![BetterEvaluation logo]

3 August 2020

To: Mr Romlie Mokak, Commissioner

Thank you for the opportunity to comment on the draft Indigenous Evaluation Strategy and its companion documents (Guide and Background Paper). Overall, we commend the new Strategy for its clear focus on key principles. Placing the needs, perspectives, priorities and knowledges of Aboriginal and Torres Strait Islander People at the centre of the Strategy is critically important for evaluation to play a constructive role in improving outcomes for Aboriginal and Torres Strait Islander People. The other principles are also important – in particular, the clear message that evaluations need to be useful, credible, ethical and transparent.

BetterEvaluation is a not-for-profit organisation and registered charity that operates globally. Our mission is to work collaboratively with our global community to create, share and support use of knowledge about how to better plan, manage, conduct and use evaluation. Our work is guided by principles [https://www.betterevaluation.org/en/principles] and we have a particular interest in supporting efforts to improve the evaluation involving Aboriginal and Torres Strait Islander People such as through our project focusing on identifying and sharing good practice in this context [https://www.betterevaluation.org/en/themes/evaluation-practice-aboriginal-and-torres-strait-islander-settings].

Our comments and suggestions for revisions to the Strategy and its companion documents relate to areas where we believe they are not consistent with the guiding principles or not sufficiently clear about how to enact them. These concerns relate to three main areas:

1. *The primacy of Aboriginal and Torres Strait Islander People's perspectives, priorities and knowledges* in evaluations that relate to them – we are concerned of the lack of explicit standards and benchmarks for putting this overarching principle into practice across *all* stages of the evaluation process, from prioritizing what to evaluate to making sure agencies act on evaluation findings appropriately. We fear that, without setting clear expectations, the cultural change in government's actions that is envisaged by the Strategy will not occur with the depth

and speed needed in order to gain parity on key health and welfare indicators between Indigenous and non-Indigenous Australians.

2. *The need for the Strategy, guidance and background document to cover all the types of evaluation that it identifies and which are needed to improve performance* – despite other types of evaluation being named, the Strategy seems to focus exclusively on impact evaluations designed to inform policy decisions, leaving out other types of evaluations that are critically important to guide and learn from implementation, including in situations where adaptive management is needed.

3. *The need to include the full range of designs and approaches that can be used for high quality impact evaluation.* The Strategy focuses only on 'rigorous' evaluations and the background paper that this refers to a narrow set of experimental and quasi-experimental designs which will only be suitable in a limited range of situations, leaving out systematic non-experimental impact evaluations which are likely to be highly appropriate in many situations, including those that are affected by complexity.

These areas are further detailed below. In addition, we have identified some other resource materials which would be useful to add to the guidance document.

We are available for further clarification or more detailed comments. Please do not hesitate to contact us.

With very best wishes,

Dr Greet Peersman, Director M&E capacity strengthening;
Prof Patricia Rogers, Chief Executive Officer;
Alice Macfarlan, Knowledge Platform Manager;
BetterEvaluation Inc., Level 6, 126 Wellington Parade, East Melbourne 3002

BetterEvaluation.org [https://www.betterevaluation.org/en]

1. **Ensuring centrality of Aboriginal and Torres Strait Islander People throughout the evaluation process**

We are delighted to see and fully endorse the intent to "centre Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges in evaluations of policies and programs that affect them" [Strategy, p.6] as the overarching principle in the principles-based approach of the Strategy [Strategy, p.10].

While this principle is at the core of the Strategy [Strategy, p.8] and provides a unique opportunity for the Australian Government to 'do things differently', the Strategy (and its companion documents) does not reflect nor draw adequately on the existing strengths of Aboriginal and Torres Strait Islander people, communities and organisations to take charge of their own destinies –which may be quite different for different Indigenous cultures around Australia.

The Strategy refers to government "working with" Aboriginal and Torres Strait Islander people [Strategy, p.4] to achieve the "shared goal of improving the lives of Aboriginal and Torres Strait Islander people" [Strategy, p.4]. While working towards a shared goal is needed, it certainly does not require "for those affected by a government policy or program to work with government to find ways to achieve mutually beneficial outcomes" [e.g., Strategy, p.5]. If indeed, the Strategy is about 'centring Aboriginal and Torres Strait Islander people', then the outcomes to be achieved are those that are important to those 'affected' (from their own perspectives).

The Strategy states that "evaluation can help to build trust in government" and acknowledges that "trust between government and Aboriginal and Torres Strait Islander people, communities and organisations is particularly important in light of Australia's colonial history and its ongoing impacts" [Strategy, p.5]. Both the Strategy and the Guidance refer to "meaningful" partnership, "meaningful, genuine, effective, appropriate" engagement, "shared" decision making [e.g., Strategy, p.4,10,11,21,22; Guide p.13,14,19,26]. However, none of these crucial concepts that –in practice– can be the 'make or break' of achieving 'centring Aboriginal and Torres Strait Islander people', in general, and, trust relationships, in particular– are defined nor is there any acknowledgement of the inherent power differences between government actors and many Aboriginal and Torres Strait Islander individuals, communities and organisations.

If the Strategy and the Guidance does not explicitly address how to deal with these issues in practice, the intended 'culture change' within government is unlikely to be achieved, and, enduring positive outcomes in the lives of lives of Aboriginal and Torres Strait Islander people are unlikely to be achieved.

We urge that the Strategy provides incentives and also clear and explicit standards for accountability in terms of implementation of the principles. In this regard, we welcome the 'maturity approach to evaluation' but note that this needs to include the benchmarks to be reached, starting from what needs to be achieved 'right now' and 'over time'. There is no time to waste if we are to break the current trajectory of 'significant gaps in major health and well-being outcomes between Indigenous and non-Indigenous Australians'. For example, consistent with the principle of 'first do no harm', cultural competency should be an entry requirement for any government actor working in the Indigenous space –regardless of what level the work takes place; not something that government agencies need to achieve for their staff 'over time'. We note the intent to detail the maturity model before finalizing the Strategy [Strategy, p.20] and we urge that the Commissioner takes the opportunity to include explicit standards and benchmarks for various levels of maturity and when to need to be reached as part of a maturity rubric.

While the 'centrality of Aboriginal and Torres Strait Islander' is reflected in 'Questions to consider in….' [add relevant pages] in the Guide, we are particularly concerned that Fig 2 [Guide, p.4] does not reflect

this core principle at all. There is one instance of noting 'identify and engage with those affected' in the policy/program cycle' (see highlighted below), falling far short of the engagement that needs to happen during every stage of the cycle and associated evaluative activity. Specified and appropriate engagement needs to be done from the early tasks including identifying priority issues to address, relevant evidence to compile and developing program logics to determining information needs (primary users and uses) to the final stages of actual use in making decisions about existing and new policies and programs.

Figure 2   **Evaluation should be considered at every stage of policy and program design and implementation**
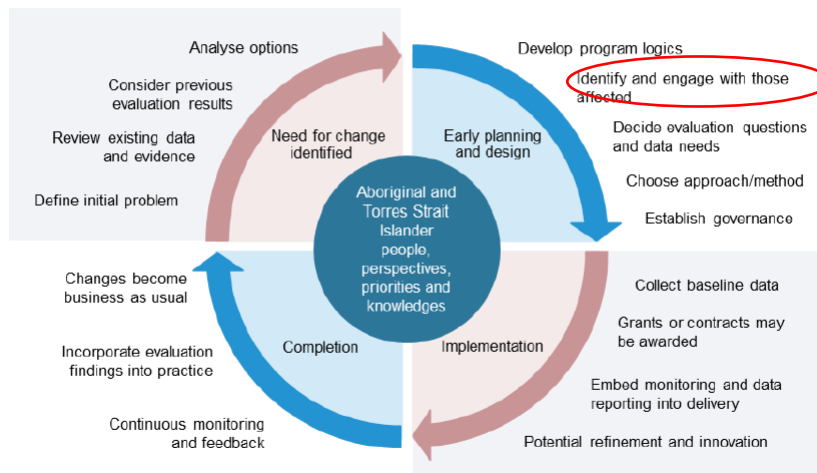
In light of the above comments, we strongly recommend that:

- all references to 'mutually beneficial outcomes' are removed;
- there is explicit endorsement of and a detailed description of processes for decision making that favours Aboriginal and Torres Strait Islander people priorities [see, for example, https://www.betterevaluation.org/en/rainbow_framework/manage/establish_decision_making _processes]
- explicit standards are defined for what constitutes 'genuine', 'meaningful', 'effective', 'shared' or the like partnerships, engagement and decision making' cognisant of existing trust and power relationships (i.e., whose voice counts) [see, for example, 'transformative participation' [1]; using 'cultural strengths'[2]; the work by Robert Chambers and colleagues[3]; https://www.unicef-irc.org/publications/750-participatory-approaches-methodological-briefs-impact-evaluation-no-5.html; https://www.betterevaluation.org/en/themes/evaluation-practice-aboriginal-and-torres-strait-islander-settings]
- explicit benchmarks and timelines are included for putting the principles into practice in the maturity model (ideally expressed as a rubric);
- Fig 2 in the Guide be revised to indicate when, how and with whom there needs to be engagement to ensure the overarching principle of 'centrality of Aboriginal and Torres Strait Islander People perspectives, priorities and knowledges' is implemented across all stages of the cycle.

---

[1] Cornwall A (2003). Whose voices? Whose choices? Reflections on gender and participatory development. World Development 31(8), pp. 1325–1342.

[2] Fletcher G (2019). 'Our Men Our Healing': Using Cultural Strengths to Restore Collective Wellbeing for Indigenous Australians. In: Engendering transformative change in international development. Routledge.

[3] See https://www.participatorymethods.org/

2. *The need for the Strategy, Guide and Background Paper to cover all the types of evaluation that it identifies* **and which are needed to improve performance**

The Strategy rightly points out that there are different types of evaluation which are used in different ways to improve decisions and actions in order to improve results. However, the Strategy, Guidance and Background Paper focus almost exclusively on supporting a few summative evaluations intended to answer policy questions about "what works". This runs the very real risk that the Strategy will lead to less resourcing, guidance and incentives for other types of evaluation – especially, evaluation that focuses on supporting implementation. Instead, the Strategy needs to clearly address ways of strengthening the management, conduct and use of different types of evaluation to improve performance and outcomes.

Implementation of policy is rarely a simple matter of ensuring compliance with policy. In most cases, various degrees of knowledge translation and adaptation are needed to effectively implement policy in different settings.

There is increasing recognition that classical linear, and siloed approaches to planning and delivery are less appropriate in conditions of volatility, uncertainty, complexity and ambiguity, or, where seeking to develop whole-of-government approaches to 'wicked problems'. Classical policy-level evaluation designs that aim to answer questions about 'what works' and then scale this up are also not useful in these instances. Instead, evaluation needs to be embedded in planning and implementation processes to ensure that delivery is adapted to local needs, opportunities and priorities, moving from finding out 'what works' to supporting local adaptation to make things work better in all contexts.

This will include generating or drawing on good descriptive quantitative and qualitative data about needs and factors that are producing problems; information about the availability of resources including existing infrastructure and capital (including human and social capital) that can be leveraged; details of how previous interventions have been implemented; information about what different people value in terms of results and processes; and, the identification of ethical issues . The evidence for policy making therefore needs to also encompass statistical databases; qualitative needs analyses; reports from previous projects, similar projects and pilot projects; opinion surveys; and expert reviews.

We recommend drawing from the recent HM Treasury (2020) *Magenta Book: Supplementary Guide: Handling Complexity in Policy Evaluation* which sets out a useful set of evaluation questions and ways of answering them (Table 4, p.45-46).

Table 4: Answering evaluation questions

Evaluation question	Approach / method	Benefits
What is important to different groups, who can champion change?

What levers are generating change, what may be inhibiting change? | Most significant change

Participatory system mapping | Most significant change is an iterative participatory process that aims to clarify the values held by different stakeholders
Participatory system mapping brings stakeholders together to build a system map and develop trust and mutual understanding
Structured conversations about whether and how the policy is delivering change, can be used to develop the theory of change
If begun at the option appraisal stage, forms a consistent framework for design, monitoring and evaluation through piloting and full implementation |
| How well was the policy implemented? How can this be improved? | Big data and associated methods | Might ultimately allow local emergence of system dynamics that subsequently spread throughout systems to be understood
Can provide near real time data to support learning |
| | Participatory, adaptive approaches | Generates trust and shared understanding, champions and agents for change |
| Is the policy making a difference, by how much?

Is it delivering value for money? | Experimental approaches | Provides robust evidence of whether a policy has made a difference, and to what extent, in a specific context |
| | Statistical association approaches | Weaker than experimental designs, but can provide a quantitative measure of the extent of impact, where it is not possible to define a counterfactual |
	Predictive modelling approaches	Using computational system modelling to predict a counterfactual has the advantage of being able to account for contextual factors and deal with the emergent properties of complex systems
Is the policy making a difference, how? What conditions are needed?	Generative causation approaches	Explores the causal mechanisms (decision making and behaviours) or processes leading to change and the impact of contextual factors
	Configurational approaches	Recognises that different combinations of factors can lead to change and that these may include factors external to the policy
How sustainable is change likely to be? How can we adapt the policy to work elsewhere? How can the policy be improved?	Generative causation approaches	Provides learning that is transferable to other contexts, and provides a basis for discussing sustainability, by uncovering the underpinning processes or mechanisms that lead to change and exploring the influence of context
	Computational system modelling	Using information available on system behaviour to date, modelling can confirm whether the policy is contributing to change and how, and project forward to explore whether change is likely to be sustained in different scenarios

We would be happy to provide additional detailed suggestions on how to revise the documents to reflect the issues we have raised. In particular, we have identified the following sections which would benefit from revision:

Draft Strategy

Explicitly name the scope of the strategy in its intent to cover all types of evaluation

Page 7 – Current diagram

Figure 1. The objective of the Strategy is to improve the lives of Aboriginal and Torres Strait Islander people
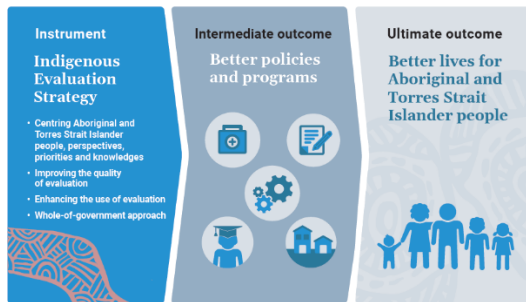
Spelling out the intermediate step between the strategy and better policies and programs would make clear the intended scope of the Strategy to cover all types of evaluations and its emphasis on improvement.

Page 7 – Suggested edit to diagram text – addition of an extra box

Page 10 – Section on choosing what to evaluate
Make it clear that this decision needs to be not only understood in terms of selecting a small number of interventions for large, discrete policy evaluations but also about supporting the appropriate types of evaluation to support improvements in a range of interventions.

Page 11 – Existing table
Make it clear that the intent is to improve evaluations generally, not only a small number of priority policy-level evaluations and avoid the use of the term 'rigorous evaluation' which suggest that credibility is the only important principle, ignoring utility, ethics and transparency, and also sometimes being understood as referring to a narrow range of experimental and quasi-experimental evaluation designs.

Table 1. Centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges in practice

| What to evaluate | – Aboriginal and Torres Strait Islander people are engaged to decide what policies and programs achieve the greatest impact on their lives and should be subject to rigorous evaluation. |

Page 11 – Suggested edit to table text

What to evaluate	Aboriginal and Torres Strait Islander people are engaged to decide what policies and programs **have potential to** achieve the greatest impact on their lives, and which should be subject to **which types of evaluation, including summative policy-level evaluation and formative implementation-level evaluation.**

Draft Guide

Page 5 - We support the statement about synthesis of high quality studies which does not limit the methods to be used for this synthesis. We also suggest adding specific reference to realist syntheses which explicitly provide information on what works in different contexts.

Box 1 Review and synthesis of existing evidence

Working out what is already known about a particular form of intervention is not always easy, particularly when there are volumes of research evidence. And because not all evidence is of equal value, some way of differentiating between high and low-quality studies is needed. Evidence synthesis methodologies aggregate evaluation findings and review them in a systematic way (and generally also assess and rate the strength of the evidence). Systematic reviews aggregate results that fit a pre-specified eligibility criteria in order to answer a specific research question. They aim to minimise bias by using explicit, systematic methods that are documented in advance with a protocol.

Systematic reviews search all the available research evidence for a particular question, which counters problems of selection bias that come from only identifying studies that are readily accessible or only published on major databases. They can help highlight gaps in the evidence (and identify evaluation questions that need to be answered) and methodological issues that may need to be considered in the design of an evaluation.

A meta-analysis pools statistical results from multiple studies as a way to increase explanatory power and improve estimates of the size of effects. **A realist synthesis combines diverse evidence to provide information about what works for whom in what contexts, making it more useful for situations where the same policies and programs don't work the same everywhere or for everyone**.

---

3. Need to include the full range of designs and approaches that can be used for high quality impact evaluation.

We support that the Strategy states explicit support for situationally-appropriate choices of methods:

Page 12

> Mixed methods (combining qualitative and quantitative methods) can maximise the strengths and compensate for limitations of any single method or approach.
>
> Credibility is grounded on rigorous methodology. The Strategy does not endorse particular evaluation approaches or methods. Agencies and evaluators should adopt methods that:
>
> – are rigorous and fit-for-purpose
>
> – answer the questions that policy makers and Aboriginal and Torres Strait Islander people alike need answered
>
> – suit the context in which the policy or program is operating, the size and importance of the policy or program, and the timeframe and resources available for evaluation.

However, this position is not reflected in the Background Paper which currently only refers to a narrow range of experimental and quasi experimental designs that use a counterfactual approach to causal inference, which will not be appropriate in all situations:

> *p.125 One of the key challenges for impact evaluation is coming up with an estimate of the actual effects of a policy or program. To do this, evaluators need to estimate the effects of a program on policy and program participants (the 'treatment group'), and compare this with non-participants (the 'control group') who represent the counterfactual (box 4.3). Measuring the counterfactual requires:*
>
> *⬚ data of sufficient quality and quantity to support the analysis*
>
> *⬚ the counterfactual to be genuinely comparable to the intervention group*
>
> *⬚ that the effect of the policy or program is sufficiently large to be distinguished from 'noise' in the data (HM Treasury 2020c, p. 34).*
>
> *Experimental approaches, such as randomised controlled trials (RCTs), construct a control group through random assignment to assess the net impact of an initiative. …*
>
> *Evaluations using quasi-experimental and other quantitative designs also measure impact but are often considered weaker forms of evidence*

Pages 128-129 in the Background Paper only lists quasi-experimental evaluation designs.

Over the past 20 years, there has been increasing attention internationally to the need for a wider range of evaluation designs beyond counterfactual designs. It is particularly unfortunate that none of these have been included in the Background Paper or reflected in the range of methods and designs listed for impact evaluations. In particular, the Background Paper ought to have drawn upon:

- Australian Public Service Commission (2007). Tackling wicked problems, a public policy perspective.
- The report commissioned by the UK Department for International Development on *'Broadening the range of designs and methods for impact evaluation'* (Stern et al. 2012) which discusses different types of causal questions and also a wider range of methods and the basis for them, and the subsequent report *'Impact Evaluation: A Guide for Commissioners and Managers'* (Stern 2015), (based on the DFID report) which set out five different bases for causal inference, in addition to using a counterfactual.
- Rogers et al, 2015 '*Choosing appropriate designs and methods for impact evaluation'*, Office of the Chief Economist, Department of Industry, Innovation and Science, Australian Government, which outlines different methods and designs that can be used in different circumstances for answering descriptive, causal and evaluative questions, including a number of non-experimental options for answering causal questions without a counterfactual.
- HM Treasury (2020) *Magenta Book: Supplementary Guide: Handling Complexity in Policy Evaluation.* While this is a recent publication, and, might not have been available when the background paper was drafted, it provides guidance for evaluation that will be highly relevant for many programs and policies that affect Aboriginal and Torres Strait Islander people.

The Background Paper and the Guide does not include any non-experimental approaches to policy evaluation which excludes many potentially viable and more appropriate approaches to impact evaluation.

There are a number of non-experimental methods and approaches which should be specifically named and discussed in the guidance:

- **Contribution Analysis**: assessing whether the program is based on a plausible theory of change, whether it was implemented as intended, whether the anticipated chain of results occurred and the extent to which other factors influenced the program's achievements.
    - http://betterevaluation.org/plan/approach/contribution_analysis
- **Process tracing**: a case-based approach to causal inference which focuses on the use of clues within a case to adjudicate between alternative possible explanations. Process tracing involves four types of causal tests:
    - 'straw in the wind', which lends support for an explanation without definitively ruling it in or out,
    - 'hoop', failed when examination of a case shows the presence of a necessary causal condition,when the outcome of interest is not present. Common hoop conditions are more persuasive than uncommon ones
    - 'smoking gun', passed when examination of a case shows the presence of a sufficient causal condition. Uncommon smoking gun conditions are more persuasive than common ones
    - 'doubly definitive' passed when examination of a case shows that a condition is both necessary and sufficient support for the explanation.
    - https://www.betterevaluation.org/en/evaluation-options/processtracing

- **Collaborative Outcomes Reporting:** mapping existing data against the theory of change, and then using a combination of expert review and community consultation to check for the credibility of the evidence.
    - https://www.betterevaluation.org/en/plan/approach/cort
- **Comparative Case Studies** - which use purposeful samples and process tracing to build and test theories and when and for whom a program works - see especially the UNICEF brief
    - https://www.betterevaluation.org/resources/guide/comparative_case_studies
- **Qualitative Impact Assessment Protocol** - a systematic approach that addresses some of the weaknesses of key informant attribution through double blinding and combines it with contribution analysis
    - https://www.betterevaluation.org/en/plan/approach/QUIP
- **Qualitative Comparative Analysis** - uses matrix displays and Boolean algebra to distinguish various complex forms of causation, including, configurations of causal conditions, not just single causes, equifinality, where there is more than one way in which an outcome can happen, causal conditions which are necessary, sufficient, both or neither, asymmetric causes – where the causes of failure may not simply be the absence of the cause of success.
    - https://www.betterevaluation.org/en/evaluation-options/qualitative_comparative_analysis
- **Realist Evaluation -** An approach especially to impact evaluation which examines what works for whom in what circumstances through what causal mechanisms, including changes in the reasoning and resources of participants.
    - https://www.betterevaluation.org/en/approach/realist_evaluation

The Background Paper also fails to discuss the risks and limitations of RCTs and quasi-experimental designs. A number of these were outlined in two papers that are cited as sources but not used (Rogers, P. 2009a, 'Learning from the evidence about evidence-based policy', Strengthening evidence-based policy in the Australian Federation: Roundtable proceedings, vol 1, Productivity Commission, Canberra. Rogers, P. 009b, 'Matching impact evaluation design to the nature of the intervention and the purpose of the evaluation', Journal of Development Effectiveness, vol. 1, no. 3, pp. 217–226.) – in particular:

> *a statistically significant difference between a treatment group and control group is not necessarily sufficient evidence to say that a policy will work when translated into wider practice. Interventions that have been found to be effective might not be feasibly implemented in other settings due to a lack of skills, expertise or resources needed to properly implement the evidence-based intervention or adequate regulatory and supervisory processes to ensure adequate implementation. Even where they can be implemented well, there can be differential effects — what works on average can be ineffective or even harmful for some. Other unintended effects may only be evident over time, and some pilots cannot be scaled up effectively — for example, programs for the longterm unemployed may be effective on a small scale, but when scaled up end up just shuffling job queues unless additional employment opportunities are created.*

There are serious risks in the over-advocacy for RCTs in the Background Paper and, hence, implied in the Guide:
- **Negative impacts of the evaluation process** – use of experimental designs risks undermining the sense of agency of local communities and implementing agencies as they are required to follow centralised protocols for implementation and engagement. Also, in the particular context of the Evaluation Strategy and the need for building of trust relationships where these have been eroded or have been challenged over a long time, a focus on experimental designs may exacerbate issues around trust in government.

- **Focus on average effects** – The emphasis on average effects can be unhelpful in supporting adaption for local conditions or to meet different needs of different groups and individuals. This risks further disadvantaging the most marginalised whose experiences might be different to the average.

- **Restricting policy options** – If only certain approaches to policy evaluation are considered credible (experimental and quasi-experimental designs where it is possible to identify a plausible counterfactual), then only policies and interventions that can use them can generate credible evidence, and,, therefore only those types can be supported, not whole-of-government, systemic or adaptive interventions.

- **Invalid use** – The notion of a hierarchy of evidence can lead to use of experimental or quasi-experimental designs when a credible counterfactual cannot be identified or created, leading to invalid comparisons and conclusions.

- **Acceptance of poor quality evidence** – The notion of a hierarchy of evidence can lead to uncritical acceptance of findings from particular research designs regardless of the actual quality of evidence, including instances of *poor measurement, poor adherence to randomisation, inadequate statistical power, unidentified differential effects, inappropriate comparisons, conducting numerous statistical analyses and only reporting statistically significant ones, differential attrition between control and treatment groups, unplanned crossover, and unacknowledged poor quality implementation of the intervention random error, treatment leakage, incomplete causal package, lack of blinding.*

Box 4.5 in the Background Paper only discusses quasi-experimental designs needs to be supplemented by an annotated list of rigorous non-experimental designs, drawing on the approaches listed above. It should also draw on *Table 4 Answering evaluation questions* from the HM Treasury (2020) Magenta Book supplement on *Handling complexity in policy evaluation* and *Table 3.1: Different types of impact evaluation questions and relevant methods* from the Dept of Industry (2015) report on *Choosing Appropriate Designs and Methods for Impact Evaluation*.

Additional resources to include in the Guide to support implementation of the Strategy

Building evaluation into policy and program design

- BetterEvaluation, Rainbow Framework:
  - DEFINE what is to be evaluated:
    https://www.betterevaluation.org/en/rainbow_framework/define
  - Establish decision-making processes:
    https://www.betterevaluation.org/en/rainbow_framework/manage/establish_decision_making_processes]
- BetterEvaluation, The Manager's Guide to Evaluation:
  - Step 1: Decide how decisions about the evaluation will be made:
    https://www.betterevaluation.org/en/commissioners_guide/step1
  - Step 2: Scope the evaluation:
    https://www.betterevaluation.org/en/commissioners_guide/step2
- UNICEF Impact Evaluation Series: Brief 5 – Participatory Approaches: https://www.unicef-irc.org/publications/750-participatory-approaches-methodological-briefs-impact-evaluation-no-5.html

Evaluation planning, design and conduct

- BetterEvaluation, Evaluation practice in Aboriginal and Torres Strait Islander settings
  https://www.betterevaluation.org/en/themes/evaluation-practice-aboriginal-and-torres-strait-islander-settings
- BetterEvaluation, The Manager's Guide:
  https://www.betterevaluation.org/en/managers_guide
- Cornwall A (2003). Whose voices? Whose choices? Reflections on gender and participatory development. World Development 31(8), pp. 1325–1342. Retrieved from:
  https://www.sciencedirect.com/science/article/abs/pii/S0305750X0300086X
- Fletcher G (2019). 'Our Men Our Healing': Using Cultural Strengths to Restore Collective Wellbeing for Indigenous Australians. In: Engendering transformative change in international development. Routledge.
- Institute of Development Studies: Participatory Methods website:
  https://www.participatorymethods.org/
- BetterEvaluation, Rainbow Framework:
  https://www.betterevaluation.org/rainbow_framework
  - Especially: UNDERSTAND CAUSES of Outcomes and Impacts
    https://www.betterevaluation.org/en/rainbow_framework/understand_causes
- BetterEvaluation, Evaluation Approaches: https://www.betterevaluation.org/en/approaches
- UNICEF, Impact Evaluation Series (in particular Brief 3: Causal Attribution):
  https://www.betterevaluation.org/en/resources/guide/unicef_impact_evaluation_series
- Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., & Befani, B. Department for International Development, (2012). *Broadening the range of designs and options for impact evaluations* (Working Paper 38). Retrieved from website: http://www.dfid.gov.uk/Documents/publications1/design-option-impact-eval.pdf

Reporting evaluation findings

- BetterEvaluation, Rainbow Framework: REPORT and SUPPORT use of findings: https://www.betterevaluation.org/en/rainbow_framework/report_support_use
- BetterEvaluation, Manager's Guide:

    o Step 8. Guide production of quality report(s): https://www.betterevaluation.org/en/commissioners_guide/step8

    o Step 9. Disseminate reports and support use of evaluation: https://www.betterevaluation.org/en/commissioners_guide/step9

Building capability and a culture of evaluation

- BetterEvaluation, Rainbow Framework: Strengthen Evaluation Capacity: https://www.betterevaluation.org/en/rainbow_framework/manage/strengthen_evaluation_capacity