

ANZSOG CASE PROGRAM – TEACHING NOTE

Please cite this Teaching Note as: *Bukowitz, R & O'Loughlin, T. (2022). Governing by Algorithm? Child Protection in Aotearoa New Zealand. [Teaching Note]. Australia and New Zealand School of Government, John L. Alford Case Library: Canberra.*

Governing by Algorithm? Child Protection in Aotearoa New Zealand

An ANZSOG Teaching Note by Rachel Bukowitz & Tim O'Loughlin

Keywords: Public policy, data analytics, algorithmic decision-making, predictive risk models, change management, social license, child protection

This Teaching Note references and provides a guide to: Bukowitz, R & O'Loughlin, T. (2022). *Governing by Algorithm? Child Protection in Aotearoa New Zealand*. Australia and New Zealand School of Government, John L. Alford Case Library: Canberra.

Summary

This note provides supplementary material to assist teachers' use of the case study 'Governing by Algorithm? Child Protection in New Zealand'. The main focus is on the issues the authors see as key for public sector management. There is no prior knowledge of management theory required, though a working understanding of Mark Moore's public value theory will help.

Learning Objectives

- To understand the particular and unique challenges for governments of using data analytics tools for service delivery functions
- To introduce students to management and normative issues arising from the use of tools
- To understand the complicating factors which arise when the tools are applied to the exercise by governments of their coercive powers

This case was commissioned by ANZSOG for the John L. Alford Case Library. Views expressed in it are not necessarily those of ANZSOG. This work is licensed under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Licence, except for logos, trademarks, photographs and other content marked as supplied by third parties. No licence is given in relation to third party material. Version 08032019. Distributed by the Case Program, The Australia and New Zealand School of Government, www.anzsoq.edu.au



- To learn to apply public administration management techniques to developing strategies for successful implementation.

The key issue here is what governments need to do to exploit the vast opportunities that dig data and data analytics offer for better policy-making and better service delivery. Creating social licence for each application is a pre-requisite for success. The case study concerns the detail of one such application which did not succeed in a specific jurisdiction, though it has succeeded in others. The teaching note seeks to use the circumstances of the case to draw out more generalised lessons for application in other cases.

Course Audience:

- Post-graduate and undergraduate students studying data science applications for public policy
- Students of public sector management
- IT students studying adoption of big data analytics tools by governments
- Practicing public sector managers at middle and senior levels

Summary of the Case

In 2014, the new Minister for Social Development in New Zealand, Anne Tolley, was presented with a briefing paper from the National Children's Director at the Ministry for Social Development (MSD) titled 'Vulnerable Children Predictive Modelling: Design for Testing and Trialling.' The paper included a proposal for a two-year observational study 'to assess whether children identified by the Predictive Modelling as at high risk of an adverse outcome/s did in fact suffer that outcome'. Though it was not the MSD's intent, Tolley viewed the 'observational study' along with an additional comment that 'the PM score would be calculated at birth for a known cohort of children and then these children's outcomes and service contacts observed' as the government suggesting a suspension of child protective services as a means to test the accuracy of the predictive risk model (PRM). Consequently, Tolley called for the PRM's implementation to stop immediately and released the proposal, along with her annotations, to the media. The resulting coverage sparked debate on the ethics of using the tool and whether it had social license for its implementation. A number of public sector management issues also emerged such as agencies capacities; change management requirements; and the timing and form of the required consultative processes both within and outside of government. It is these issues rather than the predictive ability of the data analytics tool itself which seemed to determine the outcome of the implementation process.

Background

Australia's Chief Scientist, Alan Finkel (2019) has used the pioneering work of the Government of Victoria in becoming the world's first jurisdiction to regulate the practice of IVF as a lesson for governments looking to harness the power of AI. He saw it as a challenge for Australia to again show 'the world how to advance the cause of scientific discovery while staying true to the ideals of a prudent and virtuous society.... We can define our own future by being world-leaders in the field of A.I. ethics and human rights.'

Finkel's challenge for AI applies equally to the broader uses by governments of algorithmic tools processing massive and varied datasets in pursuit of better policy and more efficient service delivery. Yet there are many pitfalls. Elsewhere, we (O'Loughlin and Bukowitz 2021) have identified numerous failures, most of which have seen governments abandon the use of the specific tools altogether. These include the UK's attempt to use algorithmic prediction to substitute for holding physical exams; the abandonment of predictive policing algorithms by many jurisdictions in the US; and a number of jurisdictions dumping child protection predictive tools, such as the tool described in the case study. Even simple data matching tools have come to grief – Ireland's cervical cancer testing scandal; the UK's abandoned central health repository data records storage; Australia's twice abandoned and unlawful Robodebt welfare fraud detection tool; and a similar exercise which brought down the Dutch Government in 2020.

For our future research, we have looked more closely at 14 such failures, finding that the most common explanation for failure by far is an inability on the part of government to establish the data analytics tool as a public good, such as the case in New Zealand with the child protection algorithm. The least common predictor of failure is a lack of technical proficiency. For example, in New Zealand, the public debate over the use of the tool was focused on normative arguments, not its technical capabilities.

It is also clear that the risks of failure are elevated when data analytics tools are used by governments in the exercise of their coercive powers. These risks become greater still when the tools also involve prediction. The presence of both factors makes achieving social licence particularly problematic. Three examples: assessing recidivism for bail applications; identifying individuals likely to commit crime; and rating children likely to experience abuse and neglect.

It was against this background that the case study was prepared. At one level, it might be seen as just a bad day in the office where a minister interprets a sentence as meaning something that was not intended (though our view is that her interpretation was warranted). However, our research into use of such tools by governments suggests the case yields deeper considerations associated with the specific difficulties of using big data and its algorithmic analytic tools for public benefit. A brief summary of the criticisms the tools have attracted is provided as further background to assist the analysis of the case study.

- The operation of such tools is necessarily opaque. In its search for improved predictive accuracy, the machine is perpetually changing variable weightings and interactions. So, the clarion call for transparency in the design of the tools is unrealisable and has become clichéd. Moreover, the use of the concept is being distorted by developers and governments using varying definitions of transparency to suit their purposes, e.g., transparency about the process used; transparency of values used in the tool's construction; and transparency about the relative predictive accuracy of the tool against the alternatives. (O'Loughlin and Bukowitz 2021). None of these measures disclose how the tools actually work. In the case of the child protection algorithm, public officials cannot explain the exact calculations that generate the risk scores. As such, social workers, as well as parents or guardians of a child that is deemed at high risk by the tool, have no insight into the precise innerworkings of how the score was determined.
- Too little attention is being given to the normative values deeply embedded in their construction of such tools and how those values are being decided. Using the simple example of a tool using blood flow volumes to assess heart rate pathologies, where should developers set the false positive and false negative thresholds? Doctors abhor false negatives so naturally look for the tool to err on the side of false positives. But what are the implications of excessive numbers of false positives for the financial sustainability of the health system and for future scientific discovery? (Kraemer et. al) Taking a more sensitive example, what level of risk should the tool used for predicting recidivism classify as acceptable? Who sets the thresholds and on what basis? Who should be making these decisions? How is society engaged in the process of determining an outcome of a decision that ultimately belongs with it? (Eaglin 2017) To take the child protection example, how is a change of partner for the primary caregiver coded? What is the correct balance of the machine using dynamic factors that the caregiver may have some agency over compared with static factors which are immutable? Resolving these value questions falls to society which is too often being bypassed by software developers and their sponsors (Eaglin 2017, Citron 2007).
- The tools used in many States of the US for assessing bail applications and for identifying individuals are prone to suffering from significant and persistent racial bias (Larson et. al. 2016, Richardson 2019). Though algorithms are often thought to be precise and certain, their outputs are derived from the data that is inputted in them. As such, all algorithms are subject to bias if their source data is biased.
- The predictive management tool for child protection has been criticised as a neo-liberal tool which individualises problems for which society is responsible for creating, chiefly poverty (Eubanks 2018, Keddell 2018).
- Public officials may defer excessively to machine projections. Keddell (2019) argues in relation to the use of the child protection algorithm that 'a social worker who relies on such a tool, but cannot explain, nor understand its inner 'black box' workings, is vulnerable to approbation from professional regulators, managers within the system and families they are working with.....in a high stakes environment and high blame environment, the opposite (to professionals overriding the tool) could also occur—the human actor may defer to the algorithm' (pp.280 – 81).

Properly operationalised, data analytics tools offer substantial benefits and efficiency for governments. The question being asked here is how can public managers best operationalise those tools, particularly those being applied to the more sensitive areas of government operations? Two frameworks are suggested for application to the case study to shed some light on how this might be achieved.

Suggested Teaching Approach

Framework 1: Public Value Theory

Moore's theory of public value (1995) offers the strategic triangle as a technique for operationalising policy. It requires the public manager to meet three tests - creating something that adds public value; is legitimate and politically sustainable; and is operationally feasible for the agency. The strategic triangle can be used as a basis to guide the classroom discussion of the case by allowing students to test the child protection algorithm against the components of the strategic triangle to see where it falls short, and what, if any, measures could be implemented to allow the algorithm to overcome the noted shortcomings.

Moore sees the test of adding public value as being met when the benefits exceed the costs. This is largely a qualitative test. The benefits are often intangible and, where they are not, frequently difficult to measure. For the costs, the dollars are easy to measure, but Moore adds another form of cost that is much harder to measure—this is the cost of the use by government of its moral authority to intervene in peoples' lives. Such interventions involve a diminution of the freedom of individuals so that diminution counts as a cost.

1. Public value	First an icebreaker question. Put the class in Anne Tolley's shoes. • Like Anne, would you intervene and stop the implementation of the child protection algorithm, or would you have allowed the algorithm's implementation to proceed? • Is there an alternative to those two options? Ask students to explain their decisions. The icebreaker can be followed by a full class discussion or small group discussions can commence. The discussions can be guided by the following questions, each of should take about 5 minutes to discuss. There is an abundance of discussion questions provided relating to Moore's strategic triangle and change management; the instructor or the students can choose which questions they would like to discuss given time constraints. If discussions are done in small groups, consider leaving time for a full group debrief at the end to share ideas brought up in small groups. All questions specifically suited for graduate students are noted below, those without notations are suited for both undergraduate and graduate students. • Would the use of the tool have added public value? - Possible student response—No, the tool does not meet the criteria outlined by Moore's strategic triangle. • How would the benefits be assessed? - Possible student responses—Difficult to quantify but could consider the number of children protected from harm due to algorithm's use, the money/resources saved on cases that algorithm deemed low risk that otherwise would have been investigated and the benefits of the earlier intervention made possible by the tool's early identification of children at risk. (Note: The ethical review carried out by Dare (2013) has a shot at quantifying the number of additional children identified by use of the PRM) • How would the costs be assessed? • Would the cost of government intervention be more or less than the existing costs? - Possible student responses—Difficult to quantify but could consider the number of false positives and false negatives. Students may argue either way about the cost of government intervention. • Does the comment from Adams that 'we don't feel we are ready for that' suggest that MSD did not yet accept that public value would be created? - Possible student response—Yes, if public value was created, she would have been ready to use the tool. • Graduate-level question: Separately, Fung and Moore (2012) have argued that public managers have to learn how to 'call a public into existence that can understand and act on its own interests.' Does Adams comment that it was necessary to first understand 'the interaction that comes out of the model and the people at the front
------------------------	--

	line that will have to use that information' suggest that the 'public' here is not only the clients but the front-line social workers as well? - Possible student response—Yes, the 'public' extends beyond the clients to all affected or potentially affected parties, which includes social workers. Another public might be the broader social welfare community – people working in other areas of social service delivery.
2. Legitimacy & Political Sustainability	In addition to the general questions listed above, questions on each aspect of the strategic triangle may be asked: • The case shows that the use of the tool had political cover from the Green and White Papers as well as the broad support for new forms of intervention from the Prime Minister, Deputy Prime Minister and outgoing Minister. Should this have been enough to satisfy this test? If not, why not? - Possible student responses— No, the Green Paper/White Paper process does not deliver social licence because the process does not really engage users of the system or the broader community. Public acceptance must be garnered beyond the traditional Green Paper/ White Paper process. • At what point should the support of the community be sought? A similar tool was introduced in Allegheny County using an implementation plan which required early community engagement (see Exhibit A). - Possible student responses— As early as possible, well before implementation. The public should be informed of even the potential of using the tool before any action towards implementation is taken. However, finding the right moment is a matter of fine judgement as some prior development is necessary. As Boyle commented in interview, 'you have to have something to consult on'. • Boyle outlined in the case study the stages which needed to be reached before such communication could take place. Tolley also doubted if social licence had been created. • What more needed to be done? - Possible student responses— Earlier public consultation, reframed language on how the tool would be used, complete abandonment of the idea of the observational study, consideration of cultural implications of the tool on the Māori population, etc. • Graduate-level question: For Moore, one of the requirements for satisfying this test is to close legitimacy gaps. Those gaps will appear at points in the authorising environment. What legitimacy gaps may have existed here? • Who were the actors in this authorising environment capable of closing any gaps and how might their support have been enlisted? - Possible student responses— Legitimacy gaps include the opaqueness of the tool's innerworkings; the potential deference of social workers to the algorithmically generated results; and the lack of public consultation on the use of the tool. The authorising environment includes the families that could be rated by the tool; social workers that would use the tool; government officials that would manage the tool's implementation; media; and interest groups such as UNICEF NZ.
3. Operational Feasibility	This test requires agencies to have the capabilities to operate the policies or the ability to acquire the resources needed to operate them. • Does Adams' comment about not being ready to use the tool for prediction signal a concern over capability? - Possible student responses—Possibly. Given her follow-up comment about wanting to first understand, 'the interaction that comes out of the model and

	the people at the front line that will have to use that information' suggests that she does not fully understand how the tool will work for social workers. • Is using it first by field workers for existing cases the preferred approach for addressing any such deficiency? - Possible student responses— Yes, that is how it can be tested. Or alternatively, students might argue that no, the tool should not be used on existing cases because there are moral and ethical reasons against this method. • What are the alternative strategies assuming that delay is inimical to successful implementation? - Possible student responses— Use past, closed cases to test the tool to see if the tools' results align with the decision made by social workers and the ultimate result of the case.
--	---

Framework 2: Change management theory

Boyle's comment in the case that the introduction of such a significant new technology 'requires deep change management' invites analysis of the connection between technology, change and politics.

1	Using Pfeffer's (1992) theories of power in management, a basic framework for considering the change management requirement could trace the following steps: • Changes in administrative processes shift power because some people lose out and others gain • Politics always accompanies shifts in power • A distinctive feature of technology-driven process changes is that they tend to shift power closer to the centre • When power is centralised, the politics favour the dominant coalition
2	Questions arising from the case study for this might be: • Dare's comment that 'politics and science collided and politics won' assumes that politics can be kept out of science. Is this possible or even desirable? - Possible student responses: Students may argue either way to this question. The instructor should encourage students to think of real examples to substantiate their answers and consider the implications if science and politics were completely separated. • Graduate-level question: Tsoukias et. al. (2013) have commented on the impact that the increasing mistrust of science is having on the ability of governments to convince citizens of the public good of data analytics projects. Do you agree? • What might the remedy have been in this case had the implementation of the PRM reached the next stage? - Possible student responses: Students may argue either way for the first question. The instructor should ask the student to provide specific examples to support their stance either way. For question two, students may suggest regular public-facing communication about the tool, a dedicated public servant to answer questions about the tool, easy-to-understand fact sheets to answer FAQs about the tool, etc.

Epilogue

Both the case study and this teaching note have been prepared to highlight aspects of public administration associated with the use of data analytics tools by governments. In this case, the controversy saw the abandonment of the PRM and its replacement by multi-agency Social Service Teams.

Prior experience with teaching this case is that students tend to gravitate to the political issues and lessons for public sector management. This can come at the expense of the important tool design and associated ethical issues coming from it. For teachers looking to go beyond the case study to explore these issues, it is recommended that they consider using Viathianathan's original paper and Dare's full ethical review.

Further Resources

Citron, D. (2007). 'Technological Due Process', *Washington University Law Review* (85):1249-1313.

Cohen, J. (2013). 'What privacy is for', *Harvard Law Review* 126(7):1904 -1932.

Dare, T. (2013). *Predictive risk modelling and child maltreatment: An ethical review*. Auckland, New Zealand, University of Auckland. <https://www.msd.govt.nz/documents/about-msd-and-our-work/publicationsresources/research/predictive-modelling/00-predictive-risk-modelling-and-child-maltreatment-anethical-review.pdf>

Eaglin, J M. (2017). 'Constructing Recidivism Risk', *Emory Law Journal*, 67(1): 59-122.
<https://www.proquest.com/scholarly-journals/constructing-recidivismrisk/docview/1965559997/se-2?accountid=9902>

Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor*. New York, NY: St Martin's Press.

Finkel, A. (2019). 'Harnessing the power of artificial intelligence to benefit all', Speech to the Go8 Artificial Intelligence Collaboration and Commercialisation Summit, 31 October, Melbourne.
<https://www.chiefscientist.gov.au/sites/default/files/201912/Group%20of%208%20Artificial%20Intelligence%20Summit%20Melbourne%202019.pdf>

Fung, A & Moore, M. (2012). 'Calling Publics into Existence: The Political Arts of Public Management' in Donahue J and Moore M (eds) *Ports in a Storm: Public Management in a Turbulent World*. Brookings.

Larson, J, Mattu, S, Kirchner L, & Angwin J. (2016). 'How We Analyzed the COMPAS Recidivism Algorithm' *ProPublica*. <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>

Kraemer, F, Van Overveld, K, & Peterson, M. (2011). Is there an ethics of algorithms? *Ethics and Information Technology*, 13(3): 251-260. <https://link.springer.com/content/pdf/10.1007/s10676-010-9233-7.pdf>

Keddell, E. (2018). The vulnerable child in neoliberal contexts: the construction of children in the Aotearoa New Zealand child protection reforms. *Childhood*, 25(1): 93-108.

Keddell, E. (2019). Algorithmic Justice in Child Protection: Statistical Fairness, Social Justice and the Implications for Practice. *Social Sciences*, 8(10): 281. <https://doi.org/10.3390/socsci8100281>

O'Loughlin, T & Berkowitz, R. (2021). 'A New Approach Towards Social Licensing of Data Analytics in the Public Sector', *Australian Journal of Social Issues*, 00:1-15. <https://onlinelibrary.wiley.com/doi/10.1002/ajs4.161>

Moore, M. (1995). *Creating Public Value: Strategic Management in Government*. Cambridge: Harvard University Press.

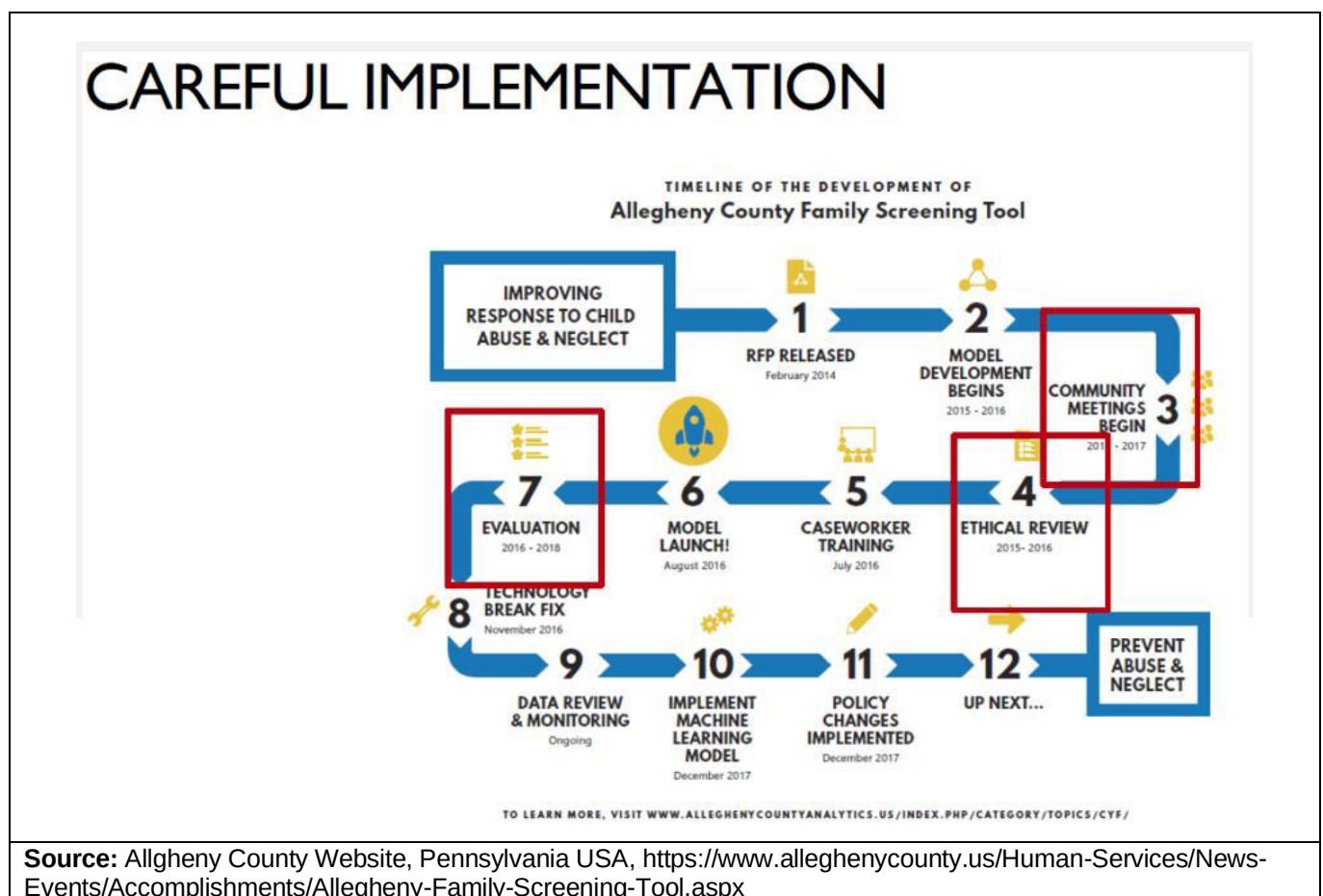
Pfeffer, J. (1992). *Managing with power: Politics and influence in organizations*. Cambridge: Harvard Business Press.

Richardson, R, Schultz, J. M, & Crawford, K. (2019). Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice. *NYUL Review Online*, 94: 192-23.
https://www.nyulawreview.org/wp-content/uploads/2019/04/NYULawReview-94-Richardson_etal-FIN.pdf

Tsoukias, A, Montibeller, G, Lucertini, & Belton, V. (2013). Policy analytics: an agenda for research and practice, *EURO Journal on Decision Processes*, 1(1):115-134. <https://link.springer.com/content/pdf/10.1007/s40070-013-0008-3.pdf>

Vaithianathan, R. (2012). *Can administrative data be used to identify children at risk of adverse outcomes*. Auckland, New Zealand: Business School, Department of Economics, University of Auckland.
<https://www.msd.govt.nz/documents/about-msd-and-our-work/publicationsresources/research/vulnerable-children/auckland-university-can-administrative-data-be-used-to-identify-children-at-risk-of-adverse-outcome.pdf>

Exhibit A



According to the Allegheny County Website:

“The design and implementation of the *The Allegheny Family Screening Tool* was a multi-year process that included careful procurement, community meetings, a validation study, and independent and rigorous process and impact evaluations. In addition, the resultant model was subjected to an ethical review prior to implementation. Tim Dare, Professor of Philosophy at the University of Auckland and Eileen Gambrell, Professor of Child and Family Studies at the University of California-Berkeley conducted the independent ethical analysis. Their report, *Ethical Analysis: Predictive Risk Models at Call Screening for Allegheny County*, concluded the tool is ethically appropriate, particularly because its accuracy exceeded the alternatives at the time and there would be ethical issues in not using the most accurate measure. DHS issued Ethical Analysis: Predictive Risk Models at Call Screening for Allegheny County - DHS Response thereafter.”

For further information see <https://www.alleghenycounty.us/Human-Services/News-Events/Accomplishments/Allegheny-Family-Screening-Tool.aspx>

Feedback on Teaching Note use

To assist with continual improvement, if you've used this Teaching Note, we'd greatly appreciate your feedback and suggestions for improvement. To do so, please email caselibrary@anzsog.edu.au with the name of the case and your comments. Many thanks in advance.

This work is licensed under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Licence, except for logos, trademarks, photographs and other content marked as supplied by third parties. No licence is given in relation to third party material. Version 08032019. Distributed by the Case Program, The Australia and New Zealand School of Government, www.anzsog.edu.au



Australia &
New Zealand
School Of
Government

ANZSOG