

John Yeabsley

Building a Working Relationship Policy Advice Quality – the fundamental capability for public management

Abstract

High-calibre written advice is the basis of government decision making. It produces good decisions and fosters trust in advisors which benefits the process. Data series of average scores based on two batches of assessments of policy advice quality over the last 16 years show little sign of sustained quality improvements. There remains a challenge.

Keywords policy quality, written advice, relationship building

What is the issue?

High-quality policy advice is an essential component of effective modern government.

—Mintrom, 2011

To cope with the bewildering array of tasks entailed by the process of government, the relatively small political administration is supported by non-political workers. This is even so when the task is to make decisions. The political component makes the decision, but is typically advised by the workers. This division allows the advisors to build up expertise and gather relevant

information, thereby contributing to good decision making. But it is not a costless way of organising a public service, as it creates its own problems.

It has long been recognised that in the New Zealand public service (based on the Westminster system), there is a fundamental divide between the interests of political (transient) ministers and their more permanent public service advisors.¹ And that this divergence of motivations (as well as background, training and roles) creates ongoing difficulties – especially of

communication² – which can have an impact on trust. And trust is an important factor itself in reaching good decisions.

Moreover, any lowering of trust can not only undermine decision making, but be inefficient, as it raises ‘transaction costs’ – the resources it takes to complete effective communication. And in the usual situation of straitened finances this pressure can lead to a poorer advice system. This, in turn, undermines the overall quality of public decision making.

So, addressing this divide is more than a matter of mere drafting nicety or being polite. It is a vital component of an effective decision-support mechanism.

The specifics

While this is a general problem in the process of advising (Maister, Green and Galford, 2021), it seems to be particularly cogent for public advisors. The generic issues of what economists call agent–principal problems are challenging at the (decision-making) boundary between politics and action.

It is fundamental to public management to have an effective and soundly functioning system to build trust between the key component partners; or, at least, between the minister and the policy advisors.³ An authoritative text on this type of relationship sums up the essentials as: ‘Trust must be earned and deserved’ (ibid.) But how?

Figure 1: Components of the Quality Policy Framework



Source: Department of the Prime Minister and Cabinet Policy Project

Another theoretical approach to advising (Morris, 2001) says that the key to convincing the politician to trust the advice offered is to find a way to demonstrate that the advisor is taking the trouble to work with the advisee, usually said to be doing things that cost the advisor. (This might be, in a small example, to blatantly adopt the politician's language.) So, whenever capabilities are being reviewed against the needs of the day, the continued high-grade capacity to effectively brief ministers is at the base of the whole operation. And this depends crucially on the mutual trust that is associated with reinforcing the relationship.

In most cases, this must be created from scratch. Our Westminster-style public service, with its merit-based appointment system, means key relationships are forged on the job. This is especially true for three areas tightly related to the effective delivery of public services:

- confidence in the advisors' professionalism – being able to take the information and comments provided at face value underpins an efficient communication system;
- decision support – 'ministerial policy making' is a stream of decisions at the heart of public management, resting on

credible analysis and evidence about risky projects; and

- information provision – in particular, updates on the effectiveness of policies and projects (including budget expenditure) and the general performance of the public delivery system.

This article looks at the quality of the advising capability of the New Zealand policy advice system over time. The focus is on the contribution this makes to building a trusting and effective relationship. The article starts by looking at the data on the quality of policy advice, at least as it is reflected in written communication. The evidence shows little sign that things have been improving overall. On this basis, it is concluded that while the quality of such advice continues to be sound and broadly does the job, there has long been room for improvement. Quality of advice is a key factor in ongoing trust and hence efficiency, so this is a capability with a potential upside. The final section suggests what might be done.

The capacity to produce high-quality policy advice – what do we know?

To discuss the question addressed in this article, we draw on some unique data. This comes from NZIER's experience in

reviewing the quality of policy advice in New Zealand public agencies for more than 25 years. It makes use of detailed assessment records from the last 15 or so years.

As the material on the Policy Project page of the Department of the Prime Minister and Cabinet website explains, policy advice is the product involved in the process by which ministers are informed, especially when it serves as a basis for decisions (see particularly Department of the Prime Minister and Cabinet, 2014). There is a specific focus on written advice material. It is classically the bulk of the information flowing to ministers. It is, too, the lasting record, and, as such, has a unique importance for legal and formal reasons. From a practical perspective, it is readily available for third-party assessment.

For more than 25 years, NZIER has undertaken independent audits of the quality of advice in the policy system. The framework was developed using the 1993 seminars and papers collected and discussed in Gary Hawke's thoughtful book on policy advice (Hawke, 1993).⁴ Initially, there was one client, but the service has steadily extended to serve a wider set of public agencies, including local bodies (not included here).

As time went on, the methodology moved from purely descriptive assessments to quantitative scores using a rating schedule, with a mark plus structured commentary for each paper. There was always a focus on systemic diagnosis and improvement, looking to encourage quality enhancement.

How the assessments were done – there is an emphasis on consistency

NZIER assessments are undertaken by experienced policy advisors and managers. All assessors have experience in producing and supervising advice. Many of our core team were previous deputy chief executives with high-level experience in policy work at ministerial level in various agencies.

The approach used a standard framework (see Appendix B). It was built around the judging of the extent to which the advice was fit for purpose. This allowed the assessments to adjust naturally to cope with different types of advice within a common structure. To provide useful feedback to the authors, the assessment sheets considered the papers under the following high-level headings: customer focus, credible and robust analysis,

and clear and concise presentation. These headings were further divided into individual areas to shape comments in detail. But, crucially, the score and the overall feedback were based on fitness for purpose, while including advice on how the piece might have been improved.

This NZIER assessment system lasted until the emergence of an alternative coordinated by the Department of the Prime Minister and Cabinet in its role of supporting the head of the policy profession. This new approach identifies four high-level elements of quality policy advice: context, analysis, advice and action (see Figure 1). More detailed explanation and helpful commentary is presented on the Policy Project website, including a helpful guide to the marking system.

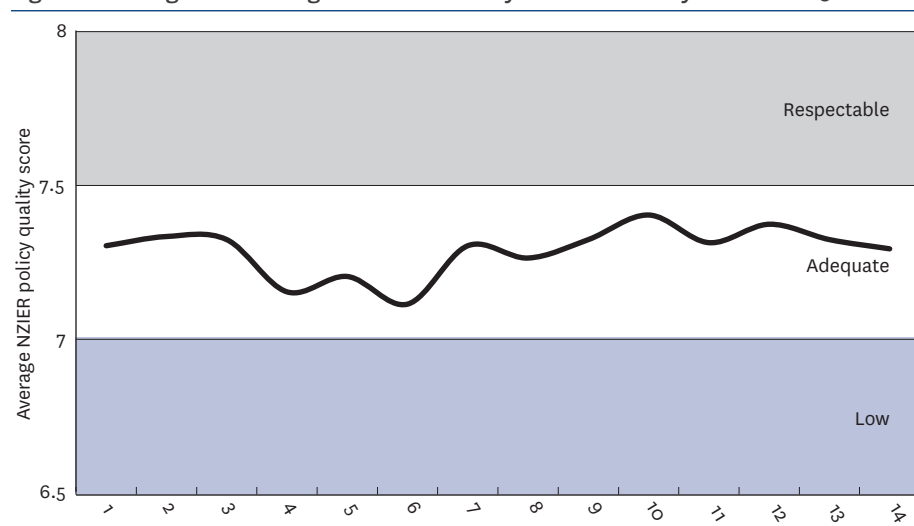
To maintain comparability with agencies not using NZIER, we shifted to the new system in 2019. We have detailed records of the agency assessments going back to 2008, and our scoring has always been based on a panel process to keep it consistent.

These days, the sample of papers is divided among our reviewers. Each assesses their papers carefully against the Policy Quality Framework and determines initial scores. That reviewer then writes an assessment of the paper: this gives a rationale for the score, comments on the key criteria for the score, points to areas of good practice and identifies possible improvements. Assessments are cross-checked via a moderation session where every paper is reviewed collectively. This can involve significant debate among assessors, but is a useful mechanism that supports consistency. As part of this process, a reasonable proportion of the sample will be reassessed by cross-marking.

These ratings provide a simple summary of the overall quality of individual papers. This means the average score is a well-founded indicator of the system quality.

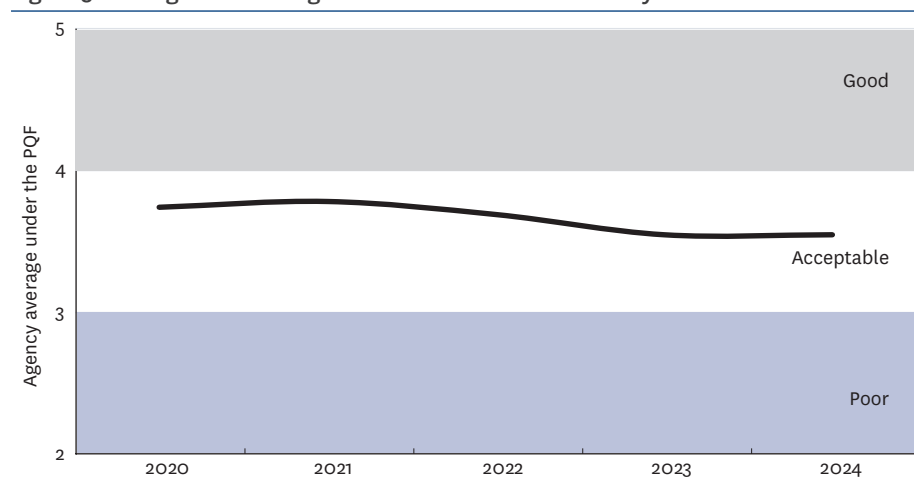
Looking at the results reported, there are differences between the make-up of the set of agencies represented over the years. We do not see this as a source of persistent bias. Obviously, this is a judgemental system. Fundamentally, it rests on the ability of the assessors to be consistent and reasonably uniform in their scoring. We believe that the measures and systems we use reduce the potential for outliers.

Figure 2: Average scores of agencies assessed by NZIER over the years 2008–19



Source: NZIER data

Figure 3: Average scores of agencies assessed in the last five years to DPMC standards



Source: NZIER data

This all suggests that the scores reported here are broadly indicative of the quality of advice over the timespan shown.

Setting matters

The task of developing advice is closely linked, of course, to the challenge of analysing and considering public policy-related issues. There is a substantial literature on policy analysis; less has been published on the difficult job of successfully preparing advice that works. It is, however, a challenging role, with many moving parts.

Trying to bring ministers to the point of being able to make an informed decision about a risky venture demands more than just piling up facts and analysis. To be helpful in the usual situation of New Zealand politicians (pressed for time, with a pile of other reading) means drafting for the reader's needs. It is very demanding – a highly skilled task.

What do we know about quality?

Three sources provide triangulation

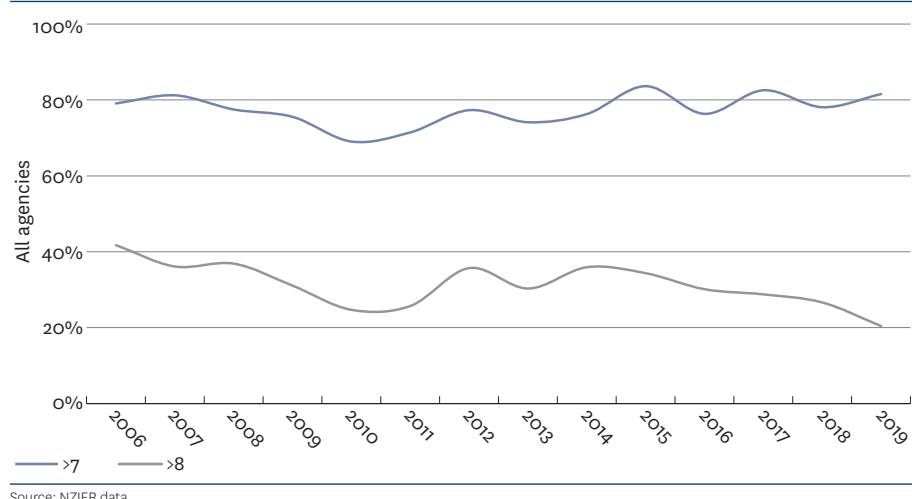
We use three different measures to consider what has been happening to advice quality.

Average scores show a high-level picture

These scores are the overall annual assessments and are in two groups. First, the 11 years of the NZIER system: the average findings are summarised in Figure 2. This has the average score lying between 7 and 7.5 for the whole period, with no clear sustained trend. As the chart shows, the results are, in NZIER jargon, 'adequate' rather than 'respectable'. But the average does stay comfortably clear of the 'low' zone.

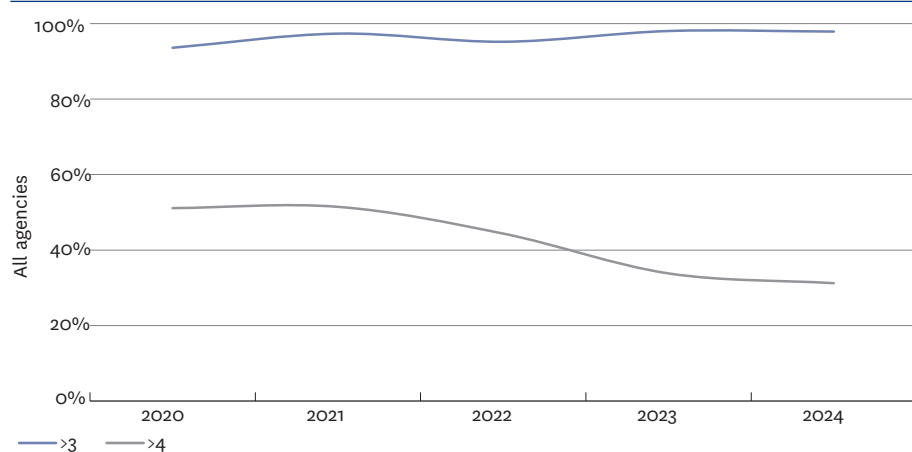
Then there are the more recent five years, using the new scoring system created by the Policy Project. These results are in Figure 3. The labels used here are those included in the marking schedule. Note that the critical value is 3 (margin of the 'poor' zone), as the scores below mean

Figure 4: Performance up to 2019 (meeting requirements is >7 and exceeding >8)



Source: NZIER data

Figure 5: Performance 2020–24 (meeting requirements is >3, exceeding is >4)



Source: NZIER data

(according to the marking schedule) that the paper should not have been signed out of the shop – it is not fit for the purpose.

The Figure 3 results seem to be broadly continuing the picture shown over a longer period in Figure 2. Over the last five years, the average-quality paper has been sound rather than startling. But signs of quality output are on show: there is clear daylight between the mean score and the critical 'poor' value. Unfortunately, there is no obvious sign of the typical advice quality improving over time; if anything, there is something of a decline shown over the five years.

This suggests that the job is challenging in the current setting (as has broadly prevailed over the years between 2006 and 2019). There are many practical and theoretical reasons why this is true.

- The problems are inherently difficult: not only are they often under-researched and typically lacking simple, neat solutions, but their possible remedies are frequently contentious,

with different social subgroups having strong views one way or another.

- In a technical sense, these issues are 'non-Paretian' questions. The outcome will inevitably tend to make some groups better off and others worse off compared with the status quo.
- The issues are often singular, 'one-offs' with limited previous experience to draw on. Even when problems recur, there is typically a new twist that needs to be taken into account.
- The available background information and data, both in New Zealand and overseas, is usually sparse and not readily accessible. (This is a product of a small set of advisors dealing with a large set of ministerially driven problems under frugal conditions.)
- The prior examples of other jurisdictions where any similar policies have been enacted are nearly always sufficiently different to make the drawing of analogies risky.⁵

So, accepting that the various agencies reported on here are not finding it easy is not a great surprise. This track is average, so there are agencies with lower scores every year. Thus, the typical advisory group is slogging away and producing solid output but not acing the test.

More granular results – some gain, some loss

We can look beyond the summary scores in the averages to examine the distribution of scores. First, we turn to two measures that agencies use to keep track of their own progress. These are the proportion of papers meeting requirements and the proportion exceeding requirements. The change in scoring means there are two graphs, Figure 4 and Figure 5. These results generally support the impressions about the quality of advice over time gained from the average scores. It seems that a set of reasonable quality papers has been produced over the years.

In terms of trends, in Figure 2, over the whole period, there was a decline, a recovery, and then another longish decline. Under the new marking schedule, there was a slight tendency for the proportion of papers exceeding requirements to gently slip lower. We wouldn't take too much out of this. The series clearly has a lot of variation, and there are, of course, a range of possible causes. These include the social experiences of the last few years, general pressure from work volumes, and even changing assessment standards.

Individual paper marks

A more detailed examination of the scores over the last few years (Figure 6) shows a fairly consistent pattern of scores. The main effect shown in this figure is a slight tendency to centralisation: fewer poor papers (scoring below 3) balanced by fewer high-scoring pieces (gaining 4 and above). What has emerged is a strong number (more than 40%) of papers scoring 3.5.

A snapshot from the outside

In August 2010 there was an announcement that a major review of the cost of the policy advice system would be conducted. Inevitably, this involved considering the quality of the product. The review team was chaired by Graeme Scott, a former senior New Zealand public servant and

policy advisor with recent experience as a consultant in many countries. It included Pat Duignan, another former senior Treasury officer and a member of the Commerce Commission, and former secretary of the Department of Human Services in Victoria, Australia, Patricia Faulkner. This exercise

arose from the Confidence and Supply agreement between the National Party and ACT, in which it was agreed that a series of task-forces would be ‘established to undertake fundamental reviews of all base government spending in identified sectors’.

The National Party and ACT have agreed to undertake a review of expenditure on policy advice across agencies.

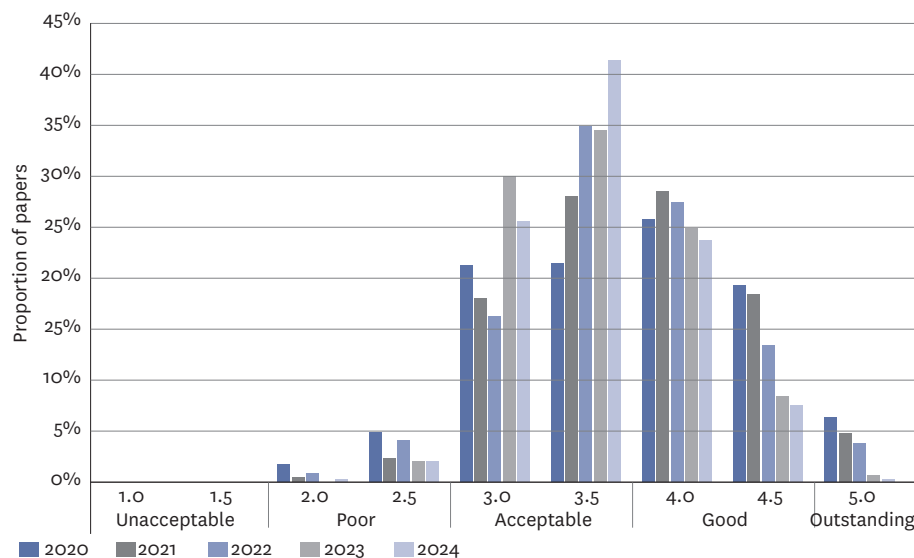
The purpose of this review is to provide recommendations to the Minister of Finance, the Minister of State Services, the Minister for Regulatory Reform and other relevant Ministers on the scope for refocusing and/or reducing total government expenditure on policy advice to ensure high professional standards, cost effectiveness and strong alignment with government policy priorities. (Review of Expenditure on Policy Advice, 2010, p.1)

In December 2010, after extensive research, what had come to be known as the Scott inquiry said:

There are parts of the policy advice system that are run down, while the best information available on quality assessment shows the general standard has a lot of room for improvement. That this can be achieved is clear from the good practices and high-quality work done in many areas. But we have been surprised at the very wide spread between the best and worst practices and standards of which we have evidence. The costs vary too widely, as does the quality. (Scott, Duignan and Faulkner, 2010, covering letter)

Of course, their focus was on expenditure, but this remark about quality was part of a careful investigation of various aspects of the advisory system.

Figure 6: Over the last few years ...



Source: NZIER data

Their recommendations were broadly accepted by the government in April 2011. The government response said the idea was ‘to adjust the system to drive sustained improvement in the quality and management of policy advice’ via a suite of actions. These were to focus on:

- producing better financial and management information to drive value for money and efficiency;
- improving the leadership and management of policy advice within agencies; and
- driving stronger central agency stewardship of the state sector to support cross-agency collaboration, performance improvement, capability building and focus on medium and longer-term policy challenges. (Treasury, 2011)

Guiding this work was put into the hands of the central agencies.

Even at the time, the set of recommendations and the associated government response was seen as, ironically, suffering from many of the same weaknesses as the system it was trying to diagnose and treat. In particular, the problems were spelled out as being a ‘preoccupation with process and risk-management’ (Mintrom, 2011, p.8). Our data above (Figures 2 and 3) suggests that, whether this comment was justified, there has been limited overall progress. And this is despite what has happened since. There has clearly been much effort – both the Department of the Prime Minister and Cabinet and Treasury

websites include helpful advice and general support. In addition, we have seen ongoing training and other supportive quality-oriented initiatives in many agencies. But for all this, our evidence suggests limited progress in achieving marked improvements in this capability.

What we have learnt over the years

There are some (relatively) easy gains, like tightening the proofreading and quality assurance stages and pushing to shorten papers. As a quick generality, these changes have (by and large) been implemented by agencies and are now in force most of the time. And, as noted above, there is a general long-running tendency to have far fewer low-scoring papers these days.

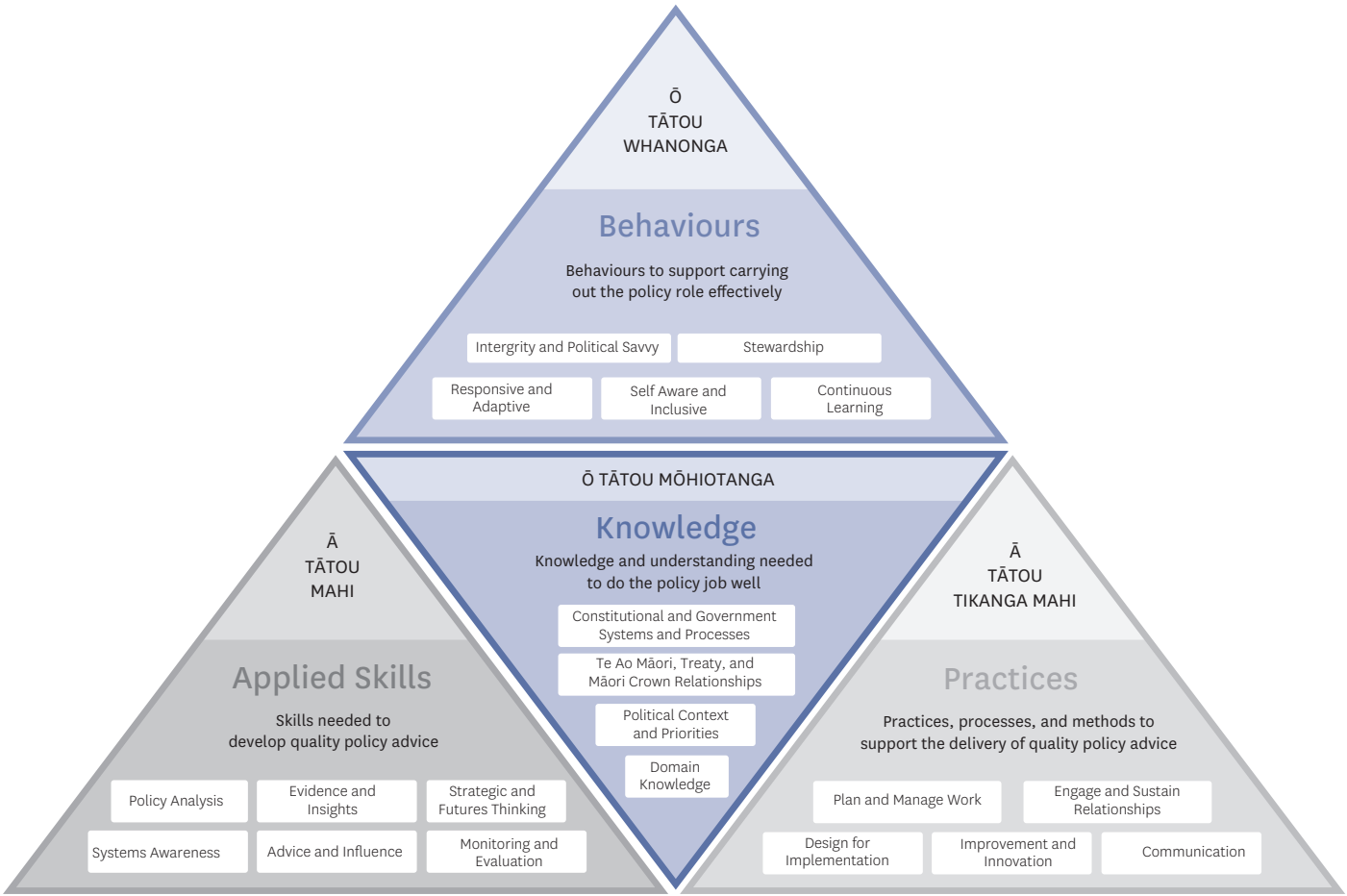
But our experience has been that even these aspects are relatively fragile. They are brought into being as part of a management initiative and belong to the organisation. Thus, unless individual analysts identify with their importance or have their jobs tied up with them, in the context of some stressful experience for the organisation or other large change⁶ it’s easy to return to simple ways.

We see the need to weave the commitment to quality advice into the policy shop’s culture – for it to become ‘the way things are done around here’. Then it is embedded.

Beyond the historical record – what do we see today?

Notwithstanding the overview of the results, our reviews lately have shown

Figure 7: The Policy Skills Framework



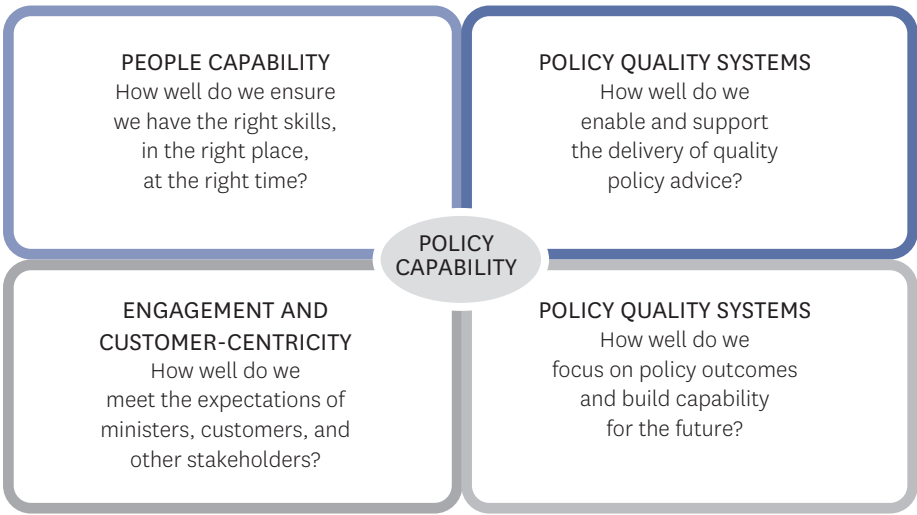
Source: Department of the Prime Minister and Cabinet Policy Project

plenty of high-calibre advice work. A couple of areas can be cited to back this up. The Covid-19 experience was both hectic and challenging. But it brought the best out of many agencies. They reacted to the situation by producing high-scoring work that kept ministers on top of the play. The topics ranged from the economic

situation to transport and supply chains. It revealed robustness and capacity to undertake quality work, including in novel areas, under pressure. Similarly, the change of government in late 2023 meant a significant number of major policy stance reversals. This was the backdrop to many pieces of free and frank

advice that were both bold and professional. Indeed, we noted instances when the agency followed up advice conveying thoughtful reasons why not to pursue a particular policy avenue, with helpful and practical advice about how best to implement it if that policy avenue was chosen.

Figure 8: The Policy Capability Framework



Source: Department of the Prime Minister and Cabinet Policy Project

Central support
As noted above, there has been a lot of effort at the centre to provide support for the agencies in their various drives to improve the quality of their advice. The Department of the Prime Minister and Cabinet website, for instance, has the Policy Skills Framework (Figure 7); the Policy Capability Framework (Figure 8); and the Policy Methods Toolbox, a series of areas of advice, including on Treaty of Waitangi analysis. These are all helpful for agencies looking to improve the quality of their policy advice over time. We have seen them used by different agencies to improve the skills of their analysts.

Conclusion

As was always the case, at any given time, some agencies are good at producing quality advice, and others are less good. The group that is better is not constant year on year, although there is often a degree of persistence. As a general remark, looking back over the years, there have been some longer-run underlying improvements in advice as a whole – generally shorter papers, for instance, although we still see outbreaks of longer briefs. However, it is not uncommon for an agency (perhaps under new management) to improve its score significantly and even repeat a good result for a few years; but then something happens, and the agency's assessment score falls away.

We have seen a substantial increase in the number of formal public policy courses being offered in academia, and the standard of general background is higher today than previously. This can be nothing but good. Nevertheless, policy advising is a craft that requires much on-the-job learning. This needs capable mentors, great research and data back-up, and a thorough commitment from agency policy leadership.

One superficially simple factor is our strongly held opinion that 'quantity is the enemy of quality'. Trying to do too much is not helpful for quality. But there is a real tension between meeting the minister's expressed needs (which might involve lots of advice) and doing good work (which usually takes time and resources).

We do not have any evidence supporting magic bullets, but our experience over the years (see NZIER, 2022) suggests that it is a constant struggle to gain and hold onto lasting improvements. One approach is to focus strongly on building trust between the minister and the team of advisors. Higher-quality advice should contribute to earning trust, possibly by showing the way the advisors were putting the minister's preferences front and centre within a careful analytical approach.

There is an unavoidable paradox here – a vital part of being a high-calibre public policy advisor is to provide free, frank and professional advice when it is appropriate. This might seem to cut across the task of reflecting the minister's preferences. But if the preparation has been sound and the

specific advice is professional, the potential for tension should be managed down.⁷ Indeed, ministers could set the tone for riskier, bolder policy and advice. Interestingly, this chimes with the recent remarks of the previous minister for the public service (McCulloch, 2024).

-
- 1 See discussion of the differences and what is involved in detail, with examples, in Prebble, 2010.
 - 2 There is a stylised treatment of this in NZIER, 2016a and 2016c.
 - 3 Other sources of advice for ministers range from chats in the Koru Lounge to intense professional lobbying. One notable development in recent years has been the rise of political advisors. These are discussed by Shaw and Eichbaum, 2019.
 - 4 This volume was based on a series of seminars and workshops involving experienced policy advisors. Along with feedback from users, it informed the development of the NZIER assessment framework as shown in Appendix B, before being overtaken by the inter-agency work on the Policy Project.
 - 5 These statements are based on NZIER assessors' collective experience of producing and assessing policy advice.
 - 6 Such situations are common in the New Zealand public sector see Hamblin and Plimmer, 2023.
 - 7 Mintrom (2011) saw this approach as being solely focused on the supply side: improving the quality of the analysts and thus of the advice. And he suggested the core problem – especially for the tougher long-term issues that confront New Zealand was on what he called the demand side – ministers.

Acknowledgements

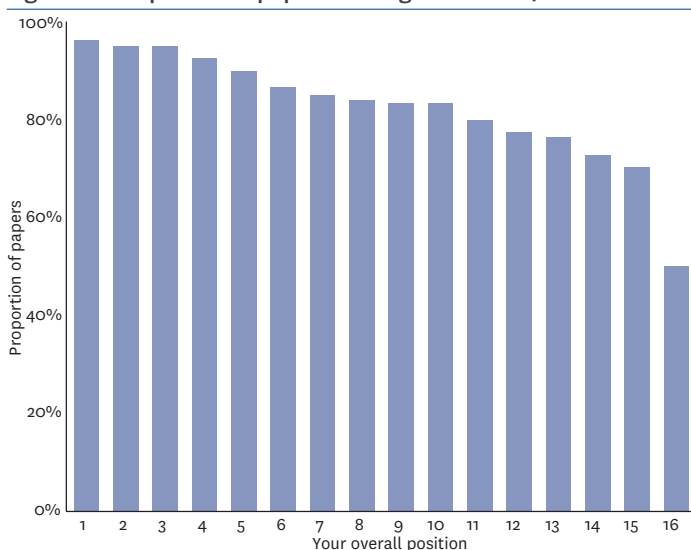
I wish to thank Cathy Scott, Todd Kriebel and Sarah Spring.

References

- Department of the Prime Minister and Cabinet (2014) 'The Policy Project: narrative and direction of travel', <https://www.dPMC.govt.nz/sites/default/files/2022-04/policy-project-narrative-direction-of-travel.pdf>
- Hamblin, R. and G. Plimmer (2023) 'Instability in New Zealand's public sector', *Policy Quarterly*, 19 (4), pp.95–105
- Hawke, G. (1993) *Improving Policy Advice*, Wellington: Institute of Policy Studies, Victoria University of Wellington
- Maister, D., C. Green and R. Galford (2021) *The Trusted Advisor*, New York: Free Press
- McCulloch, C. (2024) 'Nicola Willis asks public servants for their "best and boldest ideas"', RNZ, <https://www.rnz.co.nz/news/political/531755/nicola-willis-asks-public-servants-for-their-best-and-boldest-ideas>
- Mintrom, M. (2011) 'Policy advice and the pursuit of public value', *Policy Quarterly*, 7 (3), pp.8–12
- Morris, S. (2001) 'Political correctness', *Journal of Political Economy*, 109, pp.231–65
- NZIER (2022) 'The long march to quality', policy advice masterclass, paper 40, <https://www.nzier.org.nz/hubfs/Masterclasses/Central%20Government/Brief%2040%20The%20long%20march%20to%20quality%202022.pdf>
- Prebble, M. (2010) *With Respect: parliamentarians, officials and judges too*, Wellington: Institute of Policy Studies, Victoria University of Wellington
- Review of Expenditure on Policy Advice (2010) 'Terms of reference', <https://www.treasury.govt.nz/sites/default/files/2010-07/per-tor-aug10.pdf>
- Ryan, B. and D. Gill (2011) *Future State: directions for public management in New Zealand*, Wellington: Victoria University Press
- Scott, G., P. Duignan and P. Faulkner (2010) *Improving the Quality and Value of Policy Advice: findings of the Committee Appointed by the Government to Review Expenditure on Policy Advice*, <https://www.treasury.govt.nz/sites/default/files/2011-04/report-repa-dec10.pdf>
- Shaw, R. and C. Eichbaum (2019) 'Still friends? Revisiting New Zealand public servants' perceptions of ministerial advisors', *Political Science*, 71 (2), pp.123–39
- Treasury (2011) 'Summary of the government response to the review on policy advice', <https://www.treasury.govt.nz/publications/summary-government-response-review-expenditure-policy-advice>

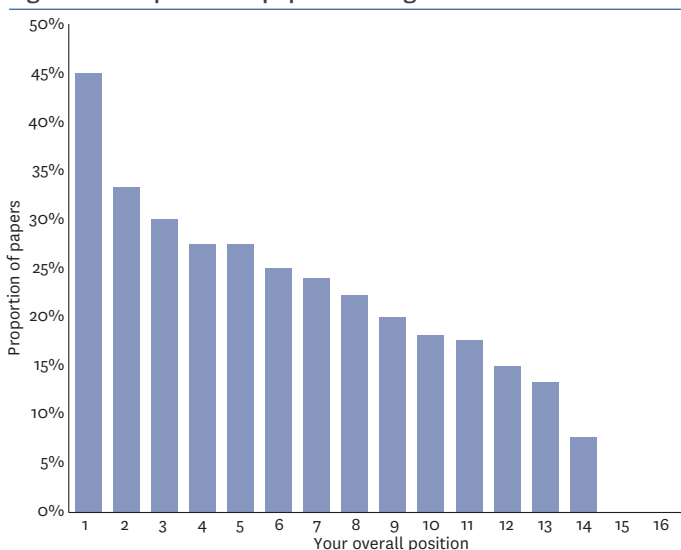
Appendix A: Supplementary data – one year’s results: 2019

Figure A1: Proportion of papers scoring more than 7



Source: NZIER data

Figure A2: Proportion of papers scoring more than 8



Source: NZIER data

Appendix B: NZIER assessment criteria

Overall assessment	
Rating	Overall fitness for purpose, key qualities and issues, and how it could have been better
Customer focus	
Purpose and context	States the objective of the briefing clearly and early Gives enough background to shape the discussion and resolution Makes linkages to wider matters, such as strategy, long-term drivers or other parts of the system
Anticipation	Addresses likely next steps and time frames Has all the necessary content to support next steps, and to pre-empt unnecessary follow-up Provides a tight summary, talking points, a ‘25 words or less’ argument, or other aids
Risk and mitigation	Identifies risks (acceptability, cost, effectiveness, implementation, stakeholder views, etc.) Indicates how risks would be managed (communication, monitoring, trials, evaluation, exit, etc.)
Credible and robust analysis	
Problem definition	Has a clear problem/opportunity definition Indicates the scale, scope and immediacy of the issue
Framework	Shows evidence of an appropriate theory, framework or logical approach to support the analysis Has logical assessment criteria and a clear explanation of how the analysis applies
Options	Shows evidence that a range of alternatives have been developed and considered Gives simple, but logical and powerful reasons for options and dismisses credible alternatives Considers the implementation requirements and challenges Shows evidence of appropriate consultation/engagement and consideration of these perspectives
Data and evidence	Uses evidence (e.g., research, data, evaluations, expert advice) and insights from different points of view (e.g., customer/client input, stakeholders’ views, agency views) to support the arguments Is clear about the strengths, sensitivities and limitations of the evidence Presents accurate numbers and calculations Uses examples or international comparisons to make points
Recommendations	Has clear recommendations that flow from the analysis Is clear about how the recommendations would be implemented, by when and by whom
Clear and concise presentation	
Language	Is concise, uses plain English and minimises jargon, with no errors Uses short sentences and paragraphs to make the reading task easier
Structure and format	Selects the medium (report, poster, presentation, one-pager) that best fits the situation Uses tables and charts that are easy to understand and read Avoids duplication or unnecessary clutter Uses meaningful subheadings as signposts and to tell a logical story