23 August 2019

Commissioner Romlie Mokak
Productivity Commission Indigenous Evaluation Strategy Inquiry

By email: indigenous.evaluation@pc.gov.au

Re: Indigenous Evaluation Strategy Inquiry

Points made in this submission arise from my research and engagement undertaken as the Inaugural Braithwaite Research Fellow at the School of Regulation and Global Governance (RegNet) at the Australian National University, which I now continue as an Australian Research Council DECRA Fellow at the Law Futures Centre and Griffith Law School at Griffith University for the project *Regulation and Governance for Indigenous Welfare: Poverty Surveillance and its Alternatives*.[1] I am also a Chief Investigator for the Australian Research Council Discovery Project *Conditional Welfare: A Comparative Case Study of Income Management Policies*.[2]

The Indigenous Evaluation Strategy Inquiry Questions to be addressed in this submission include:

- *What are the strengths and weaknesses of current evaluation systems and practices across Australian Government agencies? Can you provide examples of good and bad practice?*
- *In what ways are Aboriginal and Torres Strait Islander people and organisations contributing to policy and program evaluation?*
- *What principles should be included in an Indigenous evaluation framework to be used by Australian Government agencies?*
- *How do Australian Government agencies currently deal with ethical issues associated with evaluation?*
- *In what circumstances should evaluation projects be subject to formal ethics review? In what circumstances should evaluation projects be exempt from formal ethics review?*
- *What types of evaluation approaches and methods are currently used to evaluate Indigenous programs? How could evaluation methods be improved to ensure robust and reliable evidence is produced?*

One of the weaknesses of government evaluations of Indigenous programs in the area of welfare reform is that these are generally evaluated according to government

[1] Australian Research Council Discovery Early Career Researcher Award (DECRA) (DE180100599).
[2] Australian Research Council Discovery Project (DP) (DP180101252).

formulated and imposed criteria without respect for Indigenous self-determination and sovereignty. On the most critical view, a view that is necessary to consider in order to practice what Sandra Harding refers to as 'strong objectivity',[3] this could be seen as part of a practice of ongoing colonialism. In times of authoritarian paternalism, it could also be seen as part of a process to ensure the creation of policy-based evidence rather than evidence-based policy, an approach that can have significant costs for Indigenous program participants – especially where their participation in programs is mandatory.

Although Aboriginal and Torres Strait Islander peoples and organisations frequently contribute to evaluations, they often do so as program participants and stakeholders with no control over what evaluation questions are asked or how their answers will be used by government. Their responses can be repeatedly sought, but without bringing about their desired changes to programs.

In light of these issues and those discussed below, I welcome the establishment of an ethical framework for the evaluation of policies and programs affecting Australia's First Peoples – a framework that ensures not only that evaluations of such programs take place but that these evaluations are of high quality with sound methodology that can reliably measure program outcomes. This includes respectful and attentive listening to the concerns about programs raised by Indigenous program participants.

In the Indigenous evaluation context, there are issues with what counts as evidence and whose evidence counts the most. At times, as occurred with the Cashless Debit Card Orima evaluation discussed below, the views of Indigenous program participants encountering serious program problems were trivialised/ignored in favour of other perspectives that fit more neatly with government ideology. This is problematic in terms of ethical practice, and importantly, has left many coerced program participants experiencing protracted problems with no relief.

There needs to be clearer measures of what counts as policy success and what counts as policy failure, together with a genuine commitment to redressing harm to Indigenous program participants where this is documented in evaluation reports. The government's response to several evaluations showing harm to numerous Indigenous program participants has been woefully inadequate, for instance, those pertaining to the Cashless Debit Card and the Community Development Program.[4]

In my research on cashless welfare transfer cards I have found troubling tendencies in government commissioned evaluations and reports to date. I have outlined my

[3] Sandra Harding, *Objectivity and Diversity: Another Logic of Scientific Research* (The University of Chicago, 2015) x, xiii, 30, 44.
[4] Department of Social Services, *Cashless Debit Card Trial Evaluation: Final Evaluation Report* (Canberra: Orima Research, 2017) 6-7, 80, 82; and Commonwealth of Australia, Department of the Prime Minister and Cabinet, *The Many Pathways of the Community Development Programme – Summary Report of Community Voices and Stakeholder Perspectives from Eight Communities* (2018) 11, 39.

concerns about these at length on previous occasions to various government inquiries,[5] and therefore shall make only brief comments here.

The Cashless Debit Card (CDC), triggered by the 2014 *Forrest Review*,[6] was purportedly introduced to address substance abuse issues and to ensure that social security recipients would engage in socially responsible behaviours.[7] The program quarantines 80 per cent of a social security recipient's regular payment to the CDC.[8] The CDC applies disproportionately to Indigenous peoples.[9]

Although the government claims that the Cashless Debit Card trial is a 'success',[10] with 'positive impact'[11] they relied heavily on a flawed foundation with the evaluation conducted by Orima in doing so, an evaluation which has significant limitations. Several of these are acknowledged in their report, and they include:[12]

- recall error – because 'fieldwork was conducted 6-12 months after' the CDC was introduced,
- social desirability responses – influencing people to respond 'in a socially acceptable way',
- 'unavailability of adequate time series data to perform robust pre-Trial and post-Trial comparisons',
- 'low numbers of cases' that were volatile over the course of the study making trends difficult to identify,
- 'comparison site data were only available for a limited number of measures', and

[5] For example, Shelley Bielefeld, Submission to the United Nations Special Rapporteur on Extreme Poverty and Human Rights, *Thematic report to the UN General Assembly on digital technology, social protection and human rights*, 16 May 2019, 1-8; Shelley Bielefeld, Submission No 22 to the Senate Standing Committee on Community Affairs, *Inquiry into the Social Security (Administration) Amendment (Income Management and Cashless Welfare) Bill 2019*, 7 March 2019, 1-5; Shelley Bielefeld, Submission No 90 to the Australian Human Rights Commission, *Human Rights and Technology Issues Paper*, 16 October 2018, 1-23; Shelley Bielefeld, Submission No 68 to the Senate Standing Committee on Community Affairs, *Social Services Legislation Amendment (Cashless Debit Card Trial Expansion) Bill 2018*, 20 July 2018, 1-13; Shelley Bielefeld, Submission to the Australian National Audit Office, *The implementation and performance of the Cashless Debit Card trial*, 15 January 2018, 1-9; Shelley Bielefeld, Submission No 55 to the Senate Standing Committee on Community Affairs, *Social Services Legislation Amendment (Cashless Debit Card) Bill 2017*, 29 September 2017, 1-20; Shelley Bielefeld, Submission to the Senate Standing Committee on Community Affairs, *Social Security Legislation Amendment (Debit Card Trial) Bill 2015*, 18 September 2015, 1-15.

[6] Commonwealth of Australia, *The Forrest Review* (2014) 100-108.

[7] Shelley Bielefeld, 'Indigenous Peoples, Neoliberalism and the State: A Retreat from Rights to "Responsibilisation" via the Cashless Welfare Card', in Deirdre Howard-Wagner, Maria Bargh and Isabel Altimarino-Jiminez (eds), *The Neoliberal State, Recognition and Indigenous rights: New paternalism to new imaginings* (Australian National University Press) 147-165.

[8] unless they can get this reduced by application to a Community Panel.

[9] Aboriginal and Torres Strait Islander Social Justice Commissioner, *Social Justice and Native Title Report 2016* (Sydney: Australian Human Rights Commission) 91-92.

[10] Alan Tudge, 'Media Release: Evaluation finds "considerable positive impact" from cashless debit card', Department of Human Services, 1 September 2017, <https://www.mhs.gov.au/media-releases/2017-09-01-evaluation-finds-considerable-positive-impact-cashless-debit-card>, accessed 22 September 2017.

[11] Paul Fletcher (Minister for Families and Social Services), Parliamentary Debates, House of Representatives, 13 February 2019, 13176.

[12] Department of Social Services, *Cashless Debit Card Trial Evaluation: Final Evaluation Report* (Canberra: Orima Research, 2017) 9.

- there were 'recording and collection issues with administrative data sets which reduced their reliability'.

The combination of these factors casts doubt on the validity of many of the conclusions reached in the report.

Academic commentators with decades of policy experience warned that the Orima CDC research contains methodological flaws so serious that it cannot be relied on as a foundation for the drastic changes to social security policy introduced by government.[13]

Eva Cox observes that 'Much of the data collected from participants, if carefully examined, is flawed and some of the more qualitative responses reported are also questionable.'[14] Cox notes that qualitative data came 'mainly' from 'local leaders, often white, who supported the introduction of the trials and are not neutral observers'.[15] She states that there were 'sampling problems' in terms of participants being drawn from 'passers by in public places' – which may have affected representativeness of the sample interviewed. Cox points out that payment for participants likely led to 'contamination of responses', as could the requirement to give their ID for a government survey where there may have been concerns about confidentiality being maintained.[16] This ID practice is disturbing in terms of potential ramifications for program participants. Importantly, Cox highlights that the type of questions put to participants raise serious ethical issues because answers to these 'could have legal implications and risk child abuse interventions'. No evaluation undertaken by government should require Indigenous program participants to give their ID in ways that can be linked to other data sets—because there can be no 'free, prior and informed consent'[17] as to how this data would be used in future.

The Orima Final CDC Report states that results from quantitative surveys were 'weighted',[18] but the methods adopted for such weighting may have affected the reliability of report findings. The Orima report acknowledges that the 'weighting aligned the distribution of the CDCT participant response sample with that of the CDCT population' and that 'the reported results of each survey wave were based on balanced

[13] Eva Cox, 'Much of the data used to justify the welfare card is flawed', The Guardian, 7 September 2017, <https://www.theguardian.com/commentisfree/2017/sep/07/much-of-the-data-used-to-justify-the-welfare-card-is-flawed?CMP=share_btn_link>; Janet Hunt, 'The Cashless Debit Card Evaluation: Does it Really Prove Success?' (*Centre for Aboriginal Economic Policy Research: Topical Issue No. 2/2017*, Canberra: Australian National University, 2017).
[14] Eva Cox, 'Much of the data used to justify the welfare card is flawed', The Guardian, 7 September 2017.
[15] Eva Cox, 'Much of the data used to justify the welfare card is flawed', The Guardian, 7 September 2017.
[16] Eva Cox, 'Much of the data used to justify the welfare card is flawed', The Guardian, 7 September 2017.
[17] UNDRIP Art 19.
[18] Department of Social Services, *Cashless Debit Card Trial Evaluation: Final Evaluation Report* (Canberra: Orima Research, 2017) 19.

population estimates.'[19] Location weighting was also added to this confusion of participant weighting in the calculation of results.[20]

There are equally troubling elements present in what Orima described as the 'qualitative' aspect of their research. In their Final Evaluation Report Orima stated that:

> Where anecdotal/"hearsay" sources were cited, the qualitative research sought to validate this directly from the source. However, when this was not possible or viable, only anecdotes that were heard three times or more from different community leaders, stakeholders and/or merchants have been used as evidence in the evaluation report.[21]

This means that a thrice repeated anecdote from people in positions of power within these communities was treated as evidence in the Orima evaluation. The report's conclusions are heavily reliant upon such anecdotes. However, anecdotes from parties in power thrice repeated do not meet a robust threshold for 'evidence' – they are more accurately described as opinions. The Orima report is heavily reliant upon these opinions in reaching conclusions that drastically affect financial and social inclusion opportunities for people in need of government income support payments. If opinions are to be seen as the overriding factor rather than evidence in government policy making in this arena then the opinions of those subject to the card ought to be given equal weight. Yet a thorough reading of the Orima report reveals that this is not the case. Where feedback by CDC participants conflicted with that given by community power holders the report writers regularly found in favour of the latter. This is deeply concerning. At best the Orima CDC research might be described as providing mixed feedback about various elements of the CDC program from a range of stakeholders, at worst it could be seen as the privileging of particular anecdotes or opinions masquerading as evidence—and Indigenous peoples deserve better than anecdote-driven policy.

The government has indicated that the CDC has capacity to reduce substance abuse—but any mechanism to accurately measure this was absent from the Orima evaluation. In Wave 2 of the Orima research the majority of CDC participants reported either no change in alcohol consumption, gambling or illegal drug use since using the CDC or an increase in these behaviours.[22] These two categories combined amounted to a higher figure than the 48 per cent of CDC participants who reported doing one or more of these three behaviours less since the CDC commenced.[23] It is significant that

[19] Department of Social Services, *Cashless Debit Card Trial Evaluation: Final Evaluation Report* (Canberra: Orima Research, 2017) 19.
[20] Department of Social Services, *Cashless Debit Card Trial Evaluation: Final Evaluation Report* (Canberra: Orima Research, 2017) 21.
[21] Department of Social Services, *Cashless Debit Card Trial Evaluation: Final Evaluation Report* (Canberra: Orima Research, 2017) 24-25.
[22] Department of Social Services, *Cashless Debit Card Trial Evaluation: Final Evaluation Report* (Canberra: Orima Research, 2017) 43.
[23] Department of Social Services, *Cashless Debit Card Trial Evaluation: Final Evaluation Report* (Canberra: Orima Research, 2017) 43.

Wave 1 of the research indicated that many people subject to the card did not experience problematic consumption of these products to start with—34% abstained from them altogether, and the amount initially consumed by the 43% who said that their consumption remained the same post CDC was unclear.[24] This 43% may not have consumed these products in excess. Saying a CDC cohort now consumes less of the three prohibited products and claiming that as a positive is not meaningful if their consumption levels were not problematic to start with.

As indicated by the Australian National Audit Office (ANAO) in their 2018 CDC inquiry,[25] there is no credible evidence of the success of the Cashless Debit Card. The ANAO state that 'monitoring and evaluation' of the CDC 'was inadequate. As a consequence, it is difficult to conclude whether there had been a reduction in social harm and whether the card was a lower cost welfare quarantining approach.'[26]

In their July 2018 Report the ANAO also pointed to other weaknesses of the CDC trial, including:[27]

- Poor Risk Management: The Department of Social Services (DSS) 'did not actively monitor risks identified in risk plans'. The ANAO stated 'While a CDCT program risk register was developed, the identified risks were not actively managed, some risks were not rated in accordance with the Risk Management Framework, there was inadequate reporting of risks and some key risks were not adequately addressed by the controls or treatments identified. In particular, treatments were inadequate to address evaluation data and methodology risks that were ultimately realised.' 'There were six risks rated as medium risk whereas the risk matrix indicates that they should have been rated high.'

- No cost benefit analysis: The evaluation organised by the Department of Social Services 'did not cover some operational aspects of the trial such as efficiency, including cost.'

- Weak use of data: 'There was a lack of robustness in data collection and the department's evaluation did not make use of all available administrative data to measure the impact of the trial including any change in social harm.' Notably:
    - *merchant sales data was not used to determine whether there had been any reduction in alcohol and gambling sales in the trial areas.*
    - *Other data was also excluded from consideration, such as some state crime data.*

[24] Department of Social Services, *Cashless Debit Card Trial Evaluation: Wave 1 Interim Evaluation Report* (Canberra: Orima Research, 2017) 20.
[25] Australian National Audit Office, *The Implementation and Performance of the Cashless Debit Card Trial* (Canberra, 2018) 8.
[26] Australian National Audit Office, *The Implementation and Performance of the Cashless Debit Card Trial* July 2018, 8.
[27] Australian National Audit Office, *The Implementation and Performance of the Cashless Debit Card Trial* July 2018, 8-9, 24, 29, 33, 37, 43, 45, 59.

- *School attendance data was also excluded:* 'Anecdotal information reported to the Minister suggested an increase in school attendance, but ANAO analysis of state data available to Social Services showed that attendance was relatively stable for non-indigenous students but it had declined by 1.7 per cent for [I]ndigenous students, after the implementation of the trial compared to the same period (between May to August) in 2015.'

- Inadequate review of key performance indicators: 'There was no review of the KPIs during the trial and KPIs have not been established for the extension of the CDC.' There were also 'issues with measurement and bias' for CDC KPI's. In addition, 'Key areas of the CDCT relating to the administrative and operational aspects of the trial such as the Social Services call centre, wellbeing exemptions, community visits, levels of cash available in the community and staff training were not measured with KPIs.'

- No baseline data, and evaluation not built into program design: 'Social Services did not build evaluation into the CDCT design, nor did they collaborate and coordinate data collection to ensure an adequate baseline to measure the impact of the trial, including any change in social harm.'

- Problematic procurement of the CDC provider and evaluator: 'Aspects of the procurement process to engage the card provider and evaluator were not robust. The department did not document a value for money assessment for the card provider's IT build tender or assess all evaluators' tenders completely and consistently.' Notably, Orima as the selected evaluator initially provided a tender that was more than twice the cost of its competitors, at $922,592, and the cost of their 'evaluation was $1.6 million, over double the initial amount agreed'.

- Evidence base lacking: 'Social Services regularly reported on aspects of the performance of the CDCT to the Minister but the evidence base supporting some of its advice was lacking.'

Taken together, these factors paint a very different picture about the operation, monitoring and evaluation of the CDC than that asserted in Ministerial proclamations.

This Orima research is an example of bad evaluation practice, but rather than address its many ethical and other shortcomings, government Ministers responsible for the CDC have continued to claim this evaluation provides sufficient rationalisation for program expansion. All of which highlights the need for formal ethics review of government evaluations. It is not appropriate for evaluation projects that affect Indigenous peoples to be exempt from formal ethics review processes. Moreover, these ethics review processes should be robust.

I have also had the benefit of reading the final draft of the submission prepared by the Accountable Income Management Network, of which I am a member, and endorse its recommendations to this inquiry:

Recommendation 1
That all evaluations of programs imposing limits on a person's rights and choices are required to examine whether there are alternative ways of achieving the intended outcomes with less personal restrictions.

Recommendation 2
That appropriate baseline data be collected in a timely manner to accurately inform policy and program monitoring and evaluation.

Recommendation 3
That all policy and program design, monitoring and evaluation are conducted in line with Australian human ethics standards, particularly those specified in the Australian Institute of Aboriginal and Torres Strait Islander Studies' (AIATSIS) *Guidelines for Ethical Research in Australian Indigenous Studies*.

Recommendation 4
That evaluation agencies remain distinct from agencies involved in policy or program implementation and delivery.

Recommendation 5
That a standard procedure be developed to ensure government agencies' timely and appropriate responses to evaluation findings, and that this be made public.

Recommendation 6
That government agencies be required to respond to evaluation findings in order to meaningfully address any concerns raised therein, even if this means ending or replacing a particular policy or program.

Recommendation 7
That evaluation reports are published publicly and in a format accessible to persons subject to the relevant policy or program.

Recommendation 8
That Indigenous communities affected by government policies or programs are adequately resourced to provide input, interpret and respond to evaluation reports.

Recommendation 9
That government conduct follow-up meetings with Aboriginal and Torres Strait Islander persons and communities subject to the policy or program being evaluated to discuss evaluation findings and take on further feedback.

If I can be of any further assistance, I would be happy to oblige.

Yours sincerely,

Dr Shelley Bielefeld
ARC DECRA Fellow/Senior Lecturer
Griffith Law School
Arts, Education and Law Group
Building N61, Nathan campus, Griffith University
170 Kessels Road, Nathan, QLD, 4111, Australia

and

Visiting Scholar
School of Regulation and Global Governance (RegNet)
College of Asia and the Pacific
8 Fellows Road
The Australian National University
Acton ACT 2601 Australia