Legislation Inquiry Social Services Legislation Amendment (Cashless Debit Card Bill)

I write to you as an experienced social science researcher with over 30 years of experience in the fields of international and Indigenous development. I am as concerned about the situation of Indigenous people in Ceduna and the East Kimberley as anyone, and very much want to see their lives improve. However, I am also very much driven by evidence about what works, and as a social science researcher am concerned that the evidence provided for policy making is the most robust and credible as possible.

This legislation seeks to make possible the extension of the Cashless Debit Card trial in Ceduna and the East Kimberley and facilitate the expansion of this program geographically. My concern is whether the evidence of the evaluation supports this continuation and expansion.

My interest in this was sparked when the Wave 1 Report was released in March this year, and I decided to look at what the evaluation said. I was shocked when I read the report, as the Minister had already announced that the trial was a success and would be continued indefinitely. When I read the report, I discovered that it was extremely flawed and did not provide adequate evidence to draw the conclusions that had clearly been drawn. As I was extremely concerned at the poor quality of the evidence on which the Minister had made his decision, I wrote a critique of the Wave 1 Report, which was peer-reviewed and published by CAEPR as a Topical Issue. I attach that critique. http://caepr.anu.edu.au/Publications/topical/2017TI1.php.

Given my concerns about the quality of the Wave 1 Report and the Minister's interpretation of data from it, I was naturally interested to see whether the Wave 2 Report was a better report. In this case, the Report addressed some of my concerns (e.g. it gives better contextual information at the front), but remains problematic. In addition, where the Report's authors have qualified their positive findings with many caveats, these have been ignored by the Minister in his public statements about the evaluation.

Eva Cox[1] has highlighted many of the problems with the Wave 2 survey design, the way interviews were conducted, and the ethics of the process, all of which would suggest that the results presented should be treated with great caution. Her criticisms of the evaluation process are valid But I have tried to explore what can be drawn from the data that is presented, flawed as it is. Is there any evidence that this trial is achieving its stated objectives?

First, it is important to emphasise some problems with the evaluation design and the reporting of results, which create problems in trying to make sense of the data presented:

- People were approached for an interview, by people they would not have known, in public places, about a government program. If they agreed to be interviewed they were asked for their ID. This may well have affected any of people's answers, as Eva Cox notes. But in particular, having provided ID, their answers to questions about use of illicit drugs or any other activity that might be illegal or reportable, would almost certainly avoid revealing any such activity. The fieldwork was conducted shortly after the 2017 Federal Budget announcement of proposed drug testing of people on welfare, when this would have been

[1] https://www.theguardian.com/commentisfree/2017/sep/07/much-of-the-data-used-to-justify-the-welfare-card-is-flawed?CMP=share_btn_link

particularly sensitive.[2] The ethics of this approach is dubious, and the results likely to be of little value.

- The data from the two sites are weighted equally which favours the findings from the Ceduna sample that are slightly better than from East Kimberley. Yet the East Kimberley has by far the majority of the CDCT participants (1,247 compared with 757 in Ceduna at the outset), and their responses are thereby discounted. The sample should have been in proportion to the participant numbers in each site to give a true picture of the trial outcomes.
- While the report provides initial guidance on the confidence levels required for statistical significance of the reported findings, it rarely cautions in relation to data it provides where the statistical significance of results is very dubious due to small numbers. This can give a misleading impression about change in a number of places throughout the report. Such change may just be due to variation in the sample of respondents, and not reflect a statistically significant difference.
- The sampling approach in Wave 2 is a strange mixture of a longitudinal sample and systematic intercept sampling; whilst much is made of the longitudinal sample in the early part of the report there is absolutely no outcome data provided from that sample of 134 people who were recontacted from the Wave 1 sample. Instead this group was added to the new intercept sample from Wave 2 without explanation of the reasons for doing this. This is to add a non-random sample (people who could be contacted again) to a random sample. Further it is hard to see the Wave 1 and Wave 2 samples as comparable, when in the first Wave, 31.5% said they never drank, gambled or used illicit drugs but in Wave 2, almost 42% said they never did so. Whilst the evaluators say they applied a number of statistical procedures to deal with some of these issues, the logic and rationale for what they have done is very unclear.
- The Wave 2 data is presented differently in some respects from that in Wave 1 so that it is difficult if not impossible to make comparisons. For example, in relation to alcohol, Wave 1 reports data from participants and family members together but Wave 2 only reports data from participants, as family members were not interviewed. So the results are not comparable.
- Overall, the design of the evaluation appears to take little account of the many important principles for conducting research among Aboriginal and Torres Strait Islander communities set out in the AIATSIS Guidelines for Ethical Research, and makes no mention of them.[3] Evaluation is a form of research, and the participants in these trials are overwhelmingly Aboriginal and Torres Strait Islanders.

What was the trial supposed to achieve?

According to the Orima Initial Conditions Report (2016,p*i*), this trial is 'to deliver and manage income support payments (ISPs) with the aim of reducing levels of community harm related to alcohol consumption, drug use and gambling.' Of these, the greatest concerns the community expressed

[2] http://www.abc.net.au/news/story-streams/federal-budget-2017/2017-05-12/federal-budget-2017-pm-says-welfare-drug-test-plan-based-on-love/8520564

[3] https://aiatsis.gov.au/sites/default/files/docs/research-and-guides/ethics/gerais.pdf

before the trial began were about alcohol, with some also fearing that drug problems, notably ice, could increase in the future; and although gambling was present, there was less concern about its effects. Concerns about high levels of crime and violence were associated with alcohol in particular.

The program logic suggested that after 12 months, there should be sustained reductions in alcohol consumption, illicit drug use, and gambling resulting in less criminal and violent behaviour, fewer alcohol-related injuries and an increased sense of safety[4]. A number of performance indicators and sources of data to assess these indicators were identified. Bearing all the caveats above in mind, I have tried to understand the key results against these indicators, with a particular focus on the views and behaviours of the CDCT participants themselves.

Alcohol reduction

The Wave 2 report focusses on what people said about *change in the amount of alcohol* they consumed since joining the trial rather than their reports about *current alcohol useage*. These reports of change were positive, indicating that people thought they drank less than before the trial commenced. However, such recall over a year is not likely to be very reliable, and given the context of the interviews, people may have said what they thought the interviewer wanted to hear. The reporting of 'alcohol behaviours done lately' which might have given more reliable data than reports of change over time, is impossible to compare from Wave 1 to Wave 2. In Wave 1 data presented is for participants and family together, while in Wave 2 data is given for participants only, and only those who drink at all. Thus we cannot tell if reports of *actual behaviours* show any change. It would have been perfectly possible to present the participant only data from Wave 1 with the same for Wave 2 but that was not done.

There is also a question about the program logic behind an expected reduction in alcohol consumption between Wave 1 and wave 2 reports. The report says that people reported *a change in their alcohol consumption* between Wave 1 and Wave 2. At Wave 1, participants were already receiving their income support payments through the CDC, so their ability to purchase alcohol was already restricted. As welfare recipients it seems unlikely that they would have savings to draw on to purchase alcohol, which might reduce as time passed. So what is the program logic that would support the idea that alcohol consumption would continue to reduce many months after the CDC was first operational? That is unclear.

If self-reports of alcohol consumption may be influenced by individual's concerns that other sanctions could be introduced if their alcohol use has not dropped, participant reports of change in the community may be more likely to be accurate than their reports of their own alcohol use. Fig 12 (p.47) presents participant perceptions of change in alcohol use in the community at the two sites since the trial started. The results are very mixed. For example, in East Kimberley 20% of respondents say there has been more drinking and 18% say there has been less. In Ceduna, 14% say more, 23% say less, but 25% can't say. The largest proportion in each site say the level of drinking is the same. Non-participants in the trial have a more positive view. It is very unclear why there is such variation in these views and this is not investigated further, which it should have been.

There is also no sales data from liquor outlets checked against people's reporting, but there are anecdotes which suggest change in the right direction. In contrast to the Wave 1 report, there has been some attempt to separate the impacts of simultaneous alcohol restrictions from those of the CDCT, which suggests most of the change reported (if it is to be believed) is attributable to the CDCT. However, overall, this data raises as many questions as it answers, and if in fact there has been a

[4] Fig13 Program Logic pA8 ICR

significant decline in alcohol use, then there are further questions about the program logic behind the trial, which are explored below, as the community harms thought to be attributable to alcohol appear to persist.

Gambling

The Wave 2 report suggests that there is reduced gambling, however there were a number of qualifications to claim in the Report which were completely ignored by the Minister. These included that this did not seem to be the case in the East Kimberley, where both participants and non-participants[5] were more likely to say that they thought gambling had gone up.

In Ceduna the issue is poker machine use, and so revenue data from poker machines can provide some more objective measure of change (although clearly many people who use the poker machines are not on the CDCT). The available data on revenue from poker machine gambling however, covers an area far larger than Ceduna and reflects a 12% reduction over the twelve months following the introduction of the CDCT. The report makes clear that only 40 out of 143 of the poker machines which the data covers are in the CDCT area. This could suggest that a 12% reduction in gambling revenue over a year was not predominantly due to the CDCT, but due to other factors across the region. Or the drop may be focussed in the CDCT area. There is no further investigation about this in the evaluation report, so it is hard to draw conclusions.

What is noticeable from Fig 19 on poker machine revenue (p59) is that the level of revenue fluctuates through the year, and has increased in the three months since Jan 2017 to a level higher than in April 2016, suggesting no clear downward trend in gambling is apparent, even if expenditure on gambling has reduced. In fact a stronger downward trend was evident in 2015-16 before the trial commenced. In summary, the data presented cannot confidently support claims that gambling has significantly reduced at both sites.

Illegal Drug use

The data about illegal drug use is probably the least reliable. Importantly, the Wave 2 results may be considerably affected by the publicity about drug testing of welfare recipients, particularly just prior to the Ceduna fieldwork in May 2017. Furthermore, although self-reports suggest a drop in illegal drug use, the numbers of respondents are small and the reliability of the data in such small numbers is low. Using Orima's own guidance about the confidence one could have in the statistical significance of the results, the possible reduction may be far smaller than first appears.

Other performance indicators

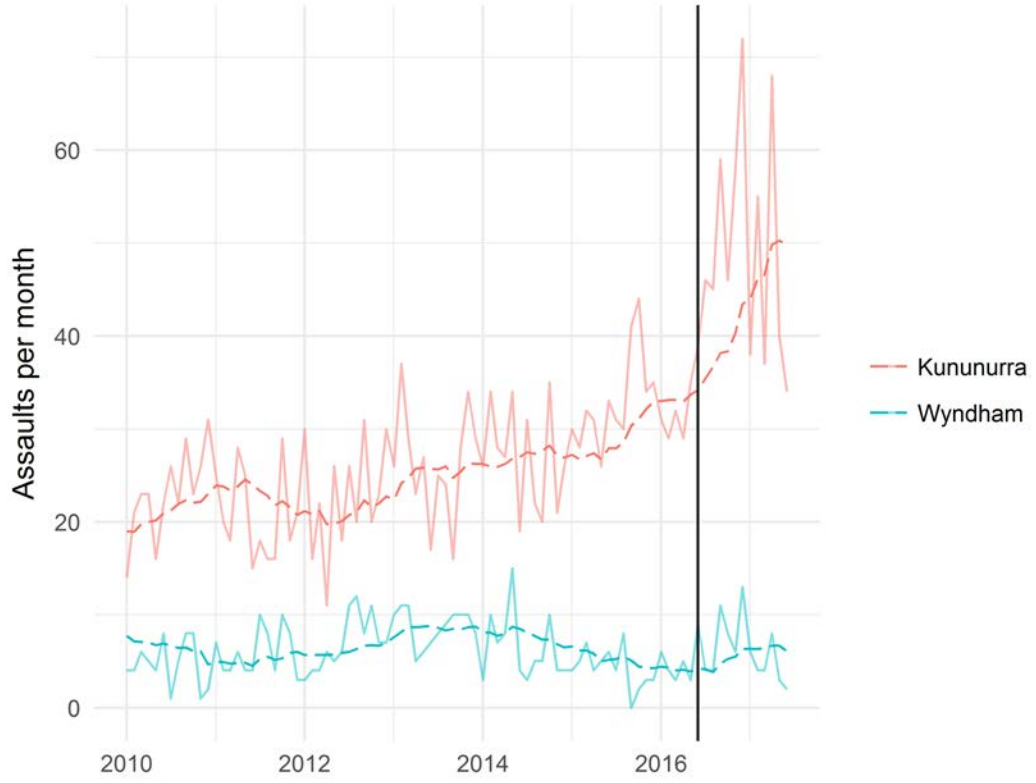
Rates of drug and alcohol related injuries and hospital admissions were listed as performance indicators and some data is presented which suggests that alcohol-related attendances at hospital emergency and outpatients departments in Ceduna have dropped. In East Kimberley the report says that there have been fewer alcohol-related pick-ups by the Community Patrol. However, there may be other explanations for the latter which are not explored and ruled out, for example whether the Community Patrol was functioning every night throughout both periods that were compared.

The percentage of respondents feeling safe was another indicator, and the report acknowledges that that there was 'no statistically significant change' between Wave I and Wave 2 data collection on participant and non-participant feelings of safety. Concerns for safety at night remained, particularly in the East Kimberley.

[5] The non-participant result was not statistically significant however.

Finally the indicators for violence and other types of crime and violent behaviour were to include police reports as well as perceptions of participants and others. No administrative data is provided for any of these, so the only data provided is perceptions of those interviewed.

Interestingly the views of CDCT participants were very mixed on this, and in the East Kimberley more participants thought that violence had increased than thought it had reduced. This is certainly borne out by data on assault offence/incidence reports from the WA Police which rise sharply around the time the CDCT began in the East Kimberley in mid-2016, as the figure below indicates.[6] This data itself needs to be treated with caution as there may have been a major change in policing behaviour that contributed to such a sharp rise in such reports, but it is consistent with the CDCT participant perception data. In relation to crime, the report itself states that administrative data did not show evidence of reduced crime since the trial began, and in fact crime increased in the East Kimberley as it did in Derby, a comparator site. This suggests that the CDC was not able to counter whatever is causing this crime.



Looking at the bigger picture

The CDCT was designed to reduce the levels of harm underpinned by the three behaviours targeted. In the early stage of the trial the community consultations identified the adverse consequences of these behaviours as relating to:

[6] https://www.police.wa.gov.au/Crime/Crime-Statistics-Portal
The vertical black line indicates 1 June, when the roll out of the CDCT in the East Kimberley was almost complete.

- Health effects
- Safety and security
- Financial problems
- Social problems such as humbugging and unemployment
- Inability to secure stable housing and overcrowding
- The impacts on the wellbeing of children.

Whilst one cannot expect major change on all these fronts in 12 months, what is of concern is that there appears to have been limited or no change in relation to many of these adverse effects identified by the communities before the trial began, *even if the reductions in the behaviours targeted are real*. As indicated above, there appears to be no change in perceptions of safety and in fact in East Kimberley perceptions of safety after dark may have worsened. Further, there seems to be no reduction in perceptions of violence or in assaults, whether domestic violence or other, and data from the East Kimberley suggest that things may have got substantially worse.

The one key area where some positive change may be emerging is in financial management – the card does appear to be helping some people manage their money better, and there are various pieces of evidence that indicate this. In all the other areas the data reveals no change or is very mixed. Health gains would be too soon to see, except where underlying health problems are now more evident, and that may be the case in a few instances according to the report.

However, the real problem, which the CDCT does nothing about, is the level of poverty people are experiencing. And as the report itself says , 'on average across the two sites, at Wave 2, participants were more likely to indicate that it *(i.e. the CDCT)* had made their lives worse than better.' (p 82). The data presented says that 23% said the trial made their lives better and 32% said it made their lives worse. It did not explore whose lives were getting better or worse. Given that many participants in each of the samples never undertook any of the three behaviours the card was targeting, I would want to know if their lives were made worse, and I would want to know if those whose lives were better were actually any of the targeted individuals. The report does not explore this, so we really do not know where any benefits are being felt or where serious problems may be occurring.

While some reports suggest parenting and family well-being may be improving, there is data which suggests this is not the full story. The report shows that around a quarter of participants run out of money for food at least every two weeks, and over half have run out of money for food in the last three months, and this may be worsening. And there are mixed findings in relation to children's wellbeing. Around 44-45% said they had run out of money to pay for essential non-food items for children (like nappies, clothes, medicine) in the last three months, and 19% had done so at least every two weeks. Such findings in themselves should raise alarm bells. If participants whose income is so firmly constrained through the CDCT cannot feed themselves or buy essentials for their children then there is a problem far larger than the card can address. In addition, parents gave mixed reports about the impact of the trial on children's lives with 17% saying it had made their child's lives better and 24% saying it had made children's lives worse (p6).

There are mixed reports about humbugging with some saying it has reduced and others experiencing more humbugging. Although there is a slight rise in people looking for work, it is hard to know if that is statistically significant, and whether it relates to the CDCT or to the pressure from the CDP program (in East Kimberley in particular). The fact is that more economic development initiatives are needed to help create suitable jobs in these locations or people will simply not be able to exit from the CDCT. The other concerns expressed at the outset of the CDCT, housing and overcrowding, are not addressed at all by the CDCT.

The use of increased services

Associated with the CDCT was funding for increased services. The report does not make clear exactly what those service increases were in each location, but does conclude that the card, rather than the services, has had the greatest impact on the result. There seem to be several reasons for this: the significant delay in providing additional services; the narrow range of services provided; and the lack of awareness on the part of trial participants of the services available. Some people had obviously found some value in the services that they had used. The contribution services might make in the future could be greater, one assumes, as they become better known, and perhaps if a broader range were provided to address the many issues identified above.

Conclusion

The Prime Minister was in Western Australia on 3 September, claiming the enormous success of the trial.

> *It's seen a massive reduction in alcohol abuse, in drug abuse, in domestic violence, in violence generally; a really huge improvement in the quality of life, not just for the families who are using the Cashless Welfare Card, but for the whole community. But above all, above all it's an investment in the future of the children.*[7]

Someone needs to tell him that the report commissioned by his Minister does not say that, and that the evaluation undertaken has serious flaws. So what to do?

There are two ways to think about what conclusions we can draw from the trial about the CDC program and its intent. First, perhaps, despite all the flaws in the evaluation, there has actually been positive change on the ground in relation to the three behaviours targeted. If that is the case, these behaviour changes do not appear to have had much impact on the harms that the program was supposed to address, particularly in relation to safety and violence which were the community's big concerns. If so, the program logic has been built on some wrong assumptions, such that despite any behaviour changes, the underlying problems remain and the program needs rethinking.

The other way of thinking about this is to suggest that perhaps the program is not reducing the alcohol, drug and gambling behaviours it was meant to target. This could be because people are finding ways around the constraints of the card, or because the problems require far more than a card to solve them. In which case the program also needs rethinking.

What is clear is that the complex and interrelated problems of drug and alcohol abuse, poverty, unemployment, poor or overcrowded housing, and violence need solutions that will work to improve the overall wellbeing of adults and children. These solutions are likely to be multi-faceted and undertaken with strong engagement of the people whose lives they are meant to improve, not imposed in a punitive way. Senator Patrick Dodson has called the trial 'a public whip'[8], and one of its influential Kimberley advocates is now saying it is not working[9].

---

[7] https://www.malcolmturnbull.com.au/media/address-to-the-wa-liberal-party-state-conference-3-september-2017

[8] https://www.theguardian.com/australia-news/2017/aug/22/pat-dodson-says-cashless-welfare-card-a-public-whip-to-control-indigenous-people

[9] https://www.theguardian.com/australia-news/2017/aug/23/aboriginal-leader-withdraws-support-for-cashless-welfare-card-and-says-he-feels-used

On the basis of the evaluation the Government cannot legitimately claim the success it is claiming and it should not roll out any more of these trials at the present time. The results are too poor and ambiguous to warrant the public expenditure.

Acknowledgments

I am grateful to Rob Bray and Francis Markham for assistance in making this analysis of the Orima Wave 2 report. However all responsibility for the accuracy of information in it and for the analysis remains mine.

8 September 2017.

Any evaluation of government programs is complex and difficult, and attributing particular outcomes to specific program interventions in particular places is always tricky. In order to deal with this problem, which is inevitable in the real world, evaluators use an approach known as Contribution Analysis. http://www.betterevaluation.org/en/plan/approach/contribution_analysis. It would have been far better if such an approach had been used for this evaluation. This is particularly the case, since from the outset in both locations there were simultaneous alcohol restrictions in place and it would be important to get a sense of the contribution to change that they might be making, as well as the contribution of additional services that were supposed to have been provided as part of the trials. Furthermore, this program had a program logic set out in the Initial Conditions Report, and Contribution Analysis is particularly well suited to a context where a program logic is clear.