



Monday August 31st, 2020

I have been in touch previously supplying you with a draft of the essay which was published today by *Inside Story* at this url:

<https://insidestory.org.au/orwell-that-ends-well/>

I would appreciate it if you would accept the essay, which is attached to this cover letter, as a formal submission on your Draft Indigenous Evaluation Strategy. I wish the Commission all the best in its further deliberations.

As we all know, these are issues our country has never mastered, or to be frank, got very close to facing up to. And even then, facing up to them is just the first step.

Regards,

Nicholas Gruen
Lateral Economics

Orwell that ends well?

NICHOLAS GRUEN

31 AUGUST 2020

Can the latest push to evaluate Indigenous programs really Close the Gap?

Right:

Prime minister Scott Morrison (right) with the lead convenor of the Coalition of Peaks, Pat Turner, and Indigenous Australians minister Ken Wyatt during a press conference on Closing the Gap on 30 July.
Lukas Coch/AAP Image



Humility is not a peculiar habit of self-effacement, rather like having an inaudible voice. It is selfless respect for reality and one of the most difficult and central of all virtues... Humility is a rare virtue and an unfashionable one... Only rarely does one meet somebody in whom it positively shines, in whom one apprehends with amazement the absence of the anxious avaricious tentacles of the self.

— Iris Murdoch

The divergence between the facts established by the intelligence services — sometimes by the decision makers themselves (as notably in the case of McNamara) and often available to the informed public — and the premises, theories, and hypotheses according to which decisions were finally made is total. And the extent of our failures and disasters throughout these years can be grasped only if one has the totality of this divergence firmly in mind.

— Hannah Arendt

I have a theory that the truth is never told during the nine-to-five hours.

— Hunter S. Thompson

1. Introduction

From 1788 till the 1960s, Europeans established themselves on Indigenous land in a brutal regime, first of dispossession and then of disregard. Yet some among them had strikingly good intentions. A year before Wilberforce took on the cause, nearly eighty years before the Emancipation Proclamation, Arthur Phillip accepted his commission insisting that slavery had

no part in the new colony. Phillip sought to treat the Indigenous people fairly, at least according to his own lights. But mutual incomprehension reigned and those with murkier intentions soon prevailed.

Today, good intentions abound, though racism often lives on in unacknowledged assumptions. Governments outlay vast sums, whether adequate or not, on specific Indigenous programs and in general expenditure on Indigenous health, education and social security. Widely supported grand gestures are announced every few years. You might think “Closing the Gap” was Kevin Rudd’s idea but it rebooted (or is that rebranded?) a Hawke government initiative of twelve years earlier. But the results are meagre.

Now comes a new cycle of activity, this one focused on whether formal evaluation processes might allow us to identify and scale up those Indigenous programs that actually “work.” Most recently, the Productivity Commission has been hard at work on a national Indigenous Evaluation Strategy, which was the immediate trigger for this essay and to which I’ll return. Will this cycle of activity produce better results than earlier efforts? I’ll explain below why I have my doubts.

To first clarify where I’m coming from, it is not from deep knowledge of Indigenous policy. My focus here is rather on a prior question: how our formal institutions of government — and most particularly our bureaucracies — might need to change to succeed where previously they have so consistently failed. To make that question concrete I draw on my experience in other intractable areas of social policy that bear family resemblances to Indigenous policy.

Programs to protect children from abuse and neglect, particularly in disadvantaged families and communities, follow the same endlessly repeated cycle of failure followed by grand plans for reform that then run into the sand before the cycle begins again. This essay focuses on how *little* the system really appreciates the distance it would need to travel to really be effective, in terms of either its own values and objectives, or those of the disadvantaged communities — including Indigenous communities — it claims to be serving.

2. The “what” and the “how,” the saying and the doing

So here’s my very simple description of the problem: despite endless pronouncements of *what* we must do, there’s minimal comprehension of *how* to do it.

This is an endemic problem. Dependable know-how itself — whether it’s improving outcomes in an Indigenous community or representing government in the High Court — is not directly *legible* to government systems. Anyone can *claim* to have that know-how but a bureaucracy needs something more dependable than that. As a consequence, it will interact with know-how as a certified, decontextualised “what.” That “what” could be a credential, the meeting of a key performance indicator, or a particular bureaucrat’s informal reputation for being a “good operator” or a “safe pair of hands.” In improving Indigenous lives, however, know-how won’t align with any such things, not least because so much of it resides among Indigenous people and communities themselves. We need to access *their* knowledge and *their* agency to improve their own lives in ways that matter to them.

The cliché used to convey this idea is “putting people at the centre” or “putting people first.” However well-intentioned such slogans are, more often than not they operate as a kind of doublethink — as if adopting the slogan were to put its intent into practice.

The philosopher Martha Nussbaum offers a story that illustrates this difference between saying and doing. She describes how a development program encounters a woman in a traditional rural community who is uninterested in education for herself or her children. Nussbaum is showing how our (reductive) framing of the other’s perspective can cut us off from the wisdom of the other’s lifeworld. “Clearly,” she writes, “a one-shot logical argument” wouldn’t be enough to engage the woman:

[S]uch a procedure would only reinforce her conviction that education has nothing to do with her. Nor would the exchange get very far if the development workers sat down with her... asking... calm and intellectual questions about what she thinks and says. But suppose, instead, they spent a long time with her, sharing her way of life and entering into it. Suppose, during this time, they vividly set before her stories of ways in which the lives of women in other parts of the world have been transformed by education of various types — all the while eliciting, from careful listening over a long period of time, in an atmosphere of trust that they would need to work hard to develop, a rich sense of what she has experienced, whom she takes herself to be, what at a deeper level she believes about her own capacities and their actualization. If they did all this, and did it with the requisite sensitivity, imagination, responsiveness, and open-mindedness, they might over time discover that she does indeed experience some frustration and anger in connection with her limited role; and she might be able to recognise and to articulate wishes and aspirations for herself that she could not have articulated to Aristotle in the classroom. In short, through narrative, memory, and friendly conversation, a more complicated view of the good might begin to emerge.

Nussbaum’s scenario is based on actual fieldwork in Bangladesh, and couldn’t be cost-effective if it involved professionals engaging rural women en masse. But, as I discovered when I was chairing the Australian Centre for Social Innovation, the *spirit* of this translational endeavour is already captured in existing and cost-effective programs in Australia.

The centre’s Family by Family program takes families who feel they’re close to crisis. A trained coach then takes each family through a structured program of mentoring by another local family that has come through similar stresses. The family seeking help chooses the family that mentors them and sets the objectives they want to work on.

The program was co-designed with families over many months, but the simplicity and obviousness of the end result gave those involved in it and many lookers-on numerous “aha” moments. Family by Family embodies the rare art of professionals vacating centrestage in a therapeutic intervention to create space for those who must do the real work. Professional knowledge, which grows with the program, is always there — but as midwife, not obstetrician.

Talking to some of these families, I was struck by their visceral engagement with the program and their mentors. To take just one example — of which there were many — one mother in the

program had received twenty-seven statutory “notifications” documenting outsiders’ suspicions that she was neglecting her kids. The relevant department was heading to court to take her four kids into guardianship. When her mentor family took her family camping, she learnt many things from them — not least to hug her kids. The department stopped proceedings against her.

The thirty-week program cost around \$13,000. If that sounds expensive, it’s a fraction of what social workers would have cost, and much more effective. Moving all four kids into care would have cost around \$224,000 per year. So, if Family by Family steered just this family from the shoals of state intervention it probably paid for its development and first couple of years of operation.

Before I saw Family by Family in action, I’d have described my outlook as that of a tragic liberal — committed to fairly generous spending on social disadvantage, but with very modest expectations of how much it could turn things around. After seeing Family by Family, the penny dropped. Ingrained patterns and social reinforcement are immensely powerful, almost immovable forces. But people’s desire to work towards better lives for themselves, their families and their communities is similarly elemental if they can somehow unlock their own agency and that of those around them.

3. Lord Acton’s fault line

After acknowledging the vast gulf between identifying the “what” and mastering the “how,” between the saying and the doing, we should . For decades I’ve referred to it in asides, but it needs to be brought centre stage so we can look it in the eye. It’s significant that it’s a joke, just as it’s significant that so many of the best insights into bureaucracy are provided by comedies like *Yes Minister*, *The Office* and *Utopia*.

More than a century ago Lord Acton quipped that rowing was the perfect preparation for public life. Why? Because you face in one direction while moving in the other. One crucial reason that we’ve made so little progress is that in a thousand ways, large and small, the actors in the system face in one direction — with their mission statements, corporate values, strategic plans, evaluation strategies and all the rest of it — while moving in the other.

Of course, they’d prefer to do a good job — most people would. But when push comes to shove, their animating imperative isn’t to keep progress going in the field. It’s to keep up appearances. Seen this way, all those grand announcements we keep making are part of the problem. They’re really directed at *our own anxieties*. They alleviate and distract us from facing our disappointment — our discomfort — that the world remains so resiliently impervious to our good intentions.

Lord Acton’s fault line appears between the two feet on which we stand — between what we say and what we do. That’s why the words we use matter so much, and why we should take George Orwell’s advice to choose the simplest and clearest words we can. As he put it:

If you simplify your English, you are freed from the worst follies of orthodoxy. You cannot speak any of the necessary dialects, and when you make a stupid remark its

stupidity will be obvious, even to yourself. Political language — and with variations, this is true of all political parties [and here we can include officialese]... is designed to make lies sound truthful... and to give an appearance of solidity to pure wind.

Since those in the system are the ones with the power, all we have to appeal to is their own self-respect — their own desire to feel better about themselves. When they say they want to change, the real question is how much. The system has said that it wants greater Indigenous agency in its programs for ages. But as I'll illustrate, our programs are so dominated by that same system's routines and perspectives that Indigenous agency barely gets a look-in. Instead it gets reduced to things that are legible to the system — such as Indigenous ethics codes and certified cultural sensitivity. These things may have some benefits. They may also have costs, which I'll discuss. But they are mostly the system saying rather than doing.

This takes us to the nub of the problem. It is only humility, or some institutionalisation of it, that can create that space within which Indigenous agency might be nurtured and grow. But “humility” itself is now turning up as a cliché in all those “how to” guides (it appears just before “nuance” and after “authenticity” — yes, authenticity really was a corporate value of PwC for a while there). So I've tried to revivify it with Iris Murdoch's magnificent words above. For the non-Indigenous among us who fancy we care, we must find ways to untangle ourselves and our institutions from the “the anxious avaricious tentacles of the self.”

4. Enter evaluation

Like a patient resisting therapy, the system constantly initiates new beginnings. But Lord Acton is never far away. At the political level leaders talk of evidence-based policy, but then shunt it aside when convenient. In fact, substantial performance evaluation was built into the structure of the Aboriginal and Torres Strait Islander Commission but sidelined after ATSIC was dismantled by John Howard's government. The failure of the Northern Territory Intervention to take an evidence-based approach is legendary, worked up as it was over a few days in Canberra in the run-up to an election and yet largely maintained by the incoming government.

More recently, while stressing his own commitment to following the evidence, newly elected prime minister Malcolm Turnbull expanded income-management schemes without mentioning that the independent evaluations were highly equivocal. However well the idea played in non-Indigenous Australia, the evaluations suggested that compulsory income management has clear, positive impacts in very few cases and gives rise to “considerable feelings of disempowerment and unfairness.” As one might expect, voluntary income management is more successful.

Now, it is one thing for senior officials not to speak publicly of their political masters' hypocrisy. But their complicity goes deeper. In 2009, a finance department review of Indigenous expenditure stressed the need for “a more rigorous approach to program evaluation at a whole of government level.” In 2016, the nation's most senior public servant, the secretary of the Department of the Prime Minister and Cabinet, Martin Parkinson, echoed those sentiments. In response to such concerns, \$40 million over four years was allocated for

evaluation. Parkinson's department was responsible for Indigenous affairs, but the Audit Office reported three years later that its performance was desultory.

As ANU researcher Michael Dillon has suggested, even the Audit Office's report was "extraordinarily hedged and timid, and failed to make a substantive assessment of the actual independence of the evaluations undertaken" by the department:

Of thirty-five evaluations on the department's 2018–19 workplan, fifteen had not commenced. Of the remaining twenty, eight had been published and twelve withheld from publication... In at least four cases (involving very significant and sensitive program evaluations) the department was waiting to brief the minister or awaiting his noting of a brief. In plain language, the minister was preventing timely publication of the evaluations.

Further, Dillon observed, Parkinson's response to the audit "fails to acknowledge or address in any way the negative content of the audit." Is it likely that the system will engineer something better if it can't acknowledge its own failure to do as it says?

Which brings us back to the Productivity Commission's Indigenous Evaluation Strategy, a draft of which was released in June. The PC has always attempted to pitch its proposals to government within the "Overton window" — that range of options that will be taken seriously by powerful people. Given that constraint, as I'll explain, I respect its compromises on policy. But the point of the PC's independence is that, however much it compromises on the policy, it spares no one, least of all itself, the truth. What the great scientist Richard Feynman wrote about science is also true of social science. For me, it's a holy grail of social policy and aligns nicely with Orwell's advice: "The first principle is that you must not fool yourself, and you are the easiest person to fool."

5. Putting Indigenous people at the centre: the words

There's a kind of ambiguity at the very heart of the PC's draft strategy that's increasingly common. It's Orwellian in the bad sense. I guess the genre was introduced into polite society by the "vision statement." Here one states an aspiration as a fact. You know the kind of thing: "PHP Residual Solutions is the world's foremost residual solutions provider." At least in its awkward baldness, it's not misleading. We all know that global domination is an aspiration, not a fact.

But this fusion of fact and fancy appears as the fundamental building block of the PC's draft strategy: "The Strategy puts Aboriginal and Torres Strait Islander people at its centre, and recognises that governments need to draw on the perspectives, priorities and knowledges of Aboriginal and Torres Strait Islander people if outcomes are to improve."

One of the ways to ensure we remain fixed to the spot with Lord Acton's fault line yawning beneath us is to encourage the idea that *saying* something is *doing* it. Does the PC know how to put Indigenous people at the centre of its strategy? Can it point us to better and worse examples of doing so? Can it highlight cautionary tales where grand claims have been made that are belied by the facts on the ground? These are some of the questions — pointed,

uncomfortable questions — that we need to answer if we're ever to step over Lord Acton's fault line and enter the promised land of "how."

At the level of programs, rather than evaluation, there are at least two perilous steps in the expedition to get from saying to doing — from signing the cheques to putting the resources of government properly at the disposal of Indigenous people and their communities:

1. We need to learn how to put Indigenous people and communities at the centre of these programs — or, to put it differently, how to realise their agency within them.
2. Then we need emerging successes to spread. That requires validated new knowledge of what's working in the field — always fragile in large organisations to say nothing of systems of organisations — to trump the institutional imperatives that so often frustrate the spread of successful practice.

To me, these are the great priorities for the Indigenous-specific programs I have focused on in this essay, though analogous priorities would apply when considering the impact of general welfare programs on Indigenous people and communities. And any evaluative strategy would emerge from an appreciation of how evaluation might contribute to their wellbeing. As progress was made it would shed light on how further priorities might be set.

But the draft strategy makes clear that this is not the kind of priority-setting the PC has in mind. Its initial priorities reproduce those of COAG's Closing the Gap report, and their foremost characteristic is their legibility to the system. They're even arranged around the system's existing organisational structure, which includes families, children and youth, health, education, economic development, housing, justice, land and waters. Makes you wonder what isn't a priority! And all of them identify a "what" rather than a "how."

6. Putting Indigenous people at the centre: the actions

How will we get Indigenous people and perspectives into the centre of evaluation? In their submission to the PC, researchers from Inala Wangarra and the University of Queensland argue that:

"Accountability" has become a lopsided concept, whereby the focus is overwhelmingly on service providers being accountable to government, and where there is no concomitant focus on the accountability of government to the most important stakeholders: Aboriginal and Torres Strait Islander peoples.

So might placing Indigenous people at the centre of an evaluation strategy involve making service providers and government policies accountable to Indigenous people? This possibility doesn't seem to have made it into the PC's strategy, even as a "what." And even if it had, I'd argue that what the PC has endorsed is likely to be implemented in a way that actively obstructs getting to the "how." The PC talks about the importance of "whole-of-government" approaches to evaluation. That sounds innocuous enough — commonsensical even. But why does it have me thinking of "whole-of-church" approaches to the solar system at the time of Galileo?

The only way I can imagine a whole-of-government agenda not doing more harm than good is if it were to imagine itself as being *at the service of* solving the concrete and urgent problems in the field — by identifying good practice in the field, for example, and coordinating the system to expand its influence.

Despite senior officials' and politicians' protestations that they aspire to encourage innovation in the field and spread and scale "what works," progress has been conspicuously lacking. Peter Shergold saw this as a major problem as he rose through the ranks of the public service, but after over a decade at its commanding heights conceded there'd been little change. As he put it in 2005:

If there were a single cultural predilection in the Australian Public Service that I could change, it would be the unspoken belief of many that contributing to the development of government policy is a higher-order function — more prestigious, more influential, more exciting — than delivering results. Perhaps it is because I have spent so much of my career in line agencies, learning to deliver Indigenous, employment, small business, and education programs that I react so strongly against this tendency.

Eight years later he confessed that little more progress had been made:

Too much innovation remains at the margin of public administration. Opportunities are only half-seized; new modes of service delivery begin and end their working lives as "demonstration projects" or "pilots," and creative solutions become progressively undermined by risk aversion and a plethora of bureaucratic guidelines.

In its preoccupation with grander narratives than identifying what works and spreading it, the PC sets its evaluation process up to be driven by the system rather than its intended beneficiaries, however much it protests that they're "at the centre." In a familiar move, the PC suggests that its strategy is driven by four principles, each identified by a pleasing adjective with them all arranged in a pleasing diagram. According to this diagram, evaluation should be "Credible, Ethical, Transparent and Useful." But these words are so general, so capaciously flaccid, that they constrain no one, like a scientific hypothesis that couldn't possibly be falsified. And so, rather than constraining (and so guiding) practice, those words will come to mean whatever people want them to mean, often in retrospect to justify whatever practice is chosen.

Note two further aspects of the high-level pronouncements echoed by the Productivity Commission. First, the PC speaks of evaluation as if its function is to bolster the accountability of those in the field to their senior managers, with evaluation's function being to objectively certify the extent to which the program meets the system's stated objectives. Second, it shows little awareness of how broad and permissive this relatively new discipline of evaluation is. In reaching for some actionable means of validating that it is embracing a thing called "evidence-based policy," evaluation is taken to be something far more settled and definitive than it is — as if getting something evaluated were like getting an auditor to check financial accounts or an engineer to check the structural integrity of a bridge.

As Michael Dillon has observed, the assumption that there are or should be simple linear relationships between objectives and performance is “problematic in cross-cultural contexts and certainly not necessarily the case in the... Indigenous domain.” In that regard the system — and the PC — seems oblivious even to the existence of “goal-free evaluation.” There, the evaluator investigates the impacts of the program without referring to — or ideally even knowing — a program’s stated goals.

In an increasingly managerial world oriented to the needs of organisations and their senior managers, this unconstrained focus deploys the evaluator’s skills in an open-minded way that can more fully reflect the interests and aspirations of other actors in the system — most particularly, intended beneficiaries of the program and the families and communities of which they are a part. Goal-free evaluation puts the evaluator in the best possible position to notice and document all consequences, both good and bad. It can also improve program hygiene just as double blindness adds to the hygiene of a randomised controlled trial.

7. The anatomy of Lord Acton’s work

Then there’s the question of exactly how, as we imagine it, evaluation will identify what is and is not working, and how these findings will find their way into improved policy and practice.

This raises several challenge at the heart of the PC’s draft strategy. First, evaluation should be independent so that it is candid. Second, it should be published, in order to help develop a “knowledge commons” around “what works” (and what doesn’t) and to strengthen incentives for policy, programs and practice to follow the evidence. Yet past behaviour shows that the system responds to such constraints by saying one thing and doing another. So why would it be any different here?

Indeed, the woods are full of regimes in which higher-order objectives are foisted on policymakers to do the Lord’s work (Lord Acton’s work that is). These systems allow those at the top to say one thing as they face towards an objective *in general*, while they do another thing that quietly prevents it happening *in particular*. And thus ensues a prosaic variant of something Oscar Wilde told us about life:

Yet each man kills the thing he loves...

The coward does it with a kiss,

The brave man with a sword!

Freedom of information regimes sit atop Lord Acton’s fault line. And the discomfort this induces is all too often relieved with strategic cowardice. Having been lowered from on high, freedom of information *faces* boldly towards transparency. At least in general and at least when it comes to the saying. When it comes to the particular, to what is actually done, officials *travel* in the other direction. Transgressions go off the record — into corridors, personal phones and email accounts — or are reclassified “cabinet in confidence” or some such. And that’s just the tip of the iceberg as far as actions that are routinely taken to delay and obfuscate transparency under FOI.

If FOI solves its problems the coward's way, regulation reviews use the sword. Today, new regulation can't be introduced without a "regulatory impact analysis" duly demonstrating that its benefits exceed its costs. Australia introduced it in 1986, and it seemed like such a good idea that it was replicated around the world — but invariably with the same (desultory) result. Here's the British Chambers of Commerce back [in 2007](#):

Both Conservative and Labour administrations approach deregulation with apparent enthusiasm, learn little or nothing from previous efforts and have little if anything to show from each initiative.

Sound familiar? Regulation review is another take on the Lord Acton quickstep. Those at the top introduce a compliance regime, but those administering it are trying to get things done for their ministers. So they obey the letter but not the spirit of the regime, and it degrades into empty box-ticking.

8. Getting past Lord Acton's fault line

To recap: as attractive as they sound, independence and transparency cannot be imposed without setting off powerful and perverse incentives. Any attempt to deal with these dilemmas must look them in the eye. I foregrounded them [in 2016](#) with my own proposal for an evaluation architecture. I called it the [evaluator-general](#) to stress the importance of independence and transparency, and also to structurally separate the delivery of services from the means by which we validate their fitness for purpose.

The organisation of the public sector already honours this principle of structural separation — between doing and validating the effects of what we're doing. Thus, the Audit Office and the Bureau of Statistics are independent information and integrity agencies whose work helps inform us of the success or otherwise of other "doing" agencies directed by ministers — such as the health department and Treasury. At the same time, we expect all these agencies to collaborate — sometimes quite closely.

My proposal for an evaluator-general provides the institutional scaffolding within which the same *close collaboration amid structural separation between doing and knowing* can be brought right down to operations in the field. That way independence and buy-in can grow quietly [from the bottom up](#) within organisations rather than being heroically imposed from the top in a grand gesture that experience suggests will fail and fail again.

My aim was to nurture the self-accountability of those out in the field — Feynman's imperative that one mustn't fool oneself — and to build system accountability on that foundation. That's how Toyota revolutionised manufacturing productivity in a way that's now imitated the world over. It found a way to build from "how." It did so by placing the workers on the line, the suppliers and the customers at the centre.

Are my ideas viable or just naive? We'll only know when we give something like them a good try. We'd need no more than a dozen or so teams to try them. In the PC's near 400-page background paper there's some reporting on these problems of independence and

transparency, but not in the context of any critical vision or clear explanation of how they can be overcome.

9. Independence-for-hire and the he-who-pays-the-piper problem

The PC's incuriosity extends to its ignoring the incentive issues arising from how evaluation is commissioned and conducted. As I've [argued](#), allowing firms in our private sector to appoint their own auditor profoundly compromises auditors' independence. By contrast, the auditing of government finances is overseen by an independent auditor-general. Still, while it's far from optimal, we've made the independence-for-hire of private sector auditors work tolerably by specifying highly prescriptive auditing standards. With evaluation, things are very different, there being any number of ways to conduct evaluations to serve numerous tastes and purposes. So evaluators' independence-for-hire provides wide scope for doing Lord Acton's work.

As I've argued [elsewhere](#), independence-for-hire sits at the heart of a "now-you-see-it-now-you-don't" catch 22 that prevents promising developments in the field even becoming visible to the system, let alone having their expansion supported by it.

It goes like this. Responding to all the stirring visions of government "scaling what works," non-government organisations seek government funding to expand their most promising programs. At this point, departments of finance oppose such funding, as well they might, until the programs are independently evaluated. They don't take responsibility by commissioning the evaluation themselves or even specifying what *kind* of evaluation they require. Thus, when the NGO returns, a few hundred thousand dollars poorer, with a Deloitte, PwC or Lateral Economics report in hand (we're cheaper!), it's ignored again because independence-for-hire isn't independence. And so the process of "scaling what works" is stopped dead in its tracks.

Though it understands the value of independence in evaluation, the PC completely flubs the "independence-for-hire" problem, simply associating contracted-out evaluation with independence. And it won't bite the bullet and recommend true independence because it knows this would be rejected out of hand. But to keep the idea of independence in play, it proposes Lord Acton's independence — an independent Office of Indigenous Policy Evaluation that will "oversee" evaluation, though the actual evaluation will continue to be conducted within the very agencies whose performance is being evaluated.

No doubt the PC hopes that this might introduce some independence into the process. But progress, if any, will be agonisingly slow. Allowing agencies to do their own regulatory impact analysis has kept the tiger of regulation review pristinely toothless for thirty-five years now in every country where it's been introduced. The old Office of Regulation Review operated within the PC itself, but the greater notional independence it had there made not the slightest bit of difference. The requisite boxes were ticked and regulations — both the good and the bad — went on piling up as normal.

10. Stated intentions and animating imperatives

It's Lord Acton pretty much all the way down. The PC's draft strategy stresses the need for evaluations to:

- be done ethically
- involve and engage Indigenous people
- be respectful of and in sympathy with Indigenous cultures and knowledges.

Now, each of these is a commendable objective as a “what.” As I keep saying, the hard part is working out the “how.” And tackling each of these matters productively requires great insight. Further (and astonishingly), the importance of each of these requirements is relatively new to the system even as a “what.” Should we really put that same system in charge of learning the “how”? What will happen is already a foregone conclusion — the PC more or less recommends it. Rather than proceed humbly, foregrounding its ignorance, the system will go through its well-worn routines. Codes of practice will be developed. I assume there'll be lots of consultation.

But these codes won't deliver what is written on the packet any more than the mission statement “putting families at the centre” would have delivered Family by Family. However well-intentioned, these codes' *animating* intent — what will matter when push comes to shove and someone might end up on the telly or in a headline — will be the institutional safety of those developing and administering the codes.

This is what happens when the system's commanding heights are put in charge of delivering something that is difficult and context-sensitive but not highly valued in our political culture. Those defending Indigenous interests would be well advised to look on the burgeoning performance regimes in numerous sectors — particularly education and university research — where more and more practitioner time is taken up complying with relentlessly expanding requirements from bureaucracies that have neither the slightest knowledge of nor regard for what's going on out in the field. As the accountability theatre ramps up, administrative numbers and salaries swell at the centre and performance declines. As Britain's Institute for Government documented in a different context, inquiries and restructurings abound and new ten-year plans are announced once every three or four years.

I recall when, in response to another paedophilia scandal, South Australia strengthened its child safety requirements. The very department whose lapses had produced the outrage refused to stagger the starting date of the new system for different community organisations. With the department's processing capacity thus overwhelmed, it took over a month to clear the new paperwork. Family by Family was paralysed. If exceptions were allowed to the deadline, they were for more important folks than us. Overnight, practices that had worked brilliantly and safely for several years — that placed families at the centre of the program — became an offence. I don't know about then, but today the department describes itself as “a customer-focused organisation that puts people first.”

In fact, an evaluation was done on Family by Family. The process was a train wreck. From memory numerous preliminary ethics processes took around nine months, though this was simply to ask families questions about their progress — as they'd been asked regularly within

the program. The evaluation ignored the program's effect on children. Why? Because getting that aspect through the ethics procedures would have been too expensive, uncertain and time-consuming. How ethical can you get?

When the evaluation finally began, the department funding the program wouldn't give evaluators the data to identify our cohort of families. So the evaluation was forced to compare impacts on *all families* in the host suburb against two other areas (one of which was bizarrely incomparable). As I recall, the result was mildly positive but inconsequential — unsurprisingly, given the small number of families involved. To use J.K. Galbraith's term, it was all “innocent fraud” — that is, all that effort and money produced an outcome that amounted to nothing. But its worthlessness was a system failure despite the best of intentions of everyone in it.

I expect that the National Health and Medical Research Council, which issued the ethics guidelines, the family services department and the university centre for family studies thought of themselves as putting people first. But far from nurturing the innovation breaking out on the edges of the system — driven by bright, idealistic, young professionals and increasingly enthusiastic families — the incumbent organisations imposed their own routines and imperatives, each one making the labyrinth denser, more bewildering, more dysfunctional, each one making it harder to put the families first.

Whether or not the evaluation report was released (I don't believe it was), we all cooperated in covering up its worthlessness, which required nothing more than not to advertise it. This is just one close-up of a phenomenon the disillusioned development economist William Easterly has called “the cartel of good intentions.” It is built on Lord Acton's fault line. But you won't see any serious engagement with any of this in the PC's material on Indigenous evaluation.

11. The perils and the promise of candour

You may think what I've written so far is scathing. Yet, as I indicated above, I think the PC makes the right basic calls in its draft strategy. Bereft as the report is of suggestions about how to bring it about, it nevertheless endorses more Indigenous involvement in evaluation. And it backs independence and transparency. In a system that's nowhere near ready to seriously engage with such things, it also makes defensible compromises in shepherding those values into policy. The real shame is that the pathologies of the existing system are deeply entrenched and yet they hardly get a look-in in the commission's analysis. So any strategy for shifting them requires something much more hard-headed — more problem-focused — than four pleasing adjectives and a well-intentioned tagline about putting Indigenous perspectives at its centre.

Here we get to Orwell's point. The greatest service the PC could do Indigenous people — the way it could really put their interests at the centre of its concerns — would be to express itself simply and candidly. Its draft strategy asserts that program participants and the broader community should “have confidence that policies and programs are being assessed objectively and independently.” Poppycock. It should stop pretending and fess up on behalf of the system.

Having recommended a highly compromised form of independence for now, it should explain that the system isn't ready for much candour right now and explain why.

Now you can see the power of Orwell's advice about speaking simply. Speaking simply makes it hard, excruciating even, for you to cover your tracks — to mask your *motives* — with the usual sophistry. Once the officialese is jettisoned (or should that be official-ease?) the discomfort that the system is defending itself against becomes *its own* discomfort in explaining the sorry situation it is dealing with. And the only way to relieve that discomfort would be to go further and sketch out a longer-term plan to reach the outcome described in the honeyed words.

12. Towards the final strategy

For the final strategy to deliver a minimum viable product, I think it needs changes to the draft.

First, it should base its policy compromise on a much harder-headed understanding of the obstacles that stand between us and the land of "how." After explaining why the whole system can't possibly embrace real independence and transparency at the moment, it should go on to sketch its own vision of how that might be grown from the bottom up. I've shown one possible model with my proposal for an evaluator-general, which involves structural separation between the system's doing on the one hand and its knowing and evaluating on the other. It needn't be grandiose and system-wide: it can be built on a small scale and grown from there. Some submissions to the PC seem to think it has merit. The PC itself gives the idea considerable elaboration, but only as reportage. If it has a better model it should set it out.

Second, if the strategy is its contribution to thought, its direct contribution to action should be to call for and begin the process of designing a new burst of energy and innovation that might grow at the margins of current activity and begin to spread through the system.

Here, the current weakness of the system lies not so much in the lack of promising experiments in the field as in the relationship between them and the system itself. The system must be able to identify, validate and acknowledge the best of those experiments. Currently, it can't do that. Evaluation can play some role in fixing that, though we should guard against something that's already clearly in evidence — the system grabbing hold of evaluation as a *deus ex machina*, it's next fad diet that will save it from itself.

And there are two far graver obstacles to progress. First, as those in the field can attest, our politicians frequently play to their own political advantage irrespective of the evidence. Second, bureaucracies have terrible trouble responding to knowledge of what's working from the field, for such bottom-up learning is countercultural in a hierarchy where power is at the top. Further, if learning were to rise from the bottom at any scale, it would involve the discomfort and uncertainty of change for large numbers of people.

The PC can do little about the first of these more serious problems. But it can hope to be influential regarding the second. I think it's possible to be very concrete and specific about what is necessary here. The system can only sustainably expand what works by bolstering the

status of the individuals and communities who have made it work and giving *them* much more authority and resources within that system.

Those at the centre of the system are just as important as the successes in the field, but there's nothing unique about them — or there shouldn't be if the system is working properly. Those in the system need to be made accountable not just for *talking* about expanding what works but for making sure it happens, despite the discomfort it will undoubtedly cause. To that end, a regular report could be recommended, by the auditor-general or some other independent guardian of integrity in the system, to document, say every two years, what progress was being made towards this goal of spreading “what works” and particularly the increasing empowerment of those who make it work.

For those of us who call ourselves Australians to properly begin the task that governor Arthur Phillip began with such high ideals and so little to show for it, we can only do it to the extent that non-Indigenous people and their institutions unloose themselves from those “anxious avaricious tentacles of the self.” To the extent we falter, the soft voice of conscience will keep whispering that destiny to us. •

This essay benefited from helpful comments on earlier drafts from Romlie Mokak, Keryn Hassall, Janina Gawler, Michael Griffith, Jon Altman, Mike Dillon, Christos Tsiolkas and Clive Kanes. As always, I am wholly responsible for the essay's remaining inadequacies. The title “Orwell that ends well” is shamelessly stolen from my friend Konstantin Kisin.

NICHOLAS GRUEN

Nicholas Gruen is CEO of [Lateral Economics](#) and a Visiting Professor at King's College London's Policy Institute. He Chaired the Australian Centre for Social Innovation from 2010 to 2016.