacities

Results from the survey measures on integrating different knowledge systems through processes (RM16, survey processes) and outputs (AO30, survey outputs), and whether intercultural capacities were improved through the project (CSE66, survey outcomes) were rated lower on average by academic respondents than by non-academic respondents, and less fit for purpose for assessing research value.
Low sample size, within-project variation and the ambiguity in the wording of these questions mean there is ambiguity here. However, it is possible that these represent differences in perceptions between project team members based primarily in research institutions and those based in government or management agencies about key measures of collaboration.

There are a range of possible interpretations of this pattern:

(1) those in research institutions may be more likely to privilege and recognise western science approaches and may be less likely to reflect in their responses to these questions how the knowledge of non-Indigenous others, such as their practitioner partners and collaborators, might be in itself an example of specialist different knowledge informing the research. This interpretation of the findings in terms of the relatively higher valuing of western science and academic approaches is in line with the fitness for purpose data showing that academic respondents were more likely to rate all three measures as having a relatively lower fitness for purpose than non-academic respondents. While interview responses did not comment widely on Indigenous knowledge and perspectives informing the research, a number of respondents from outside of research institutions did discuss the integration of practitioner knowledge and experience into the research.
(2) those in policy and practitioner agencies may have a more direct relationship with diverse locally-based perspectives, knowledge systems and people, and thus are more likely to recognise how these have shaped the research and the capacities of research teams.
(3) academics may be less likely to experience opportunities for or encouragement and support for intercultural relationships and the benefits of different knowledge systems for the development of their work and themselves.

While it is hard to draw definitive conclusions from this, this trend toward academic researchers having less cognisance of the input of different knowledge systems and opportunities for intercultural engagement in the research, and the lower valuing of these as fit for purpose measures, reflects a long-standing institutional perception (now shifting) that privileges the perspectives of western science and research. This points to the importance of having research projects that work closely with on-ground partners, who may be more likely to bring their own specialist knowledge and diverse perspectives, or to have direct relationships with others who do. This potentially represents an important value that non-academic practitioners sometimes bring to research partnerships. Training for greater awareness among academic researchers of these dynamics, and greater incentives, opportunities and support in research institutions for intercultural engagement and for learning from and integrating diverse knowledge systems, could potentially help develop the opportunities this presents not just for research projects, but for research teams to learn through these experiences.

6.5 Limitations of the data and future research priorities

As noted at the outset, one of the key limitations of this study is that it was not possible to realise the early intention to include Indigenous case studies to begin to understand what an Indigenous-led approach to assessing research value could involve. This is discussed further in the following section, along with potential next steps for achieving this.

In addition to this limitation, this is in some sense a preliminary study in other ways. The relatively small sample size for the survey and the limited range of respondents with respect to some sectors (including very low participation from research team members based in non-government organisations) constrains how definitively we can draw conclusions from this study. It is unlikely our sample reached saturation in terms of the themes canvased in interviews, as each interview delivered new and different insights into the research processes. Similarly, the certainty of statistical

comparisons in the survey data is limited by the relatively small sample size overall and the low number of projects with more than one respondent, as well as confounding factors such as the ambiguous use of language in some of the questions. It is worth noting that the size of the sample within projects is necessarily always constrained by the size and nature of the project undertaken, as projects often have very small teams of people with privileged knowledge of the project's undertakings. However, greater significance in the comparative measures of fitness for purpose, variability and concordance could be achieved by adding more case study projects to the analysis.

Likewise, owing to the short window available for case study development and testing, we were not able to fully test the quantitative, documentary and community insight measures nor options for baseline setting and context setting measures. These measures have been assessed for the feasibility of applying them within a program fully resourced for gathering centralised data, such as the NESP Threatened Species Recovery Hub. But a fuller assessment of the stability and application of quantitative, documentary and community insight measures is beyond the scope of this project. This is particularly important to note with respect to quantitative and documentary measures, in the light of earlier research from the UK Research Excellence Framework demonstrating the variability in the ways different institutions assess and report against quantitative measures, which constrains any potential for comparison or aggregation (Wilsdon et al 2015). Further testing of these measures is needed to understand where there may be potential for comparison or aggregation, and any constraints on assuming that quantitative and countable figures can be compared or aggregated. In the interim, it is important that any application of countable quantitative and documentary measures be done with care where comparison or aggregation is required, and that careful explanations be applied in these cases, along with in-program measures to test stability of the measures used.

It has also not been possible to test the effectiveness of all the measures over the long term, nor whether and how the results may vary over time. For example, it is unclear whether some of the more variable measures, such as outcomes and impacts measures, would develop a greater degree of convergence and a lower variability if measured at a later point in the research cycle. In particular, some projects were just starting to produce outputs at the time of the assessment, so it proved difficult to evaluate outcomes and impact measures. It is likely we would get quite different results if this were repeated in two or three years' time, when the outcome and impact indicators which would have had more time to manifest, but we are unable to be certain how these results would differ, and whether this would change our assessment of how different measures operate, their applicability, and how different team members assess their projects against them.

Finally, we have been unable in this particular study to test the value of the measures for assessing longevity of outcomes and impacts: A universal theme in research valuation work is the longevity of impact (see O'Connor et al. 2019), and whether it dissipates without continuing effort. As one respondent suggests, *"a project like this,..., it's not a fire and forget, it's something you've got to keep nurturing*" TSR02. If a project is assessed to have improved community awareness, it is not clear whether the measures can pick up the degree to which this will last, especially if research or management actions cease.

Future research priorities

Research to further test differences in priorities, interpretations and appropriate language around key, but potentially ambiguous, concepts in the measures, particularly around 'different knowledge systems' and 'intercultural capacities', but also around the extent of collaboration and the strengthening of ties. Measures such as these represent one of the most important areas of difference between research team members inside of and beyond research teams, both in terms of how each group ranks them as fit for purpose, and in terms of how each group rates their projects against them. However, the ambiguity in the language and modest sample size here confounds any systematic, statistically significant differences between the groups. Further research to unpack these differences and to understand whether these indeed represent real differences in perspective between these two groups would help shed light on where real cultural differences persist between research institutions and other agencies. It would also potentially provide options for language around these measures that is clear and avoids the ambiguities apparent in the measures tested here.

Greater sample size and further testing of quantitative, documentary and community insight measures: Many of the findings on the measures tested in this framework are provisional, limited by the small sample size, particularly in the case study survey, and by the inability within the scope of the project to test the application of quantitative, documentary and community insight measures in a range of circumstances. Further research applying the same analysis to these measures and to a much wider numbers of projects would help unpack which of these measures are inherently variable or systematically divergent between members of research teams, projects and programs, and/or which tend to stabilise or converge over much larger sample sizes. Longer-term research could also establish and assess the application of baseline and framing measures, and any changes in the stability of measures and their responses over longer life cycles of research projects.

Testing the extension of the framework to other environmental research contexts, and comparing with other ground-up methods for assessing research value in other disciplines, to identify whether overarching measures and approaches can be used as framing measures for research valuation across all disciplines.

6.6 Towards an Indigenous-led approach to assessing research value

Any assessment of the value of research involving Indigenous communities and organisations needs to be underpinned by criteria and methods generated by the communities themselves. The original aim of the project was to create an approach to assessing impact working, from the ground up, with Indigenous partners who were interested in offering case studies. The intention was to be able to make some preliminary suggestions for assessing research impact in Indigenous contexts, building on the existing literature, and to refine and test these approaches in collaboration with Indigenous communities and organisations with an active stake in case study conservation research projects, as research users, project partners, participants, and knowledge-holders. Through this, it was hoped that we could begin to develop an understanding of some of the ways Indigenous-led approaches to research valuation take shape, and from there to look for commonalities and differences with the framework developed here, in order to understand how the two articulate or do not. The intention was to draw on insights from these case studies undertaken in partnership with Indigenous

communities, to consider where and how the wider framework needs to be modified, or replaced, in contexts where Indigenous communities or people have an interest in the research (whether or not they have been actively involved).

Within the timeframes available, particularly under the constrained conditions of COVID-19, the project was not able to achieve this. This is a significant gap in our understanding of how to support, assess, and reward research value in culturally appropriate and sensitive ways. It places limitations on the application of this framework as it exists - in particular recognising that the methods for applying the framework, the emphasis and weight given to some measures over others, and the ways of thinking and speaking about research value are not likely to be culturally appropriate or broadly applicable to contexts where Indigenous communities, organisations or individuals have an active part in the project ecosystem. In particular, the current approach, while intended as a holistic framework, has been built on western ways of breaking apart and then reconstructing the problem, on western categories of thinking (e.g. separating environmental, social and cultural impact) and on approaches such as surveys that segment research value through lists of research dimensions, that are unlikely to connect with culturally meaningful ways of communicating and approaching issues. The whole way of approaching the question and the measures provisionally assessed in this report are therefore not always likely to be the most meaningful or important measures for projects or research contexts where Indigenous communities, organisations or individuals have an active part. In these cases, it is critical that assessments of research value, and approaches to assessing value, should continue to be guided, case-by-case, by the priorities and values of these Indigenous partners to and participants in the research.

More broadly, the ongoing application of western approaches to thinking about research value is likely to continue to disadvantage Indigenous researchers and communities, even when applied more generally. Given that a great deal of environmental research takes place on Country, and all environmental research in Australia potentially impinges on Country, cultural values and heritage, careful and judicious consideration should be given to how the application of the framework presented here takes into account and is sensitive to these considerations. While there were Indigenous people and non-Indigenous participants with experience working in partnership with Indigenous communities, involved in each stage of external surveys, case studies and the expert workshop, lending what we hope are important insights and caveats to the approaches developed here, including some of those discussed throughout the report, more work is needed to ensure that any application of research valuation assessments and measures does not further perpetuate the structural marginalisation that has long characterised western research systems. On the other hand, achieving and rewarding value in research in ways that are culturally sensitive and appropriate is potentially of great benefit, both in terms of emphasising asking what research has delivered direct outcomes sought by Indigenous communities, people and Country, and whether research has helped shape broader practice, policy and investment in ways that support cultural outcomes.

There is therefore an urgent need to address the lack of attention to culturally-relevant processes for understanding research value for both Indigenous researchers and Indigenous communities. Resourcing and attention to Indigenous-led processes are needed to develop meaningful and appropriate approaches to valuing research, and assessing research value. Research in this field should be built from the ground-up, guided by Indigenous practitioners and researchers, and working

with the values and priorities of the communities involved, to shape up Indigenous-led research valuation approaches. Recognising that doing this well will potentially involve a significant investment of time, resources and careful processes, steps to start along this pathway could include:

- A broad review of existing Indigenous-led approaches to research valuation (e.g. Tsey et al 2016, 2019) and of insights on research valuation approaches suggested in the large and growing body of resources and protocols guiding culturally appropriate approaches to research practice, both in conservation and beyond (e.g. Moggridge 2020, Woodward et al 2020, AIATSIS 2012).
- Guidance from an expert group and workshops involving leading and emerging Indigenous researchers, practitioners and partners in conservation fields in Australia, to discuss whether common principles for assessing research value might be appropriate beyond the broad foundational principle of all research and research valuation needing to be guided by the priorities of Indigenous partners and parties to the research, and especially by the Traditional Owners on whose Country the research is taking place; and to the extent appropriate, to shape up a potential set of seeding questions and concepts that could form a starting point for an Indigenous-led approach to research valuation, and identify potential diverse case studies for developing and testing these seeding questions and concepts.
- Growing, testing and broadening these approaches in case studies with Indigenous communities or organisations that are active partners to research case studies, and drawing out common themes or priorities in these approaches.
- Bringing the approaches developed to 'speak back' to the framework presented here, and modifying the framework in response to allow it to be more broadly applicable to conservation research in Australia and identify and remove cases where this framework is likely to create perverse incentives or further marginalise Indigenous partners and researchers.

Any Indigenous-led processes in this space should be used as a touchstone for any broader work undertaken to test and extend this framework to disciplines other than conservation, many of which like conservation potentially or in practice have a direct bearing on Country, people and culture.

Conclusion

This is the first attempt of its kind to build and test a multimodal, integrated, qualitative and quantitative framework for assessing value in environmental research across the life course of projects and programs. The framework is designed to help guide research planning and valuation in a flexible way, and to be tailored to suit the needs and resources available to a project, program or institution. The framework has been built by focusing on conservation research projects, which has allowed us to find and integrate common indicators and measures across diverse projects. However, there are many readily identifiable parallels that could be applied to other dimensions of environmental research, and many measures likely to be readily applicable to other disciplines, particularly in the domains of inputs, processes and outputs.This is a key area for next steps in the research.

Because the framework draws on measures that are meaningful across the research cycle, elements of the valuation approach can be applied while research projects are still underway, providing an adaptable system that can help research teams, funding bodies and managers assess and potentially adjust the implementation of projects or programs while they are still underway. It can also form the basis of longitudinal and progressive analyses of research value across programs. The possibility for creating a framework that could work in this way was founded on the insight that mechanisms identified in the literature and practice of effective research engagement and implementation - such as building relationships, trust, engagement, collaborative and co-designed processes, knowledge adaptation, and adoption - could be set alongside more widely accepted measures of outcomes and impacts, to build a comprehensive picture of research value across the different program logic domains of a research project or program. This premise was developed into a draft suite of indicators and measures, and then tested with research evaluation specialists, researchers and research users with significant experience in creating and assessing valuable, engaged and impactful conservation research. The resulting assessments of importance, feasibility, and fitness for purpose, and the practical insights gained from testing the measures across nine case study projects, confirms the broad success of this approach. Results from this work allow users of the framework to take more considered and appropriate approaches to the robustness of the measures when applying these in practice. They have also helped gather new insights into these foundational elements for achieving research value: of the range of considerations at play in building engaged and co-designed research processes and relationships, of what it means to build trust, and how trust works to deliver value in practice.

However, insofar as this represents the first attempt at a comprehensive framework for assessing conservation research value, it should be seen as a first step along a broader process of extending, developing and delivering a broad array of comprehensive and fit for purpose measures of research value. The case studies here, while providing a range of useful insights, represent a small, preliminary study sample as a proof of concept for the framework. It is appropriate to extend and test the framework on a greater number of projects and a wider array of project contexts, to refine the assessments of the specific measures and the adequacy of the approach as a whole.

Most importantly, within the timeframes available, the project was not able to develop and test meaningful, Indigenous-led approaches to assessing research value, nor did it work with Indigenous

communities, organisations and people involved in research projects and programs to create the approaches and measures that are most meaningful and important to those communities, organisations or individuals, both as research users, project partners and participants, and knowledge-holders. This is a significant gap in assessing research value to address as a priority. It places limitations on the application of this framework, particularly in where Indigenous communities, researchers and organisations are directly involved, but also in the more broader sense that most environmental research in Australia takes place on Indigenous Country and therefore potentially or actually impinges on Country cultural values and heritage. Careful and judicious consideration should be given to how the framework presented here takes into account and is sensitive to these considerations when applied to any conservation research context. Further research should be supported and undertaken to address these shortfalls, and form a basis for 'speaking back' to the framework developed here, to test whether this framework or any of these measures are perpetuating structural disadvantages for Indigenous researchers or research partners.

Recommendations

The framework is designed to be drawn on by a wide range of users for an array of purposes. The measures tested here will be useful and feasible for different circumstances, and to meet the needs of a range of valuation exercises, from internal reviews of research projects or networks, through to independent evaluations during and after research programs. They have been designed to enable those seeking to understand research value to identify and tailor the best suite of measures to draw together a picture of research value for particular contexts. Using these measures appropriately, it is possible to get a picture of the value of research through the research cycle, including to assess the value, or potential value, of research while it is still underway.

Insights for assessing the value of research

1. **Focus on research value rather than impact**

   While colloquially 'research impact' has a wide range of meanings, its technical meaning within many logic frameworks is on those changes that extend beyond the immediate control of research teams. Thus a focus on 'research impact' often lends itself to narrow interpretations (e.g. of measurable environmental benefit), without focusing on the complex pathways through which change can be measured and is usually achieved. A broader focus on 'research value' allows for significantly greater insights, can be applied much more widely across research cycles, and has a better chance of incentivising the diverse kinds of practices and approaches that most commonly lead to the greatest benefits. It also recognises that value resides in many things, including in relationships, trust, effective research approaches and greater awareness of problems and challenges, as well as in long-term environmental, social, cultural and economic impacts.

2. **Approach research valuation critically and carefully when developing measures, valuation methods and approaches**

   In particular, where possible, methods for research valuation should be tested with research users, research teams and, where relevant, research management and engagement teams, for:
   a. Their importance, usefulness and fitness for purpose for reflecting research value;
   b. Their feasibility (including resource-intensiveness/cost-effectiveness) for the circumstances in which they are to be applied;
   c. Their effectiveness in delivering meaningful, useful insights;
   d. Their reliability and consistency, how they should be applied, and any cautions or caveats around their application.

3. **Use caution in applying the framework to any circumstances in which Indigenous people are or should be considered critical partners in research projects**

   Independent Indigenous-led processes are needed to develop culturally appropriate approaches to assessing the value of research undertaken by Indigenous researchers and/or undertaken with and for Indigenous communities and organisations. This is a significant undertaking and will require considerable time, care and resources to achieve in a culturally appropriate manner. In

the interim, application of any framework for research valuation needs to take care not to create perverse outcomes or further exacerbate Indigenous dispossession and disenfranchisement from knowledge-creation and research activities and funding.

a. In Indigenous-led research and research undertaken with or for Indigenous communities and organisations, it is likely to be more appropriate to undertake specific, Indigenous-led processes for developing and designing research valuation measures, depending on the resourcing, availability and willingness for Indigenous participants.
b. In the much wider range of contexts for conservation research - for example, in any general research not specifically tailored for Indigenous communities, but possibly affecting people, Country or cultural values - consideration should be given to whether Indigenous-led or engaged processes of research valuation are appropriate to those circumstances (possibly in parallel with the application of this framework) depending on the resourcing, availability and willingness of Indigenous participants to be involved in the valuation exercise. For any part of the valuation of these projects or programs that draws on this framework, care should be taken to ensure the selection of indicators does not impact negatively or create perverse outcomes for Indigenous communities or researchers.

Insights for applying this research valuation framework

4. Use the framework and measures to assess conservation research projects and programs, across their life cycle and across a diverse range of measures

The tools developed here and this framework are designed to be adaptable to the needs of a wide range of sectors - research funding bodies, research institutions, government agencies, NGOs and others - and by researchers and research teams themselves, to better assess and inform the effectiveness of research, support and improve research practice to deliver greater value, and create more appropriate systems for rewarding and incentivising research that achieves value. This research framework can be drawn on to assess research value across a diverse range of measures, to give a fuller and more comprehensive picture, and to enable valuation throughout a research program and beyond. A good starting place is to identify conceptual and implementation considerations at the start of the project, as noted in Figure i.

5. Draw on insights from the framework and tools such as the glossary to design a suite of measures suited to context, research phase and outcomes sought.

The 96 indicators developed in this study, and the related measures developed and tested against a subset of these indicators, can be used as a guide to helping standardise, assess, compare and articular research value in conservation. They can also be applied, or adapted for application, to wider environmental research fields. Drawing on a suite of these tested measures to guide research assessments can help introduce standardised assessments across a wide range of projects and programs. To get the fullest picture of research value, we recommend drawing on an integrated suite of quantitative, survey and narrative measures from across the research cycle and the range of domains of activity and impact under consideration. Application of the framework should:

a. Draw on a suite of measures tailored to the needs of the valuation exercise.

b. Ideally integrate narrative, survey and quantitative measures to develop a nuanced and more comprehensive picture of research value.
c. Focus on contribution rather than attribution. Take care not to overplay impact of particular discrete projects or activities, which would likely downplay changes achieved over years and the precursors to step changes seen at particular points in time.
d. Give consideration to:
    i. Their importance, usefulness and fitness for purpose for reflecting the value of particular research projects, programs and activities in the particular context.
    ii. Their feasibility (including resource-intensiveness/cost-effectiveness) for the circumstances in which they are to be applied.
    iii. Their effectiveness in delivering meaningful, consistent and useful insights, drawing on insights from this study for each of the measures.
    iv. Their reliability and consistency, how they should be applied, and any cautions or caveats around their application.
e. Give specific consideration to how these measures should be applied, collected and analysed. Most notably:
    i. Tailor valuation approaches to the scope of research projects and programs. Conservation research comes in many forms and not all the indicators and measures developed are suitable to all projects. In interviews, targeted early questions around objectives and the range of project participants involved can set the direction for subsequent questions. In surveys, initial questions on project objectives, the respondents' role in the project and who the project participants were could inform question selection as well as helping guide analysis of the data.
    ii. Collect valuation data from a range of participants with diverse project roles in most cases. This will provide a richer and more comprehensive picture of research impact than simply asking researchers what they believe (although some measures yield reasonably consistent responses from across project teams). Care needs to be taken however to tailor analysis to account for different participant's knowledge of a given research project e.g. those not involved in data analysis may not be able to accurately describe issues with data. Detailed assessments of each measure in the research framework provide an indication of appropriate ways to draw on insights from different members of research teams.
    iii. Apply caution in assessing, aggregating and comparing a range of survey and quantitative measures. Do not assume that numerical values can always be aggregated or compared; ensure any measures for aggregation or comparison use similar methods and are applied in similar ways. Take care in comparing or aggregating survey measures (e.g. likert-scale) that these measures are not prone to high variability or poor within-project concordance.

6. Give attention to context and baselines.

Baseline data and contextualising information is particularly important for any valuation exercise seeking to compare across projects or assess return on investment. Where baseline data have not been collected, contextualising information can help inform and interpret retrospective assessments of value provided by diverse members of research teams later in the research process.

a. Where possible, select a provisional suite of indicators and measures at the start of projects that will be used as the basis for baseline assessments and ongoing valuation. These should be applied flexibly and tailored as research projects or programs evolve, but where possible should provide a coherent core around which ongoing valuation can evolve. To strengthen the potential for these measures to guide and support the achievement and measurement of research value, selection of this preliminary suite of indicators would ideally be developed collaboratively between research program managers, research teams (or representatives) and other stakeholders (particularly research users).
b. Where possible, set baselines at the commencement of research projects or programs. Where this has not happened, take additional steps to ensure a range of perspectives and insights are provided (e.g. sampling strategies), and triangulate measures (e.g. likert-scale survey measures from research teams) with a broader evidence-base (e.g. narrative, independent survey, documentary approaches and (where possible) quantitative data).
c. Collect data on context for research programs (e.g. duration of long-standing research collaborations) to contextualise measures of research value.

7. Where needed, draw on indicators from and beyond the wider Phase 1 suite as appropriate to reflect research value in a wide range of contexts.

The framework is designed to be enabling rather than prescriptive. In particular, the framework tested a subset of indicators from the Phase 1 study. New measures could involve:

a. Developing other measures against relevant indicators not tested from the Phase1 study, and/or against other indicators identified as valuable by program and research managers, research users and research teams, if these are appropriate to particular research contexts, projects or programs.
b. Where possible, test and develop any new measures with a range of partners to research projects or programs, before applying these widely.

Insights for research practice and resourcing

8. Research funding bodies as well as research teams and partners: Invest time and resources in developing appropriate processes and relationships, including dedicated facilitation, brokering and valuation roles.

This includes setting aside and specifically earmarking resources to undertake this work:

a. Adequately funding and rewarding researchers to invest in collaborative and engaged development and implementation of research.
b. Adequately resourcing partners to participate in research development and implementation as needed and appropriate to the circumstances. In particular, set aside

funding to resource participation for under-resourced research users, such as Indigenous communities.

c. Resourcing specialist and dedicated teams to build capacity for engaged research and to provide support for gathering, testing and assessing information to support research valuation.
d. Resourcing assessments of research value both after and at appropriate points within research cycles, to allow for adaptation, feedback and incentives for generating value in research.

9. Co-develop research through early conversations and broad engagement of end-users and stakeholders

Informing policies and on-ground management of conservation research are central to protecting and improving threatened species survival. To ensure that research findings are appropriate and will be taken up by policy makers and conservation managers, early engagement of a broad set of stakeholders and end-users are needed to co-develop research with relevant findings.

10. Researchers, research users and research program managers: recognise and draw insights from the complexity and nuance behind the common wisdom of what builds effective research relationships.

This includes careful deliberation on the meaning and benefits of trusting relationships and which principles of collaborative research development and co-design are appropriate and applicable to context.

Further research

11. Undertake further research to improve research valuation methods across the research sector.

Our project has identified a range of areas for further research. These include:

a. The need for an Indigenous-led strategy for developing culturally appropriate and meaningful approaches to research valuation.
b. Testing of quantitative, documentary and community insight measures and further testing of the survey measures with a wider range of projects to improve statistical assessment of reliability, stability and convergence of these measures.
c. Specific research to test interpretations and meanings of important but potentially ambiguous concepts such as drawing on 'different knowledge systems', and to develop appropriate language and methods for asking these questions in research valuation contexts.
d. Testing the extension of the framework to other environmental research contexts, and comparing with other ground-up methods for assessing research value in other disciplines, to identify whether overarching measures and approaches can be used as framing measures for research valuation across all disciplines.

Full framework of indicators, suggested measures and considerations on their application

Situating and branching questions

Table F1. Situating and branching measures and questions

Situating measure/question	Type of measures	Assessment	Comments
Can you please describe the nature of your involvement in project x?	Interview: situating	Good	
Please indicate which project you will be referencing in this survey	Survey: situating	Good	Our case study participants were prompted with the project name, but participants could be invited to nominate.
For how long would you estimate you have been involved in this TSR Hub project? Select one of the following options: less than 1 year, 1-3 years, >3 years.	Survey: situating	Good: minor modification	Could apply shorter or longer timeframes as appropriate.
What was the geographic scope of the project at its conception and in its application	Situating	Not tested	
Was there any expansion of scope made possible because of the value or success of the project	Situating	Not tested	
What kinds of data were sought, and what third party data did the project depend on	Situating	Not tested	
How was the individual project situated within a wider body of research (e.g. was this is the first significant research in this area, or a continuation of 10 years work)	Situating	Not tested	
How was this project situated within a wider set of relationships (e.g. whether this is the first research for this group of research users in this partnership, or this is part of an ongoing program of work between these research teams and research users in this region or field).	Situating	Not tested	
Was there turnover of stakeholders or research team and did this affect project delivery or value	Situating	Not tested	

Table F2. Situating and branching measures and questions

Situating and branching measure/question	Type of measures	Assessmt	Comments
To help get an idea of who was involved in the project, I'd like to capture each organisation or group involved, starting with your own. Can I get you to draw lines to link which ones interacted with each other? On each line can you estimate how strong the links between the organisations were at the beginning vs now?	Interview: situating and branching	Good	Analysed for what this says about project; part of this question could also be used as a situating measure for analysis. This question was asked in the interview, with a suggested scale of 0 = no link, 1 is some link, 2 is strong link. If there were many entities involved, emphasis was encouraged on key organisations.
Which best describes the type of institution you were primarily affiliated with in the TSR Hub project? (please select 1) Research, Government, Non-government organisation (NGO), Indigenous, Other (please specify)	Survey: situating and branching	Good	

Table F2 cont. Situating and branching measures and questions

Situating and branching measure/question	Type of measures	Assessmt	Comments
Please describe the nature of your contributions to the project (Choose as many as appropriate): ● Designing the project (e.g. defining initial research questions, what to measure where, etc) ● Providing advice (e.g. on who to involve, on outputs, on methods) ● Collecting data/information ● Analysing data ● Communicating about the project (e.g. discussing with community/others in your organisation, etc) ● Communication of outputs (e.g. writing reports, making videos, etc) ● Using the outputs	Survey: situating and branching	Good	Could be used more fully as an underpinning measure for analysis; used as a situating question in this study but could also be applied as a branching question to determine which further questions are put - for example, only asking those close to the data analysis about the benefits of the methods used. Providing additional project role options, including 'building partnerships and collaboration', would allow more sophisticated branching or conditioning analysis, particularly around questions such as the extent to which research users were involved in the research.
Were groups or organisations outside research institutions (like universities, CSIRO) directly involved in the project? Participants could respond "yes" or "no". If "yes" was selected a supplementary question was asked to prompt for further information on the other groups or organisations involved in the project.	Survey: situating and branching	Good	
What groups or organisations (other than research institutions) were directly involved in the project? ● This question allowed for multiple answers to be selected from the following options: Indigenous organisations (e.g. Land Councils, Aboriginal corporations, ranger groups), Individual land owners,Government agencies, Conservation non-government organisations (e.g. Bush Heritage, Australian Wildlife Conservancy), Community groups, Businesses	Survey: situating and branching	Good	
What fields(s) of application was/were originally intended for the project? (e.g. did it aim to inform policy, management, planning, community participation, education, or other)?	Situating and branching	Not tested	Could particularly be used to determine the range and order of outcomes and impacts questions put to respondents. However, the branching should not preclude any unexpected benefits that might arise from the project. Thus a question logic (in a survey context), could foreground the likely most relevant questions (e.g. around on-ground management), then add a further branching question to ask 'were there any additional benefits from the project?' prompting with those areas not flagged as the primary purpose of the project (e.g. policy, planning, community, business).
What areas of impact formed the focus of this study (e.g. social, cultural, economic or environmental)?	Situating and branching	Not tested	The branching should not preclude any other areas that might arise from the project, but could be used to prioritise which impact questions are put first.
How long have you been involved with the project? During what phases in the research cycle were you involved (conception, identification of the question, planning, design, implementation (including as a partner or stakeholder), publication, dissemination, uptake)?	Situating and branching	Not tested	This should be used as a measure to condition interpretation of the responses, and could also function as a branching question for all of the later questions.

Inputs measures

Table F3. Inputs measures and questions: **research needs identified**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Research needs, gaps, unanswered questions, new areas for research identified	How were research needs identified in the project?	Interview: narrative	Tested	Use with filter	Not all project participants will know this, particularly if they were not around at project conception. A prompt may be needed to elicit whether or not different types of research partner contributed	RDR1
	Research needs or new areas of research were identified through the project	Survey: likert agreement	Tested	Good: survey across team	Academics appear to assess higher; include partners outside research institutions to assess breadth of recognition of research needs identified.	
	Documentary evidence of research gaps that have been identified and prioritised, as well as those met through the project	Documentary	Assessed	Use with caution	Best used very selectively if at all. Difficult to assess concretely, but could possibly be used where active processes for collectively identifying and assessing research gaps have been undertaken as part of projects or programs. Critical assessment warranted of whether gaps analysis prioritises effectively and identifies the most important issues.	

Table F4. Inputs measures and questions: **research users and stakeholders active involvement in research**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Need for research has been adequately established - and all perspectives gained on whether research is valuable (e.g. particular end-user agencies) Active participation of stakeholders and end-users in research New understanding of values and priorities captured (e.g. organizational agencies, citizen science, Indigenous priorities) All participants and stakeholders contribute to designing or providing input to research questions	How did non-academic (non-research)/your and other non-academic groups participate in the project? P1 Articulated the need for research? P2 Their values & priorities informed how the project was run? P3 How actively did they participate?	Interview: narrative	Tested	Use after defining "participation"	*Need to state what defines "participation": e.g. People who helped design the project, provided advice, collected data, analysed data, developed outputs, promoted outputs to end-users *Need to ask at least 1 academic & 1 non-academic participant	RDR4 RDR5 RDRX (new indicator identified at expert workshop) **RMC22 (process)**

Table F4 cont. Inputs measures and questions: **research users and stakeholders active involvement in research**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Need for research has been adequately established - and all perspectives gained on whether research is valuable (e.g. particular end-user agencies) Active participation of stakeholders and end-users in research New understanding of values and priorities captured (e.g. organizational agencies, citizen science, Indigenous priorities)	Network map: interviewees asked to identify the different groups involved in research and describe the links between them and how relationships may have changed over time Develop a map of organisations and the relationships between them (at the beginning vs end of project	Interview: interactive mapping exercise	Tested	Good (need >1 respondent)	Will get complementary responses from people with different roles in a project The network mapping exercise should be conducted early in the interview to identify which participating organisations the interviewee is aware of. Doing the mapping visually so the interviewee can see was a good prompt (eg whiteboard function in zoom).	RDR4 RDR5 RDRX **RMC22 (process)** **CSE65 (outcome)**
Need for research has been adequately established - and all perspectives gained on whether research is valuable (e.g. particular end-user agencies)	The value of the research to different stakeholders (particularly end-user agencies) was elicited	Survey: likert agreement	Tested	Good	Rated high fit-for purpose. Academic self-report consistent with non-academic assessment.	RDRX
	General inputs measures · Number, diversity, geographic and sectoral coverage of stakeholder organisations involved in research planning, relative to project impact **Specific community-based assessments** Focus group, interviews, surveys, participatory evaluations to assess whether need and value of research has been established.	Quantitative / independent survey	Assessed	Possibly use, preliminary filter for relevance, value add	· General measures potentially more widely feasible, reported from project teams, but do not indicate how effectively research user values and perspectives were integrated. · Community-based assessments likely more feasible in single-project or targeted evaluations. Could complement or provide independent support to project team (academic and non-academic) survey measures, particularly in contexts where broader diversity of perspectives gained is essential. · Assessment should both assess and allow for churn among stakeholders where possible.	

Table F4 cont. Inputs measures and questions: **research users and stakeholders active involvement in research**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Active participation of stakeholders and end-users in research	Were groups or organisations outside research institutions (like universities, CSIRO) directly involved in the project? [if yes to above] What groups or organisations (other than research institutions) were directly involved in the project?	Survey: yes/no; multiple choice	Tested	Good	Academics may recognise slightly broader range of groups involved than non-academics, likely reflecting different perspectives on project operation. Collectively, responses could be aggregated to shed insights on breadth and sectoral diversity of project.	RDR4
	Research users and others based outside research institutions actively participated in the project	Survey: likert agreement	Tested	Use with caution: survey across team	Rated moderate-high fit for purpose. Provisionally, results point to possibility of diverging awareness of stakeholder participation among academic team members, thus need for care in averaging results across project teams. Could be useful to weight assessments toward project team members more actively involved across a range of project activities.	
	General inputs measures · Number of individuals from stakeholder community/ies actively engaged in research · Number, diversity, geographic and sectoral coverage of stakeholder organisations actively engaged in research, relative to project impact **Specific community-based assessments** Focus group, interviews, surveys, participatory evaluations to assess whether and extent of active engagement has been established.	Quantitative and independent survey	Assessed	Possibly use, preliminary filter for relevance, value add	· General measures potentially more widely feasible, reported from project teams, but do not indicate how effectively stakeholders were supported to take part in researcher whether this was seen as desirable or appropriate . · Community-based assessments likely more feasible in single-project or targeted evaluations. Could complement or provide independent support to project team (academic and non-academic) survey measures, particularly in contexts where broader diversity of perspectives gained is essential. · Assessment should both assess and allow for churn among stakeholders where possible.	
New understanding of values and priorities captured (e.g. organizational agencies, citizen science, Indigenous priorities)	The values and priorities of research users informed how the project was run	Survey: likert agreement	Tested	Good: survey across team	Rated highly fit-for-purpose. Possible that academic respondents report slightly higher than other project team members, suggesting value in surveying across project team.	RDR5

Table F5. Inputs measures and questions: **research objectives and governance strongly developed and clearly articulated**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Contextually and culturally sensitive evaluation frameworks (including impact indicators) reviewed or developed for the research	The project objectives were clearly articulated at the start	Survey: likert agreement	New measure identified (Gap 1)	Not tested		RDR13
Increase in reciprocal understanding between stakeholders and end-users around the need for, and framing of research	Research users were actively part of setting the objective of the research	Survey: likert agreement	New measure identified (Gap 1)	Not tested	Will need to articulate the difference between research objectives and research questions in survey instrument	RDR2
Agreement for working together established or re-iterated between scientists, decision-makers, practitioners, Indigenous partners	Strong relationships between research partners from different groups and agencies underpinned the project	Survey: likert agreement	New measure identified (Gap 2)	Not tested		RDR3
	Evidence on the type of partnership or collaboration established	Documentary	New measure identified (Gap 2)	Not tested		RDR3

Table F6. Inputs measures and questions: **Previously unavailable datasets made available**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Existing datasets that were previously unavailable are liberated (e.g. industry or commercial in-confidence)	Were there existing data sets/information necessary to do the research? If yes, were those made available? If not, why not & what was the consequence?	Interview: narrative	Tested	Good (with prompt about role)	Need to understand interviewees' role in the project to judge the accuracy of their answer. Use prompt if required to determine role in project	RDR8
	Additional data sets from outside of the research team were made available to support the research	Survey: likert agreement	Tested	Use with caution: survey across team	Rated highly fit for purpose. Provisionally, variability of results point to possibility of diverging awareness among academic team members. Thus need for care in averaging across project teams. Could weight toward project team members more actively involved across a range of project activities.	
	· Number and geographic scope and types of datasets newly accessible to researcher and research user networks · Number and geographic scope and data types of formerly inaccessible and/or resultant datasets now in public domain or otherwise accessible (with appropriate protections) · Documentary information and (where feasible) numerical information on data collected and geographic coverage of datasets made available, including differentiation between referents for data (e.g. species, ecosystem, geodata, remote sensing, models, genetic) and types of data (statistical vs qualitative data)	Quantitative / documentary	Assessed	Good	Number of datasets newly available (e.g. to the project team or publicly) is potentially readily reportable by project teams later in or after the end of a project. Geographic scope and types of data collected would likely be documentary information that could provide additional information for context. It would take significantly more work and care to aggregate or compare information across programs on geographic coverage and content of data collected , but this could be done for like data (e.g. number of species for which time-series monitoring data has been made available).	

Processes measures

Table F7. Processes measures and questions: **research management and processes**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
All partners and stakeholders contribute to designing or providing input to research questions	How did non-academic (non-research)/your and other non-academic groups participate in the project? P1 Articulated the need for research? P2 Their values & priorities informed how the project was run? P3 How actively did they participate?	Interview: narrative	Tested	Use after defining "participation"	*Need to state what defines "participation": e.g. People who helped design the project, provided advice, collected data, analysed data, developed outputs, promoted outputs to end-users *Need to ask at least 1 academic & 1 non-academic participant	RMC22 **(see also RDR4 RDR5 RDRX inputs)**
	Extent of involvement of stakeholder community/ies engaged in research planning, including: · Number of individuals from stakeholder communities invited to design or provide input into research questions · Number of individuals from stakeholder communities adequately supported to design or provide input into research questions · Number of individuals from stakeholder communities designing or providing input into research questions Appropriateness of involvement of stakeholder community/ies engaged in research planning, including: · Proportion and number of stakeholder organisations represented appropriately · Representation of stakeholder group involvement relative to geographic and sectoral scope of research	Quantitative	Assessed	Use with caution	Likely to be feasible and appropriate in many cases. Raw numbers of individuals involved in research design likely not as useful as representation, and numbers appropriately invited and resourced, but will also give an indication of trust and appropriate processes through demonstrating active involvement. It is important that these measures on providing input to research questions do not assume a one-size fits all approach to research design, and that the nature of the input of research-users is the best and most fit-for purpose process for the circumstances. Co-design is often promoted as the primary aim, but not all groups have the best range of tools or availability at their disposal to co-design research. A different range of tools may be warranted. Not everyone has the interest to actively participate in research design activities, though they may appreciate being asked. Sometimes research users appreciate independent research, rather than co-design, to guide the questions that can best be asked at a given point in time.	
	Community survey measures asking project stakeholders and community participants to assess the quality and appropriateness of the approaches taken to inviting and involving them in designing and providing input to research questions, e.g. through interviews and focus groups.	Community insight	Assessed	Possibly use, preliminary filter for relevance, value add	Community interviews and focus groups are unlikely to be feasible except in individual project assessments and small, in-depth case studies. Likely to be more useful at mid-term evaluation stage.	

Table F7 cont. Processes measures and questions: **research management and processes**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessmt	Comments	Indicator code*
Processes to gain research permission (permits, ethics etc.) maintained, improved, made more meaningful and rigorous in relation to research content and processes, or streamlined where appropriate	Knowledge of and/or capacity to engage in ethical research by researchers and/or research users was improved as a result of the project	Survey: likert agreement	Tested	Possibly use, preliminary filter for relevance, value add	Rated by most participants as only medium fitness for purpose. Reasonable within-project concordance. Many participants rated their projects 'n/a or don't know' or 'neither agree nor disagree' against this measure, suggesting low levels of conviction about this measure.	RMC28
Contextually and culturally appropriate forms of communication among all research team developed, improved or already excellent Barriers and problems with developing effective research processes or collaborations recognised and strategies for overcoming them maintained or developed	Could you describe the communication processes within the project team? Prompt: Do you feel they were adequate? Why/why not?	Interview/ narrative	New measure identified (Gap 3)	Not tested		RMC24 RMC27
	Communication within the project team was initiated by both researchers and research users	Survey: likert agreement	New measure identified (Gap 3)	Not tested		
	Communication between researchers and research users was at an appropriate frequency for the project	Survey: likert agreement	New measure identified (Gap 3)	Not tested		
	Communication between researchers and research users was effective for addressing challenges or overcoming hurdles in the research process	Survey: likert agreement	New measure identified (Gap 3)	Not tested		
Trust built or maintained within research networks and collaborations	Do you feel trust was developed and /or improved among partners? could you give an example to illustrate?	Interview: narrative	Tested	Good (don't just ask project leader)	Trust is a complex concept and so is best suited to narrative measures that can unpack how people assess that trust is present (e.g. strong relationships, honesty and respect). Academic and non-academic perspectives should be sought	RMC18
	Trust was developed or improved among research partners	Survey: likert agreement	Tested	Good	Rated high fit for purpose, low within-project variability, close concordance between academic and non-academic respondents.	
	Before-and-after survey of project collaborators, partners, and community participants where appropriate, e.g.: · how confident would you be to work with these (researchers/project partners) again? · how likely would you be to recommend these (researchers/project partners) to others to work with?	Longitudinal survey	Tested	Possibly use, preliminary filter for relevance, value add	Could be used as a final evaluation measure in cases where a baseline has been established before projects progress far. Likely to be more feasible in individual project assessments, or small bodies of projects. Would need to allow for cases where trust is high at baseline, as well as where significant increases can be observed. Qualitative information on conditioning circumstances (e.g. external factors contributing to trust or erosion of trust) could be invited to contextualise assessments.	

Table F7 cont. Processes measures and questions: **research management and processes**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Obtaining Prior Informed Consent from Indigenous communities understood, developed, prioritised, excellent	Did the research need to go through permission processes (permits, ethics, etc)?	Survey: multiple choice, branching	Small sample	Insufficiently tested	As a branching question, may want to word more specifically as research permission processes for Indigenous engagement.	RMC20
	If Indigenous partners in network map: How was informed consent obtained from Indigenous partners?	Interview: narrative	Small sample	Insufficiently tested	Question needs a filter/prompt as won't apply to all projects. The mapping question is good for this. Requires input from more Indigenous research participants to test/assess.	
	[if indigenous partners were selected as a participating group] Obtaining Prior Informed Consent from Indigenous communities was prioritised in the project	Survey: likert agreement	Small sample	Insufficiently tested	Only applicable to 3 out of 18 respondents. Rated as highly fit-for-purpose by two out of three of these respondents.	
Processes to gain research permission (permits, ethics etc.) maintained, improved, made more meaningful and rigorous in relation to research content and processes, or streamlined where appropriate	Knowledge of and/or capacity to engage in ethical research by researchers and/or research users was improved as a result of the project	Survey: likert agreement	Tested	Possibly use, preliminary filter for relevance, value add	Rated by most participants as only medium fitness for purpose. Reasonable within-project concordance. Many participants rated their projects 'n/a or don't know' or 'neither agree nor disagree' against this measure, suggesting low levels of conviction about this measure.	RMC28
Contextually and culturally appropriate forms of communication among all research team developed, improved or already excellent Barriers and problems with developing effective research processes or collaborations recognised and strategies for overcoming them maintained or developed	Could you describe the communication processes within the project team? Prompt: Do you feel they were adequate? Why/why not?	Interview/ narrative	New measure identified (Gap 3)	Not tested		RMC24 RMC27
	Communication within the project team was initiated by both researchers and research users	Survey: likert agreement	New measure identified (Gap 3)	Not tested		
	Communication between researchers and research users was at an appropriate frequency for the project	Survey: likert agreement	New measure identified (Gap 3)	Not tested		
	Communication between researchers and research users was effective for addressing challenges or overcoming hurdles in the research process	Survey: likert agreement	New measure identified (Gap 3)	Not tested		

Table F8. Processes measures and questions: **research methods**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Best methods for data collection, management and analysis developed or maintained across disciplines and knowledge systems	What is your view on the research methods used in the project? Did it draw on different disciplines & knowledge systems?	Interview: narrative	Tested	Use with modification	Avoid using the term "knoweldge systems", instead ask did different disciplines or people from outside academia contribute. Use a prompt about the quality of the methods if required. Interviews must include academics and non-academics	RM16
	Methods used in the project drew on insights from different knowledge systems	Survey: likert agreement	Tested	Good: survey across team with modification	Rated moderate to high fitness for purpose. Academics may rate their project lower than non-academic partners against this, though within error. Question would benefit from clearer language (avoid use of the term "knowledge systems") specifying information sought. Ensure non-academic partners are surveyed where possible. See also survey measure for AO30 under Outputs.	
	Data collection methods used in the project were: a) Below standard b) Standard c) Above standard/new d) Don't know	Survey: multiple choice	Tested	Good	Rated high fit for purpose, low within-project variability, close concordance between academic and non-academic respondents.	
	Data management methods used in the project were: a) Below standard b) Standard c) Above standard/new d) Don't know	Survey: multiple choice	Tested	Good: preliminary filter for relevance	Rated moderate fitness for purpose. Moderate within-project variability, reasonable within-project concordance, though researchers may slightly under-report (within error).	
	Data analysis methods used in the project were: a) Below standard b) Standard c) Above standard/new d) Don't know	Survey: multiple choice	Tested	Good: minor modification	Highly fit for purpose, low within-project variability, strong within-project concordance. May be worth adding qualifier: "e.g. the analysis gleaned the greatest value from the available data".	
Retention, continuity and evolution of research teams and partnerships achieved	· Number of students, postdoctoral researchers and other early career researchers supported to participate · Number of early career practitioners outside of research institutions supported to participate · Number of Indigenous rangers and early career researchers supported to participate · Number of citizen scientists enabled and supported to participate	Quantitative	Assessed	Good	Likely to be feasible to measure, speaks to knowledge transfer and potentially growing capacity for quality, engaged research. Data on student numbers should differentiate between different kinds of students (e.g. honours, Masters, PhD).	RMC23

Outputs measures

Table F9. Outputs measures and questions: **production of quality outputs**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Legitimate, valuable, and rigorous knowledge developed according to the various knowledge systems involved in the research Publications relating to the research authored or co-authored by researcher or team	How would you characterise the knowledge produced by the project? · (depending on network map) How did different knowledge systems contribute to the project?	Interview: narrative	Tested	Use with modification	The first outputs question needs to be "What outputs from the project are you aware of? This can include Data, Presentations, Media, etc This should be the last outputs question as it is reflective One or two people mentioned rigour and no-one legitimacy so if these are of particular interest they may require a prompt.	AO30 AO33
	Valuable knowledge was produced through the project	Survey: likert agreement	Tested	Good	Rated as highly fit for purpose by all respondents. Low internal variability, high consistency between academic and non-academic respondents.	
	Different "knowledge systems" informed the outputs of the project	Survey: likert agreement	Tested	Use with amendments: apply comparatively	Rated as highly fit for purpose by non-academic respondents, only moderately fit for purpose by academic respondents. Moderate within-project variability among both groups. Academics ranked their project lower than non-academic partners against this, though within error. Question would benefit from clearer language specifying information sought. Ensure non-academic partners are surveyed where possible. Consider analysing data separately for comparison across team roles. Potentially include write-in field for types of knowledge systems. See also measure for RM16 under Processes.	
	What type of project outputs have you seen or are aware of? · technical publications (reports, peer-reviewed papers)	Survey: multiple choice yes/no	Tested	Good: apply comparatively, minor note on implementation	Question gives both an indication of the outputs produced, and how aware non-academic partners are of the range and types of products delivered. Relatively low within-project variability, but non-academic partners show less awareness of these products. Potentially modify to ask numbers as well as types.	
	Number of peer reviewed academic articles	Quantitative	Assessed	Good: preliminary filter to decide where relevant	Rated as among the least important indicators of research impact in Phase 1. However, may provide some information as a small component of a suite of measures of research impact, as peer review gives an indication of acceptance by research peers.	

Table F9 cont. Outputs measures and questions: **production of quality outputs**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessmt	Comments	Indicator code*
Postgraduate students completed, training and certificates completed by research participants, capacity of young researchers increased	Number of students and postdoctoral researchers trained through project; number of students and postdoctoral researchers having undertaken formal intercultural training.	Quantitative	Assessed	Good	Likely feasible and relevant to building sectoral capacity	AO31
	Numbers of people outside of traditional research institutions trained or gaining skills through direct involvement in project, e.g.: · number of Indigenous rangers trained or growing in experience and capacity, through involvement in research · number of community members, landholders, etc trained and/or grown in capacity to undertake effective actions or further research · number of citizen scientists trained and/or grown in capacity to undertake effective actions or further research	Quantitative	Assessed	Good: preliminary filter to decide where relevant	Likely somewhat feasible for use where relevant	
Clear communication of failures, limitations, and challenges in the research	The limitations, failures or challenges in the research and the reasons for these have been clearly communicated	Survey: likert agreement	New measure identified (Gap 1)	Not tested		AO41

Table F10. Outputs measures and questions: **research tailored, packaged and disseminated through stakeholder networks**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Outputs produced - general Could include information and/or prompts on each of the types of products below	What do you think about how knowledge generated by the project has been packaged, presented and disseminated?	Interview: narrative	Tested	Not recommended	Individual project participants may not be in a position to judge what appropriate dissemination is. A better question might be "Did the project generate outputs that met your needs? Then: How did outputs meet/not meet your needs?"	RD52, TD43, TD44, AO32, AO33, RD46, RD51, RD52, RD53, RD55
Tailored reports, stories or summaries prepared on research and findings for specific audiences (e.g. land managers, Indigenous communities)	Outputs from the project have been tailored to specific end-users (i.e. land managers, communities)	Survey: likert agreement	Tested	Good: survey across team	Rated highly fit for purpose. Relatively low within-project variability. Academic respondents may rate slightly higher. Survey both groups where possible. Consider analysing data separately for comparison across team roles.	RD51

Table F10 cont. Outputs measures and questions: **research tailored, packaged and disseminated through stakeholder networks**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Guidelines, guides, checklists, standards for dispersed but like-minded audiences (e.g. parties who do environmental monitoring, all parties who participate in species translocations) Tools made available to land managers, practitioners and citizens/communities for on-ground action (management plans, planning frameworks and processes) Training packages developed for practitioners, citizens, community members and stakeholder groups Data sets (including spatial) made available to other researchers, practitioners and citizens/communities Tailored reports, stories or summaries prepared on research and findings for specific audiences (e.g. land managers, Indigenous communities) Summaries for policy makers prepared (policy options papers, submissions to policy forums Contribution to public policy advisory committee(s)	What type of project outputs have you seen or are aware of? · tools for implementation (e.g. planning frameworks, guidelines, checklists) · training packages for practitioners, citizens, community members, etc · data sets · summaries and other outputs tailored for specific audiences (e.g. factsheets)	Survey: multiple choice yes/no	Tested	Good: apply comparatively, minor note on implementation	Question gives both an indication of the outputs produced, and how aware non-academic partners are of the range and types of products delivered. Low to moderate within-project variability. Non-academic partners show less awareness of these products on the whole. Consider surveying across project team to gauge awareness of products. Potentially modify to ask numbers as well as types.	RD52 TD43 TD44 AO32 RD51 RD53 RD54
	What type of project outputs have you seen or are aware of? · presentations	Survey: multiple choice y/n	Tested	Good: apply comparatively with modification	Split question about presentation between those specifically designed for research users (relevant here), and more general awareness-raising and broadcast presentations.	
	Production and public availability of outputs tailored for use by specific groups of stakeholders/research users, e.g.: · number of guidelines, standards, protocols, etc · number of tools produced · number of training packages prepared · number of datasets published · number of tailored reports, stories, summaries · number of policy summaries, submissions, etc Active dissemination and promotion of specific, tailored outputs: proportion of products above actively promoted to relevant audiences of partners outside of research institutions, stakeholders and research users for each of the above	Quantitative	Assessed	Good	Availability of specific, tailored products rated in phase 1 as consistently more important than academic outputs. Number of products of each kind, along with whether or not these have been actively disseminated to relevant audiences, provides a ready measure for aggregation across programs.	

Table F10 cont. Outputs measures and questions: **research tailored, packaged and disseminated through stakeholder networks**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Guidelines, guides, checklists, standards for dispersed but like-minded audiences (e.g. parties who do environmental monitoring, all parties who participate in species translocations) Tools made available to land managers, practitioners and citizens/communities for on-ground action (management plans, planning frameworks and processes) Training packages developed for practitioners, citizens, community members and stakeholder groups Data sets (including spatial) made available to other researchers, practitioners and citizens/communities Tailored reports, stories or summaries prepared on research and findings for specific audiences (e.g. land managers, Indigenous communities) Summaries for policy makers prepared (policy options papers, submissions to policy forums Contribution to public policy advisory committee(s)	Communication of research findings to audiences, including: Research actively shared with partners outside of research institutions, stakeholders and research-users: · Number and reach of workshops, discussions, forums specifically to tailor and enable adoption of research findings · Number and reach of training forums to support adoption of tools, new techniques developed through research · Number and reach of stakeholder-focused presentations given to support adoption · Number of participants and number and diversity of groups participating in forums Research actively shared with senior decision-makers: · Number of briefings given to Minister · Number of briefings given to senior decision-makers, Minister's advisers · Number and kind of appearances at inquiries, expert panels, advisory committees	Quantitative	Assessed	Good: preliminary filter to decide where relevant	Number and reach of targeted briefings, workshops, training forums and presentations specifically to support adoption gives a strong measure of active research dissemination and potential uptake. May be difficult to capture across wide array of projects - could focus on small numbers of projects, or centrally-organised forums across larger programs	
Presentations to schools, community groups and forums	What type of project outputs have you seen or are aware of? · presentations	Survey: multiple choice yes/no	Tested	Good: apply comparatively with modification	Split question about presentation between those specifically designed for research users, and more general awareness-raising and broadcast presentations (relevant here). Likely of greater value to specify presentations for particular groups of interest (e.g. academic audiences, potential research users, community).	RD55

Table F10 cont. Outputs measures and questions: **research tailored, packaged and disseminated through stakeholder networks**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Collaborative research publications or outputs	Can you describe how project outputs were produced, particularly who was involved?	Interview: narrative	Tested	Good: minor modification	Ask what outputs they are aware of first and then ask how they were produced. That way can ensure definition of outputs is beyond papers.	AO39
	What type of project outputs have you seen or are aware of? · collaborative outputs (may be any combination output types)	Survey: multiple choice y/n	Tested	Possibly use, preliminary filter for relevance, value add	Moderate within-project variability among both groups, but good within-project concordance. Consider surveying whole project team. Useful if information is sought on awareness of this across project team. Where information is sought on the extent to which collaborative outputs were produced, could be more effectively measured quantitatively.	
	Publications and other project outputs were produced in a collaborative manner	Survey: likert agreement	Tested	Good	Rated as highly fit for purpose by non-academic respondents, moderate-high by academic respondents. Low within-project variability among both groups and good within-project concordance.	
	· Numbers of products co-authored with research partners outside of university/research contexts · Where relevant: Number of products co-authored with Indigenous partners Including co-authored: · data · academic articles · guidelines, standards, protocols, etc · tools · training packages · accessible reports, stories, summaries · policy summaries, submissions, etc · popular articles · Number of presentations co-presented with research partners outside of university/research contexts · Where relevant: Number of presentations co-presented with Indigenous partners	Quantitative	Assessed	Good	Rated among the more important academic outputs in Phase 1. Readily measured.	

Table F11. Outputs measures and questions: **research promoted widely**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Research publications freely and openly accessible Data sets (including spatial) made available to other researchers, practitioners and citizens/ communities	Research outputs are freely accessible to all including research participants and partners	Survey: likert agreement	Tested	Good: minor modification	Rated as highly fit for purpose by both groups. Low within- project variability, high concordance between respondents. Modify question to emphasise broad accessibility over sharing with partners and research participants.	RD46 AO32
	Outputs are stored in a place that will be appropriately accessible once the project is complete	Survey: likert agreement	New measure identified (Gap 4)	Not tested		
	· Numbers and proportion of research products (of different types, as above) made freely, openly and publicly accessible, as appropriate. · Altmetric figures for research publications. Distribution and download statistics for: · research publications · tailored products: guidelines, tools, training packages · tailored summaries (policy and management-accessible information) · data sets produced	Quantitative	Assessed	Good	Relatively feasible and relevant measures.	RD46 AO32 RD52 TD43 TD44 RD51 RD53
	Number and proportion of outputs stored and accessible in a place that will be enduring and appropriately accessible once the project is complete	Quantitative	New measure identified (Gap 4)	Not tested		

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessmt	Comments	Indicator code*
Presentations to schools, community groups and forums Media and social media coverage for scientific publications authored or co-authored by researcher or team in the applicable time period Media and social media coverage for tools and other outputs (beyond research publications) by researcher or team in the applicable time period Websites, web pages, blogs produced Popular articles including magazine and newsletter articles written by researchers and journalists Generic reports or stories prepared on research and results (e.g. fact sheets, videos)	**Higher priority:** · Presentations to popular audiences including schools, community groups, forums: number and audience reach **Lower priority:** Number, download and reach statistics of articles and activities **developed by study authors** to popularise and share findings, including: · popular articles and opinion pieces · social media posts, videos, webinars and related outputs · websites, web pages and blogs	Quantitative	Assessed	Good: preliminary filter to decide where relevant	Could use where important to assessment. Presentations to schools, community groups and forums both rated moderately high importance in phase 1. Popular articles, media and social media, generic summaries and web pages generated by research team were relatively lower in importance, but should be relatively feasible.	RD55 RD45 RD47 RD48 RD49 RD50

Outcomes measures

Table F12. Outcomes measures and questions: **research awareness and citation**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Academic citations for scientific publications authored or co-authored by researcher or team in the applicable time period	Number of academic citations of academic publications linked to research	Quantitative	Assessed	Good: preliminary filter to decide where relevant	Rated low in Phase 1 as measure of research value. Use if relevant to assess research quality and acceptance in research sector. (See outputs RD46)	RU56
Research findings perpetuated via popular articles, newsletters, fact sheets, reports written by third parties	Number, download and reach statistics of articles and activities **developed by third parties** to popularise and share findings in general, including: · popular articles · social media posts, videos, webinars and related outputs · websites, web pages and blogs	Quantitative	Assessed	Good: preliminary filter to decide where relevant	Rated moderate in Phase 1 as a measure. Readily measured.	RU58
Public and grey-literature citations for scientific publications authored or co-authored by researcher or team in the applicable time period	Citations and reach of grey literature citing findings and/or publications linked to research for general purposes, including general citations in policy or management documents (excluding e.g. citations specifically reporting changes in management, practice, policy or planning, e.g. citations in management guidelines)	Quantitative	Assessed	Not recommended as a priority	Rated low in Phase 1 as a measure of research value. Currently can involve considerable work to assess, as traditional citation aggregation tools exclude grey literature.	RU57

Table F13. Outcomes measures and questions: **changes to policy and practice - general**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Changes to practice and policy - general Could prompt further on full range of specific outcomes below	Can you describe any changes to practice or policy the research may have contributed to? E.g. informed management actions, contributed to listing or policy	Interview: narrative	Tested	Good: ask non-academics & academics	Care is needed to emphasise contribution as changes will rarely if ever link solely to 1 research project. A generic question like this appropriate when projects are just finishing. Types of response will depend on lag period between research and outcomes	PPP78, AR59 AR60, AR61 AR62, OGA71 OGA72, OGA73 OGA75, OGA76

Table F14. Outcomes measures and questions: **on ground action**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Research findings, tools and guidelines adopted in practice and implemented in monitoring, management, community practices, regulation, business practices	· Please select which of the following outcomes related to on-ground practice your project has, or is likely to contribute: Improved monitoring for species, ecosystems, heritage places, etc · Please select which option(s) best applies to the scale of this outcome that your project has, or is likely to contribute: a) Change within project region or partnership	Survey: multiple choice, confidence	Tested	Use with amendments, use with caution, survey across team	Rated as moderately fit for purpose by non-academic respondents and high by academics. Moderate within-project variability among academic respondents. Academic respondents appear to rate higher (though within error). Requires further testing to understand whether these differences are systematic. Survey on-ground partners along with academic team. Split changes that have happened from likely changes, add free-text field for details, where feasible.	OGA72
	Improved monitoring for species, ecosystems, heritage places, etc: b) Change beyond project region or partnership	Survey: multiple choice, confidence	Tested	Use with amendments, survey across team	Rated as moderately fit for purpose by non-academic respondents and high by academics. Moderate within-project variability among non-academic respondents, but good within-project concordance between academics and non-academics. Survey whole project team where feasible. Split changes that have happened from likely changes, add free-text field for details, where feasible.	
Adoption of more effective techniques for conservation practice (e.g. captive breeding, reintroduction in the wild) Findings from research incorporated into real world experimental tests or trials	Improved management of species, ecosystems, heritage places, etc: a) Change within project region or partnership	Survey: multiple choice, confidence	Tested	Use with amendments, survey across team	Rated highly fit for purpose. Moderate within-project variability among non-academic respondents, but good within-project concordance between academics and non-academics. Survey whole project team where feasible. Split changes that have happened from likely changes, add free-text field for details, where feasible.	AR62 OGA71
	Improved management of species, ecosystems, heritage places, etc: b) Change beyond project region or partnership	Survey: multiple choice, confidence	Tested	Use with amendments, use with caution, survey across team	Rated highly fit for purpose. Moderate within-project variability. Academics appear to rate significantly higher than non-academics for same projects, which may be overestimation or because they are in position to observe changes beyond the project relationships. Survey on-ground partners along with academic team. Split changes that have happened from likely changes, add free-text field for details, where feasible.	
Establishment of adaptive management trials, close coupled science-practice feedback loops	Improved adaptive management or management planning for species, ecosystems, heritage places, etc: a) Change within project region or partnership	Survey: multiple choice, confidence	Tested	Use with amendments, survey across team	Rated highly fit for purpose. Moderate within-project variability among non-academic respondents, but good within-project concordance between academics and non-academics. Survey whole project team where feasible. Split changes that have happened from likely changes, add free-text field for details, where feasible.	OGA73
	Improved adaptive management or management planning for species, ecosystems, heritage places, etc: b) Change beyond project region or partnership	Survey: multiple choice, confidence	Tested	Use with amendments, survey across team	Rated highly fit for purpose. Low within-project variability. Academics appear to rate higher than non-academics for same projects, though within error. This may be overestimation or because they are in position to observe changes beyond the project relationships. Requires further testing to understand if these differences are systematic. Survey whole project team where feasible. Split changes that have happened from likely changes, add free-text field for details, where feasible.	

Table F14 cont. Outcomes measures and questions: **on ground action**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Research findings, tools and guidelines adopted in practice and implemented in monitoring, management, community practices, regulation, business practices	Improved data management, reporting and/or analysis about species, ecosystems, heritage places, etc: a) Change within project region or partnership	Survey: multiple choice, confidence	Tested	Possibly use with amend-ments if relevant, survey across team	Rated low fitness for purpose by non-academics and moderate by academics. Moderate within-project variability among academic respondents, good within-project concordance. Consider applying in select circumstances where this information is a priority. Survey whole project team where feasible. Split changes that have happened from likely changes, add free-text field for details, where feasible.	OGA72
	Improved data management, reporting and/or analysis about species, ecosystems, heritage places, etc: b) Change beyond project region or partnership	Survey: multiple choice, confidence	Tested	Possibly use with amend-ments if relevant, survey across team	Rated low fitness for purpose by non-academics and moderate by academics. Low within-project variability; low number of positive responses limits assessment. Academics appear to rate higher than non-academics for same projects, though within error. Requires further testing to understand if these differences are systematic. Apply in select circumstances where this information is a priority. Survey whole project team. Split changes that have happened from likely changes, add free-text field for details, where feasible.	
Reduction in threatening process(es)	Reduction in threatening processes a) Change within project region or partnership	Survey: multiple choice, confidence	Tested	Use with amendments, survey across team	Fitness for purpose rated moderately high by non-academic respondents but moderately low by academic respondents. Moderate within-project variability, but good within-project concordance between academics and non-academics. Survey whole project team where feasible. Split changes that have happened from likely changes, add free-text field for details, where feasible.	OGA76
	Reduction in threatening processes: b) Change beyond project region or partnership	Survey: multiple choice, confidence	Tested	Use with amendments, survey across team	Fitness for purpose rated moderate by non-academic respondents and low by academic respondents. Moderate within-project variability among academic respondents. Low number of positive responses limits assessment, but provisional results suggest value in surveying whole project team for stronger convergence. Split changes that have happened from likely changes, add free-text field for details, where feasible.	

Table F14 cont. Outcomes measures and questions: **on ground action**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessmt	Comments	Indicator code*
Adoption of more effective techniques for conservation practice (e.g. captive breeding, reintroduction in the wild)	Evidence of incorporation of research findings into monitoring plans/process, management plans, policy, trigger points, recovery plans · citation for specific implementation in management plans, grey literature	Quantitative / documentary	Assessed	Good	Should be relatively feasible to source over moderate number of projects	AR62 OGA72
Findings from research incorporated into real world experimental tests or trials Research findings, tools and guidelines adopted in practice and implemented in monitoring, management, community practices, regulation, business practices Establishment of adaptive management trials, close coupled science-practice feedback loops New skills and training developed through research implemented Increase in protected area size and/or quality Reduction in threatening processes	**Immediate adoption to support new practice** Documentary evidence of adoption of findings by partners e.g.: · plans resulting from research implemented by relevant stakeholders · evidence that research influences priorities, practices or policy (e.g. adoption in management operating procedures, planning guidelines, policy, regulation) Number of instances of changes to management, policy, actions adopting research findings, e.g.: · number and diversity of new sites established from research findings (e.g. new monitoring sites, new management areas) **Broader adoption to support new practice** Measures of adoption of findings among wider stakeholders: · number and reach of organisations adopting research tools, guidelines, findings, techniques · number and reach of people trained in and applying new techniques, approaches **Environmental value resulting from changes to practices** Measures of environmental values within scope of changed management or practices e.g.: · number of threatened or culturally significant species under area affected by positive changes to action · coverage of threatened ecological community, wetland system or place affected by changed actions arising from research · coverage of protected area/IPA land positively influenced by research	Quantitative / documentary	Assessed	Good	Measures of adoption of findings among project partners relatively readily assessed Numbers of sites positively affected moderately readily assessed for specific projects or small case studies Measures of broader adoption relatively feasible to assess for a set range of stakeholders Measures of environmental values within scope of changes implemented would require more work to estimate and aggregate, but are potentially relatively feasible where more information is required, e.g. for significant reporting milestones	AR62 OGA71 OGA72 OGA73 OGA74 OGA75 OGA76

Table F15. Outcomes measures and questions: **change in government policy**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Research findings help to inform the decision-making process and result in change to public policy, strategy and/or program design Species or ecological community listed as threatened and / or protected by legislative instrument Landscape or threatened ecological community heritage listed Results trigger inscription of national park or conservation land tenure Increase in protected area size and/or quality	Were there any other outcomes (expected or unexpected) connected at least broadly with the project? Eg investment decisions, program design, etc?	Interview: narrative	Tested	Not recommended as stand-alone question. Use as prompt	This question can prompt people to think beyond planned outcomes but often come out in earlier questions in the interview. Use as a prompt for other outcomes questions	PPP78 AR59 AR60 AR61 OGA75
	Has the project contributed to any of these other outcomes? · Helped to inform decision-making processes related to public policy, strategy and/or program design?	Survey: multiple choice, confidence	Tested	Use with amendments, use with caution, survey across team	Fitness for purpose rated as moderate (non-academics) to high (academics). Moderate within-project variability. Academic respondents appear to rate lower, but within error. Requires further testing to understand if these differences are systematic. Consider surveying whole project team, particularly government research users. Split changes that have happened from likely changes, add free-text field for details, where feasible.	PPP78 AR59 AR60 AR61 OGA75

Table F15 cont. Outcomes measures and questions: **change in government policy**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessmt	Comments	Indicator code*
Research findings help to inform the decision-making process and result in change to public policy, strategy and/or program design Species or ecological community listed as threatened and / or protected by legislative instrument Landscape or threatened ecological community heritage listed Results trigger inscription of national park or conservation land tenure Increase in protected area size and/or quality	**Changes to policy** Changes to implementation of policy instruments, including: · Listing, uplisting and/or downlisting of species and/or ecosystems as a result of research · Adoption of findings, guidelines or recommendations in policy instruments (recovery plans, conservation advices, assessments) · Integration of new spatial data (e.g. species distributions) into government planning and regulatory databases Changes to operation of policy instruments, including: · Adoption of tools, findings, guidelines or recommendations into new policy instruments (planning processes, regulatory decisions) · Integration of new methods for assessing spatial data (e.g. species distributions) into government planning and regulatory databases Research findings and tools adopted in major policy decisions including: · Adoption of findings into legislation or significant policy strategies · Inscription of protected area · New or targeted program or funding **Environmental value resulting from changes to policy** Measures of environmental values within scope of changed policy e.g.: · number of threatened or culturally significant species under area affected by positive changes to action · coverage of threatened ecological community, wetland system or place affected by changed actions arising from research · coverage of protected area/IPA land positively influenced by research	Quantitative / documentary	Assessed	Good	Measures of adoption of findings among project partners relatively readily assessed Measures of environmental values within scope of changes implemented would require more work to estimate and aggregate, but are potentially relatively feasible where more information is required, e.g. for significant reporting milestones	PPP78 AR59 AR60 AR61 OGA75

Table F16. Outcomes measures and questions: **inform decision-making in private and non-government contexts**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Research processes and findings result in beneficial change to for-profit, business and private sector policy (e.g. private companies, financial sector consortium)	Has the project contributed to any of these other outcomes? · Resulted in beneficial change in the for-profit private sector?	Survey: multiple choice, confidence	Tested	Possibly use with amendments if relevant, survey across team	Rated as low fitness for purpose by most respondents. Possibly apply in select cases where changes in business decisions are priority. Low number of positive responses limits assessment on variability and concordance; requires further testing. Split changes that have happened from likely changes, add free-text field for details, where feasible.	PPP81
	Uptake and/or reach of economic tools and measures produced contributing to greater transparency in economic decision-making e.g.: · citation in industry ESG strategies, environmental accounting assessments, etc · number, reach and economic impact of businesses adopting research tools and/or findings	Quantitative / documentary	Assessed	Good: preliminary filter to decide where relevant	Generally moderately low importance in Phase 1. Potentially readily sourced. May apply to specific studies with economic tools and products.	PPP81
Research findings result in beneficial change to non-governmental and not-for-profit policy (e.g. International Union for the Conservation of Nature, Indigenous corporations)	Were there any other outcomes (expected or unexpected) connected at least broadly with the project? Eg investment decisions, program design, etc?	Interview: narrative	Tested	Not recommended as stand-alone question. Use as prompt	This question can prompt people to think beyond planned outcomes but often come out in earlier questions in the interview. Use as a prompt for other outcomes questions	PPP82
	Has the project contributed to any of these other outcomes? · Resulted in beneficial change in the non-governmental, not-for-profit policy sector?	Survey: multiple choice, confidence	Tested	Possibly use with amendments if relevant, survey across team	Rated moderate fitness for purpose by academic respondents, low by non-academics; academics appear to rate higher (within error), possibly as a result of low number of NGO respondents. Adjusting for confidence removed differences. Requires further testing with NGO partners to understand if differences are systematic. Consider using for projects where relevant; ensure NGO partners are included. Consider splitting changes that have happened from likely changes, and adding free-text field for details, where feasible.	PPP82
	Evidence and number of cases of adoption in non-government planning or policy documents	Documentary	Assessed	Good: preliminary filter to decide where relevant	Generally moderately low importance in Phase 1. Potentially readily sourced where applicable. May apply to specific studies with intended outcomes for NGOs.	PPP82

Table F17. Outcomes measures and questions: **strengthened partnerships and responsiveness**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Mutually established and agreed upon ideas research and goals between researchers and community organisations, Indigenous communities, conservation groups, land managers	The research focus changed or was adapted as a result of emerging needs or the contributions of research users	Survey: likert agreement	New measure identified (Gap 5)	Not tested		CSE64
Links formed/improved between members of the research team, community organisations, Indigenous communities, conservation groups, land managers, for on ground action	Network map: interviewees asked to identify the different groups involved in research and describe the links between them and how relationships may have changed over time	Interview: interactive mapping exercise	Tested	Good (need >1 respondent)	Will get complementary responses from people with different roles in a project The network mapping exercise should be conducted early in the interview to identify which participating organisations the interviewee is aware of. Doing the mapping visually so the interviewee can see was a good prompt (eg whiteboard function in zoom).	CSE65 **RDR4** **RDR5** **RDRX** **RMC22** **(inputs)**
	Links between members of the research team and key stakeholders (such as Indigenous communities, conservation groups, etc) have improved	Survey: likert agreement	Tested	Good	Rated as moderate to highly fit for purpose. Low within-project variability, strong within-project concordance.	CSE65
	Evidence of new networks, partnerships and capacities built to implement effective practice guided by research, e.g.: · number and coverage of Indigenous ranger groups engaged and/or benefiting from changes arising from research · number and coverage of community or landcare groups engaged and/or benefiting from changes arising from research · number and coverage of citizen scientist groups and networks established and/or made more effective through implementing research	Quantitative	Assessed	Use with caution	May be somewhat feasible, but may in places be difficult to define boundaries of engagement or to tie directly to research project. Caution should be applied in aggregating any examples as a result of variable assessments of what constitutes groups implementing changes.	CSE65

Table F18. Outcomes measures and questions: **improved capabilities**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
New skills and training developed through research implemented Improved or increased intercultural capacities, via collective training of researchers on-ground personnel, in community organisations, Indigenous communities, conservation groups, land managers	Do you think your project helped to develop the skillset of you and/or your partners to undertake this work? If yes, what types of skills were developed?	Interview: narrative	Tested	Good (need >1 respondent)	Prompt on intercultural capacities if relevant. People primarily aware of their own capacities so need multiple respondents	OGA74 (CSE66)
Improved or increased intercultural capacities, via collective training of researchers on-ground personnel, in community organisations, Indigenous communities, conservation groups, land managers	Intercultural capacities of the research team have improved	Survey: likert agreement	Tested	Use with amendments, use with caution, survey across team	Rated as moderate-highly fit for purpose by non-academic respondents and moderate-low by academic respondents. Moderate within-project variability. Researchers appear to rate lower (but within error). Requires further testing to assess whether differences are systematic. Consider language to clarify meaning of 'intercultural capacities', and surveying across project teams.	CSE66
	· Number of Indigenous community members newly engaging in research · Number of research team members and number of early career researchers and students newly engaging in research with Indigenous partners	Quantitative	Assessed	Use with caution	Likely feasible to measure. Caution should be applied in aggregating as a result of different interpretations of this measure.	
Increased support or legitimacy for community/citizen practices that help to achieve conservation goals Involved, engaged or affected stakeholders inspired/supported to engage further in independent research (e.g. citizen science programs, Indigenous knowledge projects)	Expanded capacity of people to undertake effective actions or further research, e.g.: · evidence of Indigenous ranger groups or communities developing further work to extend project · evidence of community members, landholders, etc growing networks to share or extend on research · evidence of citizen scientists undertaking effective independent actions or further research · evidence of extension of citizen scientist-based methods developed to new research projects or contexts	Quantitative / documentary	Assessed	Use with caution	Documentary measures may be somewhat feasible, but may in places be difficult to tie to research project. Caution should be applied in aggregating any examples since assessments of what constitutes extension of practice.	AR63 CSE70

Table F18 cont. Outcomes measures and questions: **improved capabilities**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
New skills and training developed through research implemented	What has been learned through the project? Prompt: could you give an example?	Interview: Narrative	New measure identified (Gap 6)	Not tested	May need to prompt about sites of learning - collective and individual.	OGA74 (CSE66)
	Research team members have gained important skills through the project	Survey: likert agreement	New measure identified (Gap 6)	Not tested		
	The project helped to develop the skillset of you and/or your partners to undertake work of this kind	Survey: likert agreement	New measure identified (Gap 6)	Not tested	Need to think about whether you only ask people to speak about themselves	
	The project has developed early career researchers to be skilled in - the topic area - research engagement - intercultural capacities	Survey: likert agreement	New measure identified (Gap 4/6)	Not tested	Relates to RMC23 *Retention, continuity and evolution of research teams and partnerships achieved* but focus here is on life beyond the project rather than through it	
	The project has developed early career researchers to be trusted experts in the topic area	Survey: likert agreement	New measure identified (Gap 4/6)	Not tested	Relates to RMC23 *Retention, continuity and evolution of research teams and partnerships achieved* but focus here is on life beyond the project rather than through it	
	I have learnt... Research users and research team members have learnt... more about the topic through the research	Survey: likert agreement	New measure identified (Gap 6)	Not tested		
	I have learnt ... Research team members have learnt ... more about the wider policy and management challenges and context to which the project relates	Survey: likert agreement	New measure identified (Gap 6)	Not tested		

Table F19. Outcomes measures and questions: **objectives and legacy**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
(New)	The original objectives of the research have been met . **If answer no:** The reasons for not meeting the objectives have been appropriately assessed and responded to	Survey: likert agreement	New measure identified (Gap 1)	Not tested	logic chain might be useful here to reduce ambiguity	OGA / PPP / CSE (new)
(New)	What do you think the key legacy of the project will be? prompt: How well do you think legacy has been planned for?	Interview: Narrative	New measure identified (Gap 4)	Not tested		OGA / PPP / CSE (new)
(New)	Legacy of projects has been adequately planned for including planning for discontinued funding which may result in incomplete data, loss of key staff, etc	Survey: likert agreement	New measure identified (Gap 4)	Not tested		OGA (new)
(New)	Funding has been received to continue or implement the research	Survey: likert agreement	New measure identified (Gap 4)	Not tested	Relates to RDR12 *Ability to attract research income or make in-kind contributions to future research increased or already excellent*, but specifically addresses investment in this work as an outcome of the work	OGA / PPP / CSE (new)
(New)	Evidence of funding received to continue or implement the research	Documentary	New measure identified (Gap 4)	Not tested	Relates to RDR12 *Ability to attract research income or make in-kind contributions to future research increased or already excellent*, but specifically addresses investment in this work as an outcome of the work	OGA / PPP / CSE (new)

Impact measures

Table F20. Impact measures and questions: **reflection on overall project performance**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
(New)	Can you reflect on how you think the project performed against its objectives?	Interview/ narrative	New measure identified (Gap 1)	Not tested	Where there is a lag to impact or people are hesitant about attribution this might be a useful way of talking about impact/value of the project. Ask at end.	REI (new)

Table F21. Impact measures and questions: **environmental benefit**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Environmental impacts - general	Can you describe any environmental impacts related to your project?	Interview: narrative	Tested	Use with prompts & multiple people	Always explain how impact is being defined (to help avoid outcomes being listed) Time lags mean impacts may not have manifested or to have been detected A given individual may be unaware of impacts.	REI84, REI87 REI89, REI90 REI91

Table F21 cont. Impact measures and questions: **environmental benefit**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Species recovery or avoided loss, including: Population increase and/or avoided loss in threatened species Avoided loss of biodiversity	The project has or will contribute to an increase in population trajectory for a threatened or significant species	Survey: likert likelihood, confidence	Tested	Use with amendments, use with caution, apply comparatively	Rated highly fit for purpose, but generated high degree of within-project variability amongst non-academic respondents. Results suggest academics may assess lower than non-academic partners (though within large error range), including when degree of confidence is taken into account. Survey approach should seek responses across project teams, but may not deliver convergence. Apply with caution, avoid averaging across project teams. Split impacts that have happened from those projected. Request further information through free text where feasible.	REI87 REI90
	The project has or will contribute to avoiding loss of biodiversity	Survey: likert likelihood, confidence	Tested	Use with amendments, survey across team	Rated moderately to highly fit for purpose, but generated high degree of within-project variability amongst non-academic respondents. Academic and non-academic assessments within projects reasonably concordant, particularly once degree of confidence taken into account. Split impacts that have happened from those projected. Consider requesting further information through free text where feasible.	
	Measures of project contribution to population increase in a threatened species, avoided declines in threatened species population or distribution or avoided biodiversity loss. These could include: · Measurable improvement in population and/or condition metrics against baselines or from long-term monitoring trends · Likelihood of species persistence as assessed through Population Viability Analysis · Improvements in species or ecosystem trajectories, avoided decline or likelihood of avoided decline as assessed through expert elicitation	Modelled or monitored	Assessed	Possibly use, preliminary filter for relevance, value add	Quantitative measures of population may be resource intensive, and only feasible if measures were built into project design or developed for related purposes (e.g. listing). More feasible as in-depth measures for a specific project or small suite of projects, but aggregate approaches could be used in assessment, e.g., (Bayraktarov et al., 2021) for large-scale or intensive, locally focused research programs (such as action-research in one conservation reserve), where monitoring data are available. Where relevant, direct measures via monitoring (e.g. using before-after-impact-control design), may be desirable, but is dependent on baseline and presence of controls. May not likely see improvements during assessment timeframe except locally, where specific, targeted interventions are particularly intensive and successful, e.g., (Stojanovic et al., 2019). Projections e.g. via population viability analysis may be of benefit where data and analysis capacity allow. Avoided loss is harder to quantify against baseline, but more likely to be achieved within timeframes, and can usefully be assessed via expert elicitation (see Geyle, Garnett, Legge, & Woinarski, 2019, Maron & Evans 2018).	REI87

Table F21 cont. Impact measures and questions: **environmental benefit**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Improved condition of places, including: Improvement in the maintenance and/or condition of an ecosystem, wetland, marine environment	The project has or will contribute to an improvement in condition of an ecosystem, wetland or marine environment	Survey: likert likelihood, confidence	Tested	Use with amendments, use with caution, apply comparatively	Rated moderately to highly fit for purpose, but generated high degree of within-project variability amongst non-academic respondents. Academics rate significantly lower than non-academic partners (divergence in the same direction but within error once degree of confidence taken into account). Survey approach should seek responses across project teams, but may not deliver convergence. Apply with caution, avoid averaging across project teams. Split impacts that have happened from those projected. Request further information through free text where feasible.	REI89
	Measures of project contribution to an ecosystem, wetland or marine environment, including measures and evidence of: · Improvement or avoided decline in function or condition of an ecosystem, wetland or marine environment · Improvement or avoided decline in extent of an ecosystem or threatened ecological community	Modelled or monitored	Assessed	Possibly use, preliminary filter for relevance, value add	Similar considerations to population increase or avoided loss. Need a strong condition framework, indicators and specific metrics for measurement, e.g., (Sharp & Gould, 2014), but unlikely to see measurable change in time frames of projects. May need to consider feasibility of alternative approaches (e.g. expert elicitation of avoided decline) where data do not provide strong indication or are not available.	REI89 REI91

Table F22. Impact measures and questions: **enhanced governance and capacities for locally-defined priorities and inclusive decision-making**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Improved capacities of management, governance and institutions to engage in effective, equitable and informed deliberation and decision-making around conservation issues	Can you describe any changes that have taken place in how decision-making is undertaken in the region? Have any of these involved changes to how agencies (such as planners or management agencies) operate? Do you think these changes will be lasting?	Narrative	No	Not tested	Need to carefully consider who is asked this question. Probably only suitable for agency staff themselves or those who describe close interactions with such agencies.	REI96
	The project has or will contribute to changes in the social or cultural processes or systems in place for managing species, ecosystems or places	Survey: likert likelihood, confidence	Tested	Use with amendments, use with caution, apply comparatively	Rated moderately fit for purpose. Moderate degree of within-project variability amongst non-academic respondents. Academic respondents assessed substantially lower than non-academic respondents, more so once degree of confidence was taken into account, perhaps reflecting different perspectives and insights. Academics may underreport; surveying those directly involved in managing species, ecosystems or places may deliver more consistent results. Avoid averaging across project teams. Split impacts that have happened from those projected. Consider requesting further information through free text where feasible.	

Table F22 cont. Impact measures and questions: **enhanced governance and capacities for locally-defined priorities and inclusive decision-making**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Improved capacities of management, governance and institutions to engage in effective, equitable and informed deliberation and decision-making around conservation issues	Project contributions to improvements in capacities of agencies to engage in effective, equitable and informed decision-making: · Evidence of the application of good practice processes from the research to other decisions, or other decision-making contexts · Evidence of changes to policies, programs or processes improving decision-making for managing species, ecosystems or places as a result of research.	Quantitative/ documentary	Assessed	Good: preliminary filter to decide where relevant	Measures may not apply across many projects, but where present, documentary evidence would provide a strong independent indication of contribution of research to enhancing capacities for deliberation and decision-making.	REI96
	Project contributions to improvements in capacities of agencies to engage in effective, equitable and informed decision-making: · Use of community surveys, focus groups or participatory evaluations, where relevant.	Community insight	Assessed	Possibly use, preliminary filter for relevance, value add	Community approaches require specialist expertise to administer and place an additional burden on stakeholder communities. However, in projects where these are important anticipated outcomes, these measures could be built into project design.	

Table F23. Impact measures and questions: **social, cultural and economic benefit**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Social, cultural and economic impacts - general	Can you describe any societal, cultural or economic impacts the project may have contributed to?	Interview: narrative	No	Inadequately tested	The 3 interviewees who were asked this, struggled to answer. How do individuals know the answer to this? Our cases did not have clear pathways to economic impacts and so we have not tested the best way to elicit these.	REI94 REI95
A better informed society with greater social licence for conservation Influence on public policy debate	The project has or will contribute to greater public awareness of the importance of, options for or challenges of conservation	Survey: likert likelihood, confidence	Tested	Possibly use where relevant with amendments, survey across team	Non-academic respondents rated fitness for purpose low-moderate. Academic and non-academic responses reasonably concordant, but survey generated a low-moderate degree of within-project variability. Split impacts that have happened from those projected. Consider requesting further information through free text where feasible.	REI94 **PPP77 (outcome)**

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
A better informed society with greater social licence for conservation Influence on public policy debate	Project contribution to greater public awareness of the importance of, options for or challenges of conservation, including: Influence on public policy debate: · Documentary evidence of citation in reports from public inquiries Substantial direct engagement with findings by third parties, e.g.: · Documentary evidence of citation in opinion pieces and other public contributions to public debate that engages with research, written by people other than research partners · Significant syndication of media opinion pieces · Significant degree of likes and reshares of social media opinion pieces · Reach and engagement figures where available	Quanti-tative/ docu-mentary	Assessed	Good: preliminary filter to decide where relevant	Could involve documentary and/or quantitative measures. Measures unlikely to apply across many projects, but where evidence is present, would provide a strong independent indication of contribution to wider public discussion. Examples and numbers of citations in public policy discussions, such as inquiry reports, and direct independent media and social media citations of the research for the purposes of drawing on it to inform public discourse could be ascertained across whole programs. Syndication, or likes and re-shares of social media pieces, could provide proxy measures in the absence of direct reach and engagement figures. Reach and engagement figures likely to be harder to assess, as these are not author-driven, but could data could be requested if needed for significant assessments in particular cases.	REI94 **PPP77 (outcome)**
Improvement in human wellbeing derived from species and/or ecosystems, including access to resources, health, livelihoods, self-determination, social relations, and cultural and spiritual satisfaction: · Social and cultural measures of contribution to improvement in human wellbeing derived from species and/or ecosystems, including access to resources, health, livelihoods, self-determination, social relations, and cultural and spiritual satisfaction. · Economic contribution to improvement in human wellbeing derived from species and/or ecosystems, including access to resources and livelihoods.	The project has or will contribute to positive social or cultural outcomes for people	Survey: likert likelihood confi-dence	Tested	Not recommended	Rated as moderately low fitness for purpose. Responses showed moderate to high degree of within-project variability. Respondents had a consistently low degree of confidence in their responses. Academics rate as significantly less likely than non-academic partners (including once degree of confidence taken into account). Suggests a likert-scale survey approach may not be the most appropriate, and alternative approaches to assess this should be sought. Where social and cultural outcomes are important to assessment and a survey approach is adopted, should seek responses across project teams, and may not deliver convergence. Apply with caution, avoid averaging across project teams. Split impacts that have happened from those projected and requesting further information through free text where feasible.	REI95
	The project has or will contribute to positive economic change	Survey: likert likelihood confi-dence	Tested	Possibly use where relevant with amendments use with caution, survey across team	Rated as low fitness for purpose by both academic and non-academic respondents. Many respondents rated this indicator relatively low, but academic respondents rated this slightly higher (though within error), with results showing a moderate degree of within-project variability among academic respondents. Within-project variability and differences reduced once degree of confidence was taken into account. Possibly apply, with degree of confidence qualifier, in select cases where economic change is a priority. Split impacts that have happened from those projected and consider requesting further information through free text.	Not in original study

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Whether tested	Assessment	Comments	Indicator code*
Improvement in human wellbeing derived from species and/or ecosystems, including access to resources, health, livelihoods, self-determination, social relations, and cultural and spiritual satisfaction: · Social and cultural measures of contribution to improvement in human wellbeing derived from species and/or ecosystems, including access to resources, health, livelihoods, self-determination, social relations, and cultural and spiritual satisfaction. · Economic contribution to improvement in human wellbeing derived from species and/or ecosystems, including access to resources and livelihoods.	**Documentary evidence of changes to social or cultural programs** · Inclusion of project in cultural, educational, health, job-creation or wellbeing contexts **Contribution to economic measures or decision-making** · Uptake and/or reach of economic tools and measures (e.g. evidence of use of research findings in independent environmental accounting assessments) **Documentary evidence of economic and social value for local communities, e.g.:** · jobs created · programs created · enhanced access to resources, e.g. regulatory changes, new industries · economic value of new or enhanced industries **External community satisfaction or wellbeing surveys, interviews or focus groups** Could explore as relevant: · program satisfaction measures as a proxy for direct social measures · health and psychological wellbeing benefits · educational benefits · wellbeing benefits derived from greater access to resources and livelihoods · improved social relations · sense of self-determination · sense of place · increased cultural and spiritual satisfaction **Measures or proxies of current or potential changes to ecosystem services, e.g.:** · water quality and water provisioning services · carbon sequestration · pollination services	Quanti-tative / docu-Mentary / commu-nity insight / modelled or monitored	Assessed	Possibly use, preliminary filter for relevance, value add	Documentary and quantitative evidence for improvements to human wellbeing derived from species and/or ecosystems is considerably harder to measure, attribute to research, or determine with any certainty. Direct evidence of uptake of inclusion of research in cultural and social programs or processes, or of uptake and reach of tools and measures into economic programs or processes, are unlikely to apply to many projects. However, where evidence is present, would be relatively feasible to gather, and provide a strong independent indication of contribution of research to social, cultural and/or economic well being. Evidence of jobs or economic programs created, changes to regulation or new industries are likely to be rare outcomes of conservation projects, but likewise, where present, would be relatively feasible to gather, and provide a strong independent indication of contribution of research to social, cultural and/or economic well being. Surveys of wider community to establish independent insights into social and cultural benefits and improvements in human wellbeing derived from species and/or ecosystems should only be used selectively, as they require specialist expertise to administer and place an additional burden on stakeholder communities. However, in projects were these are important anticipated outcomes, these measures could be built into project design, (see for example Ward et al., 2021). Measures of current or potential changes to ecosystem services require specialist skills and intensive resources to assess, and may not be seen within the life of research projects. However, where such outcomes are anticipated or desired as part of research (such as practice-change projects to achieve economic benefits), a processes for assessing these benefits could be build into research design. Proxies, such as water quality estimates in catchments with productive landscapes, could be used as a substitute where full environmental economics measures are not feasible. May be difficult to see benefit within assessment timeframe. Could be projected/modelled (e.g. Keith et al 2017).	REI95

Appendix A: Summary of indicators and measures tested/assessed

Table A1: Summary of input indicators and measures: research needs identified

Indicator descriptor	Measures/questions addressing indicators	Type of measure	Indicator code*
Research needs identified			
Research needs, gaps, unanswered questions, new areas for research identified	How were research needs identified in the project?	Interview: narrative	RDR1
	Research needs or new areas of research were identified through the project	Survey: likert agreement and fit for purpose	
	Documentary evidence of research gaps that have been identified and prioritised, as well as those met through the project	Documentary (assessed, not tested)	

Table A2: Summary of input indicators and measures: research user and stakeholder involvement

Indicator descriptor	Measures/questions addressing indicators	Type of measure	Indicator code*
Research users and stakeholders active involvement in research			
Need for research has been adequately established - and all perspectives gained on whether research is valuable (e.g. particular end-user agencies) Active participation of stakeholders and end-users in research New understanding of values and priorities captured (e.g. organizational agencies, citizen science, Indigenous priorities) All participants and stakeholders contribute to designing or providing input to research questions	How did non-academic groups participate in the project? · P1 Articulated the need for research? · P2 Their values & priorities informed how the project was run? · P3 How actively did they participate?	Interview: narrative	RDR4 RDR5 RDRX (new indicator identified at expert workshop) RMC22
Need for research has been adequately established - and all perspectives gained on whether research is valuable (e.g. particular end-user agencies) Active participation of stakeholders and end-users in research New understanding of values and priorities captured (e.g. organizational agencies, citizen science, Indigenous priorities)	Network map: interviewees asked to identify the different groups involved in research and describe the links between them and how relationships may have changed over time Develop a map of organisations and the relationships between them (at the beginning vs end of project	Interview: interactive mapping exercise	RDR4 RDR5 RDRX (new indicator identified at expert workshop) RMC22
Need for research has been adequately established - and all perspectives gained on whether research is valuable (e.g. particular end-user agencies)	**General measures** · Number, diversity, geographic and sectoral coverage of stakeholder organisations involved in research planning, relative to project impact **Specific community-based assessments** Focus group, interviews, surveys, participatory evaluations to assess whether need and value of research has been established.	Quantitative / community insight (assessed, not tested)	RDRX (not in original study, new indicator identified at expert workshop)
	The value of the research to different stakeholders (particularly end-user agencies) was elicited	Survey: likert agreement and fit for purpose	

Table A2 (cont): Summary of input indicators and measures: research user and stakeholder involvement

Indicator descriptor	Measures/questions addressing indicators	Type of measure	Indicator code*
Research users and stakeholders active involvement in research			
Active participation of stakeholders and end-users in research	**General measures** · Number of individuals from stakeholder community/ies actively engaged in research · Number, diversity, geographic and sectoral coverage of stakeholder organisations actively engaged in research, relative to project impact **Specific community-based assessments** Focus group, interviews, surveys, participatory evaluations to assess whether and extent of active engagement has been established.	Quantitative / community insight (assessed, not tested)	RDR4
	Were groups or organisations outside research institutions (like universities, CSIRO) directly involved in the project? [if yes to above] What groups or organisations (other than research institutions) were directly involved in the project?	Survey: yes/no; multiple choice	
	Research users and others based outside research institutions actively participated in the project	Survey: likert agreement and fit for purpose	
New understanding of values and priorities captured (e.g. organizational agencies, citizen science, Indigenous priorities)	The values and priorities of research users informed how the project was run	Survey: likert agreement and fit for purpose	RDR5

Table A3: Summary of inputs indicators and measures: data accessibility

Indicator descriptor	Measures/questions addressing indicators	Type of measure	Indicator code*
Previously unavailable datasets made available			
Existing datasets that were previously unavailable are liberated (e.g. industry or commercial in-confidence)	Were there existing data sets/information necessary to do the research? If yes, were those made available? If not, why not & what was the consequence?	Interview: narrative	RDR8
	Additional data sets from outside of the research team were made available to support the research	Survey: likert agreement and fit for purpose	
	· Number and geographic scope and types of datasets newly accessible to researcher and research user networks · Number and geographic scope and data types of formerly inaccessible and/or resultant datasets now in public domain or otherwise accessible (with appropriate protections) · Documentary information and (where feasible) numerical information on data collected and geographic coverage of datasets made available, including differentiation between referents for data (e.g. species, ecosystem, geodata, remote sensing, models, genetic) and types of data (statistical vs qualitative data)	Quantitative / documentary (assessed, not tested)	

Table A4: Summary of processes indicators and measures: research management & conduct

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Indicator code*
Research management and conduct			
All participants and stakeholders contribute to designing or providing input to research questions	How did non-academic (non-research)/your and other non-academic groups participate in the project? P1 Articulated the need for research? P2 Their values & priorities informed how the project was run? P3 How actively did they participate?	Interview: narrative	RMC22 **(see also RDR4 RDR5 RDRX inputs)**
	Extent of involvement of stakeholder community/ies engaged in research planning, including: · Number of individuals from stakeholder communities invited to design or provide input into research questions · Number of individuals from stakeholder communities adequately supported to design or provide input into research questions · Number of individuals from stakeholder communities designing or providing input into research questions Appropriateness of involvement of stakeholder community/ies engaged in research planning, including: · Proportion and number of stakeholder organisations represented appropriately · Representation of stakeholder group involvement relative to geographic and sectoral scope of research	Quantitative (assessed, not tested)	
	Community insight measures asking project stakeholders and community participants to assess the quality and appropriateness of the approaches taken to inviting and involving them in designing and providing input to research questions, e.g. through interviews and focus groups.	Community insight (assessed, not tested)	
Trust built or maintained within research networks and collaborations	Do you feel trust was developed and /or improved among partners? could you give an example to illustrate?	Interview: narrative	RMC18
	Trust was developed or improved among research partners	Survey: likert agreement and fit for purpose	
	Before-and-after survey of project collaborators, partners, and community participants where appropriate, e.g.: · how confident would you be to work with these (researchers/project partners) again? · how likely would you be to recommend these (researchers/project partners) to others to work with?	Longitudinal survey (assessed, not tested)	
Obtaining Prior Informed Consent from Indigenous communities understood, developed, prioritised, excellent	Did the research need to go through permission processes (permits, ethics, etc)?	Survey: multiple choice, branch qn	RMC20
	If Indigenous partners in network map: How was informed consent obtained from Indigenous partners?	Interview: narrative	
	[if indigenous partners were selected as a participating group] Obtaining Prior Informed Consent from Indigenous communities was prioritised in the project	Survey: likert agreement and fit for purpose	
Processes to gain research permission (permits, ethics etc.) maintained, improved, made more meaningful and rigorous in relation to research content and processes, or streamlined where appropriate	Knowledge of and/or capacity to engage in ethical research by researchers and/or research users was improved as a result of the project	Survey: likert agreement and fit for purpose	RMC28

Table A5: Summary of processes indicators and measures: research methods

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Indicator code*
Research methods			
Best methods for data collection, management and analysis developed or maintained across disciplines and knowledge systems	Q9. What is your view on the research methods used in the project? Did it draw on different disciplines & knowledge systems?	Interview: narrative	RM16
	Methods used in the project drew on insights from different knowledge systems	Survey: likert agreement and fit for purpose	
	Data collection methods used in the project were: a) Below standard b) Standard c) Above standard/new d) Don't know	Survey: multiple choice	
	Data management methods used in the project were: a) Below standard b) Standard c) Above standard/new d) Don't know	Survey: multiple choice	
	Data analysis methods used in the project were: a) Below standard b) Standard c) Above standard/new d) Don't know	Survey: multiple choice	
Retention, continuity and evolution of research teams and partnerships achieved	· Number of students, postdoctoral researchers and other early career researchers supported to participate · Number of early career practitioners outside of research institutions supported to participate · Number of Indigenous rangers and early career researchers supported to participate · Number of citizen scientists enabled and supported to participate	Quantitative (assessed, not tested)	RMC23

Table A6: Summary of output indicators and measures: quality outputs

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Indicator code*
Production of quality outputs			
Legitimate, valuable, and rigorous knowledge developed according to the various knowledge systems involved in the research Publications relating to the research authored or co-authored by researcher or team	How would you characterise the knowledge produced by the project? · (depending on network map) How did different knowledge systems contribute to the project?	Interview: narrative	AO30 AO33
	Valuable knowledge was produced through the project	Survey: likert agreement and fit for purpose	
	Different "knowledge systems" informed the outputs of the project	Survey: likert agreement and fit for purpose	
	What type of project outputs have you seen or are aware of? · technical publications (reports, peer-reviewed papers)	Survey: multiple choice yes/no	
	Number of peer reviewed academic articles	Quantitative (assessed, not tested)	
Production of quality outputs			
Postgraduate students completed, training and certificates completed by research participants, capacity of young researchers increased	Number of students and postdoctoral researchers trained through project; number of students and postdoctoral researchers having undertaken formal intercultural training.	Quantitative (assessed, not tested)	AO31
	Numbers of people outside of traditional research institutions trained or gaining skills through direct involvement in project, e.g.: · number of Indigenous rangers trained or growing in experience and capacity, through involvement in research · number of community members, landholders, etc trained and/or grown in capacity to undertake effective actions or further research · number of citizen scientists trained and/or grown in capacity to undertake effective actions or further research	Quantitative (assessed, not tested)	

Table A7: Summary of outputs indicators and measures: research packaged and distributed

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Indicator code*
Research tailored, packaged and disseminated through stakeholder networks			
Outputs produced - general Could include information and/or prompts on each of the types of products below	What do you think about how knowledge generated by the project has been packaged, presented and disseminated?	Interview: narrative	RD52, TD43, TD44, AO32, AO33, RD46, RD51, RD52, RD53, RD55
Tailored reports, stories or summaries prepared on research and findings for specific audiences (e.g. land managers, Indigenous communities)	Outputs from the project have been tailored to specific end-users (i.e. land managers, communities)	Survey: likert agreement and fit for purpose	RD51
Guidelines, guides, checklists, standards for dispersed but like-minded audiences (e.g. parties who do environmental monitoring, all parties who participate in species translocations) Tools made available to land managers, practitioners and citizens/communities for on-ground action (management plans, planning frameworks and processes) Training packages developed for practitioners, citizens, community members and stakeholder groups Data sets (including spatial) made available to other researchers, practitioners and citizens/communities Tailored reports, stories or summaries prepared on research and findings for specific audiences (e.g. land managers, Indigenous communities) Summaries for policy makers prepared (policy options papers, submissions to policy forums Contribution to public policy advisory committee(s)	What type of project outputs have you seen or are aware of? · tools for implementation (e.g. planning frameworks, guidelines, checklists) · training packages for practitioners, citizens, community members, etc · data sets · summaries and other outputs tailored for specific audiences (e.g. factsheets)	Survey: multiple choice yes/no	RD52 TD43 TD44 AO32 RD51 RD53 RD54
	Production and public availability of outputs tailored for use by specific groups of stakeholders/research users, e.g.: · number of guidelines, standards, protocols, etc · number of tools produced · number of training packages prepared · number of datasets published · number of tailored reports, stories, summaries · number of policy summaries, submissions, etc Active dissemination and promotion of specific, tailored outputs: proportion of products above actively promoted to relevant audiences of partners outside of research institutions, stakeholders and research users for each of the above	Quantitative (assessed, not tested)	
	Communication of research findings to audiences, including: Research actively shared with partners outside of research institutions, stakeholders and research-users: · Number and reach of workshops, discussions, forums specifically to tailor and enable adoption of research findings · Number and reach of training forums to support adoption of tools, new techniques developed through research · Number and reach of stakeholder-focused presentations given to support adoption · Number of participants and number and diversity of groups participating in forums Research actively shared with senior decision-makers: · Number of briefings given to Minister · Number of briefings given to senior decision-makers, Minister's advisers · Number and kind of appearances at inquiries, expert panels, advisory committees	Quantitative (assessed, not tested)	

Table A7 (cont): Summary of outputs indicators and measures: research packaged & distributed

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Indicator code*
Research tailored, packaged and disseminated through stakeholder networks			
Presentations to schools, community groups and forums	What type of project outputs have you seen or are aware of? · presentations	Survey: multiple choice yes/no	RD55
Collaborative research publications or outputs	Can you describe how project outputs were produced, particularly who was involved?	Interview: narrative	AO39
	What type of project outputs have you seen or are aware of? · collaborative outputs (may be any combination output types)	Survey: multiple choice yes/no	
	Publications and other project outputs were produced in a collaborative manner	Survey: likert agreement and fit for purpose	
	· Numbers of products co-authored with research partners outside of university/research contexts · Where relevant: Number of products co-authored with Indigenous partners Including co-authored: · data · academic articles · guidelines, standards, protocols, etc · tools · training packages · accessible reports, stories, summaries · policy summaries, submissions, etc · popular articles · Number of presentations co-presented with research partners outside of university/research contexts · Where relevant: number of presentations co-presented with Indigenous partners	Quantitative (assessed, not tested)	

Table A8: Summary of outputs indicators and measures: research promoted widely

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Indicator code*
Research promoted widely			
Research publications freely and openly accessible Data sets (including spatial) made available to other researchers, practitioners and citizens/communities	Research outputs are freely accessible to all including research participants and partners	Survey: likert agreement and fit for purpose	RD46 AO32
Research publications freely and openly accessible	· Numbers and proportion of research products (of different types, as above) made freely, openly and publicly accessible, as appropriate. · Altmetric figures for research publications. Distribution and download statistics for: · research publications · tailored products: guidelines, tools, training packages · tailored summaries (policy and management-accessible information) · data sets produced	Quantitative (assessed, not tested)	RD46 RD52 TD43 TD44 AO32 RD51 RD53
Presentations to schools, community groups and forums Media and social media coverage for scientific publications authored or co-authored by researcher or team in the applicable time period Media and social media coverage for tools and other outputs (beyond research publications) by researcher or team in the applicable time period Websites, web pages, blogs produced Popular articles including magazine and newsletter articles written by researchers and journalists Generic reports or stories prepared on research and results (e.g. fact sheets, videos, animations)	**Higher priority:** · Presentations to popular audiences including schools, community groups, forums: number and audience reach **Lower priority:** Number, download and reach statistics of articles and activities **developed by study authors** to popularise and share findings, including: · popular articles and opinion pieces · social media posts, videos, webinars and related outputs · websites, web pages and blogs	Quantitative (assessed, not tested)	RD55 RD45 RD47 RD48 RD49 RD50

Table A9: Summary of outcomes indicators and measures: research awareness and citation

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Indicator code*
Research awareness and citation			
Academic citations for scientific publications authored or co-authored by researcher or team in the applicable time period	Academic citations of academic publications linked to research	Quantitative (assessed, not tested)	RU56
Research findings perpetuated via popular articles, newsletters, fact sheets, reports written by third parties	Number, download and reach statistics of articles and activities developed **by third parties** to popularise and share findings in general, including: · popular articles · social media posts, videos, webinars and related outputs · websites, web pages and blogs	Quantitative (assessed, not tested)	RU58
Public and grey-literature citations for scientific publications authored or co-authored by researcher or team in the applicable time period	Citations and reach of grey literature citing findings and/or publications linked to research for general purposes, including general citations in policy or management documents (excluding e.g. citations specifically reporting changes in management, practice, policy or planning, e.g. citations in management guidelines)	Quantitative (assessed, not tested)	RU57

Table A10: Summary of outcomes indicators and measures: changes to practice and policy

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Indicator code*
Changes to practice and policy: general			
Changes to practice and policy - general Could prompt further on full range of specific outcomes below	Can you describe any changes to practice or policy the research may have contributed to? E.g. informed management actions, contributed to listing or policy	Interview: narrative	PPP78, AR59 AR60, AR61 AR62, OGA71 OGA72, OGA73 OGA75, OGA76
On-ground action			
Research findings, tools and guidelines adopted in practice and implemented in monitoring, management, community practices, regulation, business practices	· Please select which of the following outcomes related to on-ground practice your project has, or is likely to contribute: Improved monitoring for species, ecosystems, heritage places, etc · Please select which option(s) best applies to the scale of this outcome that your project has, or is likely to contribute: a) Change within project region or partnership · Outcomes can manifest in the months and years after a research project has finished. Can you indicate your level of confidence in your responses about outcomes based on the evidence available to you? ('confidence' measure applies to all related measures)	Survey: multiple choice, confidence and fit for purpose	OGA72
	Improved monitoring for species, ecosystems, heritage places, etc: b) Change beyond project region or partnership	Survey: multiple choice, confidence and fit for purpose	
Adoption of more effective techniques for conservation practice (e.g. captive breeding, reintroduction in the wild) Findings from research incorporated into real world experimental tests or trials	Improved management of species, ecosystems, heritage places, etc: a) Change within project region or partnership	Survey: multiple choice, confidence and fit for purpose	AR62 OGA71
	Improved management of species, ecosystems, heritage places, etc: b) Change beyond project region or partnership	Survey: multiple choice, confidence and fit for purpose	
Establishment of adaptive management trials, close coupled science-practice feedback loops	Improved adaptive management or management planning for species, ecosystems, heritage places, etc: a) Change within project region or partnership	Survey: multiple choice, confidence and fit for purpose	OGA73
	Improved adaptive management or management planning for species, ecosystems, heritage places, etc: b) Change beyond project region or partnership	Survey: multiple choice, confidence and fit for purpose	
Research findings, tools and guidelines adopted in practice and implemented in monitoring, management, community practices, regulation, business practices	Improved data management, reporting and/or analysis about species, ecosystems, heritage places, etc: a) Change within project region or partnership	Survey: multiple choice, confidence and fit for purpose	OGA72
	Improved data management, reporting and/or analysis about species, ecosystems, heritage places, etc: b) Change beyond project region or partnership	Survey: multiple choice, confidence and fit for purpose	
Reduction in threatening process(es)	Reduction in threatening processes a) Change within project region or partnership	Survey: multiple choice, confidence and fit for purpose	OGA76
	Reduction in threatening processes: b) Change beyond project region or partnership	Survey: multiple choice, confidence and fit for purpose	

Table A10 (cont): Summary of outcomes indicators and measures: changes to practice and policy

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Indicator code*
On-ground action (continued)			
Adoption of more effective techniques for conservation practice (e.g. captive breeding, reintroduction in the wild) Findings from research incorporated into real world experimental tests or trials Research findings, tools and guidelines adopted in practice and implemented in monitoring, management, community practices, regulation, business practices Establishment of adaptive management trials, close coupled science-practice feedback loops New skills and training developed through research implemented Increase in protected area size and/or quality Reduction in threatening processes	Evidence of incorporation of research findings into monitoring plans/process, management plans, policy, trigger points, recovery plans · citation for specific implementation in management plans, grey literature	Quantitative / documentary (assessed, not tested)	AR62 OGA72
	Immediate adoption to support new practice Documentary evidence of adoption of findings by partners e.g.: · plans resulting from research implemented by relevant stakeholders · evidence that research influences priorities, practices or policy (e.g. adoption in management operating procedures, planning guidelines, policy, regulation) Number of instances of changes to management, policy, actions adopting research findings, e.g.: · number and diversity of new sites established from research findings (e.g. new monitoring sites, new management areas) **Broader adoption to support new practice** Measures of adoption of findings among wider stakeholders: · number and reach of organisations adopting research tools, guidelines, findings, techniques · number and reach of people trained in and applying new techniques, approaches	Quantitative / documentary (assessed, not tested)	AR62 OGA71 OGA72 OGA73 OGA74 OGA75 OGA76
Inform decision-making: government			
Research findings help to inform the decision-making process and result in change to public policy, strategy and/or program design Species or ecological community listed as threatened and / or protected by legislative instrument Landscape or threatened ecological community heritage listed Results trigger inscription of national park or conservation land tenure Increase in protected area size and/or quality	Were there any other outcomes (expected or unexpected) connected at least broadly with the project? Eg investment decisions, program design, etc?	Interview: narrative	PPP78 AR59 AR60 AR61 OGA75
	Has the project contributed to any of these other outcomes? · Helped to inform decision-making processes related to public policy, strategy and/or program design?	Survey: multiple choice, confidence and fit for purpose	
	Evidence of changes to implementation of policy instruments, including: · Listing, uplisting and/or downlisting of species and/or ecosystems as a result of research · Adoption of findings, guidelines or recommendations in policy instruments (recovery plans, conservation advices, assessments) · Integration of new spatial data (e.g. species distributions) into government planning and regulatory databases Evidence of changes to operation of policy instruments, including: · Adoption of tools, findings, guidelines or recommendations into new policy instruments (planning processes, regulatory decisions) · Integration of new methods for assessing spatial data (e.g. species distributions) into government planning and regulatory databases Evidence of findings and tools adopted in major policy decisions including: · Adoption of findings into legislation or significant policy strategies · Inscription of protected area · New or targeted program or funding	Quantitative / documentary (assessed, not tested)	

Table A10 (cont): Summary of outcomes indicators and measures: changes to practice and policy

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Indicator code*
Inform decision-making:for-profit and non-government sectors			
Research processes and findings result in beneficial change to for-profit, business and private sector policy (e.g. private companies, financial sector consortium)	Has the project contributed to any of these other outcomes? · Resulted in beneficial change in the for-profit private sector?	Survey: multiple choice, confidence and fit for purpose	PPP81
	Uptake and/or reach of economic tools and measures produced contributing to greater transparency in economic decision-making e.g.: · citation in industry ESG strategies, environmental accounting assessments, etc · number, reach and economic impact of businesses adopting research tools and/or findings	Quantitative / documentary (assessed, not tested)	
Research findings result in beneficial change to non-governmental and not-for-profit policy (e.g. International Union for the Conservation of Nature, Indigenous corporations)	Were there any other outcomes (expected or unexpected) connected at least broadly with the project? Eg investment decisions, program design, etc?	Interview: narrative	PPP82
	Has the project contributed to any of these other outcomes? · Resulted in beneficial change in the non-governmental, not-for-profit policy sector?	Survey: multiple choice, confidence and fit for purpose	
	Evidence and number of cases of adoption in non-government planning or policy documents	Documentary (assessed not tested)	
Environmental consequences of changes to practices or policy			
Outcomes resulting from changes to practice and policy - general	Measures of environmental values within scope of changed management or practices e.g.: · number of threatened or culturally significant species under area affected by positive changes to action · coverage of threatened ecological community, wetland system or place affected by changed actions arising from research · coverage of protected area/IPA land positively influenced by research	Quantitative (assessed, not tested)	AR59, 60, 61, 62 OGA71, 72, 73, 74, 75 76, PPP78, 81, 82
Improved capabilities			
Links formed/improved between members of the research team, community organisations, Indigenous communities, conservation groups, land managers, for on ground action	Network map: interviewees asked to identify the different groups involved in research and describe the links between them and how relationships may have changed over time	Interview: interactive mapping exercise	CSE65
	Links between members of the research team and key stakeholders (such as Indigenous communities, conservation groups, etc) have improved	Survey: likert agreement and fit for purpose	
	Evidence of new networks, partnerships and capacities built to implement effective practice guided by research, e.g.: · number and coverage of Indigenous ranger groups engaged and/or benefiting from changes arising from research · number and coverage of community or landcare groups engaged and/or benefiting from changes arising from research · number and coverage of citizen scientist groups and networks established and/or made more effective through implementing research	Quantitative (assessed, not tested)	
New skills and training developed through research implemented	Do you think your project helped to develop the skillset of you and/or your partners to undertake this work? If yes, what types of skills were developed?	Interview: narrative	OGA74 (CSE66)

Table A11: Summary of outcomes indicators and measures: improved capabilities

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Indicator code*
Improved capabilities (continued)			
Increased support or legitimacy for community/citizen practices that help to achieve conservation goals Involved, engaged or affected stakeholders inspired/supported to engage further in independent research (e.g. citizen science programs, Indigenous knowledge projects)	Expanded capacity of people to undertake effective actions or further research, e.g.: · evidence of Indigenous ranger groups or communities developing further work to extend project · evidence of community members, landholders, etc growing networks to share or extend on research · evidence of citizen scientists undertaking effective independent actions or further research · evidence of extension of citizen scientist-based methods developed to new research projects or contexts	Quantitative / documentary (not tested)	AR63 CSE70
Improved or increased intercultural capacities, via collective training of researchers on-ground personnel, in community organisations, Indigenous communities, conservation groups, land managers	Intercultural capacities of the research team have improved	Survey: likert agreement and fit for purpose	CSE66
	Number of Indigenous community members newly engaging in research Number of research team members and number of early career researchers and students newly engaging in research with Indigenous partners	Quantitative (not tested)	CSE66

Table A12: Summary of impacts indicators and measures: environmental benefits

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Indicator code*
Environmental benefits			
Environmental impacts - general	Can you describe any environmental impacts related to your project?	Interview: narrative	REI84 REI87 REI89 REI90 REI91
Species recovery or avoided loss, including: Population increase and/or avoided loss in threatened species Avoided loss of biodiversity	The project has or will contribute to an increase in population trajectory for a threatened or significant species	Survey: likert likelihood, confidence and fit for purpose	REI87 REI90
	The project has or will contribute to avoiding loss of biodiversity	Survey: likert likelihood, confidence and fit for purpose	
	Measures of population increase in a threatened species; measures of avoided declines in threatened species population or distribution, avoided biodiversity loss. These could include: · Measurable improvement in population and/or condition metrics against baselines or from long-term monitoring trends · Likelihood of species persistence as assessed through Population Viability Analysis · Improvements in species or ecosystem trajectories, avoided decline or likelihood of avoided decline as assessed through expert elicitation	Modelled or monitored (assessed, not tested)	
Improved condition of places, including: Improvement in the maintenance and/or condition of an ecosystem, wetland, marine environment	The project has or will contribute to an improvement in condition of an ecosystem, wetland or marine environment	Survey: likert likelihood, confidence and fit for purpose	REI89
	Measures of project contribution to an ecosystem, wetland or marine environment, including measures and evidence of: · Improvement or avoided decline in function or condition of an ecosystem, wetland or marine environment · Improvement or avoided decline in extent of an ecosystem or threatened ecological community	Modelled or monitored (assessed, not tested)	REI89 REI91

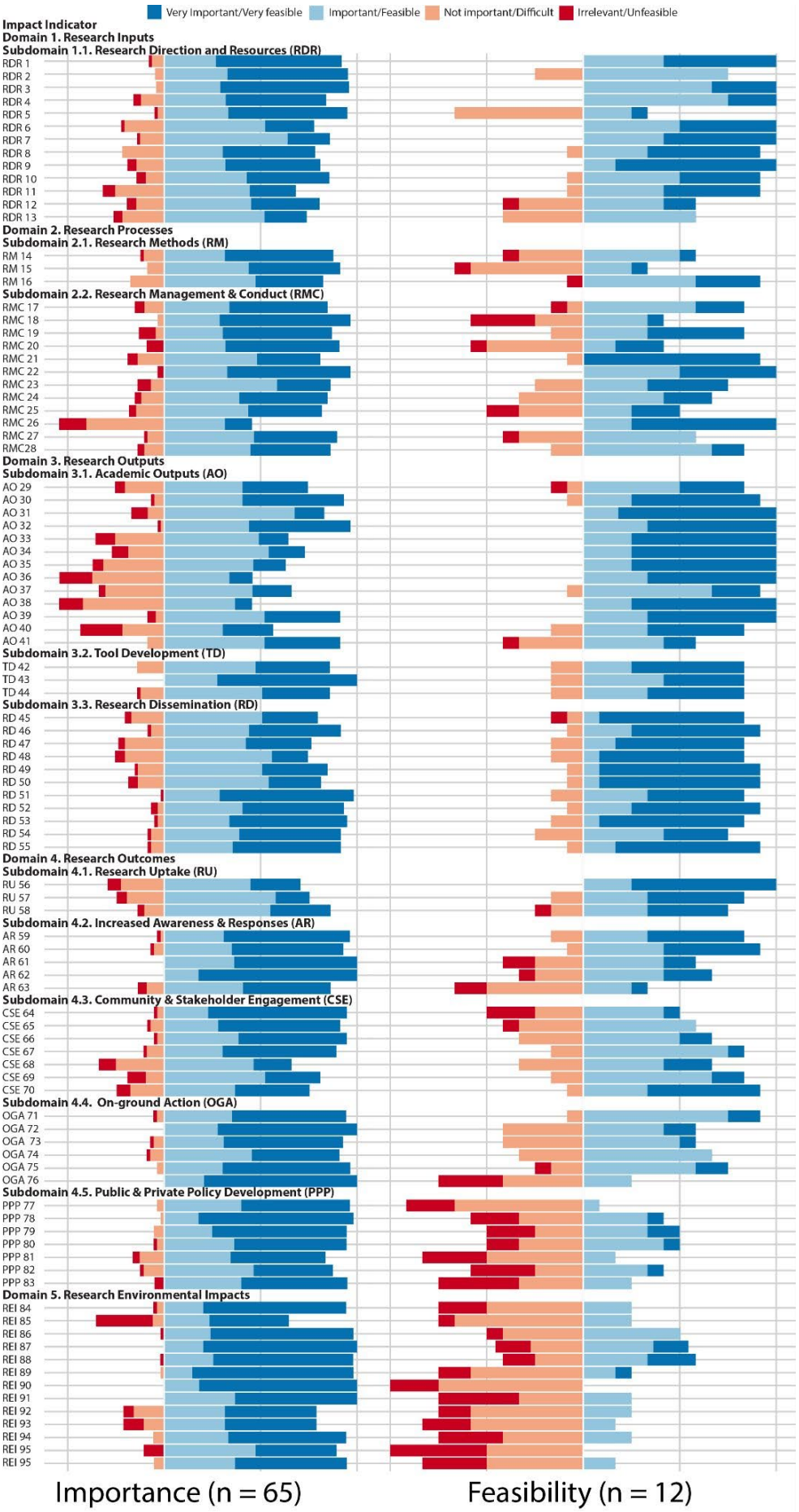
Table A13: Summary of impacts indicators and measures: governance and local capacities

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Indicator code*
Enhanced governance and capacities for locally-defined priorities and inclusive decision-making			
Improvement in the abilities of parties involved in or affected by conservation research to pursue locally defined priorities, objectives and goals (e.g. Indigenous communities, community conservation groups, natural resource management authorities)	Has the project contributed to improving the capacity of local communities to pursue their/your own locally-defined priorities, objectives and goals? In what ways?	Narrative	REI92
	The project has or will contribute to improvement in my group, community or organisation's capacity to pursue locally-defined priorities, objectives, and goals	Survey: likert likelihood, confidence, fit for purpose	
	Project contributions to improvement in local capacity to pursue locally-defined priorities, objectives, and goals? Evidence of: · application of research approaches or findings to other local priorities; · local plans developed that draw from research; · higher levels of participation in decision-making bodies of research participants.	Quantitative / documentary (assessed, not tested)	
	Project contributions to improvement in local capacity to pursue locally-defined priorities, objectives, and goals? · Use of community surveys, focus groups or participatory evaluations, where relevant	Community insight (assessed, not tested)	
Improved capacities of management, governance and institutions to engage in effective, equitable and informed deliberation and decision-making around conservation issues	Can you describe any changes that have taken place in how decision-making is undertaken in the region? Have any of these involved changes to how agencies (such as planners or management agencies) operate? Do you think these changes will be lasting?	Narrative	REI96
	The project has or will contribute to changes in the social or cultural processes or systems in place for managing species, ecosystems or places	Survey: likert likelihood, confidence, fit for purpose	
	Project contributions to improvements in capacities of agencies to engage in effective, equitable and informed decision-making: · Evidence of the application of good practice processes from the research to other decisions, or other decision-making contexts · Evidence of changes to policies, programs or processes improving decision-making for managing species, ecosystems or places as a result of research.	Quantitative / documentary (assessed, not tested)	
	Project contributions to improvements in capacities of agencies to engage in effective, equitable and informed decision-making: · Use of community surveys, focus groups or participatory evaluations, where relevant	Community insight (assessed, not tested)	
Social, cultural and economic benefits			
Social, cultural and economic impacts	Can you describe any societal, cultural or economic impacts the project may have contributed to?	Interview: narrative	REI94 REI95
A better informed society with greater social licence for conservation Influence on public policy debate	The project has or will contribute to greater public awareness of the importance of, options for or challenges of conservation	Survey: likert likelihood, confidence, fit for purpose	REI94 PPP77
	Project contribution to greater public awareness of the importance of, options for or challenges of conservation, including: Influence on public policy debate: · Documentary evidence of citation in reports from public inquiries Substantial direct engagement with findings by third parties, e.g.: · Documentary evidence of citation in opinion pieces and other public contributions to public debate that engages with research, written by people other than research partners · Significant syndication of media opinion pieces · Significant degree of likes and reshares of social media opinion pieces · Reach and engagement figures where available	Quantitative/ documentary (assessed, not tested)	

Table A13 (cont): Summary of impacts indicators and measures: social, cultural and economic benefits

Indicator descriptor	Measures/questions addressing indicators	Type of measures	Indicator code*
Social, cultural and economic benefits			
Improvement in human wellbeing derived from species and/or ecosystems, including access to resources, health, livelihoods, self-determination, social relations, and cultural and spiritual satisfaction: · Social and cultural measures of contribution to improvement in human wellbeing derived from species and/or ecosystems, including access to resources, health, livelihoods, self-determination, social relations, and cultural and spiritual satisfaction. · Economic contribution to improvement in human wellbeing derived from species and/or ecosystems, including access to resources and livelihoods.	The project has or will contribute to positive social or cultural outcomes for people	Survey: likert likelihood, confidence and fit for purpose	REI95
	The project has or will contribute to positive economic change	Survey: likert likelihood, confidence and fit for purpose	Not separately identified in original study
	Documentary evidence of changes to social or cultural programs · Inclusion of project in cultural, educational, health, job-creation or wellbeing contexts **Contribution to economic measures or decision-making** · Uptake and/or reach of economic tools and measures (e.g. evidence of use of research findings in independent environmental accounting assessments) **Documentary evidence of economic and social value for local communities, e.g.:** · jobs created · programs created · enhanced access to resources, e.g. regulatory changes, new industries · economic value of new or enhanced industries **External community satisfaction or wellbeing surveys, interviews or focus groups** Could explore as relevant: · program satisfaction measures as a proxy for direct social measures · health and psychological wellbeing benefits · educational benefits · wellbeing benefits derived from greater access to resources and livelihoods · improved social relations · sense of self-determination · sense of place · increased cultural and spiritual satisfaction **Measures or proxies of current or potential changes to ecosystem services, e.g.:** · water quality and water provisioning services · carbon sequestration · pollination services	Quantitative / documentary community insight / modelled or monitored (assessed, not tested)	REI95 Economic indicators not separately identified in original study, partially indicated in REI95

Appendix B: Importance and feasibility of full suite of indicators from Phase 1



Summary of indicators from Phase 1 study, grouped by importance and feasibility scores. Importance and feasibility scores are

averages of responses by questionnaire participants to the questions “How important do you think this impact indicator is?” (1 = irrelevant, 2 = not important, 3 = important, 4 = very important); and “How feasible do you think it is to measure this impact indicator?” (1 = unfeasible, 2 = difficult, 3 = feasible, 4 = very feasible). Scores are shaded according to a scale from red (lowest average score returned for an individual impact indicator = 2.41; 1.50) to green (highest average score returned for an individual impact indicator = 3.82; 3.83).

Table B1. Indicators ranked in top 50% for importance and feasibility		Importance (n = 65)	Feasibility (n = 12)
	Subdomain 1.1 Research Direction and Resources		
RDR 1	Research needs, gaps, unanswered questions, new areas for research identified	3.56	3.58
RDR 3	Agreement for working together established or re-iterated between scientists, decision-makers, practitioners, Indigenous partners	3.63	3.33
RDR 4	Active participation of stakeholders and end-users in research	3.32	3.25
	Subdomain 2.2 Research Management & Conduct		
RMC 19	Data storage, management, ownership, access, curation processes outlined, understood and agreed to by all parties	3.35	3.33
RMC 22	All participants and stakeholders contribute to designing or providing input to research questions	3.57	3.50
	Subdomain 3.1 Academic Outputs		
AO 30	Legitimate, valuable, and rigorous knowledge developed according to the various knowledge systems involved in the research	3.44	3.58
AO 32	Data sets (including spatial) made available to other researchers, practitioners and citizens/communities	3.47	3.67
	Subdomain 3.2 Tool Development		
TD 43	Tools made available to land managers, practitioners and citizens/communities for on-ground action (management plans, planning frameworks and processes)	3.72	3.25
	Subdomain 3.3 Research Dissemination		
RD 46	Research publications freely and openly accessible to all including research participants and partners	3.37	3.58
RD 51	Tailored reports, stories or summaries prepared on research and findings for specific audiences (e.g. land managers, Indigenous communities)	3.66	3.33
RD 52	Guidelines, guides, checklists, standards (for dispersed but like-minded audiences (e.g. parties who do environmental monitoring, all parties who participate in species translocations)	3.42	3.58
RD 53	Summaries for policy makers prepared (policy options papers, submissions to policy forums, etc)	3.54	3.58
RD 54	Contribution to public policy advisory committee(s)	3.42	3.08
RD 55	Presentations to schools, community groups and forums	3.46	3.67
	Subdomain 4.2 Increased Awareness & Responses		
AR 59	Species or ecological community listed as threatened and / or protected by legislative instrument	3.60	3.33
AR 60	Landscape or threatened ecological community heritage listed	3.49	3.42
	Subdomain 4.4 On-ground Action		
OGA 71	Findings from research incorporated into real world experimental tests or trials	3.52	3.08

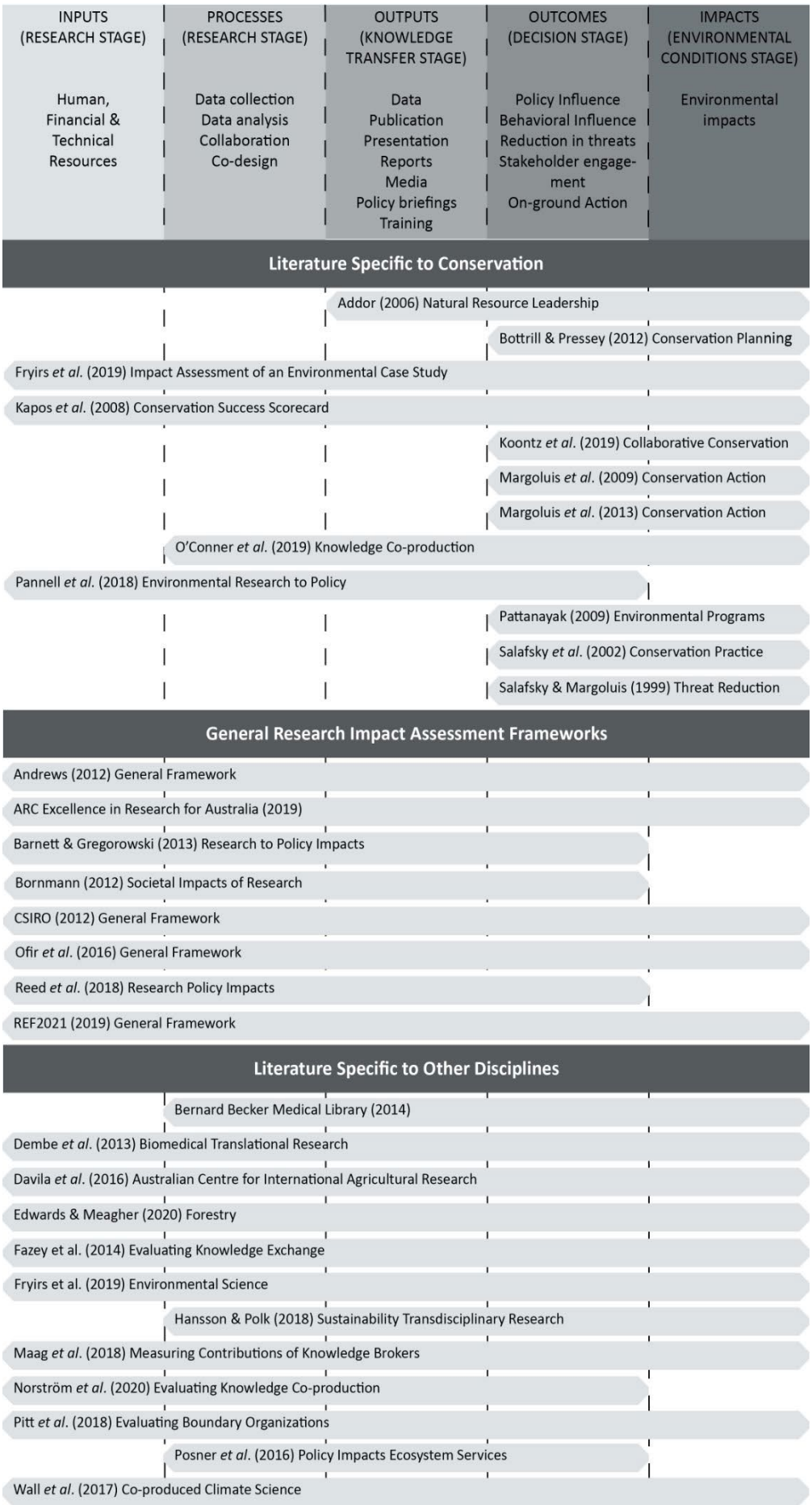
Table B2. Indicators ranked moderate for importance (50-75%), higher for feasibility (top 50%).		Importance (65)	Feasibility (12)
	Subdomain 1.1 Research Direction and Resources		
RDR 8	Existing datasets that were previously unavailable are liberated (e.g. industry or commercial in-confidence)	3.26	3.50
RDR 9	Qualified research or technical personnel recruited and including a variety of contextual and culturally relevant credentials	3.25	3.83
RDR 10	Skills and qualifications of existing and new research team improved or already excellent	3.24	3.33
	Subdomain 2.1 Research Methods		
RM 16	Best methods for data collection, management and analysis developed or maintained across disciplines and knowledge systems	3.18	3.17
	Subdomain 2.2 Research Management & Conduct		
RMC 17	Effective research networks and collaborations, opportunities for joint research	3.31	3.00
RMC 21	Contextually and culturally appropriate ways of enabling all participants to co-develop research and build partnerships are conducted	3.09	3.83
RMC 28	Processes to gain research permission (permits, ethics etc.) maintained, improved, made more meaningful and rigorous in relation to research content and processes, or streamlined where appropriate	3.24	3.00
	Subdomain 3.1 Academic Outputs		
AO 39	Collaborative research publications or outputs	3.26	3.67
	Subdomain 3.2 Tool Development		
TD 42	Tools and methods made available to other researchers and research groups	3.25	3.42
TD 44	Training packages developed for practitioners, citizens, community members and stakeholder groups	3.19	3.33
	Subdomain 3.3 Research Dissemination		
RD 49	Popular articles including magazine and newsletter articles written by researchers and journalists	3.17	3.75
	Subdomain 4.1 Research Uptake		
RU 58	Research findings perpetuated via popular articles, newsletters, fact sheets, reports written by third parties	3.14	3.08

Table B3. Indicators ranked high to moderate for importance (top 75%), lowest for feasibility (below 3).		Importance (65)	Feasibility (12)
	Subdomain 1.1 Research Direction and Resources		
RDR 2	Increase in reciprocal understanding between stakeholders and end-users around the need for, and framing of research	3.58	2.75
RDR 5	New understanding of values and priorities captured (e.g. organizational agencies, citizen science, Indigenous priorities)	3.55	2.42
RDR 12	Ability to attract research income or make in-kind contributions to future research increased or already excellent	3.11	2.67
	Subdomain 2.1 Research Methods		
RM 14	Tools, procedures and methods to co-design research are improved or already excellent	3.42	2.58
RM 15	New or improved methods for conducting studies spanning multiple disciplines and knowledge systems developed and or implemented	3.39	2.33
	Subdomain 2.2 Research Management & Conduct		
RMC 18	Trust built or maintained within research networks and collaborations	3.64	2.17
RMC 20	Obtaining Prior Informed Consent from Indigenous communities understood, developed, prioritised, excellent	3.41	2.58
RMC 24	Contextually and culturally appropriate forms of communication among all research team developed, improved or already excellent	3.27	2.92
RMC 25	Quality or quantity of teamwork across scientific disciplines and knowledge systems improved or already excellent	3.16	2.58
RMC 27	Barriers and problems with developing effective research processes or collaborations recognised and strategies for overcoming them maintained or developed	3.31	2.50
	Subdomain 3.1 Academic Outputs		
AO 41	Clear communication of failures, limitations, and challenges in the research	3.30	2.67
	Subdomain 4.2 Increased Awareness & Responses		
AR 61	Results trigger inscription of national park or conservation land tenure	3.64	2.58
AR 62	Adoption of more effective techniques for conservation practice (e.g. captive breeding, reintroduction in the wild)	3.82	2.83
AR 63	Increased support or legitimacy for community/citizen practices that help to achieve conservation goals	3.27	2.25
	Subdomain 4.3 Community & Stakeholder Engagement		
CSE 64	Mutually established and agreed upon ideas research and goals between researchers and community organisations, Indigenous communities, conservation groups, land managers	3.65	2.33
CSE 65	Links formed/improved between members of the research team, community organisations, Indigenous communities, conservation groups, land managers, for on ground action	3.53	2.50
CSE 66	Improved or increased intercultural capacities, via collective training of researchers on-ground personnel, in community organisations, Indigenous communities, conservation groups, land managers	3.49	2.83
CSE 67	Participation by government, community organisations, Indigenous communities, NGOs, land managers to co-design further research questions and methods	3.46	2.92
	Subdomain 4.4 On-ground Action		
OGA 72	Research findings, tools and guidelines adopted in practice and implemented in monitoring, management, community practices, regulation, business practices	3.72	2.75
OGA 73	Establishment of adaptive management trials, close coupled science-practice feedback loops	3.53	2.67
OGA 74	New skills and training developed through research implemented	3.35	2.67
OGA 75	Increase in protected area size and/or quality	3.63	2.83
OGA 76	Reduction in threatening process(es)	3.79	1.92

	Subdomain 4.5 Public & Private Policy Development		
PPP 77	Influence on public policy debate	3.53	1.83
PPP 78	Research findings help to inform the decision-making process and result in change to public policy, strategy and/or program design	3.79	2.25
PPP 79	Legislation implemented or changed	3.64	2.42
PPP 80	Regulations and/or Natural Resource Management Monitoring, Evaluation, Reporting and Improvement Framework (MERI Framework) improved	3.51	2.42
PPP 81	Research processes and findings result in beneficial change to for-profit, business and private sector policy (e.g. private companies, financial sector consortium)	3.29	1.83
PPP 82	Research findings result in beneficial change to non-governmental and not-for-profit policy (e.g. International Union for the Conservation of Nature, Indigenous corporations)	3.27	2.17
PPP 83	Research processes and findings contribute to and result in beneficial change in community deliberation and decision making	3.45	1.83
	Research Environmental Impacts		
REI 84	Improvement in condition of the environment	3.67	2.00
REI 85	Species and country healthy as measured by Indigenous-led definitions	3.16	2.17
REI 86	Increase in threatened species habitat/threatened ecological community area or quality	3.70	2.42
REI 87	Population increase or avoided declines in a threatened species or ecological community	3.80	2.55
REI 88	Improvement in the threatened status of a species or threatened ecological community	3.69	2.67
REI 89	Improvement in the condition of an ecosystem, wetland, marine environment	3.82	2.17
REI 90	Avoided loss of biodiversity	3.82	1.75
REI 91	Improvement or maintenance of the functioning of an ecosystem, wetland, marine environment.	3.63	1.83
REI 92	Improvement in the abilities of parties involved in or affected by conservation research to pursue locally defined priorities, objectives and goals (e.g. Indigenous communities, community conservation groups, natural resource management authorities)	3.21	2.08
REI 93	Improvement in support for performance, experience, and/or rejuvenation of contextually and culturally appropriate relationships between people and species and ecological communities	3.16	1.92
REI 94	A better informed society with greater social licence for conservation	3.56	1.92
REI 95	Improvement in human wellbeing derived from species and/or ecosystems, including access to resources, health, livelihoods, self-determination, social relations, and cultural and spiritual satisfaction	3.21	1.50
REI 96	Improved capacities of management, governance and institutions to engage in effective, equitable and informed deliberation and decision-making around conservation issues	3.53	1.83

Table B4. Indicators ranked lowest for importance (bottom 25%)		Importance (65)	Feasibility (12)
	Subdomain 1.1 Research Direction and Resources		
RDR 6	Research methods or techniques maintained, developed or extended	3.02	3.50
RDR 7	Methods or resources for collecting, storing, analysing data maintained or developed	3.06	3.58
RDR 11	Research equipment and facilities improved or already excellent	2.86	3.42
RDR 13	Contextually and culturally sensitive evaluation frameworks (including impact indicators) reviewed or developed for the research	2.91	2.58
	Subdomain 2.2 Research Management and Conduct		
RMC 23	Retention, continuity and evolution of research teams and partnerships achieved	3.07	3.17
RMC 26	Research team members abilities and opportunities to participate in other forums, contexts and roles, including organisational boards, societies, and across research, policy and practice, developed or improved	2.46	3.75
	Subdomain 3.1 Academic Outputs		
AO 29	New techniques and methods developed for generating and capturing knowledge	3.03	3.08
AO 31	Postgraduate students completed, training and certificates completed by research participants, capacity of young researchers increased	2.90	3.82
AO 33	Number of publications relating to the research authored or co-authored by researcher or team, and (if appropriate) their impact factors	2.69	3.75
AO 34	Presentations at regional, national and international workshops, conferences and gatherings	2.83	3.75
AO 35	Additions made to scientific collections	2.74	3.75
AO 36	Increased number of grant submissions by research team	2.41	3.67
AO 37	Researchers recognised for achievements and/or leadership in the field	2.83	3.17
AO 38	Study, project or research team recognised by award(s)	2.42	3.75
AO 40	Increased contributions to teaching by research participants	2.61	3.33
	Subdomain 3.3 Research Dissemination		
RD 45	Media and social media coverage for scientific publications authored or co-authored by researcher or team in the applicable time period	3.05	3.50
RD 47	Media and social media coverage for tools and other outputs (beyond research publications) by researcher or team in the applicable time period	3.07	3.50
RD 48	Websites, web pages, blogs produced	2.88	3.58
RD 50	Generic reports or stories prepared on research and results (e.g. fact sheets, videos, animations)	3.03	3.75
	Subdomain 4.1 Research Uptake		
RU 56	Academic citations for scientific publications authored or co-authored by researcher or team in the applicable time period	2.90	3.75
RU 57	Public and grey-literature citations for scientific publications authored or co-authored by researcher or team in the applicable time period	2.88	3.33
	Subdomain 4.3 Community and Stakeholder Engagement		
CSE 68	Emails, phone calls, social media messages enquiring about research outputs	2.77	2.92
CSE 69	Testimonies and reflections from relevant engaged and affected stakeholders about the conduct, value and significance of the research	3.00	3.00
CSE 70	Involved, engaged or affected stakeholders inspired/supported to engage further in independent research (e.g. citizen science programs, Indigenous knowledge projects)	3.07	3.50

Appendix C: Summary of related frameworks for assessing impact



Appendix D: Survey instrument for testing indicators

Q1 Please indicate which Threatened Species Recovery (TSR) Hub project you will be referencing in this survey. e.g. Project 7.1 – Methods for measuring uptake and outcomes from environmental research. If you are unsure, please refer to our previous correspondence where we have noted which project we are interested in.

Q2 Which best describes the type of institution you were primarily affiliated with in the TSR Hub project? (please select 1)

- Research
- Government
- Non-Government organisation
- Indigenous
- Other(please identify)

Q3 For how long would you estimate you have been involved in this TSR Hub project? (please select 1)

- Less than 1 year
- 1-3 years
- >3 years

Q4 Please describe the nature of your contributions to the TSR Hub project (pick as many as are applicable)

- Designing the project (e.g. defining initial research questions, what to measure where, etc)
- Providing advice (e.g. on who to involve, on outputs, on methods)
- Collecting data/information
- Analysing data
- Communicating about the project (e.g. discussing with community/others in your organisation, etc)
- Communication of outputs (e.g. writing reports, making videos, etc)
- Using the outputs

Q5 Were groups or organisations outside research institutions (like universities, CSIRO) directly involved in the project?

- Yes
- No

Q6 [if yes to Q5] What groups or organisations (other than research institutions) were directly involved in the project?

- Indigenous organisations (e.g. Land Councils, Aboriginal corporations, ranger groups)
- Individual land owners
- Government agencies
- Conservation non-government organisations (e.g. Bush Heritage, Australian Wildlife Conservancy)
- Community groups

- Businesses

DOMAIN 1: INPUTS

Q7 To what level do you agree / disagree with the following statements regarding the inputs to your project?
[Strongly Agree (5) Agree (4) Neither agree or disagree (3) Disagree (2) Strongly Disagree (1) N/A or Don't know (-)]

- Research needs or new areas of research were identified through the project
- The values and priorities of research users informed how the project was run
- The value of the research to different stakeholders (particularly end-user agencies) was elicited
- Research users and others based outside research institutions actively participated in the project
- Additional data sets from outside of the research team were made available to support the research

Q8 How fit for purpose do you think each of these input indicators are in terms capturing the value of the research project? *Note: an indicator may be important even if you think your project performed poorly against it

High Medium Low Not Applicable

- The research needs or new areas of research were identified.
- The research needs of different stakeholders (particularly end-user agencies) were elicited.
- [If 'yes' to Q5] Stakeholders [participants based outside research institutions] actively participated in the project
- Stakeholder values and priorities informed how the project ran
- Existing data sets to the research were made available

Q9 Please add any comments on the value of input indicators in the text box:

DOMAIN 2: PROCESSES

Q10 Did the research need to go through permission processes (permits, ethics, etc)?

- Yes
- No
- Unsure/Not Applicable

Q11 To what extent do you agree/disagree with the following statements regarding the research processes of your project:
[Strongly Agree (5) Agree (4) Neither agree or disagree (3) Disagree (2) Strongly Disagree (1) N/A or Don't know (-)]

- Knowledge of and/or capacity to engage in ethical research by researchers and/or research users was improved as a result of the project
- Methods used in the project drew on insights from different knowledge systems
- Trust was developed or improved among research partners

- [if Q6 showed 'Indigenous organisations (e.g. Land Councils, Aboriginal corporations, ranger groups)' directly involved in the project,and if yes to Q10]
  Obtaining Prior Informed Consent from Indigenous communities was prioritised in the project

Q12 Please pick the option that best characterises the methods used in your project:

- Data collection methods used in the project were:
  - Below standard
  - Standard
  - Above standard/new
  - Don't know
- Data management methods used in the project were:
  - Below standard
  - Standard
  - Above standard/new
  - Don't know
- Data analysis methods used in the project were:
  - Below standard
  - Standard
  - Above standard/new
  - Don't know

Q13 How fit for purpose do you think each of these research process indicators are in terms of capturing the value of the research project? *Note: an indicator may be important even if you think your project performed poorly against it

High Medium Low Not Applicable

- Processes to gain research permission (permits, ethics etc.) improved as a result of the project
- Methods used in the project drew on insights from different "knowledge systems"
- Trust built or maintained within research networks and collaborations
- [Provisional on selecting "Indigenous organisations" at Q6] Obtaining Prior Informed Consent from Indigenous communities was prioritised
- Standard of data collection methods
- Standard of data management methods

- Standard of data analysis methods

Q14 Please add any comments on the value of process indicators in the text box below:

DOMAIN 3: OUTPUTS

Q15 What type of project outputs have you seen or are aware of? (Choose as many as apply)

- technical publications (reports, peer-reviewed papers)
- presentations
- tools for implementation (e.g. planning frameworks, guidelines, checklists)
- data sets
- summaries and other outputs tailored for specific audiences (e.g. factsheets)
- training packages for practitioners, citizens, community members, etc
- collaborative outputs (may be any combination of the outputs displayed in this section)

Q16 To what extent do you agree / disagree with the following statements regarding the outputs from your project:
[Strongly Agree (5) Agree (4) Neither agree or disagree (3) Disagree (2) Strongly Disagree (1) N/A or Don't know (-)]

- Valuable knowledge was produced through the project
- Different "knowledge systems" informed the outputs of the project (agents, practices, and institutions that organize the production, transfer and use of knowledge i.e. policy, management, societal, Indigenous)
- Publications and other project outputs were produced in a collaborative manner
- Outputs from the project have been tailored to specific end-users (i.e. land managers, communities)
- Research outputs are freely accessible to all including research participants and partners

Q17 How fit for purpose do you think each of these research output indicators are in terms of capturing the value of the research project? *Note: an indicator may be important even if you think your project performed poorly against it

High Medium Low Not Applicable

- Valuable knowledge was produced through the project
- Different knowledge systems informed the outputs of the project
- Publications and other project outputs were produced in a collaborative manner
- Outputs from the project have been tailored to specific end-users (e.g. land managers, communities)
- Research outputs are freely accessible to all including research participants and partners

Q18 Please add any comments on the value of output indicators in the text box below:

DOMAIN 4: OUTCOMES

19. Please select which of the following outcomes related to on-ground practice your project has, or is likely to contribute. (Choose as many as apply)

- Improved monitoring for species, ecosystems, heritage places, etc
- Improved management of species, ecosystems, heritage places, etc
- Improved adaptive management or management planning for species, ecosystems, heritage places, etc
- Improved data management, reporting and/or analysis about species, ecosystems, heritage places, etc
- Reduction in threatening processes (6)

20. Please select which option(s) best apply to the scale of the outcomes your project has, or is likely to contribute (You can select one or both)

[Change within project region or partnership; Change beyond project region or partnership]

- Improved monitoring for species, ecosystems, heritage places, etc
- Improved management of species, ecosystems, heritage places, etc
- Improved adaptive management planning for species, ecosystems, heritage places, etc
- Improved data management, reporting and / or analysis about species, ecosystems, heritage places, etc
- Reduction in threatening processes

21. Has the project contributed to any of these other outcomes? (Choose as many as apply)

- Helped to inform decision-making processes related to public policy, strategy and/or program design?
- Resulted in beneficial change in the for-profit private sector?
- Resulted in beneficial change in the non-governmental, not-for-profit policy sector?

22. Outcomes can manifest in the months and years after a research project has finished. Can you indicate your level of confidence in your responses about outcomes based on the evidence available to you? limited = a hunch based on your knowledge of the context; medium = verbal or other indications the outcome will occur, such as citation of the change in a proposal or plan; robust = outcome has already occurred, process is in train for it to occur. (e.g. legislation being reviewed)

[Limited Medium Robust]

- Improved monitoring for species, ecosystems, heritage places, etc
- Improved management of species, ecosystems, heritage places, etc
- Improved adaptive management planning for species, ecosystems, heritage places, etc
- Improved data management, reporting and/or analysis about species, ecosystems, heritage places, etc
- Reduction in threatening processes

- Helped to inform decision-making processes related to public policy, strategy and/or program design?
- Resulted in beneficial change in the for-profit private sector?
- Resulted in beneficial change in the non-governmental, not-for-profit policy sector?

23. To what level do you agree/disagree with the following statements regarding the outcomes of your project?
[Strongly Agree (5) Agree (4) Neither agree or disagree (3) Disagree (2) Strongly Disagree (1) N/A or Don't know (-)]

- Links between members of the research team and key stakeholders (such as Indigenous communities, conservation groups, etc) have improved
- Intercultural capacities of the research team have improved

24. How fit for purpose do you think each of these research outcome indicators are in terms of capturing the value of the research project? *Note: an indicator may be important even if you think your project performed poorly against it

High Medium Low Not Applicable

- Improved monitoring for species, ecosystems, heritage places, etc
- Improved management of species, ecosystems, heritage places, etc
- Improved adaptive management planning for species, ecosystems, heritage places, etc
- Improved data management, reporting and / or analysis about species, ecosystems, heritage places, etc
- Reduction in threatening processes
- Helped to inform decision-making processes related to public policy, strategy and / or program design
- Resulted in beneficial change in the for-profit, private sector
- Resulted in beneficial change in the non-governmental, not-for-profit policy sector
- Links between members of the research team and key stakeholders (such as Indigenous communities, conservation groups, etc) have improved
- Intercultural capacities of the research team have improved

25. Please add any comments on the value of outcome indicators in the text box below:

DOMAIN 5: IMPACTS

26. How likely do you think the following impacts are for your project?
[Strongly Agree (5) Agree (4) Neither agree or disagree (3) Disagree (2) Strongly Disagree (1) N/A or Don't know (-)]

- The project has or will contribute to avoiding loss of biodiversity
- The project has or will contribute to an increase in population trajectory for a threatened or significant species
- The project has or will contribute to an improvement in condition of an ecosystem, wetland or marine environment

· The project has or will contribute to greater public awareness of the importance of, options for or challenges of conservation

· The project has or will contribute to changes in the social or cultural processes or systems in place for managing species, ecosystems or places

· The project has or will contribute to improvement in my group, community or organisation's capacity to pursue locally-defined priorities, objectives, and goals

· The project has or will contribute to positive social or cultural outcomes for people

· The project has or will contribute to positive economic change

27. Impacts can manifest in the months and years after a research project has finished. Can you indicate your level of confidence in your responses about impacts based on the evidence available to you? limited = a hunch based on your knowledge of the context; medium = verbal or other indications the impact has occurred or will occur; robust = impact has already occurred, (e.g. changes have been observed or measured)

[Limited Medium Robust]

· The project has or will contribute to an increase in population trajectory for a threatened or significant species

· The project has or will contribute to avoiding loss of biodiversity

· The project has or will contribute to an improvement in condition of an ecosystem, wetland or marine environment

· There has been significant improvement in your group to define priorities, objectives, and goals

· The project has or will contribute to greater public awareness of the importance of options for or challenges of conservation

· The project will contribute to changes in the social or cultural processes or systems in place for managing species, ecosystems or places

· The project has or will contribute to positive social or cultural outcomes for people

· The project has or will contribute to positive economic change

28. How fit for purpose do you think each of these research impact indicators are in terms of capturing the value of the research project? *Note: an indicator may be important even if you think your project did poorly against it - The project has or will contribute to avoiding loss of biodiversity

High Medium Low Not Applicable

· The project has or will contribute to avoiding loss of biodiversity

· The project has or will contribute to an increase in population trajectory for a threatened or significant species

· The project has or will contribute to an improvement in condition of an ecosystem, wetland or marine environment

· The project has or will contribute to greater public awareness of the importance of options for or challenges of conservation

· The project has or will contribute to changes in the social or cultural processes or systems in place for managing species, ecosystems or places

· The project has or will contribute to improvement in my group, community or organisation's capacity to pursue locally-defined priorities, objectives, and goals as a result of the project

· The project has or will contribute to positive social or cultural outcomes for people

· The project has or will contribute to positive economic change

Q29. Please add any comments on the value of impact indicators in the text box below:

Appendix E: Interview guide

Domain	Indicator	Question/prompt
		Q1. Did you complete the survey? Is it OK to focus on **project X** today? (pref focus on same project as referenced in survey)
All	-	**Q2.** Can you please describe the nature of your involvement in **project X**?
Inputs & outcomes View options /annotate	CSE65	**Q3**. To help get an idea of who was involved in the project I'd like to capture each organisation or group involved, starting with your own. Now can I get you to draw lines to link which ones interacted with each other. Now on each line can you estimate of how strong the links between the organisations were at the beginning vs [now] 0=no link; 1 is some link 2 is strong link [if many focus on key ones]
inputs	RDR4 RDRX RDR5	**Q4.** How did non-academic (non-research)/your and other non-academic groups participate in the project? P1 Articulated the need for research? P2 Their values & priorities informed how the project was run? P3 How actively did they participate?
inputs	RDR1	**Q5.** How were research needs identified in the project?
process	RMC18	**Q6.** Do you feel trust was developed and /or improved among partners? P1 could you give an example to illustrate?
process	RMC20	**Q7. If Indigenous partners in original map.** How was informed consent obtained from Indigenous partners?
inputs	RDR8	**Q8.** Were there existing data sets/information necessary to do the research? If yes, were those made available? If not, why not & what was the consequence?
process	RM16	**Q9.** What is your view on the research methods used in the project? P1 Did it draw on different disciplines & knowledge systems?
outputs	AO30	**Q10.** How would you characterise the knowledge produced by the project? P1 (**depending on map**) How did different knowledge systems* contribute to the project?

outputs	AO39	*Outputs:products of research eg Data Presentations, docs, Media, Students, Workshops* Q11. Can you describe how project outputs were produced, particularly who was involved?
outputs	TD43	Q12. What do you think about how knowledge generated by the project has been packaged, presented and disseminated?
outcomes	CSE66	*OK moving to outcomes WE see these as Influence on: policy, action, and behaviour* Q13. Do you think your project helped to develop the skillset of you and/or your partners to undertake this work? If yes, what types of skills were developed?
outcomes	AR59-62 PPP78	Q14. Can you describe any changes to practice or policy the research may have contributed to? E.g. informed management actions, contributed to listing or policy.
outcomes	PP78, 81, 82	Q15. Were there any other outcomes (expected or unexpected) connected at least broadly with the project? Eg investment decisions, program design, etc?
impacts	REI87 REI89 REI90	We consider impacts to be broader than change in the immediate partners or organisations ie An effect on, change or benefit to the environment, society, culture, economy and/or research Q.16 Keeping this in mind, can you describe any environmental impacts re your project, for example: P1 Population increase or avoided declines P2 Improvement in ecosystem condition P3 Avoiding biodiversity loss
impacts	REI94 REI95	Q17. Can you describe any societal, cultural or economic impacts the project may have contributed to? P1 better informed society with greater social licence for conservation P2 Improved human wellbeing
overarching		Q18a) Which did you feel to be the most relevant questions for your research? Why? 18b Which did you feel to be the least relevant? Why?
overarching		Q19. Do you have any suggestions as to how we could improve upon our approach to measuring research impact? P1 Things we've missed out? P2 How it is designed (does it need to be tailored/co-designed?) P3 Different methods to collect data?
Overarching CSE69		Q20. Do you have any other observations or suggestions you would like to make?

Acknowledgements

We would like to acknowledge and thank all the people who have contributed their time, expertise, research case studies and insights throughout the development of this project, particularly through workshops, surveys, interviews, and the project steering committee. We would like particularly to acknowledge Leonie Valentine for co-developing the workshop that initiated the project, Carina Wyborn for detailed critiques of the suite of impact indicators, Dan Rogers and Vanessa Westcott for critical input at the scoping stages, and the Science Partnerships Section in the Australian Government Department of Agriculture, Water and the Environment for guidance in the development and conceptualisation of the project.

Ethics statement

The ethical aspects of this research have been approved by the ANU Human Research Ethics Committee (Protocol 2019/895).

References

AIATSIS. (2012). *Guidelines for Ethical Research in Australian Indigenous Studies*. Retrieved from Canberra: https://aiatsis.gov.au/sites/default/files/2020-09/gerais.pdf

Australian Research Council. (2017). *Engagement and Impact Assessment Pilot 2017 Report*. Australian Government Retrieved from https://www.arc.gov.au/engagement-and-impact-assessment/ei-pilot-overview

Australian Research Council. 2019a. Engagement and Impact Assessment 2018–19 National Report. Australian Research Council.

---. 2019b. Engagement and Impact Assessment 2018-19: Impact Studies. (14 January 2020; https://dataportal.arc.gov.au/EI/Web/impact/ImpactStudies)

Barnett, C., & Gregorowski, R. (2013). Learning about theories of change for the monitoring and evaluation of research uptake. *IDS Practice Paper In Brief 4*.

Bayraktarov, E., Ehmke, G., Tulloch, A. I. T., Chauvenet, A. L., Avery-Gomm, S., McRae, L., . . . Possingham, H. P. (2021). A threatened species index for Australian birds. *3*(2), e322. doi:https://doi.org/10.1111/csp2.322

Charterpoint. (2018). *National Environmental Science Program: Mid-Term Evaluation Final Report*. Retrieved from https://www.environment.gov.au/system/files/pages/2f561690-b47e-4bf2-b028-d18739b3486f/files/nesp-mid-term-evaluation-final-report.pdf

CSIRO. 2020. Impact evaluation guide. The Commonwealth Scientific and Industrial Research Organisation.

Conservation Measures Partnership. 2018. CMP Conservation Actions Classification v 2.0. CMP. https://www.conservationmeasures.org/about-cmp/

Chyung, S Y, Roberts, K, Swanson, I, & Hankinson, A. (2017). Evidence-based survey design: The use of a midpoint on the Likert scale. Performance Improvement, 56(10), 15-23. doi:10.1002/pfi.21727

Davila F, Sloan T, van Kerkhoff L. 2016. Knowledge Systems and RAPID Framework for Impact Assessments. *Australian Centre for International Agricultural Research*.

Davies, S. R. (2013). The rules of engagement: Power and interaction in dialogue events. *Public Understanding of Science*, 22(1), 65-79.

Evely, Anna C., Michelle Pinard, Mark S. Reed, and Ioan Fazey (2011). 'High Levels of Participation in Conservation Projects Enhance Learning'. Conservation Letters 4, no. 2: 116–26. https://doi.org/10.1111/j.1755-263X.2010.00152.x.

Deeming S, Searles A, Reeves P, Nilsson M. (2017). Measuring research impact in Australia's medical research institutes: a scoping literature review of the objectives for and an assessment of the capabilities of research impact assessment frameworks. Health Research Policy and Systems 15: 22.

Donovan, C. (2011). State of the art in assessing research impact: introduction to a special issue. Research Evaluation, 20(3), 175-179.

Geyle, H. M., Garnett, S. T., Legge, S. M., & Woinarski, J. C. Z. (2019). *Report to Office of the Threatened Species Commissioner: 3-year review of progress on priority bird and mammal species* Retrieved from https://www.environment.gov.au/system/files/resources/03dafd44-3db0-40b0-abfc-7420da95ce7c/files/3-year-review-progress-priority-bird-mammal-species.pdf

Gibbons, P., Zammit, C., Youngentob, K., Possingham, H. P., Lindenmayer, D. B., Bekessy, S., . . . Restoration. (2008). Some practical suggestions for improving engagement between researchers and policy-makers in natural resource management. *9*(3), 182-186.

Hanski, I. (2000). Extinction debt and species credit in boreal forests: modelling the consequences of different approaches to biodiversity conservation. *Annales Zoologici Fennici, 37*(4), 271-280.

Hansson, S., & Polk, M. (2018). Assessing the impact of transdisciplinary research: The usefulness of relevance, credibility, and legitimacy for understanding the link between process and impact. *Research Evaluation, 27*(2), 132-144. doi:10.1093/reseval/rvy004 %J Research Evaluation

Hourihan, M., & Parkes, D. (2016). *Federal R&D budget trends: A short summary*. Retrieved from https://www.aaas.org/sites/default/files/2019-01/AAAS%20RD%20Primer%202019_2.pdf

Jackson, S. T., & Sax, D. F. (2010). Balancing biodiversity in a changing environment: extinction debt, immigration credit and species turnover. *Trends Ecol Evol, 25*(3), 153-160. doi:10.1016/j.tree.2009.10.001

Jarvis, Rebecca M., Stephanie B. Borrelle, Natalie J. Forsdick, Katharina-Victoria Pérez-Hämmerle, Natalie S. Dubois, Sean R. Griffin, Angela Recalde-Salas, et al. "Navigating Spaces between Conservation Research and Practice: Are We Making Progress?" *Ecological Solutions and Evidence* 1, no. 2 (2020): e12028. https://doi.org/10.1002/2688-8319.12028.

Kapos, V., Balmford, A., Aveling, R., Bubb, P., Carey, P., Entwistle, A., . . . Manica, A. (2008). Calibrating conservation: new tools for measuring success. *1*(4), 155-164. doi:https://doi.org/10.1111/j.1755-263X.2008.00025.x

Keith, H., Vardon, M., Stein, J., Stein, J., Lindenmayer, D.B. (2017). Experimental Ecosystem Accounts for the Central Highlands of Victoria - Summary document. National Environmental Science Program Threatened Species Recovery Hub.

Kuruvilla S, Mays N, Pleasant A, Walt G. (2006). Describing the impact of health research: a Research Impact Framework. BMC Health Services Research 6: 134.

Kuussaari M, Bommarco R, Heikkinen RK, Helm A, Krauss J, Lindborg R, Ockinger E, Pärtel M, Pino J, Rodà F, Stefanescu C, Teder T, Zobel M, Steffan-Dewenter I. Extinction debt: a challenge for biodiversity conservation (2009). Trends Ecol Evol. Oct; 24(10):564-71. doi: 10.1016/j.tree.2009.04.011. Epub 2009 Aug 6. PMID: 19665254.

Lacey, J., Howden, M., Cvitanovic, C., & Colvin, R. M. (2018). Understanding and managing trust at the climate science–policy interface. Nature Climate Change, 8(1), 22-28.

Lavery, T. H., Morgain, R., Fitzsimons, J. A., Fluin, J., Macgregor, N. A., Robinson, N. M., . . . Vuong, H. (2021). Impact Indicators for Biodiversity Conservation Research: Measuring Influence within and beyond Academia. BioScience, 71(4), 383-395.

Maron, M. and Evans, M. (2018). Estimating the benefits of offsetting: Expert elicitation. National Environmental Science Program Threatened Species Recovery Hub.

Mervis J. (2020). NSF rolls out huge makeover of science statistics. Science 367:352–353.

Miller, Clark A., and Carina Wyborn (2020). "Co-Production in Global Sustainability: Histories and Theories." *Environmental Science & Policy*, Into the fray. Strategic perspectives on biodiversity sciences and politics, 113 (November 1, 2020): 88–95. https://doi.org/10.1016/j.envsci.2018.01.016.

Moggridge, B. 2020. Indigenous Engagement Protocols for Threatened Species Researchers. National Environmental Science Program Threatened Species Recovery Hub.

Morris ZS, Wooding S, Grant J. 2011. The answer is 17 years, what is the question: understanding time lags in translational research. Journal of the Royal Society of Medicine 104:510–520.

Muashekele, Chris, Heike Winschiers-Theophilus, and Gereon Koch Kapuire. "Co-Design as a Means of Fostering Appropriation of Conservation Monitoring Technology by Indigenous Communities." In *Proceedings of the 9th*

International Conference on Communities & Technologies - Transforming Communities, 126–30. C&T '19. New York, NY, USA: Association for Computing Machinery, 2019. https://doi.org/10.1145/3328320.3328383.

Nel, Jeanne L., Dirk J. Roux, Amanda Driver, Liesl Hill, Ashton C. Maherry, Kate Snaddon, Chantel R. Petersen, Lindie B. Smith-Adao, Heidi Van Deventer, and Belinda Reyers. "Knowledge co-production and boundary work to promote implementation of conservation plans." *Conservation Biology* 30, no. 1 (2016): 176–88. https://doi.org/10.1111/cobi.12560.

O'Connor, R A, Nel, J L, Roux, DJ, Lim-Camacho, L., van Kerkhoff, L., & Leach, J. (2019). Principles for evaluating knowledge co-production in natural resource management: Incorporating decision-maker values. Journal of Environmental Management, 249, 109392. doi:10.1016/j.jenvman.2019.109392

Otter.ai (2021). Transcription software. Retrieved from: https://otter.ai/

Ovseiko PV, Oancea A, Buchan AM. 2012. Assessing research impact in academic clinical medicine: a study using Research Excellence Framework pilot impact indicators. BMC Health Services Research 12: 478.

Pannell DJ, Alston JM, Jeffrey S, Buckley YM, Vesk P, Rhodes JR, McDonald-Madden E, Nally S, Goucher G, Thamo T. 2018. Policy-oriented environmental research: What is it worth? Environmental Science and Policy 86:64–71.

Papageorgiou, A., McLaren, T., Schoep, T., & on behalf of the AAMRI Research Impact Working Group. (2021). *The AAMRI Research Impact Framework Final Report*. Retrieved from https://www.aamri.org.au/wp-content/uploads/2021/05/AAMRI_Impact_Report-2021_WEB.pdf

Penfield T, Baker MJ, Scoble R, Wykes MC. 2014. Assessment, evaluations, and definitions of research impact: A review. Research Evaluation 23: 21–32.

Radicchia F, Fortunatoa S, Castellano C. 2008. Universality of citation distributions: Toward an objective measure of scientific impact. Proceedings of the National Academy of Sciences 105: 17268–17272.

Reale, E., Avramov, D., Canhial, K., Donovan, C., Flecha, R., Holm, P., . . . Van Horik, R. (2017). A review of literature on evaluating the scientific, social and political impact of social sciences and humanities research. *Research Evaluation, 27*(4), 298-308. doi:10.1093/reseval/rvx025 %J Research Evaluation

Reed, M. S., L. C. Stringer, I. Fazey, A. C. Evely, and J. H. J. Kruijsen. "Five Principles for the Practice of Knowledge Exchange in Environmental Management." *Journal of Environmental Management* 146 (December 15, 2014): 337–45. https://doi.org/10.1016/j.jenvman.2014.07.021.

Research Excellence Framework (REF). 2011. Decisions on Assessing Research Impact. Available at: https://www.ref.ac.uk/2014/media/ref/content/pub/decisionsonassessingresearchimpact/01_11.pdf

Research Excellence Framework (REF). 2019. Guidance on submissions. Available at: https://www.ref.ac.uk/publications/guidance-on-submissions-201901/

Rickards, L., Steele, W., Kokshagina, O. and Morales, O. (2020) Research Impact as Ethos.RMIT University, Melbourne, Australia. https://cur.org.au/cms/wp-content/uploads/2020/09/rickards-et-al-2020-research-impact-as-ethos.pdf

Scharnhorst A, Garfield E. (2010). Tracing scientific influence. Dynamics of Socio-Economic Systems 2: 1–33.

Schreier, M. (2012). Qualitative content analysis in practice. London: Sage publications.

Sharp, S., & Gould, L. (2014). *ACT region vegwatch manual : vegetation and habitat condition assessment and monitoring for community / Sarah Sharp and Lori Gould*. Canberra: Molonglo Catchment Group.

Stojanovic, D., Cook, H. C. L., Sato, C., Alves, F., Harris, G., McKernan, A., . . . Heinsohn, R. (2019). Pre-emptive action as a measure for conserving nomadic species. *83*(1), 64-71. doi:https://doi.org/10.1002/jwmg.21575

Tsey, K., Lawson, K., Kinchin, I., Bainbridge, R., McCalman, J., Watkin, F., . . . Rossetto, A. (2016). Evaluating Research Impact: The Development of a Research for Impact Tool. *4*(160). doi:10.3389/fpubh.2016.00160

Tsey, K., Onnis, L.-a., Whiteside, M., McCalman, J., Williams, M., Heyeres, M., . . . Kinchin, I. (2019). Assessing research impact: Australian Research Council criteria and the case of Family Wellbeing research. *Evaluation and Program Planning, 73*, 176-186. doi:https://doi.org/10.1016/j.evalprogplan.2019.01.004

Hel, Sandra van der (2016). "New Science for Global Sustainability? The Institutionalisation of Knowledge Co-Production in Future Earth." *Environmental Science & Policy* 61 (July 1, 2016): 165–75. https://doi.org/10.1016/j.envsci.2016.03.012.

Van Kerkhoff, Lorrae, and Louis Lebel (2015). "Coproductive Capacities: Rethinking Science-Governance Relations in a Diverse World." *Ecology and Society* 20, no. 1 (January 28, 2015). https://doi.org/10.5751/ES-07188-200114.

van Kerkhoff, L. (2018). *Research impact over the long term: narratives, numbers and networks*. Paper presented at the Australasian Aid Conference, Australian National University, Canberra. https://devpolicy.org/2018-Australasian-Aid-Conference/Presentations/LorraevanKerkhoff.pdf

Vieira ES, Gomes JANF. 2010. A research impact indicator for institutions. Journal of Informetrics 4: 581–590

Wardropper CB, Dayer AA, Goebel MS, Martin VY (2021). Conducting conservation social science surveys online. Conservation Biology;1–9. https://doi.org/10.1002/cobi.13747.

Watt, I. (2016). Review of Research Policy and Funding Arrangements-Report, Department of Education and Training. https://apo.org.au/sites/default/files/resource-files/2016-05/apo-nid63694.pdf

Watts, K., Whytock, R. C., Park, K. J., Fuentes-Montemayor, E., Macgregor, N. A., Duffield, S., & McGowan, P. J. K. (2020). Ecological time lags and the journey towards conservation success. *Nature Ecology & Evolution, 4*(3), 304-311. doi:10.1038/s41559-019-1087-8

Ward, N., Garrard, G. E., Emily A. Gregg, May, B., Wandin, D., Harrison, M., . . . Bekessy, S. A. (2021). "Totemic species" can be an effective lens for engaging students with Indigenous knowledge and biodiversity conservation "Manuscript submitted for publication".

Weiss , A. P. (2007). Measuring the Impact of Medical Research: Moving From Outputs to Outcomes. *American Journal of Psychiatry, 164*(2), 206-214. doi:10.1176/ajp.2007.164.2.206

Wilsdon J, Allen L, Belfiore E, Campbell P, Johnson B. (2015). The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management. https://responsiblemetrics.org/wp-content/uploads/2019/02/2015_metrictide.pdf

Woodward, E., Hill, R., Harkness, P. and R. Archer (Eds) (2020). Our Knowledge Our Way in caring for Country: Indigenous-led approaches to strengthening and sharing our knowledge for land and sea management. Best Practice Guidelines from Australian experiences. NAILSMA and CSIRO. https://www.csiro.au/en/research/indigenous-science/Indigenous-knowledge/Our-Knowledge-Our-Way/OKOW-resources

Wyborn, Carina (2015). "Connecting Knowledge with Action through Coproductive Capacities: Adaptive Governance and Connectivity Conservation." *Ecology and Society* 20, no. 1 (January 28, 2015). https://doi.org/10.5751/ES-06510-200111.

Further information:
http://www.nespthreatenedspecies.edu.au

This project is supported through funding from the
Australian Government's National Environmental Science Program.

Threatened Species Recovery Hub

National Environmental Science Programme