AIATSIS

AUSTRALIAN INSTITUTE OF ABORIGINAL AND TORRES STRAIT ISLANDER STUDIES

Sharing success, measuring impact

Annotated bibliography

Jacqueline Battin and Allister Mills

Contents

Introduction

Assessing research impact is a critical element of AIATSIS' aspirations to collaborate with Indigenous partners and create meaningful change within communities. Through assessing the impact of AIATSIS' research, we can understand whether we are meeting the priorities of the Indigenous communities we work with, and our research can evolve to better support these priorities. However, how impact is defined and assessed is not clear and nor is there an established methodology for assessing the unique research that is carried out at AIATSIS. We have received positive feedback for our research activities, but this feedback is mostly anecdotal and has not been able to articulate the interlinkages between research, storytelling and its contribution to cultural confidence or its reinvigoration of cultural practices, governance forms and languages.

In reviewing the literature of this space, we found a great deal of variation in perspectives. We examined impact evaluation guides of both research bodies and social change organisations, papers that compared impact evaluation frameworks and methods, debates that surround the space (such as whether qualitative or quantitative methods should be used), and case studies that utilised different methodologies. We did not restrict our research to impact evaluation or to research impact, but also looked at the literature on evaluation in general, as this will help in developing a methodology for assessing our own research impact.

This annotated bibliography lists a selection of research impact literature and is organised into three parts. First, why is it important to evaluate research impact? Why should we be investing time and money into assessing it? While much of the literature on research impact addresses these preliminary questions, the items in this section are particularly useful in understanding current thinking surrounding research impact and why there is an imperative to understand and gather evidence of it. The literature points to the recognition of thinking broadly about research impact to encompass societal, cultural, health and wellbeing, and environmental changes beyond monetary or academic valuations oriented to publication outcomes.

Second, we explore literature on incorporating Indigenous perspectives in research impact evaluation. As indicated above, it is important to develop a methodology that appropriately focuses on the perspectives and experiences of Indigenous stakeholders with whom AIATSIS collaborates. How do we do this? How are other Indigenous organisations addressing this question? What are the issues involved in this process? How can we ensure our methodology meets ethical guidelines? There are few examples of literature that look at evaluating the impact of Indigenous research (see Tsey et al. 2016 for a notable exception), particularly papers that use a decolonised methodology – a methodology that engages with Indigenous peoples as co-researchers, focuses on their interests and priorities, and values Indigenous approaches and ways of knowing and doing. The examples provided in this bibliography are relevant to AIATSIS' research and explore multiple elements of the research process, focusing on knowledge production and reproduction via collections and archives as well as focusing more broadly on research as a form of narrative itself.

Third, what approaches and methodologies are used to measure impact and how valuable are they? What are the various approaches to evaluating impact and what methods are used for collecting data? Here, we examine literature that describes current approaches and methods to evaluating impact. These approaches include:

- Logic models and theories of change
- Economic valuations (particularly cost-benefit analysis and Social Return on Investment (SROI))
- Tools to measure wellbeing
- Surveys and questionnaires
- Participatory methods
- Interviews, narratives, and storytelling
- Most Significant Change (MSC)

As there is quite a lot of overlap, we have grouped overarching frameworks and data collection methods together. We begin by considering the more conventional approaches, and then move to approaches that might serve to address their weaknesses. We then list a number of papers which examine multiple methods.

We found that much of the literature highly regards logic models and theories of change (using diagrams to set out the inputs, outputs, outcomes, and impacts of a project), economic valuations (assigning monetary values on outcomes), and quantitative methods to assess impact (for example, see CSIRO 2015; Panel on Return on Investment in Health Research 2009; Productivity Commission 2010). While there are a number of benefits to using these approaches (see Table 1 below), there are numerous issues in applying them to the research AIATSIS carries out (see Arvidson et al. 2010; Fujiwara & Campbell 2011; Marsh et al. 2016; Maughan 2012; Wavell et al. 2002). For example, they restrict the ability of the researchers to capture the perspectives of the stakeholders and the people most impacted by the research, as they often utilise predetermined top-down categories of impacts rather than create the necessary room for respondents to express their experiences, and they simplify or homogenise these experiences.

In seeking a methodology for evaluating research impact that is in accordance with the principles of the Guidelines for Ethical Research in Australian Indigenous Studies (GERAIS) and in collaboration with communities (see Smith 2012),[1] we found that qualitative and participatory approaches were more appropriate. While many approaches seem to choose indicators to measure outcomes through researchers hypothesising potential impacts amongst themselves, Tsey et al. (2016) and Wavell et al. (2002) argue that indicators should be chosen and defined according to the perspectives of the people impacted by the project. Tsey et al. emphasise indicators and 'assessment domains' must reflect the values, interests, and aspirations of Indigenous peoples, and Wavell et al. argue identifying these indicators is best achieved through qualitative approaches. A number of authors (Costantino & Greene 2003; Barnes in Productivity Commission 2013; Smith 2012; Thompson et al. 2010) agree that stories and qualitative approaches value the perspectives of research participants. Moreover, these methods seem to address the shortcomings present in conventional methods; they enable collaboration with Indigenous partners in designing the evaluation, allow for complexities, and provide insight into the lived experiences and multifaceted meanings of the research (Abma 2005; Costantino & Greene 2003; Marsh et al. 2016; W.K. Kellogg Foundation 2004). In particular, MSC (see Davies & Dart 2005) and storytelling/narrative research (see Costantino & Greene 2003; Marsh et al. 2016; Riessman 1993) seem to have significant potential in evaluating the research impact of AIATSIS. However, no one approach is a panacea. See below for a summary of the strengths and weaknesses of each approach.

Table 1: Strengths and weaknesses of approaches and methods to impact evaluation

Approach or method	Strengths	Weaknesses
Logic model/theory of change	• Clarifies components of the project • Clarifies cause and effect • Facilitates discussion about expectations (benefit to project design and implementation) • Can capture intangible impacts and long-term impacts	• Interpretation of reality; may only capture expected impacts • Simplifies the project and its impacts; does not capture complexity or heterogeneity • Impacts are not usually identified by the people impacted • Omits unintended impacts • Usually inflexible

[1] See also Taylor, R 2003, 'An Indigenous perspective on evaluations in the intercultural context: how far can one throw a Moree boomerang?', *Evaluation Journal of Australasia*, vol. 3, no. 2, pp. 44–52; Hurworth, R & Harvey, G (eds.) 2012, 'Indigenous evaluation' [Special issue], *Evaluation Journal of Australasia*, vol. 12, no. 1 (these papers are not included in the annotated bibliography).

Approach or method	Strengths	Weaknesses
Economic valuations	• Speaks the language of funders and investors • Widely used; well-established and regarded as reliable • Identifies counterfactual or baseline; robustness of results can be tested through sensitivity analysis • Can be used to compare impact across a number of projects	• Depends on the quality of many assumptions • Difficulties in valuing intangible impacts • Difficulties in reflecting the complexities and heterogeneity of impacts • Impacts are not usually identified by the people impacted • Usually does not capture cause and effect and does not facilitate learning and improvement • Resource-intensive • Requires skilled evaluators
Surveys/questionnaires	• Generates quantitative data, which are regarded as objective and reliable • Flexible – can also generate qualitative data through open-answer questions • Can capture intangible impacts • Not resource-intensive	• Usually does not capture cause and effect • Closed-answer questions limit the ability for respondents' perspectives to be captured as categories of impact are pre-determined by researchers • May obscure significant but scarcely experienced impacts
Participatory methods	• Values and gains insight into the perspectives of stakeholders; captures what is most meaningful to them • High levels of engagement and collaboration with stakeholders at all stages of evaluation • Captures complexities • Can capture intangible impacts • Flexible; tailored to each group of participants • Reflexive and critical	• Often not regarded as reliable or objective • Resource-intensive • Difficult to execute well • May misrepresent Indigenous perspectives by reconstructing them in a non-Indigenous logic
Interviews/storytelling	• Values and gains insight into the perspectives of stakeholders; captures what is most meaningful to them • Captures complexities, depth, and heterogeneity of impact • Can capture cause and effect; attribution more easily assessed • Enables learning and improvement • Can capture intangible impacts • Flexible • Can produce both qualitative and quantitative data (through analysis of interviews)	• Often not regarded as reliable or objective • Resource-intensive • Requires skilled interviewers • Produces complex data that are difficult and time-consuming to analyse • Likely to be small in scale, which reduces perception of reliability • Results are difficult to compare across projects

Approach or method	Strengths	Weaknesses
Most Significant Change	• Values and gains insight into the perspectives of stakeholders; captures what is most meaningful to them • Captures actual impacts, not potential or expected impacts • Captures complexities • Captures cause and effect; attribution more easily assessed • Can capture negative impacts • Enables learning and improvement • Can capture intangible impacts • Flexible; easy to modify • Can produce both qualitative and quantitative data (through analysis of stories)	• Captures outliers of impact, rather than average experiences of impact • Often resource-intensive • Organisational hierarchy may not be appropriate for selection process • Selection process can omit important feedback

As the above table reflects, a number of challenges in evaluating research impact persist. For example, several papers (Arvidson et al. 2010; Maughan 2012; Productivity Commission 2010; Tsey et al. 2016) stress the practical challenges of constraints in time and expenses. This is a challenge across all methods; for example, while technical expertise is needed for economic valuations, qualitative methods require a great deal of time. Other challenges are more deep-seated, such as the difficulty in measuring intangible impacts (Marsh et al. 2016; Tsey et al. 2016). Arvidson et al. (2010) and the Productivity Commission (2010) explore the argument that market proxies are unable to reflect the true value of intangible outcomes, and others (Costantino & Greene 2003; Marsh et al. 2016; Stake 1976) explore how evaluation of intangible outcomes can evolve to be more reflective of reality.

Several papers suggest there is a tendency for impact evaluations to essentialise impacts, emphasising the need to allow for complexities (Marsh et al. 2010; Stake 1976; W.K. Kellogg Foundation 2004). Crookston et al. (2016) and Marsh et al. (2010) argue that qualitative methods are needed to capture these complexities. CSIRO (2015) frames this in terms of identifying 'distribution' – the heterogeneity of impacts, recognising that research projects often engender both 'winners' and 'losers'. However, capturing negative impacts is a significant challenge. Davies and Dart (2005), Onciul (2015), and Willetts and Crawford (2007) explore this challenge, offering methodological adjustments to mitigate it. Conversely, Wavell et al. (2002) argue that it is the choice of indicators that is crucial to ensuring all impacts are assessed.

However, many are sceptical of the reliability and validity of qualitative methods and data, whereas quantitative methods are usually viewed as more reliable (Butler et al. 2011; Fujiwara & Campbell 2011; Marsh et al. 2016; Maughan 2012; Riessman 1993; Wavell et al. 2002). Wavell et al. (2010) emphasise that evaluations need to gather ‘hard evidence’ of impacts, rather than evidence that they *could* occur. Riessman (1993) explores a number of ways narrative research can establish validity, concluding that no matter the approach, the methodology and data should be transparently presented so that the audience may verify the conclusions drawn (see also Productivity Commission 2010). A related challenge is minimising the effect of various biases on evaluation, no matter the method used (Davies & Dart 2005; Fujiwara & Campbell 2011; Kahneman & Kruger 2006). These include strategical bias (respondents deliberately changing their answers to attempt to achieve a more desirable outcome for themselves) and interviewer bias (respondents giving answers that deviate from their true opinions due to the influence of the interviewer), as well as contextual effects such as question order and the current mood of the respondent.

Finally, a number of papers examine the issue of attribution (the extent to which impacts can be ascribed to a research project) and establishing cause and effect (Arvidson et al. 2010; CSIRO 2015; Davis et al. 2008; Maughan 2012; Panel on Return on Investment in Health Research 2009; Productivity Commission 2010; Tsey et al. 2016; Wavell et al. 2002).[2] The difficulty of establishing the degree of attribution is largely due to the complexities in determining the counterfactual or baseline and comparing these to the outcomes of the project while accounting for external influences. Solutions to this challenge vary greatly, depending on the overall approach (ibid.), but methods which capture cause and effect are, by nature, more able to determine the degree of attribution.

There are gaps in this bibliography that derive from our goal to select the most relevant articles for AIATSIS community engaged research. First, while we reviewed the most relevant approaches to evaluating the impact of Indigenous research, the field of impact evaluation is extensive and multifaceted, and thus numerous approaches were omitted. We decided to leave out frameworks commonly used by universities, such as the Research Quality Framework, as most of these frameworks define impact in terms of citations and publications, rather than looking at the broader social, environmental, cultural and economic impacts. Second, we focused on data collection methods, and only touched on data analysis and interpretation methods (see Riessman 1993 for methods for analysing narrative research), which is an important part of developing a methodology for assessing research impact. Third, further research is needed to develop an approach for specifically identifying negative impacts, and potentially for understanding the impact of inquiring into negative impacts. Finally, ensuring that lessons generated from impact evaluations are implemented can be a challenge, and this requires further research (this is

[2] Establishing cause and effect is explored in detail in White, H 2009, *Theory-based impact evaluation: principles and practice*, Working Paper 3, International Initiative for Impact Evaluation, New Delhi (this paper is not included in the annotated bibliography).

explored a little in Abma 2005; Productivity Commission 2010; Willetts & Crawford 2007).

In summary, this annotated bibliography provides a synthesis of approaches to impact evaluation particularly relevant to research with Indigenous communities. We have examined a range of methods and frameworks that are widely used or have potential to contribute to the development of a methodology for impact evaluations of AIATSIS' research, recognising that there are strengths and weaknesses to any approach. We have framed the emerging field of impact assessment in terms of how we can meaningfully engage with Indigenous communities to assess impact, so that we may better understand whether we are meeting the priorities of the Indigenous communities we work with, and how our research can evolve to better support these priorities.

1. Why is it important to evaluate research impact?

Department of Innovation Industry, Science and Research 2011, *Focusing Australia's publicly funded research review: maximising the innovation dividends review key findings and future directions*, Australian Government.

> This is a report on a review of the performance of publicly funded research in Australia. It is largely concerned with topics such as quality, productivity, innovation, and coordination of research, but it also argues for the importance of developing a tool to evaluate the wider benefits of research. The report demonstrates the Australian government's general position on evaluating research impact – that it is important to evaluate the wider benefits of publicly funded research through a mechanism similar to Excellence in Research for Australia (ERA). It recognises the value of the ERA in assessing quality of research, but recognises that there is no process for measuring broader economic, social, and environmental benefits. The Department recommends a research impact assessment tool be developed separate from ERA (which is currently in development).

Productivity Commission 2010, *Contribution of the not-for-profit sector*, Research Report, Canberra.

> This is a report on the Productivity Commission's study on the impacts of the not for profit sector and how to improve the sector's measurement of impact. In its assessment of the impact of the sector, it focuses on economic contribution and it predominantly uses monetary valuation methods. However, it does recognise that this fails to measure broader community benefits. Chapter 3 explores various reasons why it is useful to measure impact, and then develops a contribution measurement framework (where contribution is defined as inputs, outputs, outcomes, and impacts). It proposes a logic model approach (see page 40), and suggests four broad categories of outputs and outcomes, as well as six broad categories of impacts, or wellbeing domains. It does not indicate a preference to certain methods, as the methods used should change depending on context and purpose, but it is largely focused on valuing outcomes in monetary terms, and so the methods it mentions are all valuation techniques. Appendix B discusses issues with valuation and market proxies, provides examples of outcome maps, lists wellbeing measurement tools, analyses and compares the approaches to measurement in more detail, and provides useful guidance for choosing an approach.

Tsey, K, Lawson, K, Kinchin, I, Bainbridge, R, McCalman, J, Watkin, F, Cadet-James, Y & Rossetto, A 2016, 'Evaluating research impact: the development of a

research for impact tool', *Frontiers in Public Health*, vol. 4, viewed 5 May 2017, <http://journal.frontiersin.org/article/10.3389/fpubh.2016.00160/>.

Tsey et al. have developed a model for evaluating the impact of Indigenous health research called the Research for Impact Tool. This paper describes the process of developing the tool and explores the challenges and issues in the application of the tool in the Indigenous health space. The authors argue for the importance of evaluating research impact and developing a reliable evaluation framework. They argue that other models for evaluating research impact are inadequate as the Indigenous sector needs a framework that meets ethical principles in Indigenous research. The Research for Impact Tool, they claim, is designed to meet such ethical principles and emphasises the impacts that participants' experience. They also emphasise designing the model in such a way to ensure it is used. Thus, the authors argue the model aligns with grant criteria of the major research funding bodies and with Indigenous expectations. They define research impact in terms of societal, cultural, health and wellbeing, and environmental changes, rather than focusing on academic metrics such as number of citations. The model is notable for this way of defining research impact and for its focus on best-practice engagement with Indigenous people. However, the model itself needs to be further expanded. The model does not provide guidance on the specifics of each step, such as how to identify and measure the impacts, but it does encourage the use of 'appropriate participatory learning-by-doing approaches'. This is because, as the authors assert, there are 'no easy template solutions' to assessing research impact. Furthermore, the model does not address many of the challenges and issues the authors raised (for example, determining attribution, data collection methods, or choosing indicators).

2. Incorporating Indigenous perspectives in research impact evaluation

Campbell, G 2014, 'Song as artefact: the reclaiming of song recordings empowering indigenous stakeholders-and the recordings themselves' in A. Harris (ed.) *Circulating cultures: exchanges of Australian Indigenous music, dance and media*, ANU Press, Canberra, Australia, pp. 101-127.

> In this chapter the author discusses the emotional, socio-political, legal, and ethical issues surrounding the repatriation of digital archival material, specifically songs, to Tiwi Islanders. Though Campbell hopes to revitalise song culture in the Tiwi Islands, she acknowledges the potential negative impact that repatriation might have on the dynamic practice of culture. She considers how repatriation of legacy song recordings can contribute to a phenomenon of rote replication of song rather than being a tool with which to improve one's skill at traditional improvised composition. Campbell also discusses the complications which arise through the use of songs which are considered copyright under Australian legislation to the recorders, rather than the singers, as well as the bureaucratic process of gaining permission from archives to use the songs in pursuits other than simply listening to them. She discusses an outcome in which a Tiwi elder had to sign a request form, and then authorise his own request with another form. Despite these issues, some Tiwi People feel that the repatriated songs are a meaningful resource for maintaining and continuing existing and new forms of Tiwi music making.

Conaty, G & Carter, B 2005, 'Our story in our words: diversity and equality in the Glenbow Museum' in R. Janes & G. Conaty (eds.) *Looking reality in the eye: museums and social responsibility*, University of Calgary Press, Calgary, pp.43-58.

> This chapter discusses the importance of a participatory model of community engagement, rather than simply the advisory or consultative models which are frequently implemented in many museums and cultural institutions today. Conaty and Carter use the example of co-curating the *Nitsitapiisinni: Our Way of Life* exhibition at the Glenbow Museum with the local Blackfoot communities in order to describe their method of cooperatively developing and implementing the exhibition. They describe the significance of understanding the multiplicity of goals involved in creating the exhibition, and recognising that the goals of the Blackfoot co-developers was as important as that of the museum staff. Additionally, Conaty and Carter discuss the necessity to allow varied viewpoints within groups of Indigenous co-developers, and to accommodate the discussion and decision-making process within Indigenous groups before the ideas are presented to a museum team. The attitude towards the finalised exhibition was

overwhelmingly positive and many First Nations visitors expressed a sense of pride in finding a museum exhibition expressed in a First Nations voice, and emphasise its significance for developing self-esteem in First Nations youth.

Kovach, M 2015, ‘Emerging from the margins: Indigenous methodologies’, in S Strega & L Brown (eds), *Research as resistance: revisiting critical, Indigenous and anti-oppressive approaches*, 2nd edn, Canadian Scholar’s Press, Toronto, Canada.

> Kovach explores Indigenous research and its interaction with dominant understandings of research. She discusses Indigenous research as a form of emancipatory research and argues that taking control of research is critical to the goals of self-determination. Yet, there are substantial difficulties in carving space for Indigenous methodologies and changing the perceptions of other researchers regarding their legitimacy. Kovach raises the issues of ‘Indigenizing Western models of research’ and using the English language to communicate Indigenous concepts. She explores what Indigenous research entails, such as the role of experience and storytelling, and the importance of relationship between researcher and participants. While this paper is focused on Canadian Indigenous methodologies, it points to the importance of conducting appropriate research with Indigenous peoples and the issue of perceptions of legitimacy within the research impact field.

Lonetree, A 2011, ‘Museums as sites of decolonization: truth telling in national and tribal museums’, in S. Sleeper-Smith (ed.) *Contesting knowledge: museums and indigenous perspectives*, University of Nebraska Press, Lincoln, pp. 322-336.

> In this chapter, Lonetree critiques the lack of ‘truth telling’ in the National Museum of the American Indian (NMAI), despite the role of Native American knowledge systems in influencing the development of the NMAI, or the important collaborative methodology with Indigenous communities in the Western Hemisphere. Lonetree contends that the museum shies away from ‘hard truths’ and the historical exhibitions of the NMAI “fail to present a clear and coherent understanding of colonialism and its ongoing effects”. Lonetree points instead to the Saginaw Chippewa’s Ziibiwing Centre for Anishinabe Culture and Lifeways as a tribal museum which discusses the ongoing problems caused by colonialism. The Tribally owned and operated museum privileges oral tradition, and by confronting hard truths about the ongoing effects of colonisation promotes healing and understanding for their community.

Nakata, M 2006, ‘Australian Indigenous studies: a question of discipline’, *The Australian Journal of Anthropology; Oxford*, vol. 17, no. 3, pp. 265–275.

Situated in the broader question of how Indigenous researchers approach ‘Indigenous Studies’, Nakata explores ways in which Indigenous and non-Indigenous approaches may intersect to negotiate meanings which allow them to work together. He explores the position that adequately portraying the experiences of Indigenous peoples requires new frameworks based on Indigenous ways of knowing, but argues for going beyond a ‘superficial application’ of Indigenous paradigms. He rejects the reduction of the two systems to a black and white opposition, but argues that the intersection is a ‘tangled web’ in which there are both clear boundaries and blurred boundaries. From this position, Nakata argues, negotiated methodologies will be more useful in representing Indigenous realities.

Onciul, B 2015, *Museums, heritage, and Indigenous voice: decolonising engagement*, Routledge, New York.

Onciul covers a broad range of topics using four case studies centred around the Blackfoot Confederacy in Canada. Onciul utilises a combination of interviews with Blackfoot community members and museum personnel, and museological theory to discuss community engagement and engagement zones, and the decolonisation of museums to better present Indigenous cultures. The case studies display strategic use of Blackfoot identity in order to create a single Blackfoot community with which to communicate a counter-narrative to that told by the dominant culture. In this way, the exhibitions may not conform to the ‘truth telling’ described by Lonetree in the above item, instead creating a narrative which develops mediated cultural pride among Blackfoot Community members, especially youths, and to potentially help decolonise museum-Blackfoot relations. Onciul used elements from participatory methods for capturing Indigenous expressions of impact in the case studies, providing the space and process for communities to reflect negatively on their engagement and participation in the public production of knowledge. In discussing the varying degrees of success in the decolonisation of the museum in each case study, Onciul argues that some forms of engagement can have a negative impact on community relations, cultural confidence and knowledge production. She discusses the limiting effects of casting community stakeholders as beneficiaries, as this obscures the possibility for the analysis of negative impacts for community members.

Productivity Commission 2013, *Better Indigenous policies: the role of evaluation,* Roundtable Proceedings, Productivity Commission, Canberra.

This edited volume includes contributions from a number of authors focused on five key themes including the mechanics of Indigenous evaluation, institutionalising better evaluation practices and the use of evidence, evaluation and broader Indigenous policy, and contrasting international experiences. Of note is the chapter by Les Malezer noting that evaluation methodologies 'fail to include Indigenous people's expectations, perspectives and participation' and that evaluations need to be focused on the need of Indigenous peoples, rather than on one's own values and interests. In describing Māori approaches to evaluation, Helen Moewaka Barnes argues that positivist quantitative methods may be inappropriate, whereas 'qualitative methods can be seen as giving voice to people and as resonating with descriptions of Māori culture as oral, holistic and relationship-based' (p. 167). The authors stress the need for partnership between evaluators and Indigenous communities in developing evaluations, arguing for the incorporation of Indigenous perspectives into evaluation frameworks.

Smith, LT 2012, *Decolonizing methodologies: research and Indigenous peoples*, Zed Books, New York and London.

Smith's book, written from the perspective of a Māori academic, highlights how the research context is 'a significant site of struggle between the interests and ways of knowing of the West and the interests and ways of knowing of the Other' (p. 31). She argues that Indigenous perspectives are rarely approached as a starting point in research practice and provides a provocative account of 'why' Indigenous methodologies are significant and equally 'how' they can be included without exploiting Indigenous knowledges. She posits research as a form of knowledge production – both a source of colonisation and decolonisation – highlighting the importance of early and effective partnership with Indigenous peoples in the process. She highlights the role of consultations, open debates, free-response interviews, storytelling, and flexible tools in ethical and respectful research. The book provides valuable case studies to illustrate the challenges of repositioning research practice to support Indigenous knowledges and perspectives.

Treloyn, S, Martin, MD & Charles, RG 2016, 'Cultural precedents for the repatriation of legacy song records to communities of origin', *Australian Aboriginal Studies*, 2:94–103.

While this is not an impact study as such, this article explores the potential impacts of legacy song repatriation programs. In this instance, the repatriation and return of songs back to community was a community-led initiative

involving an Indigenous knowledge content management system and online repository of Junba songs and dances. From reflecting on her conversation with senior Ngarinyin and Wunambal elder and singer, Matthew Dembal Martin, Treloyn identifies both positive and negative impacts of song repatriation. The potential negative effects are the presence of researchers inhibiting the ability of others to learn the songs, and the distortion of performance practice 'by reinforcing notions of fixed authoritative versions of songs over the traditionally … fluid approaches to performance' (p. 96). However, as Martin reveals, the positives far outweigh the negatives. Repatriation of songs facilitates Wurnan (a deeply important system that involves sharing, teaching, learning and continuing song traditions, as a set of reciprocal responsibilities to others), increases connection to Country, brings the spirits of the singers/dancers/composers in the recordings back to Country, and more. All of these effects increase wellbeing (or *liyan*). Treloyn highlights the importance of local perspectives in assessing impact and of Martin's stories in expressing the spiritual significance of song repatriation.

Thompson, SL, Chenhall, RD & Brimblecombe, JK 2010, ‘Indigenous perspectives on active living in remote Australia: a qualitative exploration of the socio-cultural link between health, the environment and economics’, *BMC Public Health*, vol. 13, pp. 473–484.

This is a study aiming to explore local perspectives and experiences of physical activity in two remote Indigenous communities. The authors argue that programs which aim to promote health through physical activities should incorporate Indigenous meanings of health and physical activity – performing activities on country. This paper is notable for its collaborative approach to research and its aim to understand the meanings and experiences Indigenous people in the communities attach to physical activities. It is also notable for its use of qualitative and participatory methodology (primarily semi-structured interviews) to ensure Indigenous perspectives were recognised. Another method used to collect data was respondents creating paintings to represent physical activity. Data were thematically coded and the interpretations of the stories and paintings were established as reliable through relaying the interpretations back and confirming them as representative of the respondents' views.

Tsey, K, Lawson, K, Kinchin, I, Bainbridge, R, McCalman, J, Watkin, F, Cadet-James, Y & Rossetto, A 2016, ‘Evaluating research impact: the development of a research for impact tool’, *Frontiers in Public Health*, vol. 4, viewed 5 May 2017, <http://journal.frontiersin.org/article/10.3389/fpubh.2016.00160/>.

Refer to p.9 for annotation.

3. What approaches and methodologies are used to measure impact and how valuable are they?

Logic models and theories of change

CSIRO 2015, *Impact evaluation guide*, CSIRO, Canberra.

This guide describes the Commonwealth Scientific and Industrial Research Organisation's (CSIRO) approach to evaluating research impact, which is largely based on logic models and cost-benefit analysis The guide outlines the importance of impact assessment and the need for a common framework, the principles of the framework (e.g. all outcomes and impacts should be monetised or quantified, all assumptions and decisions should be documented, etc.), and the steps to be followed in carrying out the assessment (a summary of these steps is on page 5). The steps are fairly rigorous; the counterfactual, attribution, adoption, distribution (which is often ignored), discounting, etc. are all given serious consideration. When cost-benefit analysis is not appropriate, other methods are used, but the majority of these are valuation methods (a useful overview is included in Appendix E). Social Impact Assessment and Most Significant Change are mentioned briefly as methods to use when others fail. Concerns with this guide include its assumption that valuation methods are the best way to evaluate impact, and its lack of any rigorous methods to initially identify impacts (which are instead to be identified through consideration and background research).

Davis, J, Gordon, J, Pearce, D & Templeton, D 2008, 'Guidelines for assessing the impacts of ACIAR's research activities', *ACIAR Impact Assessment Series*, Report No.58, viewed 28 November 2016, <http://aciar.gov.au/publication/ias058>.

The Australian Centre for International Agricultural Research (ACIAR) conducted reviews of impact assessment studies and from this developed a set of guidelines to use as a basis for future assessments of research impact. This paper outlines their approach to evaluating research impact, details the steps of the process, and discusses analysis of social impacts, including a basic critique of the Willingness to Pay (WTP) and Willingness to Accept approaches. The approach uses cost-benefit analysis and a logic model framework, including identification of the causal chain between outputs, outcomes, and impacts. It emphasises identifying both intended and unintended outputs and identifying the counterfactual. When impacts do not have market prices, WTP methods are used to value the impacts. The guidelines are largely concerned with economic outcomes and impacts (for example, changes in supply and demand) and how to value these, only briefly touching on environmental and social impacts. They also assume the output of research is a good, limiting the applicability of the approach.

Panel on Return on Investment in Health Research 2009, *Making an impact: A preferred framework and indicators to measure returns on investment in health research*, Canadian Academy of Health Sciences, Ottawa.

The Canadian Academy of Health Sciences propose a framework for measuring impact of health research based on the logic model approach and its earlier Payback approach. This report offers guidelines on implementing the framework, including how to address various issues, such as attribution and the halo effect. It also provides a range of indicators (both quantitative and qualitative) relevant to health research. Many of these indicators are concerned with bibliometric data and prevalence of health conditions; however, indicators which measure broader impacts are included on pages 30-32. The report discusses various areas that are less discussed in the general literature, such as issues with indicators (for example, their validity and relevance), and ensuring the evaluation is comprehensive enough to meet several purposes. As for methods, the report discusses bibliometric analysis, surveys, economic analyses, case studies, and peer reviews, providing advantages and disadvantages of each method. It recommends using a number of different methods. It acknowledges the criticism of economic analyses but contends it is a powerful tool for advocacy and accountability. The report recognises a wide range of perspectives to measuring research impact, but is only concerned with health research, offering little guidance on specific methods and techniques to be used for data collection.

Pasanen, T and Shaxson, L 2016, *How to design a monitoring and evaluation framework for a policy research project*, Methods Lab, Overseas Development Institute, London.

This paper provides practical guidance for designing a monitoring and evaluation framework for research projects which seek to affect policy through their outcomes. Drawing on Ingie Hovland’s work, Prasanen and Shaxson propose six performances areas of monitoring and evaluation: strategy and direction, management, outputs, uptake, outcomes and impact, and context. It recommends the development of a theory of change, and suggests that Outcome Mapping is likely to be useful to this development. Through the development of a theory of change it is possible to understand the types of outputs which will be useful in practically implementing the desired policy change, the applicable methods of monitoring and evaluating uptakes, and methods of monitoring and evaluating outcomes and impacts. They also

acknowledge that the context of the project should be monitored and evaluated at regular intervals throughout the project, but that there should be a focus on this at the beginning and end of the project.

Productivity Commission 2010, *Contribution of the not-for-profit sector*, Research Report, Canberra.

Refer to p.8 for annotation.

Social Ventures Australia (SVA) Consulting 2016, *Social return on investment analysis of the Birriliburu and Matuwa Kurrara Kurrara Indigenous protected areas*, Department of the Prime Minister & Cabinet.

Refer to p.20 for annotation.

W.K. Kellogg Foundation 2004, *W.K. Kellogg Foundation evaluation handbook*, W.K. Kellogg Foundation, Battle Creek, Michigan.

This is a handbook for project managers and evaluators working for the W.K. Kellogg Foundation. It describes the philosophy of the Foundation's approach and the process they use. Their approach emphasises the purpose of evaluation being learning and improvement, using multiple methods, using participatory processes, allowing for flexibility, and building capacity through reflexivity. Like in much of the literature, evaluation is considered to be an ongoing process throughout the life of the project, not an event at the end, in order for meaningful data to be collected. The handbook argues conventional evaluation approaches are inappropriate for human-services fields, and proposes several consequences of using a conventional (usually quantitative) approach: it espouses the belief that there is only one way to evaluate, important questions are not asked or examined, it fails to capture impacts of complex system change and community initiatives, and, because conventional approaches present truth as neutral and objective, evaluators forget that their work is political and value-laden. They instead recommend evaluators to learn about and reflect on alternative approaches. The handbook advocates the use of logic models; this is for several reasons: they focus staff on outcomes throughout the project, link activities to outcomes, help to focus evaluation on measuring each set of events, can be used to test assumptions, and clarify the project between stakeholders. Despite recognising that every evaluation should be different and that there is no silver bullet, the handbook then provides steps for implementing an evaluation. In these steps, they emphasise the importance of identifying multiple and diverse stakeholders; discuss several methods such as observation, interviews, focus groups, and questionnaires; and identify techniques for data analysis. The greatest concern with this guide is the assumption that logic

models are applicable and effective in all evaluation contexts. However, the flexibility and philosophy of reflexivity of this guide means it is can be used to complement a range of approaches and methods.

Economic valuations

Arvidson, M, Lyon, F, McKay, S & Moro, D 2010, ‘The ambitions and challenges of SROI’, *Third Sector Research Centre.* Working Paper 49, viewed 21 November 2016, <http://eprints.mdx.ac.uk/7104/>.

> This paper examines the Social Return on Investment (SROI) approach and presents several issues with the method in regards to: the judgement and assumptions that underpin the choice of indicators, the focus on valuing the impact at the expense of not understanding how it was produced (cause and effect), the method's inability to record and value less tangible benefits, financial proxies not capturing the social value of an impact in terms of improvement of personal utility (quality of life), putting a price on volunteering, capturing deadweight and attribution, using the results and the ratio, and the high cost required to use the approach. The paper also provides a good description of SROI and cost-benefit analysis, and compares them. While the paper is not comprehensive, it raises key issues with SROI that need to be further considered.

CSIRO 2015, *Impact evaluation guide*, CSIRO, Canberra.

> Refer to p.15 for annotation.

Davis, J, Gordon, J, Pearce, D & Templeton, D 2008, ‘Guidelines for assessing the impacts of ACIAR’s research activities’, *ACIAR Impact Assessment Series*, Report No.58, viewed 28 November 2016, <http://aciar.gov.au/publication/ias058>.

> Refer to p.15 for annotation.

Fujiwara, D & Campbell, R 2011, *Valuation techniques for social cost-benefit analysis: stated preference, revealed preference and subjective well-being approaches: a discussion of the current issues*, HM Treasury and Department for Work and Pensions, London.

> This paper discusses the stated and revealed preference methods and the life satisfaction (LS) methods (such as Wellbeing Valuation) for impact valuation. It discusses the strengths and weaknesses of each approach, and then gives recommendations for their use. Notably, Fujiwara and Campbell argue that the strength of the preference methods is that they are widely used and researched and hence more robust, whereas their weaknesses include the

various biases they present (such as hypothetical, strategical, and interviewer), and the disparity between Willingness to Pay and Willingness to Accept. They argue the strengths of LS methods include: it does not rely on people having well-defined pre-existing rational preferences, data are translated into monetary terms by analyst not interviewee, and it is not prone to as many biases. Its weaknesses, however, include that it is in development and not as robust, it is open to context effects (question order and current mood) and interviewer bias, it is difficult to estimate marginal utility of income and of the non-market good, and it may be difficult to attribute change in wellbeing to non-market good. The paper concludes that LS methods should be used as a complement to preference methods as it often gives "implausibly high" valuations. However, this seems to overlook the point that the two approaches are measuring different things (the value of market goods similar to the non-market good, versus the amount of income that produces the same effect as the non-market good) and therefore their valuations shouldn't be compared.

Nicholls, J, Lawlor, E, Neitzert, E & Goodspeed, T 2009, *A guide to social return on investment*, UK Cabinet Office of the Third Sector.

This guide describes SROI, discusses what it should be used for, outlines the process in detail of using it, and provides examples and other resources. Throughout the explanation of the process, the paper gives tips and proposes practical questions to answer and activities to complete in order to implement the method. If SROI is to be used, this is a very useful document. However, please note issues with SROI in the other items of this section.

Social Ventures Australia (SVA) Consulting 2014, *Kanyirninpa Jukurrpa: evaluative social return on investment report: social, economic and cultural impact of Kanyirninpa Jukurrpa's On-Country programs*, viewed 28 November 2016, <http://socialventures.com.au/assets/2014-KJ-SROI-Report-FINAL.pdf>.

This report is a SROI analysis of the social, economic and cultural outcomes of Kanyirninpa Jukurrpa's (KJ) On-Country programs. The logic model that KJ uses is informed by the Martu worldview, and its activities address the intersection of Martu interests and whitefella interests. This report examines the complexities and methodologies of identifying outcomes, determining indicators, and valuating the outcomes. Outcomes are identified through interviews and consultations with stakeholders. While a good example of using SROI in an Indigenous context, the indicators that measure intangible impacts appear to be based on many assumptions.

Social Ventures Australia (SVA) Consulting 2016, *Social return on investment analysis of the Birriliburu and Matuwa Kurrara Kurrara Indigenous protected areas*, Department of the Prime Minister & Cabinet.

> This report, similar to the previous item, is a SROI analysis of the social, economic, cultural, and environmental outcomes of the Birriliburu and Matuwa Kurrara Kurrara Indigenous protected areas. It examines the methods used to identify outputs, derive outcomes, find indicators, measure each outcome, and value each outcome. It largely uses the revealed preferences approach, but also uses interviews. The interview guide in Appendix C is likely to be useful and could be used in conjunction with other interview techniques. While a good example of using SROI in an Indigenous context, the indicators that measure intangible impacts appear to be based on many assumptions. However, this report is useful outside the context of SROI as it uses logic models and methods of measuring change that may serve as examples. The step of economic valuation may be easily left out if necessary.

Trotter, L, Vine, J, Leach, M & Fujiwara, D 2014, *Measuring the social impact of community investment: a guide to using the Wellbeing Valuation approach*, Housing Associations' Charitable Trust (HACT). AND Fujiwara, D 2014, *Measuring the social impact of community investment: the methodology paper*, Housing Associations' Charitable Trust (HACT).

> The first item is a guide to the Wellbeing Valuation method, a life satisfaction (LS) method, which measures the impact of a project by how much it increases people's wellbeing. It does this through analysing existing surveys, isolating the effect of a particular factor on a person's wellbeing, and through using a tool called the Social Value Bank, identifying the equivalent amount of money needed to increase someone's wellbeing by the same amount. The Social Value Bank is particular to the United Kingdom, but an Australian Social Value Bank is due to be released in 2017 which is sold as a cost effective mechanism for DIY SROI. The guide claims the strength of the Wellbeing Valuation method is that it is consistent and robust, particularly when compared to other valuation approaches. It explains the method and how to use it (providing surveys in Appendix B), identifies limitations of the approach, and presents the values of the Social Value Bank. It is important to note that these values are not extensive and many potential outcomes that AIATSIS's research projects produce will not be listed.
>
> The second item is a companion methodology paper to the guide. It details the methodology behind the valuation of the outcomes in the Social Value Bank. It also touches on the methodology of revealed and stated preference methods, comparing these to the Wellbeing Valuation approach, and argues

that Wellbeing Valuation addresses the problems that arise in these other methods. The paper then discusses the method's approach to several issues, such as selection bias, to ensure the method is robust.

Verwayen, H, Wilms, J & Fallon, J 2012, *Workers underground: an impact assessment case study – Europeana 1914-1918*, Viewed 15 March 2017 < http://pro.europeana.eu/files/Europeana_Professional/Publications/workers-underground-an-impact-assessment-case-study-europeana-1914-1918.pdf>

This case study reviews the process for assessing the social, economic, and innovative impact for the *Europeana 1914-1918* project. It utilises a model which balances both social and economic return on investment. The evaluation used a number of theoretical lenses to frame questions in understanding perceived value from specific perspectives; utility (perceived value of the products use), existence (perceived value of the products existence), legacy (perceived value of the inheritance or bequeathing of knowledge and resources within the project), learning (perceived value of an enhanced sense of cultural heritage, education or knowledge as a result of the product), and community (perceived value of the experience of being part of a community engaging with the product). These lenses allowed the project developers to understand visitor expectations to a museum exhibition, and to improve the product in areas in which these expectations were not met. Additionally, though somewhat tenuously, asking visitors and non-visitors for a perceived economic value for the existence of the product allowed them to attempt to place an economic value on the social return, though with the caveat that some respondents stopped participating in the survey when asked these questions.

Tools to measure wellbeing

International Wellbeing Group 2013, *Personal Wellbeing Index: 5th edition*, Melbourne: Australian Centre on Quality of Life, Deakin University.

As a part of an evaluation, particularly one that involves Wellbeing Valuation, tools such as the Personal Wellbeing Index (PWI) can be used to measure subjective wellbeing. This paper discusses the PWI tool, its history and its validity, and provides guidelines for its use and the survey itself. The International Wellbeing Group assumes that this tool is sufficient to measure changes in personal wellbeing, even though wellbeing is known to be relatively stable due to the tendency for people to self-regulate their wellbeing. Furthermore, the paper does not address bias and other effects (such as current mood of respondent) or issues like attribution. They claim that the tool

is valid in cross-cultural contexts. However, the items on the survey are chosen as significant indicators of wellbeing from a non-Indigenous perspective of what constitutes wellbeing (for example, standard of living, achievements, and future security), and each indicator is given the same weight. See below for Mission Australia's response to this issue.

Kahneman, D & Krueger, AB 2006, ‘Developments in the measurement of subjective well-being’, *Journal of Economic Perspectives*, 20(1): 3–24.

This paper examines the validity of the tools like PWI. It discusses various biases and contextual effects (such as current mood), the relationship between wellbeing and various respondent characteristics (such as income, age, disability, etc.), and adaptation and self-regulation issues. It also touches on why this approach may more accurately reflect preferences than revealed preference methods, includes discussion of tools more focused on emotions, and proposes the U-index as a tool that minimises biases and other issues. Kahneman and Kreueger conclude that tools that measure wellbeing are useful and reliable, and that many biases and contextual effects like current mood average out in larger samples. This indicates that tools such as PWI may be useful when collecting data on wellbeing from a large number of respondents, but is unlikely to be as reliable or useful when measuring the wellbeing of smaller groups.

Mission Australia 2015, *Impact measurement and client wellbeing*, viewed 22 November 2016, <http://www.missionaustralia.com.au/publications/annual-reports/doc_download/449-impact-measurement-and-client-wellbeing-report-2015>.

Mission Australia used PWI to measure impact, but to ensure the determinants of wellbeing they were particularly concerned with (e.g. independence) were included in the analysis, they added questions from other surveys to the PWI survey. This report outlines their approach and experiences with using the tool. A similar strategy may be utilised to improve the PWI tool for particular contexts.

Productivity Commission 2010, *Contribution of the not-for-profit sector*, Research Report, Canberra.

Refer to p.8 for annotation.

Surveys and questionnaires

Crookston, M, Oliver, G, Tikao, A, Diamond, P, Liew, CL & Douglas, S-L 2016, *Kōrero Kitea: ngā hua o te whakamamatitanga: the impacts of digitised te reo archival collections*, Victoria University of Wellington, Wellington.

> Kōrero Kitea is a research project that explores some of the ways digitised te reo Māori collections are being used and what impact that use has. This report examines the findings of the project, discusses the methodology used, outlines the key drivers in the design of the project, includes the survey used, and provides a literature review of impact studies in the memory industry. The project is important to consider in the context of evaluating the impact of use, digitisation, and repatriation of intangible cultural heritage (particularly language). It includes measuring negative impacts, which is often neglected in impact evaluations. The methodology involves preliminary discussions with users and an online survey. The survey identified possible impacts and asked about these, as well as asking participants to identify other impacts. The authors argue the use of open-ended questions worked well to elicit narratives and complex answers. They identify a series of important aspects of impact evaluation design: value of narratives, long time-frame, use of a number of methods, user-focus, using culturally appropriate methods, and allowing respondents to answer in te reo. The report does not consider the effects of possible bias on the evaluation, but otherwise presents a notable impact evaluation design.

Bailey, J, Yang, H-Y, Penhall, S & Shehadeh, T 2015, *International arts activity - Australian arts sector: survey detailed report*, Australia Council for the Arts, Sydney, viewed 6 December 2016, <http://www.australiacouncil.gov.au/research/current-projects/international-arts-activity-australian-arts-sector-report/>.

> The Australia Council for the Arts conducted a survey on the international activities of artists and art organisations, including what perceived impact these activities have. This paper reviews the survey and its findings. It is included in this bibliography due to its transparency; the methodology of the survey is discussed on pages 7-8, the findings of the questions on impact are on pages 38-43, and the survey itself is on pages 87-102 (question on outcomes on page 94). A survey like this may be useful when there is a need to collect a range of quantitative data from various stakeholders.

Bradburn, NM, Sudman, S & Wansink, B 2004, *Asking questions: the definitive guide to questionnaire design - for market research, political polls, and social and health questionnaires*, Rev. ed, Jossey-Bass, San Francisco.

> This book is a guide to formulating questions for interviews and questionnaires. Part Two in particular provides checklists, tips, step-by-step guides, and examples that would be useful. Most of the book is concerned with closed, fixed-response interviews, but in regards to open-ended questions, tips and a brief analysis of advantages and disadvantages of open-ended questions are given on pages 153-156. While most of the book focuses on questions that might be asked in a marketing survey, questions used to evaluate performance are discussed on pages 213-242. The focus here is on evaluating employees, teachers, etc. but the material is still applicable. Part 3 may also be useful as it provides recommendations concerning formatting (and why formatting is important), coding, and delivery methods, as well as a step-by-step guide for preparing a questionnaire on pages 315-316 and a section for frequently asked questions. While some of this book provides out-dated information and is not critically reflexive, much of its content and advice is practical and straightforward and should be very useful for designing interviews and questionnaires.

Participatory methods

Butler, S, Nichols, K, Dart, J & Boulet, M 2011, 'Walking the line between impartiality and participation', paper presented at Australasian Evaluation Society International Conference, Sydney.

> One challenge in participatory evaluation is the need to balance participatory methods with the wish for robust, independent evaluation. This paper discusses this challenge by using an evaluation of the impact of a leadership program as an example. The paper outlines each step of the evaluation process and identifies the benefits and lessons learned from each step. To achieve balance between independence and participation, some steps were participatory (planning, some analysis, and re-design) and some steps were independently conducted by an external evaluator (data collection, preliminary analysis, and reporting). It was found that the planning stage being participatory was important to achieve a sense of ownership, while the re-design stage being participatory was important to ensure lessons were utilised. Data collection being independent was important to ensure informants talked openly and honestly. Data were collected through semi-structured interviews, review of data, and an internet survey. The participatory analysis approach was informed by Appreciative Inquiry, Most Significant

Change (MSC), and Collaborative Outcomes Reporting. This approach demonstrates a way to use selection process of the MSC technique where participants perform the selection, rather than the organisational hierarchy. However, it was not discussed whether this produced a conflict of interest. The approach also provides a good example of how to balance the needs of being robust and using participatory methods.

Gosling, L & Edwards, M 2003, *Toolkits: a practical guide to planning, monitoring, evaluation and impact assessment*, Save the Children, London.

This is a handbook published by Save the Children that discusses the principles, processes, and tools of participatory monitoring and evaluation. Throughout the book it provides examples and asks key questions to consider. Notably, it includes specific information on involving children in collection methods, which is seldom explored elsewhere in the literature. The description of the processes to use is thorough, including many questions to consider in order to shape the evaluation. Chapter 9 discusses how impact assessment in particular is conducted, using a theory of change model. The authors then describe a range of tools and methods for monitoring and evaluation, and identify strengths and weaknesses of each. These tools include observation, interviews, diagrams and maps, surveys, and cost-effectiveness analysis (including why it is more appropriate to use for social benefits than cost-benefit analysis). However, many of the weaknesses of these tools are not identified. While there is useful content in this book, it is repetitive and unnecessarily long. Further, many of the weaknesses of the participatory approach are thought to be resolved by considering questions about them, rather than implementing a tool to ensure they are resolved.

Guijt, I, Arevalo, M & Saladores, K 1998, 'Participatory monitoring and evaluation: tracking change together', *International Institute for Environment and Development*, 31, PLA Notes, pp. 28–36.

This paper discusses the methods of participatory monitoring and evaluation and issues with them. Guijt et al examine the need for participatory methods, various interpretations of what participation entails, what it is used for, its core principles, and what influences people's participation in monitoring and evaluation. The methods they discuss include meetings (to collect baseline data, using participatory tools such as mapping, transect walks, etc.), modified Logframes, open-ended questions, impact flow diagrams, and building on culturally valid (not just culturally sensitive) frameworks and methods. It discusses issues regarding methodologies (for example, combining conventional methods with participatory methods), institutionalisation and

scaling-up (for example, reconciling bottom-up methods with top-down bureaucracy), and documentation.

Thompson, SL, Chenhall, RD & Brimblecombe, JK 2010, ‘Indigenous perspectives on active living in remote Australia: a qualitative exploration of the socio-cultural link between health, the environment and economics’, *BMC Public Health*, vol. 13, pp. 473–484.

Refer to p.14 for annotation.

UNDP 1997, *Who are the question-makers?*, Office of Evaluation and Strategic Planning, viewed 20 November 2016, <http://web.undp.org/evaluation/documents/who.htm>.

This is a handbook for participatory monitoring and evaluation. It gives an overview of the approach and its development, offers a rationale for its use, discusses the differences between participatory approaches and more conventional ones (including the differences between various levels of participation), and outlines a participatory evaluation framework. Descriptions of a number of participatory methods are provided in Annex II, including beneficiary assessment, focus groups, Logframe, interviews, social mapping, and transect walks. A number of guides for using these tools are also mentioned for further reading. The handbook is concise and informative, but it lacks detail. Moreover, it lacks impartiality, dismissing conventional and quantitative methods without considering their strengths.

W.K. Kellogg Foundation 2004, *W.K. Kellogg Foundation evaluation handbook*, W.K. Kellogg Foundation, Battle Creek, Michigan.

Refer to p.17 for annotation.

Interviews, narratives and storytelling

Abma, T 2005, ‘Responsive evaluation: its meaning and special contribution to health promotion’, *Evaluation and Program Planning*, 28(3): 279–289.

Abma, well known in the evaluation field for her work in exploring the role of narrative and dialogue in Responsive Evaluation, uses a case study of an injury prevention program to illustrate her approach. She explains the core concepts of Responsive Evaluation and discusses why it is valuable. For example, she argues the evidence it produces is more likely to be implemented in improving the program as it is context-bound, more quickly available, and more readily accepted due to the sense of ownership stakeholders have over it. Abma’s approach is similar to Stake’s (see below) in that it is focused on issues, stakeholders are actively involved in the process, and the design and criteria emerge as issues are identified. Abma develops the approach by adding processes for stakeholders to interact and engage in dialogue, and she emphasises the importance of storytelling and narrative to serve as a vehicle for learning and dialogue. She also emphasises the role Responsive Evaluation can play in challenging power relations. In the case study data were collected through conversational interviews, focus groups, and participant observation. While this approach has many strengths, it has been developed with the assumption that evaluation is conducted independently. Thus, if this approach is to be used by the researchers involved in the initial project, it must be appropriately adapted.

Costantino, TE & Greene, JC 2003, 'Reflections on the use of narrative in evaluation', *American Journal of Evaluation*, 24(1): 35–49.

This paper tells the story of an evaluation the authors undertook on a program in rural US. They approached the evaluation through a qualitative case study within an interpretive, phenomenological methodology. The methods they used include reviewing program documents; group interviews with staff, participants and other stakeholders; observation; phone interviews; and email correspondence. The interviews were semi-structured with open-ended questions. The respondents reframed the questions and responded by telling stories that gave insight into the program's true nature. Through analysing the stories and examining the structure, function, and relevance of the stories, Costantino and Greene identified the impacts of the program and found that what was meaningful to the participants in the program was embedded in the stories they told. They conclude that evaluation that uses narrative can give voice to participants' perspectives, connect the various stakeholders, increase understandings, portray contextualised meanings, and reflect the true complexity of the program and its benefits.

Marsh, DE, Punzalan, RL, Leopold, R, Butler, B & Petrozzi, M 2016, 'Stories of impact: the role of narrative in understanding the value and impact of digital collections', *Archival Science*, vol. 16, no. 4, pp. 327–372.

Marsh et al. interviewed a number of heritage professionals to understand their perspectives and experiences of impact evaluation. They argue that the systematic collection of stories is a valuable method in assessing the impact of digitised ethnographic collections. The authors regard conventional models for impact assessment, particularly those that only generate quantitative data, as inadequate due to the complex and distinctive nature of ethnographic collections. They also argue impact evaluation models need to be 'flexible enough to incorporate the incredibly varied and contextual nature of meaning

and use' (p. 358). This article is particularly notable for its exploration of the issues in impact assessment, such as the inadequacies of both quantitative and qualitative data, the susceptibility for evaluations to essentialise impacts, and the difficulties in choosing questions. The paper also demonstrates the value of stories and storytelling. The authors argue open-ended, free response questions enable respondents to identify intangible impacts and impacts that would otherwise not be identified. However, Marsh et al. do not specifically explore the perspectives of Indigenous people or the cultural custodians of digital collections.

Patton, MQ 2002, *Qualitative research & evaluation methods*, 3rd edn, Sage Publications, Thousand Oaks.

This is a comprehensive book on qualitative methods for evaluation. It is highly regarded as a guidance document for qualitative evaluation. It discusses a range of conceptual issues in qualitative inquiry, designing qualitative studies, fieldwork and observation methods, interviewing guidelines and methods, and qualitative analysis, interpretation and reporting. The chapter on interviewing examines various interview structures, open-ended questions, various question types, issues in wording questions, group interviews, cross-cultural interviewing, and ethical considerations, among other topics. Throughout, it includes practical checklists and guides. It is unbiased and does not participate in the debate of quantitative versus qualitative methods, unlike much literature. It also includes an extensive list of references for further reading. It is a valuable resource if qualitative methods (particularly interviews and observation) are to be used.

Riessman, CK 1993, *Narrative analysis*, vol. 30, Qualitative Research Methods, Sage Publications, Newbury Park, CA.

This book introduces and illustrates forms of narrative analysis. Riessman first discusses why and when narrative research is useful and discusses the underlying philosophy of narrative research – supporting varied perspectives as a legitimate form of evidence. She illustrates the interpretative and constructed nature of each level of narrative research: experiencing the event, telling the story about the event, transcribing the story, analysing the transcript, and reading the report. Riessman emphasises the importance of being aware of these levels. She provides guidance in developing questions to elicit narratives, arguing interviews should be conversational with minimal structure. A number of practical models of analysis (and transcription) are examined, most of which aim to begin analysis by identifying the structure of the narrative in differing ways. Riessman also explores the issue of validity and validation, concluding that however the researcher validates their

interpretation, they should describe their methods and make primary data available, so that others may determine the research as trustworthy.

Smith, L 2010, '"Man's inhumanity to man" and other platitudes of avoidance and misrecognition: an analysis of visitor responses to exhibitions marking the 1807 bicentenary', *Museum and Society*, 8(3): 193–214.

Smith conducted face-to-face interviews with visitors to British museums on their views of the 1807 bicentenary exhibitions. The interviews included questions to determine whether their views were changed and what personal impact the exhibition had on them. Smith discusses her methodology for these interviews on pages 196. A sample questionnaire can be found at the link below. A number of these questions can be adapted to similar contexts where the impact of an exhibition is to be evaluated. It may be very time-consuming and work intensive to conduct these interviews with the public, but Smith argues elsewhere that face-to-face interviews provide information on the way in which people answered questions, which would not be available in a written survey.[3]

Sample questionnaire:
http://www.history.ac.uk/1807commemorated/audiences/audience.html

Stake, R 1976, 'To evaluate an arts program', *Journal of Aesthetic Education*, 10(3/4): 115–133.

In this article, Stake introduces Responsive Evaluation as an alternative approach to evaluation. This approach is not restricted to assessing impact, but is used for evaluation more broadly. Stake argues responsive evaluation is often better to use than 'preordinate evaluation' as its findings are more useful and relevant, and it doesn't oversimplify the program under evaluation, but appreciates the complexity of issues and the multiplicity of perspectives, and thus has a better grounding in reality. Stake also argues it improves communication with the audience by showing what the program 'was like'. Responsive evaluation involves a flexible and adaptive approach, with no set timeline of activities or fixed plan, but the evaluator conducts activities as they become relevant and responds to issues as they emerge. Responsive evaluation looks at the different perspectives of stakeholders and is inclusive of multiple and contradictory sources of information. Stake does not provide much guidance on specific methods to collect data, as he asserts the particularities of the methodology should adapt to the specific context, but he does use interviews as the primary method. One concern with this approach

[3] Smith, L 2015, 'Theorizing museum and heritage visiting', in K Message and A Witcomb (eds.), *The international handbook of museums studies*, Vol. 1, John Wiley & Sons.

is that it may not be practical. He assumes there is infinite time and resources for the evaluation, and that all people involved in the program are willing to invest much time and effort to participate. Further, the article is fairly abstract and vague. However, it sufficiently illustrates the beginnings of Responsive Evaluation.

Most Significant Change

Choy, S & Lidstone, J 2011, ‘Most significant change technique: a supplementary evaluation tool’, paper presented at 14th Annual Conference of the Australian Vocational Education and Training Researchers Association, Melbourne.

The Most Significant Change technique (MSC) is a specific narrative tool which collects interviews and stories of change and then systematically selects the most significant of these stories. This conference paper reports on the use of MSC as a supplement to more conventional methods in evaluating a leadership capacity building course. The purpose of the evaluation was learning and improvement. This paper outlines the reasons they used MSC and its advantages, briefly provides their methodology, and overviews examples of responses. It concludes that MSC was a valuable tool to use as it clearly identified the prevailing values of the contexts of the respondents, it encouraged respondents to critically evaluate the changes they experienced, and it created a rich picture of the impact. The report mentions that they only interviewed students to evaluate the program, and that they should have interviewed all stakeholders. Although not acknowledged in the paper, all responses given were individual changes, highlighting the need to ask about various domains of change if one wishes to capture wider impacts.

Davies, R & Dart, J 2005, ‘The “most significant change” (MSC) technique’, viewed 14 November 2016, <https://www.kepa.fi/tiedostot/most-significant-change-guide.pdf>.

This is a guide on the MSC technique by its creators. MSC is usually used in a development or social change program context, but it can be easily applied to a range of contexts. Its purpose is to learn what impact a project has had and how the impact occurred. This guide describes the technique, outlines the process to use it, addresses various concerns brought against it, and discusses the theory and history behind it. It discusses why it is a valuable tool, when to use it and when to not use it, and explains that it should not be used as the only method of evaluation, but be used to complement other methods. Throughout explaining the process it gives valuable tips and examples. It addresses a wide range of concerns, including how to capture stories of negative changes, or how to mitigate biases and problems with voice and power. It compares MSC to other approaches, and discusses how it might be used in combination with other approaches or under frameworks like logic models. Finally, it provides samples of stories, collection and reporting formats, and a facilitation guide. Overall, this guide is comprehensive and critically reflexive, and an essential source of information if MSC is being considered as a method to capture research impact.

Willetts, J & Crawford, P 2007, 'The most significant lessons about the most significant change technique', *Development in Practice*, 17(3): 367–379.

> This article gives insight into the challenges and complexities of the MSC technique based on its use in evaluating two development projects in Laos. The article speaks positively of the technique and identifies various lessons they learned to successfully implement it. They argue that the predominant benefits of MSC were that it enabled learning and reflection, and that it filled the gaps that other methods left (such as capturing negative impacts). The main challenges and lessons learned include: need to ensure interviewees are representative of all stakeholders, need for better training of staff, challenge of many reported stories being inaccurate, need to clarify the required detail for stories when collecting them, need for transparency, challenge of increased workload for stories that needed to be translated into English, need to implement a process to ensure negative stories are used, and the need for incentives and accountability to ensure lessons are implemented. This article argues that MSC is valuable as it provides insight into impact, but it should be used with other methods.

Multiple methods

Gosling, L & Edwards, M 2003, *Toolkits: a practical guide to planning, monitoring, evaluation and impact assessment*, Save the Children, London.

> Refer to p.25 for annotation.

Guthrie, S, Wamae, W, Diepeveen, S, Wooding, S & Grant, J 2013, *Measuring research: a guide to research evaluation frameworks and tools*, RAND Corporation, Santa Monica, California.

> This work analyses a number of common frameworks and methods used to measure the outcomes and impacts of research. It compares 11 methods, of which 9 may be used to assess impact: case studies, data mining, data visualisation, document review, economic analysis, interviews, logic models, site visits, and surveys. For each method, a description, advice for its use,

and examples are provided, and analysis is undertaken for a limited number of them. Summaries of these frameworks and methods are in Appendix A, with detailed information in Appendices C and D. Guthrie et al argue that there is no silver bullet and all approaches have weaknesses, and therefore the design of a framework should depend on the purpose of the evaluation. They assert that the four most common purposes are advocacy, accountability, analysis (learning and improvement) and allocation/funding. In Chapter 3, Guthrie et al give guidance on developing a new research evaluation framework. They propose a series of practical questions to consider surrounding the purpose, characteristics, pitfalls, context, methods, aggregation, and implementation of the framework to be developed.

Maughan, C 2012, *Monitoring and evaluating social impacts in Australia*, CRC-REP Working Paper CW003, Ninti One Limited, Alice Springs.

Maughan analyses four frameworks in this paper: ‘Monitoring, Evaluation, Reporting and Improvement’, logic models, theory of change, and Social Accounting and Audit. She then discusses three methods: Global Reporting Initiative, Social Return on Investment (SROI), and Most Significant Change. For each framework and method, the report explains how it works, argues what its benefits and limitations are, and gives examples of it in use. The paper is notable for its balanced assessment of each framework and method. For example, Maughan identifies several weaknesses of the logic model approach, an approach that is highly regarded by most literature. She argues that logic models are not necessarily a true depiction of reality, they are inflexible, and they can simplify the project and ignore unintended impacts. Similarly, in her analysis of SROI she argues it speaks the language of funders and investors and can often give more credibility to an evaluation, but it is resource intensive, depends on the quality of the many assumptions it requires, and doesn’t capture cause and effect.

Wavell, C, Baxter, G, Johnson, I & Williams, D 2002, *Impact evaluation of museums, archives and libraries: available evidence project*, A report for Resource: the Council for Museums, Archives and Libraries, Robert Gordon University, Aberdeen.

This report examines impact evaluation literature in the Galleries, Libraries, Archives, and Museum (GLAM) sector and proposes a series of recommendations for impact assessments. It examines a large amount of literature, and summarises various studies to illustrate different approaches and findings of impact evaluations. While much of the report is concerned with identifying the impacts the GLAM sector has had, it is notable for its exploration of key issues in impact evaluation, such as gathering hard evidence of intangible impacts (rather than merely demonstrating the *potential* for impacts), establishing the validity of qualitative data, choosing appropriate and reliable indicators, and establishing causality and attribution. The report also outlines various methods, but does not discuss them in detail. Wavell et al. present various positions on the value of qualitative data, concluding that qualitative methods are valuable in initially choosing the indicators, which can then be measured quantitatively.

W.K. Kellogg Foundation 2004, *W.K. Kellogg Foundation evaluation handbook*, W.K. Kellogg Foundation, Battle Creek, Michigan.

Refer to p.17 for annotation.