8 November 2021

Ref: OIA-2021/22-0419

Dear

Official Information Act request relating to the Policy Quality Framework

Thank you for your Official Information Act 1982 (the Act) request received on 11 October 2021. You requested:

1) *For 2019/20, please provide any benchmarking document/report which shows each Government agency's policy quality score, as scored using the Policy Quality Framework.*

2) *For 2019/20, please provide any benchmarking document/report which shows each Government agency's ministerial satisfaction survey score.*

3) *For 2020/21, please provide any benchmarking document/report which shows each Government agency's policy quality score, as scored using the Policy Quality Framework.*

4) *For 2020/21 please provide any benchmarking document/report which shows each Government agency's ministerial satisfaction survey score.*

5) *A list of every Government agency/department required to assess the quality of its policy advice using the Policy Quality Framework.*

Request parts 1 & 2

I have decided to release the following documents to you, in full:

Item	Date	Document Description/Subject
1.	25/02/2021	Policy Performance 2019/20 slide pack provided to the Policy Profession Board for its meeting on 25 February 2021.
2.	25/02/2021	Policy Profession Board paper that accompanied the 25 February 2021 Policy Performance 2019/20 slide pack.
3.	02/03/2021	Policy Performance slide pack for a Tier 2 Policy Leaders' Network meeting on 2 March 2021.

These documents were prepared for the Policy Profession Board, as part of a 2019/20 policy information request sent by the Policy Project to agencies.

Please note the following caveats about the released information:

- While all agencies are using the same Policy Quality Framework-based scoring scale, there will be some individual variation between agency review panels and over time in how they interpret and apply the Policy Quality Framework and the scoring scale. The Policy Project is currently considering options to foster greater consistency of scoring across agencies (e.g. implementing some form of moderation process for the coming years). Viewing the results as a league table is therefore not helpful as you may not be comparing 'like with like'.

- The policy quality scores included in the slide packs do not take into account the full context of each agency's policy performance relative to that of others (e.g. relative workloads, complexity of subject matter, etc).

- The sample size and types of papers considered may also vary significantly between agencies, depending on the number of policy advice papers prepared by an agency, and the different types of papers included in the sample, which vary according to agencies' roles. Some papers may contain quite complex policy advice, while others may be more procedural.

- The context and time constraints under which policy advice papers are developed may also vary considerably from paper to paper. This is likely to impact on scores.

- The scope for individual variation in ministerial satisfaction scoring is also considerable. The approach ministers take to scoring their agencies will vary according to the personal style of the minister and different ministerial expectations, as well as the quality of the policy advice and related services provided by their agency/agencies.

- The policy quality scores for the Department of Prime Minister and Cabinet do not cover the performance of the Department's Policy Advisory Group (PAG), as the PAG's advice is not submitted for the annual policy quality review.

Request parts 3 & 4

The Policy Project did not send a policy information request to agencies for the 2020/21 year and does not hold a collated set of information from agencies on policy quality and Ministerial satisfaction scores for that year. Accordingly, I am refusing these requests under section 18(g)(i) of the Act, as the information is not held by the Department of the Prime Minister and Cabinet and I have no grounds for believing that the information would be held by another agency or minister of the Crown.

All government agencies with a policy appropriation are required to report a policy quality score and a ministerial satisfaction score in their Annual Reports, however. If you are still interested in pursuing this information, I suggest that you look in each individual agency's 2021 Annual Report for its 2020/21 results.

Request part 5

Appendix 1 of the Policy Profession Board paper that accompanied the 25 February 2021 Policy Performance 2019/20 slide pack (document 2 above) includes a list of the agencies that the Policy Project included in its 2019/20 policy information request.

Cabinet's decision in 2012 also included the Education Review Office (ERO) and the New Zealand Defence Force (NZDF) in the list of agencies that had to report on policy performance scores. However, ERO no longer has a policy appropriation so is not required to report a policy score. NZDF's policy quality assessment information was not collected as part of the 2019/20 policy information request as that request focussed on core public agencies not crown entities.

You have the right to ask the Ombudsman to investigate and review my decision under section 28(3) of the Act.

This response will be published on the Department of the Prime Minister and Cabinet's website during our regular publication cycle. Typically, information is released monthly, or as otherwise determined. Your personal information, including name and contact details, will be removed for publication.

Yours sincerely
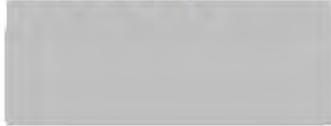
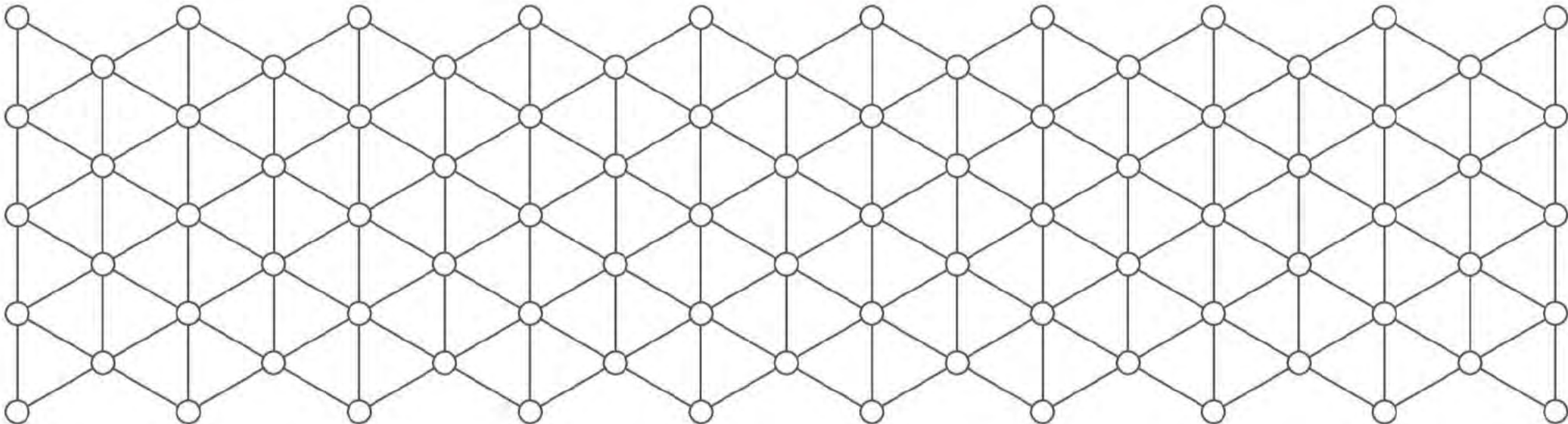Anneliese Parkin
Deputy Chief Executive, Policy

Policy Performance 2019/20

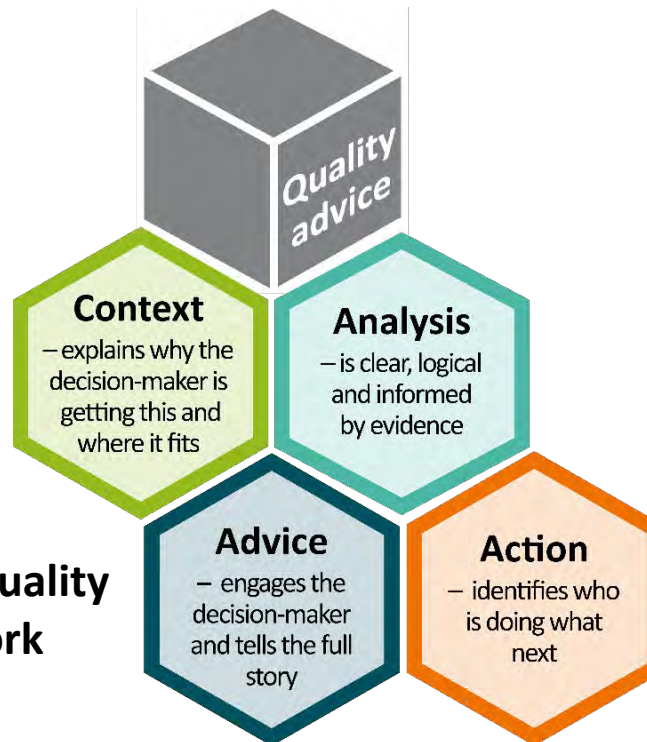## Policy Profession Board

25 February 2021

Overview

Analysis of information request data – on the quality of policy advice and on ministerial satisfaction with policy services

In September 2020, the Policy Project asked all 28 public service agencies with a policy function to provide information on:

- a range of quality of policy advice indicators from policy quality review panels and ministers
- their use of the revised Policy Quality Framework (PQF) and how to improve it, and the Ministerial Policy Satisfaction Survey.

This is the third information request the Policy Project has undertaken since 2017 – and the first since all agencies adopted the PQF standardised scoring system for quality of advice assessment. In 2019/20 the revised PQF was also incorporated in the revised Ministerial Policy Satisfaction Survey. This year's results can therefore act as a baseline for assessing policy performance trends in subsequent years. They are also of interest in their own right.

The Policy Quality Framework

Quality advice

Context – explains why the decision-maker is getting this and where it fits

Analysis – is clear, logical and informed by evidence

Advice – engages the decision-maker and tells the full story

Action – identifies who is doing what next

Objectives

➢ The objectives of the Policy Performance 2019/20 exercise are to analyse information received from public service agencies on policy quality assessments and ministerial satisfaction, to enable the Policy Project to:

- Better understand the performance of the policy system.

- Identify areas of common development needs – so the Policy Project can design and provide agencies with practical support to address these needs.

- Understand how to improve the Policy Quality Framework and Ministerial Policy Satisfaction Survey resources that the Policy Project has provided for the policy system.

- Identify where we most need to support policy practitioners in using the Policy Quality Framework (and related Policy Skills and Capability Frameworks).
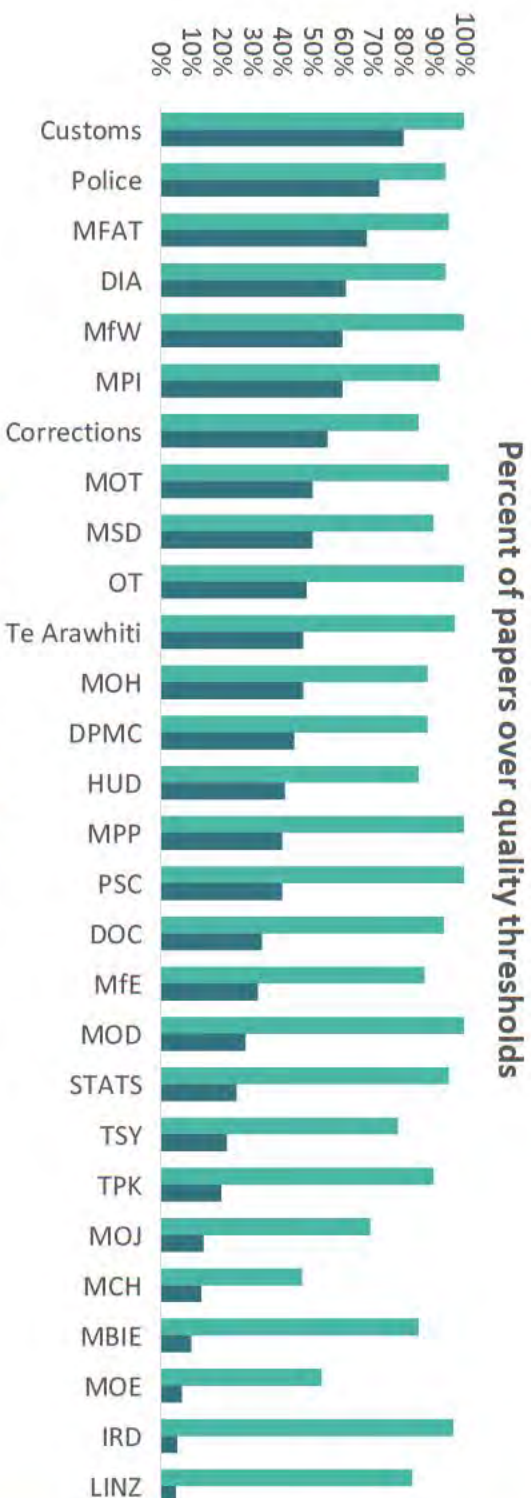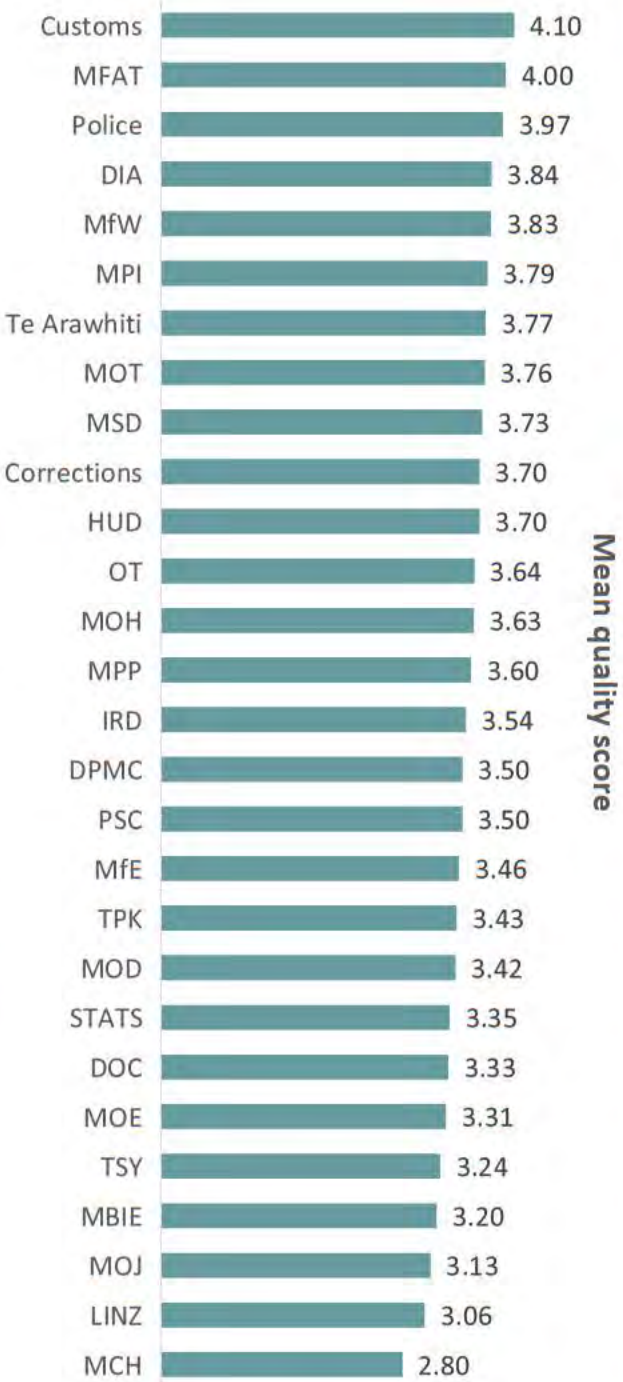
What this presentation covers

This presentation provides a high-level overview of policy performance results and key insights from the data received through the 2019/20 and the previous information request. It encompasses:

1. **Policy quality performance**: comparative policy quality scores for all public service agencies in 2019/20 on three policy quality metrics, changes in mean scores since 2017/18, and next steps

2. **Panel types and frequency of quality of advice reviews:** information about the types and frequency of panel reviews of the quality of agencies' policy advice

3. **Performance insights:** implications of the quantitative policy quality metrics and thematic analysis of qualitative data from review reports for future Policy Project focus areas

4. **How agencies are using the PQF:** utilisation patterns and agency perceptions of usefulness

5. **Ministerial Policy Satisfaction Survey:** quantitative analysis of performance scores in responses to the surveys, and thematic analysis of ministers' comments about what matters to them

6. **Next steps for analysing the performance data**: seeking feedback on future data analysis priorities

7. **Implications for the Policy Project:** seeking feedback on what insights from this data analysis means for the Policy Project's future focus in supporting agencies to improve policy performance.

Mean quality score

Agency	Score
Customs	4.10
MFAT	4.00
Police	3.97
DIA	3.84
MfW	3.83
MPI	3.79
Te Arawhiti	3.77
MOT	3.76
MSD	3.73
Corrections	3.70
HUD	3.70
OT	3.64
MOH	3.63
MPP	3.60
IRD	3.54
DPMC	3.50
PSC	3.50
MfE	3.46
TPK	3.43
MOD	3.42
STATS	3.35
DOC	3.33
MOE	3.31
TSY	3.24
MBIE	3.20
MOJ	3.13
LINZ	3.06
MCH	2.80

Percent of papers over quality thresholds

- Papers scoring over 3
- Papers scoring over 4

Agencies (in order): Customs, Police, MFAT, DIA, MfW, MPI, Corrections, MOT, MSD, OT, Te Arawhiti, MOH, DPMC, HUD, MPP, PSC, DOC, MfE, MOD, STATS, TSY, TPK, MOJ, MCH, MBIE, MOE, IRD, LINZ

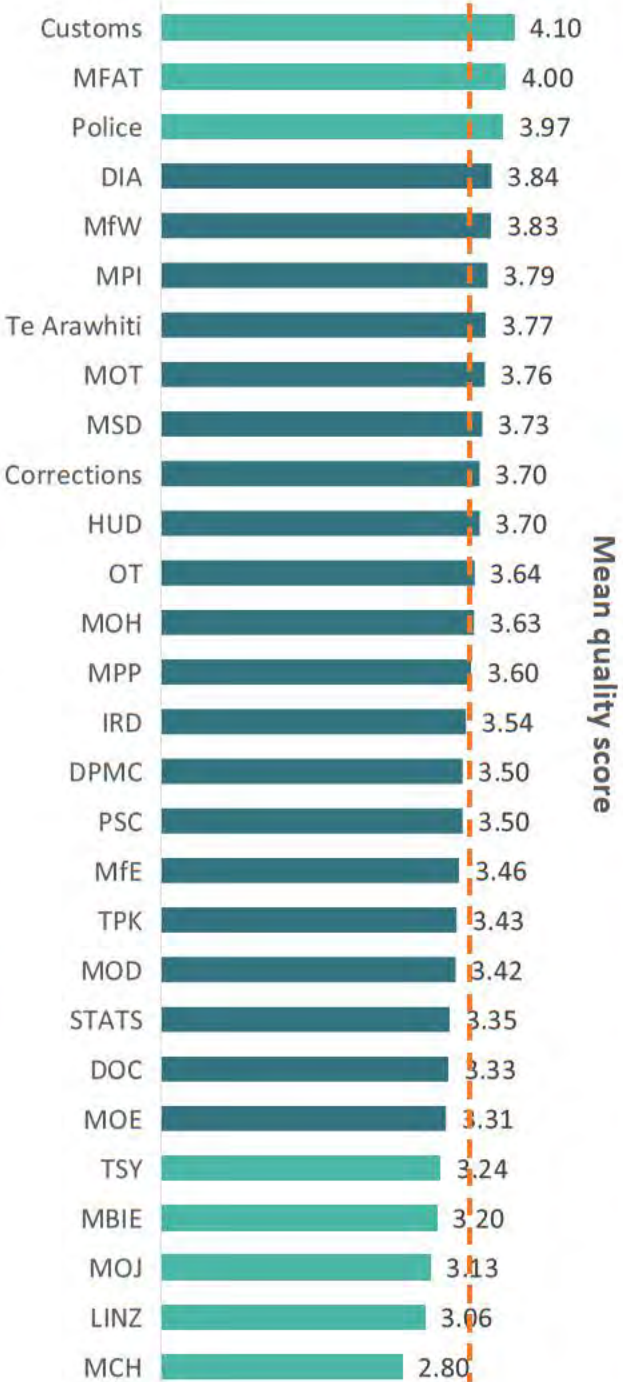(axis: 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%)
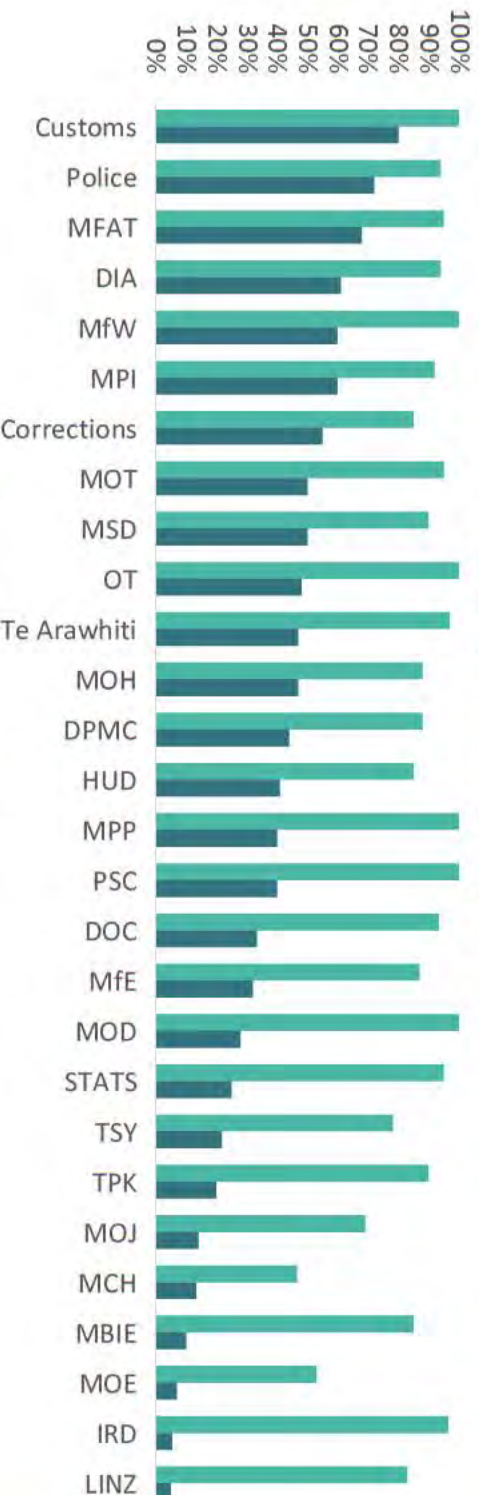
Notes on the data

- All 28 agencies provided data on their mean policy quality score for a sample of papers assessed using the PQF.

- Papers scoring 1 or 2 in the PQF scale are assessed as below minimum standards. 3 is acceptable, 4 good, and 5 outstanding.

- The mean score ranges from 2.8 to 4.1. The public service mean of all agency mean scores is 3.55, with a standard deviation of 0.297.

- The clear majority of agencies (20 of 28) are within one standard deviation of the mean, with three above and five below. That is, there are a large group of agencies clustered around the mean, with few outliers.

4

1a. Policy quality performance – assessed by review panels

Mean quality score

Agency	Score
Customs	4.10
MFAT	4.00
Police	3.97
DIA	3.84
MfW	3.83
MPI	3.79
Te Arawhiti	3.77
MOT	3.76
MSD	3.73
Corrections	3.70
HUD	3.70
OT	3.64
MOH	3.63
MPP	3.60
IRD	3.54
DPMC	3.50
PSC	3.50
MfE	3.46
TPK	3.43
MOD	3.42
STATS	3.35
DOC	3.33
MOE	3.31
TSY	3.24
MBIE	3.20
MOJ	3.13
LINZ	3.06
MCH	2.80

Percent of papers over quality thresholds

- Papers scoring over 3
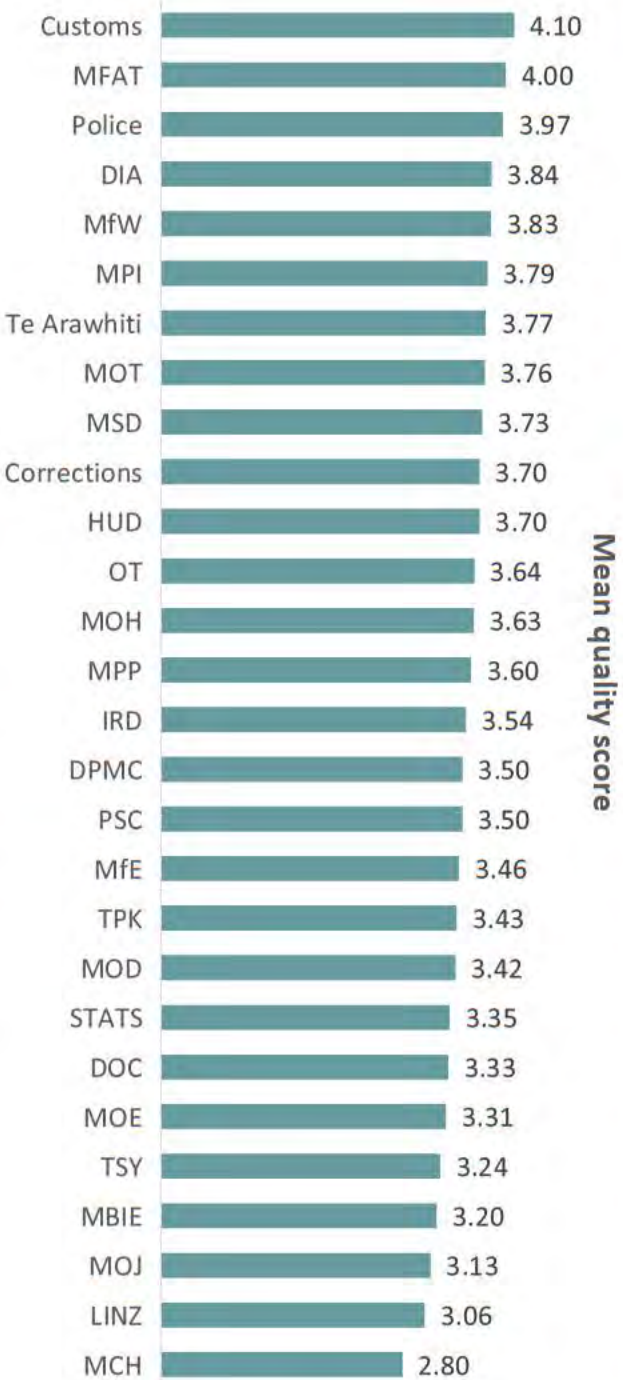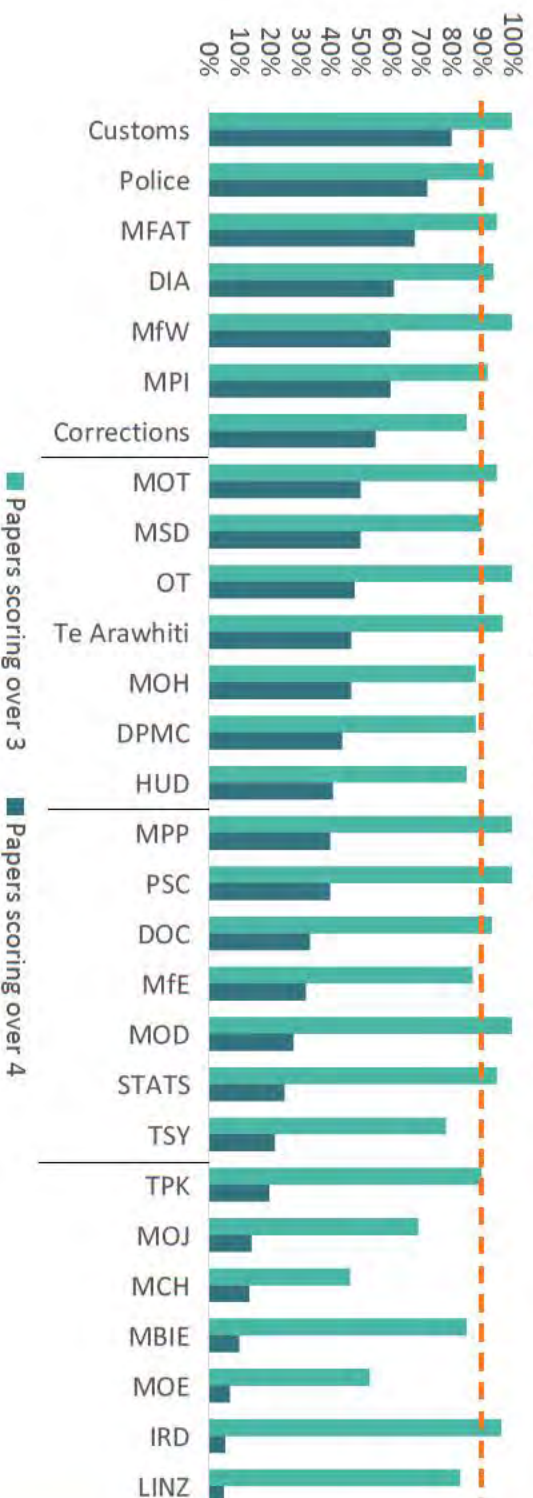- Papers scoring over 4

Notes on the data

- All 28 agencies provided data on their mean policy quality score for a sample of papers assessed using the PQF.

- Papers scoring 1 or 2 in the PQF scale are assessed as below minimum standards. 3 is acceptable, 4 good, and 5 outstanding.

- The mean score ranges from 2.8 to 4.1. The public service mean of all agency mean scores is 3.55, with a standard deviation of 0.297.

- The clear majority of agencies (20 of 28) are within one standard deviation of the mean, with three above and five below. That is, there are a large group of agencies clustered around the mean, with few outliers.

5

1a. Policy quality performance – assessed by review panels

Mean quality score

Agency	Score
Customs	4.10
MFAT	4.00
Police	3.97
DIA	3.84
MfW	3.83
MPI	3.79
Te Arawhiti	3.77
MOT	3.76
MSD	3.73
Corrections	3.70
HUD	3.70
OT	3.64
MOH	3.63
MPP	3.60
IRD	3.54
DPMC	3.50
PSC	3.50
MfE	3.46
TPK	3.43
MOD	3.42
STATS	3.35
DOC	3.33
MOE	3.31
TSY	3.24
MBIE	3.20
MOJ	3.13
LINZ	3.06
MCH	2.80

Percent of papers over quality thresholds

Legend:
- Papers scoring over 3
- Papers scoring over 4

Agencies (top to bottom): Customs, Police, MFAT, DIA, MfW, MPI, Corrections, MOT, MSD, OT, Te Arawhiti, MOH, DPMC, HUD, MPP, PSC, DOC, MfE, MOD, STATS, TSY, TPK, MOJ, MCH, MBIE, MOE, IRD, LINZ

Notes on the data

- All 28 agencies provided data on their mean policy quality score for a sample of papers assessed using the PQF.

- Papers scoring 1 or 2 in the PQF scale are assessed as below minimum standards. 3 is acceptable, 4 good, and 5 outstanding.

- The mean score ranges from 2.8 to 4.1. The public service mean of all agency mean scores is 3.55, with a standard deviation of 0.297.

- The clear majority of agencies (20 of 28) are within one standard deviation of the mean, with three above and five below. That is, there are a large group of agencies clustered around the mean, with few outliers.
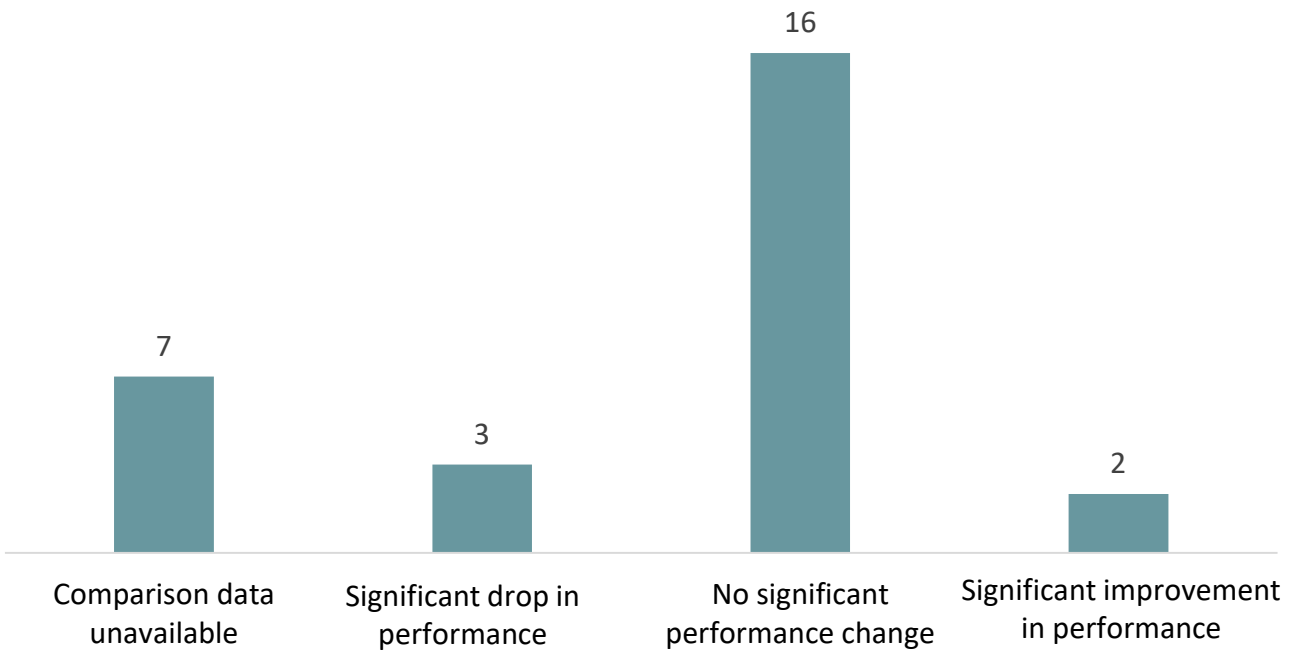
6

1b. Policy quality performance – tracking over time

28 agencies provided data on their mean policy quality in the 2020 year. Of these, 21 had previously provided us with data for 2017/18 that could be compared with their scores using the revised PQF for 2019/20. The public service-wide mean score in 2019/20 was 3.55.

Analysis of the differences in mean quality scores between 2017/18 and 2019/20 for the 22 agencies for which reasonably comparable data was achievable reveals the following trends. The overwhelming majority of agencies – 16 out of 21 (76%) – showed no significant shift in performance between 2017/18 and 2019/20 (which we have defined as having a difference between mean scores for these years of less than half a point). Three agencies had a drop in performance of more than half a point (LINZ, MCH, MOJ). Two agencies improved by more than half a point (MPP and the Treasury).

Shifts in quality of policy advice performance 2018/2020

7	3	16	2
Comparison data unavailable	Significant drop in performance	No significant performance change	Significant improvement in performance

Notes on the data

➤ 2017/18 mean scores for 21 agencies were converted to revised PQF scores using a methodology developed by NZIER in 2020 and reviewed by the Policy Project.

➤ We haven't converted scores for agencies that had their panel use bespoke scoring systems in 2018, as this wouldn't be sufficiently accurate for a reliable comparison (MfE, MBIE, MoE, IRD Oranga Tamariki)

➤ In addition, the two public service departments that didn't exist in 2017/18 – HUD and Te Arawhiti – couldn't be included in the comparison.

➢ The first Policy quality performance slide assessed systemic quality of policy advice performance using three performance measures. Each measure provides useful information for assessing a different aspect of performance – and any one of them alone will not give a full picture. Other metrics could also be used (e.g. the percentage of papers scoring less than 3).

➢ The analysis revealed that an agency's performance on one metric is a relatively poor predictor of performance on another. For example, two agencies with similar mean scores may have a very different spread of paper scores, with one agency having several low scoring papers and several high scoring papers, while another agency has most papers scoring near the middle of the range.

➢ Under current Treasury reporting requirements, agencies must use the refreshed Policy Quality Framework from 2019/20 to assess the quality of their policy advice papers, and must report on this in Annual Reports – and select their own policy quality measures and targets.

➢ Current guidance from the Policy Project on performance metrics in Annual Reports recommends that agencies report on both:

  • the mean score of their assessed sample, and

  • the percentage of papers that meet a performance target set by the agency (for example, 50% of papers scoring 4 or over).

For discussion

Should guidance be revised to recommend a more consistent approach to reporting against benchmarks?

➢ Should Policy Project guidance enable a fuller picture and easier comparison between agencies in future, by recommending that all agencies:

  a. report on:
     – the mean score
     – the percentage of papers scoring over 3 ('adequate')
     – the percentage scoring over 4 ('good')

  b. while also emphasising that each agency still sets its own targets for each measure?

➢ Note: policy performance reviews for future years could examine whether trends start to emerge between different performance metrics over time.
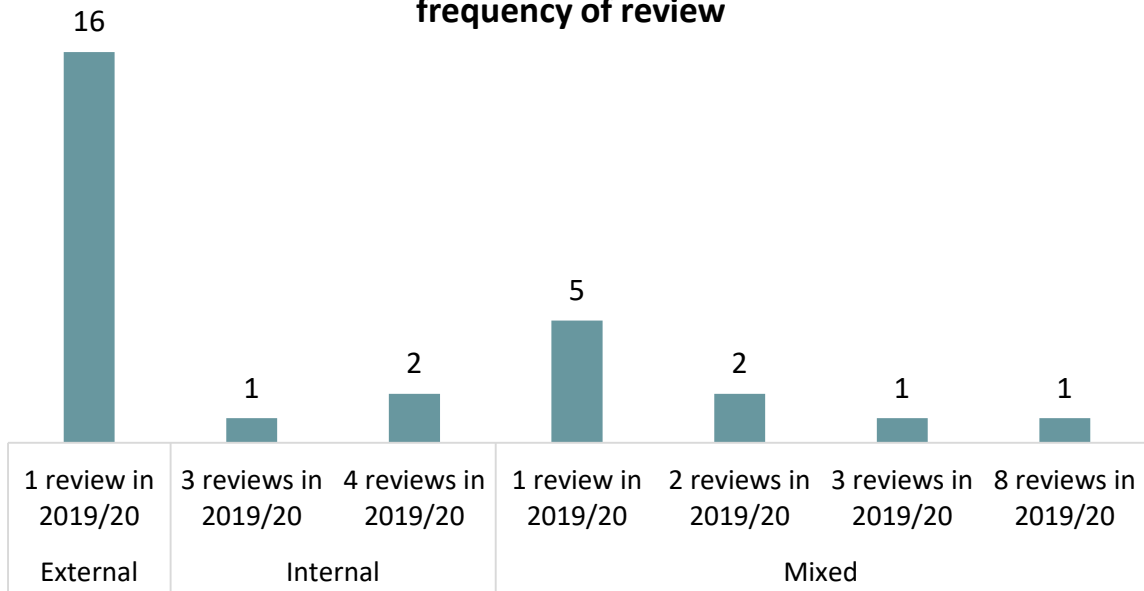
Most agencies use NZIER and do one annual review

Agencies are free to choose who does their policy quality assessment and how frequently. In 2019/20, 16 out of 28 agencies (57%) used NZIER to undertake policy quality assessments in a single annual review. We refer to this type of policy quality assessment as undertaken by an 'external' panel.

Nine agencies used panels composed of staff from the agency and other external expert members. We refer to this type of review as being by a 'mixed' panel. The remaining four agencies used panels drawing only from internal staff – referred to as 'internal'. Two of the internal panels also reported to us that they used external moderators.

Number of agencies using each type and frequency of review

Type	Frequency	Count
External	1 review in 2019/20	16
Internal	3 reviews in 2019/20	1
Internal	4 reviews in 2019/20	2
Mixed	1 review in 2019/20	5
Mixed	2 reviews in 2019/20	2
Mixed	3 reviews in 2019/20	1
Mixed	8 reviews in 2019/20	1

Commentary and next steps

➢ Internal and mixed panels can potentially take a more strategic and nuanced view than external panels (e.g. considering the overall PQF score across a set of related papers, or average scores by subject matter area or group). Using an internal or mixed panel also helps departments grow internal capability in quality assessment of policy advice papers, with internal panel members becoming champions for quality policy advice in their teams.

➢ However, external panels and external members on panels (whether NZIER or otherwise) can provide a new perspective and draw on insights from reviewing a wide range of papers across government. Both have value.

➢ Agencies that are only doing an internal review may wish to consider using an external review at periodic intervals as a benchmarking exercise, and/or consider how to use external moderation (subject to capacity and demand, the Policy Project could have a role in this – having already undertaken this once for MoJ).

➢ Agencies that are only doing an external review may wish to consider how to develop internal quality assurance processes that draw on the PQF and provide for quicker feedback than can be provided though a single, external, yearly review.

➢ The Policy Project can provide cross-government training based on our *A guide on panels and processes for assessing policy advice papers*, and tailored support for agencies wanting to establish internal review processes.

9

3a. Performance insights – lifting the floor

Focus Area 1

Lifting the floor – addressing the lowest performing papers
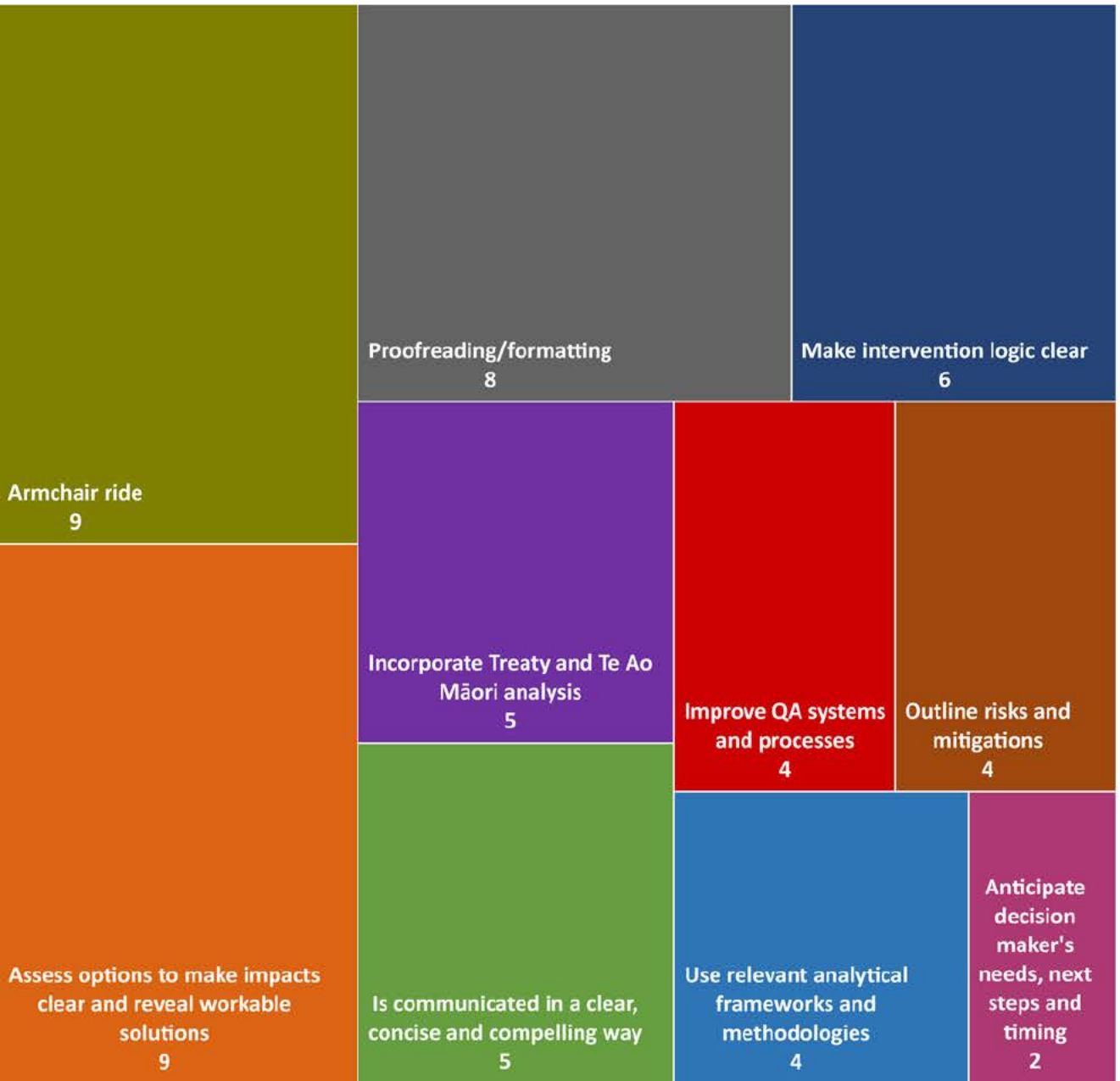
Key points

- **Low performance is relatively common**: All agencies reported on the minimum quality score in their sample of papers. 22 of the 28 agencies (79%) had some papers scoring below a 3 (the standard for 'acceptable' under the PQF). Of these, 8 had a minimum score of 2.5, 13 had a minimum score of 2, and one had a minimum score of 1.5.

- **Papers scoring under 3 should not be signed out:** The rating scale of the PQF builds in the concept that papers scoring below 3 should not have been signed out. Under the PQF, a paper scoring 1 "creates serious risk of poor decision making", and needs "fundamental rework". A paper scoring a 2 "creates a risk of poor decision making" and needs "substantial improvement in important areas".

- **Low scoring papers create risks:** Low-scoring papers can have a particularly detrimental effect on a minister's confidence in an agency's policy capability. If a poor quality paper is the first in a series of papers on a related topic, there is a risk of poor decision making that has flow-on effects for a whole stream of advice.

Possible next steps (for discussion):

➢ The Policy Project: could include discussion of the incidence of low scoring papers and how to reduce this through more agency use of ex-ante quality assurance and other means on the agenda at the forthcoming Policy Quality Review Panel Forum.

➢ Agencies with low-scoring papers: could undertake further work on the contexts in which they are signed out, to identify and then plan how to address underlying factors. This could include identifying the types of papers that are most likely to be low scoring, and whether there are agency-specific or system-wide improvements that could lift the performance of these papers (for example, improved peer review processes).

➢ The Head of the Policy Profession: could ask Tier 2 Policy Leaders to reiterate the importance of policy staff using the PQF checklist before papers are signed out. This could be done when they're considering the system-wide report on Policy Performance 2019/20 at their 27 April meeting.

Armchair ride
9

Assess options to make impacts clear and reveal workable solutions
9

Proofreading/formatting
8

Incorporate Treaty and Te Ao Māori analysis
5

Is communicated in a clear, concise and compelling way
5

Make intervention logic clear
6

Improve QA systems and processes
4

Outline risks and mitigations
4

Use relevant analytical frameworks and methodologies
4

Anticipate decision maker's needs, next steps and timing
2

Notes on the data

➤ Thematic analysis of the narrative sections of all 28 agencies 2019/20 policy quality reports revealed the common areas for development. Narrative statements have been thematically linked and summarised as consistently as possible. The larger the box, the more agencies identified that area needed development.

➤ The majority of areas for improvement are from the 'analysis' section of the PQF, followed by the 'advice' section.

➤ The need to give ministers more of an 'Armchair ride' was regularly identified by NZIER in the 16/28 agency quality review reports they wrote. This means providing a minister with all the support the minister may need. The PQF captures three different dimension of this:

- Purpose, context, priorities and connections across government are clear.

- Enables a clear and informed decision on next steps.

- Anticipates decision maker's needs, next steps, and timing.

➤ The theme of 'Make intervention logic clear' arising mostly in NZIER reports, means developing a coherent narrative across the problem definition, options analysis, and recommended action.

11

Focus Area 2

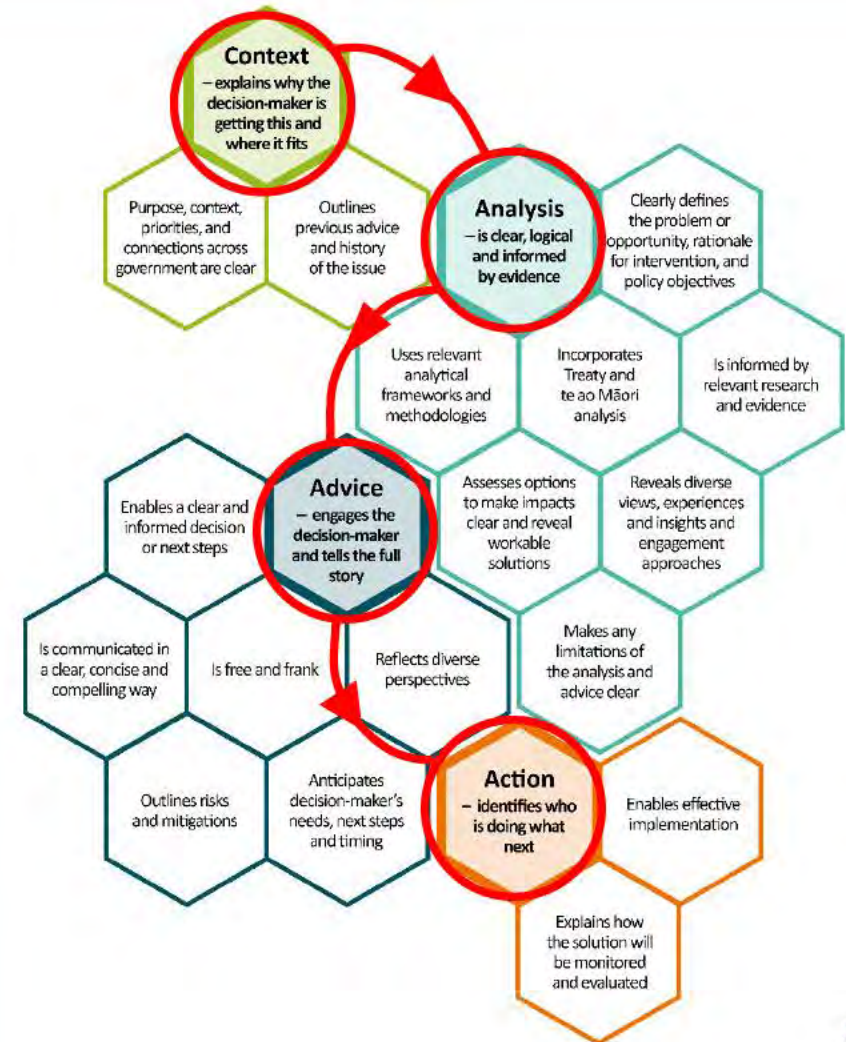### The whole is greater than the sum of the parts: drawing connections across the PQF

Key points

- Review panels consistently identified that the best papers create a logical chain from the context, through the analysis, to advice and action. This goes to the heart of creating robust policy advice.

- Some internal panels scored papers against each element of the PQF separately, and analysis of these assessments shows that a paper with a weak problem definition is likely to remain weak throughout. Similarly, if the analysis is poor, the advice will not score well.

- While each element of the PQF is an essential component for good policy advice, it can also be understood as a sequential building of robust advice from the outset.

Possible next steps (for discussion):

The Policy Project could:

- identify and disseminate annotated examples of best practice, that demonstrate how elements of the advice chain link together

- encourage agencies to spot issues in the problem definition and options analysis internally, ideally before final decision papers are written up

- include discussion of the policy quality chain of logic in training materials for new graduates and other policy professionals.
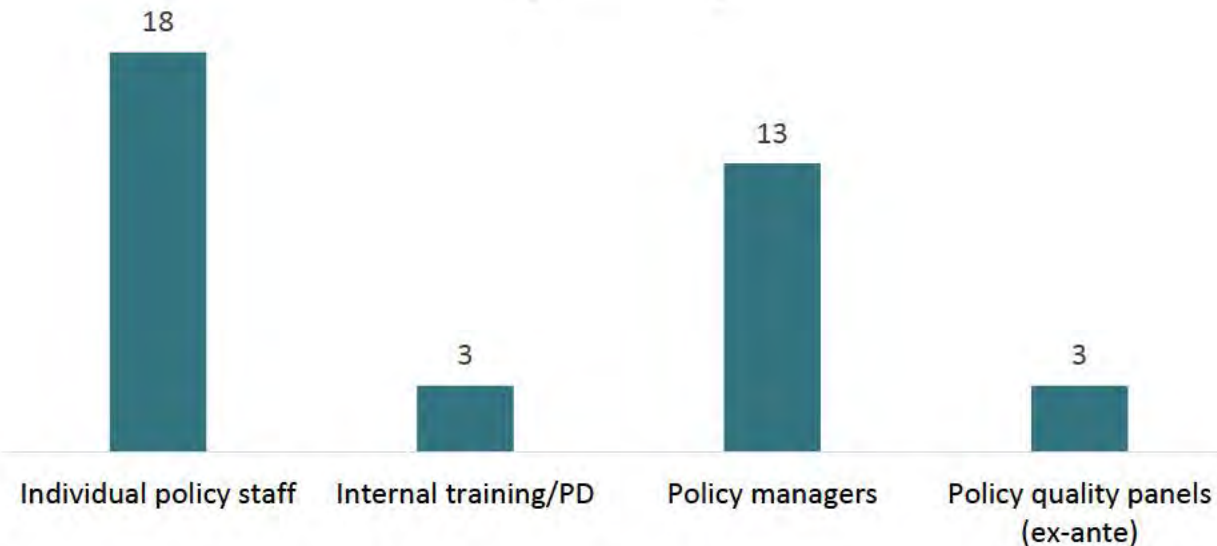

12

4. Using the Policy Quality Framework

Agencies are using the PQF and rate it highly

As a minimum, all agencies are required to use the PQF to conduct ex-post reviews of the quality of policy advice for Estimates reporting – and all 28 reported that their panel does. Other ways that agencies can use it to support them in developing quality papers include: when developing them, when peer reviewing them, when assessing them before signoff to the minister (ex-ante), etc. 27 of 28 agencies are also using the PQF for some other purposes. Individual policy staff are using the PQF to develop and review papers in 18 of 28 agencies (64%), and managers are using the PQF in assessing papers in 13 out of 28 agencies (46%).

How are agencies using the PQF?

Agencies were also asked how helpful they find the resource *Using the Policy Quality Framework to assess papers: A guide on panels and processes*. 79% of agencies rated it positively, with 15 out of 28 saying it was either "extremely helpful" or "very helpful".

PQF Assessment

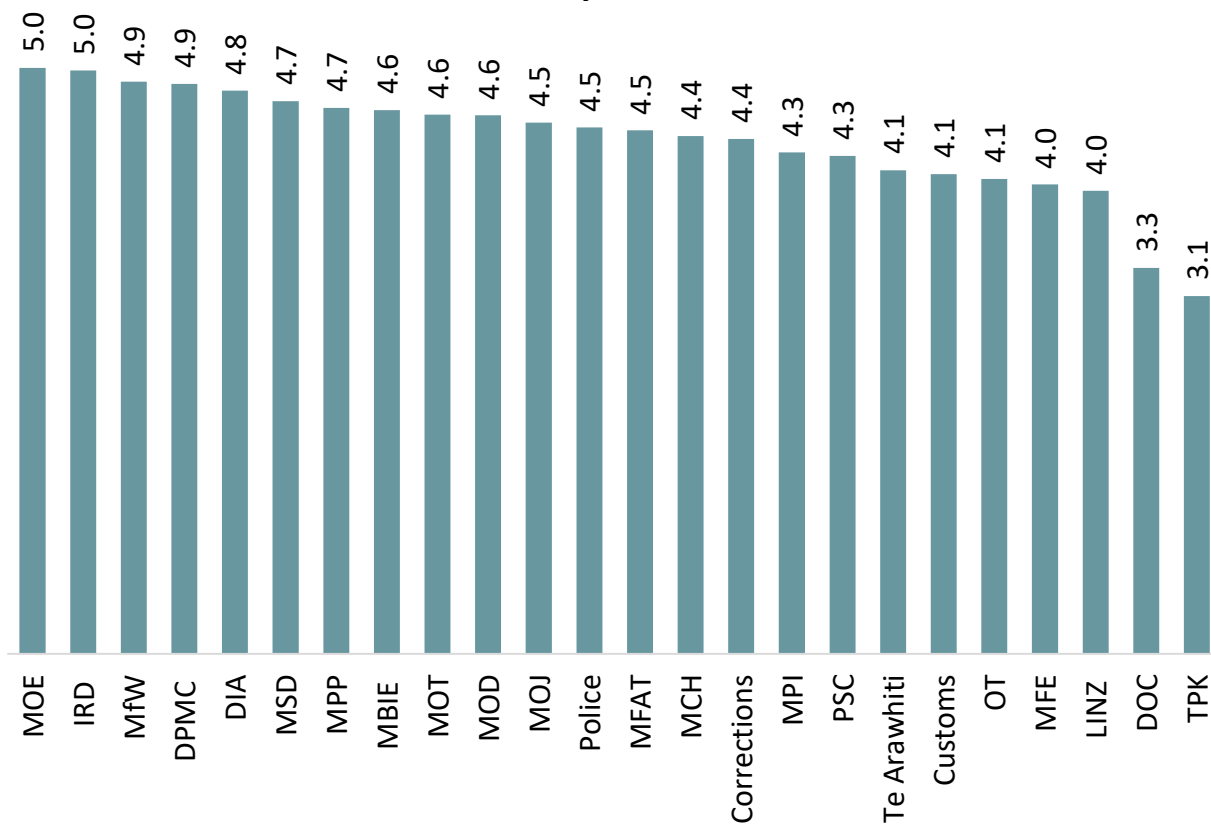| Question	Responses (out of 28)
PQF was easy to use	
Strongly agree	3
Agree	19
Neither agree or disagree	3
Disagree	1
No response	2
PQF includes all standards agency needs	
Strongly agree	2
Agree	21
Neither agree or disagree	3
Disagree	1
Strongly disagree	1
PQF could be applied to a range of papers	
Strongly agree	4
Agree	14
Neither agree or disagree	7
Disagree	3

5a. Ministerial Satisfaction scores

Average results, by agency

The Ministerial Policy Satisfaction Survey scores are reported in Annual Reports. They are calculated by averaging the scores each minister gave to subsets of 18 different aspects of their agency's performance in delivering policy services. The public service as a whole demonstrated high levels of overall ministerial satisfaction in 2019/20 – with a mean of 4.4, and a range of 3.1 to 5, out of 5.

Ministerial Policy Satisfaction Scores

Agency	Score
MOE	5.0
IRD	5.0
MfW	4.9
DPMC	4.9
DIA	4.8
MSD	4.7
MPP	4.7
MBIE	4.6
MOT	4.6
MOD	4.6
MOJ	4.5
Police	4.5
MFAT	4.5
MCH	4.4
Corrections	4.4
MPI	4.3
PSC	4.3
Te Arawhiti	4.1
Customs	4.1
OT	4.1
MFE	4.0
LINZ	4.0
DOC	3.3
TPK	3.1

Notes on the data

➢ Ministers can elect not to complete a survey. We received 54 surveys across 24 agencies (out of 28 agencies). For agencies with more than one minister completing the survey, we averaged the scores of their ministers.

➢ Three agencies asked ministers additional questions relevant to their work programme – the scores for those questions were not included in the averages reported opposite.

➢ We compared how agencies reported the results of ministerial surveys and panel's policy quality assessments in their Annual Reports. This varies widely in both cases.

➢ No clear relationship is evident between these ministerial survey scores and the ex-post policy quality scores of a sample of policy papers in slide 3. Agencies whose review panels assessed their policy advice papers as scoring highly or poorly on policy quality were not consistently scored highly or poorly by their minister(s).

➢ As both are measuring somewhat different but related aspects of performance, it's important that agencies focus on what they can learn from the quantitative and qualitative feedback both include.

➢ The written comments from ministers contain valuable insights on areas where agencies can improve policy performance (see next slide).

14

5b. Ministerial Policy Satisfaction – thematic comments analysis

Maintain open communication on any new issues, agency priorities, and timeframes
9

Concise, clear writing that identifies the key issues
7

Discuss policy early and explicitly with minister to capture minister's priorities.
7

Engage with other agencies to improve whole of government delivery on priorities
4

Recognising Treaty of Waitangi responsibilities and engaging with Māori
4

Early communication to minister of any potential stakeholder issues
3

Whole of project thinking – planning in advance for smooth handover from policy to operations, effective implementation, and public outreach/education/ communication (where relevant) at the conclusion of the policy phase
3

Cultural competency and diversity
2

Focus on prompt delivery of outcomes – remembering policy is a step towards operational or legislative changes and ministers have timeframes in mind for achieving those changes
5

Maintaining agency capacity to respond quickly to new and emerging issues
3

Ensure advice is free and frank, and allow for full engagement with minister on difficult or contentious issues
2

Identify the appropriate level of detail needed – manage agency's workload and minister's time well
2

Timely response to requests from private secretaries and ministerial advisors
2

Note on the data

➤ Thematic analysis of the open-ended survey responses has identified areas ministers value – whether they were mentioned as a strength or an area for improvement. Narrative statements have been summarised as consistently as possible, using terminology in the Policy Quality Framework and Policy Capability Framework.

Insights for improvement

➤ The ministerial survey provides particularly useful detail for policy advisers on how to give better effect to the quality indicator of "anticipates decision-makers needs, next steps, and timings". The Policy Project could consider drawing on this to create additional resources on this aspect.

➤ It's clear that ministers see agencies as full service providers and expect close working relationships between policy and operational service delivery, and engagement across government.

➤ There's close alignment between matters ministers value, and matters identified in the Policy Quality and Policy Capability frameworks.

15

6. Further analysis of 2019/20 policy performance data?

Which options below (or what else) do you consider top priority next steps, and why?

➢ **Possible areas for further data analysis by the Policy Project in the coming months**:

   a) Detailed analysis of Ministerial Policy Satisfaction Survey scores, focused on how ministers score agencies on PQF elements – and how those scores compare with agency's panel scores.

   b) Triangulating the results with Cabinet Paper qualitative assessments by the CE and DCE (Policy), DPMC.

   c) Analysing patterns in agency targets for policy quality scores, and whether they are achieved.

   d) Assessing the impacts of COVID-19 on quality score patterns.

   e) Identifying aspects of the PQF that agencies consider need changing, including further analysis of open-ended responses from agencies on the PQF.

   f) Analysing agency feedback on Ministerial Policy Satisfaction Survey aspects needing change.

   g) Detailed analysis of different uses of internal or mixed panels and external moderation.

   h) Considering the implications of any of the above for whether refinements are needed to the *Policy Quality Framework: A guide on panels and processes for assessing policy advice papers.*

   i) Analysis of agency differences in policy pay scales and remuneration policy (from information request) – and in actual remuneration (drawn from PSC's Human Resource Capability Survey.

7. How can the Policy Project help improve performance?

Discussion – key questions and future directions

➢ Discussion of the implications of this analysis of policy performance in 2019/20 for the Policy Project's 2020/21 work programme is necessarily preliminary – as that also needs to be informed by the results of the interim evaluation. The evaluation will help us assess whether we are collecting the right data about policy performance in the right way, and what to change.

➢ **We would appreciate your feedback on the following questions**:

    a) Do you agree that in supporting agencies to improve the quality of policy advice, the Policy Project should increase its focus on encouraging them to:

- lift the floor – identify and address factors that result in low scoring papers?
- ensure their policy advisers and managers understand the importance of drawing connections across the PQF and that the whole is greater than the sum of the parts?

    b) What further steps could the Policy Project take to support agencies to enhance their quality assurance processes:

- to embed ongoing improvement into their organisation's approach to producing policy advice and
- focus on processes to identify and respond to gaps in performance as these emerge?

Policy Performance 2019/20
Policy Profession Board meeting
25 February 2021

1. Purpose

The purpose of this paper and the attached slide pack is to:

- provide a high-level overview of policy system performance in the 2019/20 year
- seek your feedback on some suggestions for more detailed data analysis, and possible future Policy Project areas of focus to lift agency policy performance.

2. Background

Collection of policy performance information

In September 2020, the Policy Project undertook the third of its semi-regular Policy Information Requests, asking all 28 public service agencies with a policy function to provide the following data:

- A copy of the agency's policy quality assessment reports for 2018/19 and 2019/20.
- A copy of the agency's completed Ministerial Policy Satisfaction Survey(s) for 2019/20.
- Information on the frequency and format of the agency's policy quality processes (if the agency ran an agency panel to assess policy quality).
- Feedback on the use and usefulness of the 2018 refreshed Policy Quality Framework (PQF).
- Feedback on the 2018 revised Ministerial Policy Satisfaction Survey.
- Information on remuneration policy and pay scales for policy staff.[1]

The Policy Project followed up with agencies to ensure a 100% response rate. A list of the 28 public service agencies that responded is attached as Appendix 1.

Fit with Policy Project priorities, and how the information will be used to achieve them

Analysis of the performance data collected – as reported in the attached slide pack – will support the Policy Project's approach to building policy system maturity. It will particularly support the following two objectives identified in the Policy Project's four-year plan: promoting common standards and intervening at the system level.

The Policy Project will use the results of the analysis to help identify how best to lift policy performance across the public service. Therefore, the analysis is primarily focused on trends that emerge for more than one agency, with the objectives of enabling the Policy Project to:

- better understand the performance of the policy system
- identify areas of common development needs, so the Policy Project can design and provide agencies with practical support to address their policy performance improvement needs
- understand how to improve the Policy Quality Framework (PQF) and Ministerial Policy Satisfaction Survey that the Policy Project has provided for the policy system

---

[1] The slide pack and this paper only reports on policy performance, Ministerial satisfaction, and use of the Policy Quality Framework – the data on remuneration for policy staff will be reported on separately, after we have analysed it.

- identify where we most need to support policy practitioners in using the PQF (and related Policy Skills and Policy Capability Frameworks).

Scope of paper

This paper is intended to provide background information on key aspects of the methodologies applied in the data analysis presented in the slide pack, and it provides additional context to assist you in navigating the analysis of the policy performance data. It should be read together with the slide pack.

3. Essential context

Scoring of policy papers under the Policy Quality Framework

The first section in the slide pack analyses the quantitative scores for a sample of policy advice papers delivered by all public service agencies in 2019/20 – as assessed by a panel of experienced policy practitioners.

In 2019/20 for the first time, all agencies' panels used the PQF for ex-post assessments of policy papers. Appendix 2 provides the summary version of all the elements of the PQF.

The PQF panel guidance sets out a five-point scoring scale, summarised as:

1	Unacceptable	Does not meet the relevant quality standards in fundamental ways
2	Poor	Does not meet the relevant quality standards in material ways
3	Acceptable	Meets the relevant quality standards overall, but with some shortfalls
4	Good	Meets all the relevant quality standards
5	Outstanding	Meets all the relevant quality standards and adds something extra

The full version of the PQF scoring scale (with detailed descriptors for each score) is provided in Appendix 3, along with the scoring template used by panel members for each paper.

Types of panels assessing agency quality of policy advice

Public service agencies use a variety of different variants of 'panels' – small teams of experienced policy practitioners – to score individual papers in each agency's sample of papers. Collectively, these scores determine each agency's overall quality of policy advice score (when the relevant performance metric is applied to the scores of papers in the sample – as outlined in Section 1 of the slide pack).

When agencies use the experienced policy practitioners of the New Zealand Institute of Economic Research (NZIER) to undertake their review of the quality of policy advice – as 15 out of 28 agencies did in 2019/20 – in the slide pack we refer to these agencies as having an 'external panel'. Other agencies use either a panel of senior policy staff from within the agency, or a mixture of policy staff from the agency and external experts. These are respectively referred to as having an 'internal panel' (four agencies in 2019/20) and 'mixed panel' (nine agencies in 2019/20).

4. Caveats about the data

Before reviewing the policy performance results for agencies reported in the slide pack, it's important to be aware of the caveats relating to the quantitative data presented there – for quality of policy advice scores and ministerial satisfaction scores.

Quality of policy advice scores

The first caveat concerns the size of the sample of policy papers that different agencies' policy quality review panels choose to assess to determine an overall agency score – this varies considerably between agencies. Factors contributing to sample size variations include differences in the volume of papers produced, in the extent to which sub-group results are sought, and in resourcing available. For any given agency sample, this means that it's possible that a different sample of the same size, or a larger sample, might yield somewhat different agency quality of advice results – particularly if the different or bigger sample contains more outlying papers (whether of exceptionally high or exceptionally low quality).

The second caveat is it's likely the mix of types of policy or policy-rated papers included in the sample assessed by a panel may vary between agencies – as current guidance allows each agency to make the call on the types of paper from which the sample to be assessed is drawn. Factors influencing this choice likely include the types of policy work each agency tends to do (e.g. strategic policy versus more operational policy), and/or the types of policy services that their minister demands (e.g. the demand for policy advice briefings and Cabinet papers versus the demand for event or meeting briefs, talking points and speech notes). The 2019/20 NZIER reports indicate that for the agencies for which they undertake quality of policy advice assessments, sometimes the simpler papers that don't provide much policy analysis or advice are scored relatively highly. That may advantage agencies that have a higher proportion of those type of papers in their samples, by boosting their overall agency scores.

A third caveat is that while all agencies' panels are now using the same PQF-based scoring scale, there will inevitably be some individual variation in how panel members interpret both the PQF and the scoring scale. Internal and external moderation is used to varying degrees by different agencies' panels to try to reduce such individual variation.

Fourthly, the time constraints under which policy advice papers are produced vary considerably from paper to paper – which may impact the ability to produce quality policy advice. The document *Using the Policy Quality Framework to assess papers: A guide on panels and processes* is clear that constraints to quality, such as time pressure, should not affect a paper's score (p 9):

> "When scoring advice papers, don't adjust scoring based on constraints to quality. You will want to note these, but they should not affect the score. Assessors may want to reflect on which constraints most affect the quality of papers overall and any recurring themes."

In principle this advice should limit agency variation in how time constraints are factored into scores. However anecdotal evidence indicates that there may be some variation between panels in whether that advice is heeded.

Finally, the policy quality scores included in the slide pack for an agency are one way of assessing an agency's quality of policy advice, but they don't show the full context of each agency's policy performance relative to that of others. For example, there may be substantial differences in the relative workloads of agencies, and/or the relative complexity of the subject matter they work on. Under those circumstances, achieving high quality of advice scores may be more difficult: their mean score might be around 3, yet an observer understanding their context might consider they are doing well under the circumstances.

Ministerial policy satisfaction methodologies

Our caveat about the ministerial policy satisfaction scores is that the scope for individual variation in ministerial scoring of agencies' policy performance is also considerable. The approach ministers take to scoring their agencies will inevitably vary according not only to their perceptions of agency

performance, but also to the personal style of the minister and the timing of their receipt of the questionnaire. There is no moderation process.

It should also be noted that some ministers choose to not complete the survey, preferring to provide oral feedback directly to senior agency staff. It is difficult to know whether the average score, for the public service as a whole, would have been lower or higher if the performance feedback of all Ministers had been able to be collected and included in the analysis.

5. Methodology – and findings provided in the slide pack

Quality of policy advice – quantitative metrics and qualitative thematic analysis

In the slide pack, we assess agencies' quantitative policy performance according to three metrics:

- Average score of the sample of papers (out of 5).
- Percentage of the sample scoring at least 3 ('acceptable').
- Percentage of the sample scoring at least 4 ('good').

This enables a more finely grained understanding of individual agency specific and policy system performance than if only one metric were used. We haven't yet analysed patterns in the targets set by agencies relative to one or more of those metrics, or the extent to which the targets set were achieved.

To complement the quantitative analysis of quality of policy advice performance, we have also undertaken a thematic analysis of the qualitative feedback from panel reports. This identifies the extent to which there are common areas for improvement identified in the papers scored.

The results of this quantitative and qualitative analysis are provided in the Section 1 of the slide pack, while Section 2 provides analysis of panel types scoring the papers, and the frequency with which they assess policy quality. Section 3 identifies key performance insights from the analysis in Sections 1 and 2, while Section 4 provides information on how agencies are using the Policy Quality Framework – including and beyond panel assessments of the quality of advice papers.

Ministerial policy satisfaction – qualitative and quantitative methodologies

Ministerial policy satisfaction is assessed by the survey instrument attached in appendix 4. The slide pack applies a single quantitative metric (to the scores derived from the survey responses) for assessing both agency specific and system perspectives on ministerial satisfaction with policy services provided to one or more ministers by an agency. That's the average ministerial satisfaction score out of 5, calculated across all 18 of the dimensions on which the 2019/20 Ministerial Policy Satisfaction Survey asks ministers to score each agency on.

At a later stage, we could undertake more detailed analysis of quantitative survey results – for example, examining agency and policy system differences between Part 1 scores (of the minister's perceptions of the agency's performance in meeting their overall needs as a customer) and Part 2 scores (specifically focused on the minister's assessment of the quality of policy advice the agency provided – under the broad headings of the Policy Quality Framework).

We have also undertaken qualitative thematic analysis of the written comments provided by ministers, which provides useful insights into the areas that ministers value.

Section 5 of the slide pack provides the results of both the quantitative and qualitative analysis of the matters that ministers value enough to comment on their existence or otherwise.

6. Next steps

The final two sections of the slide pack – Sections 6 and 7 – seek feedback from you on two matters. The first is feedback on what you see as priorities for further analysis of the 2019/20 policy performance data collected via the recent information request. The second is feedback on the implications of our analysis to date for the future focus of the Policy Project in supporting agencies to improve their policy performance.

7. Recommendations

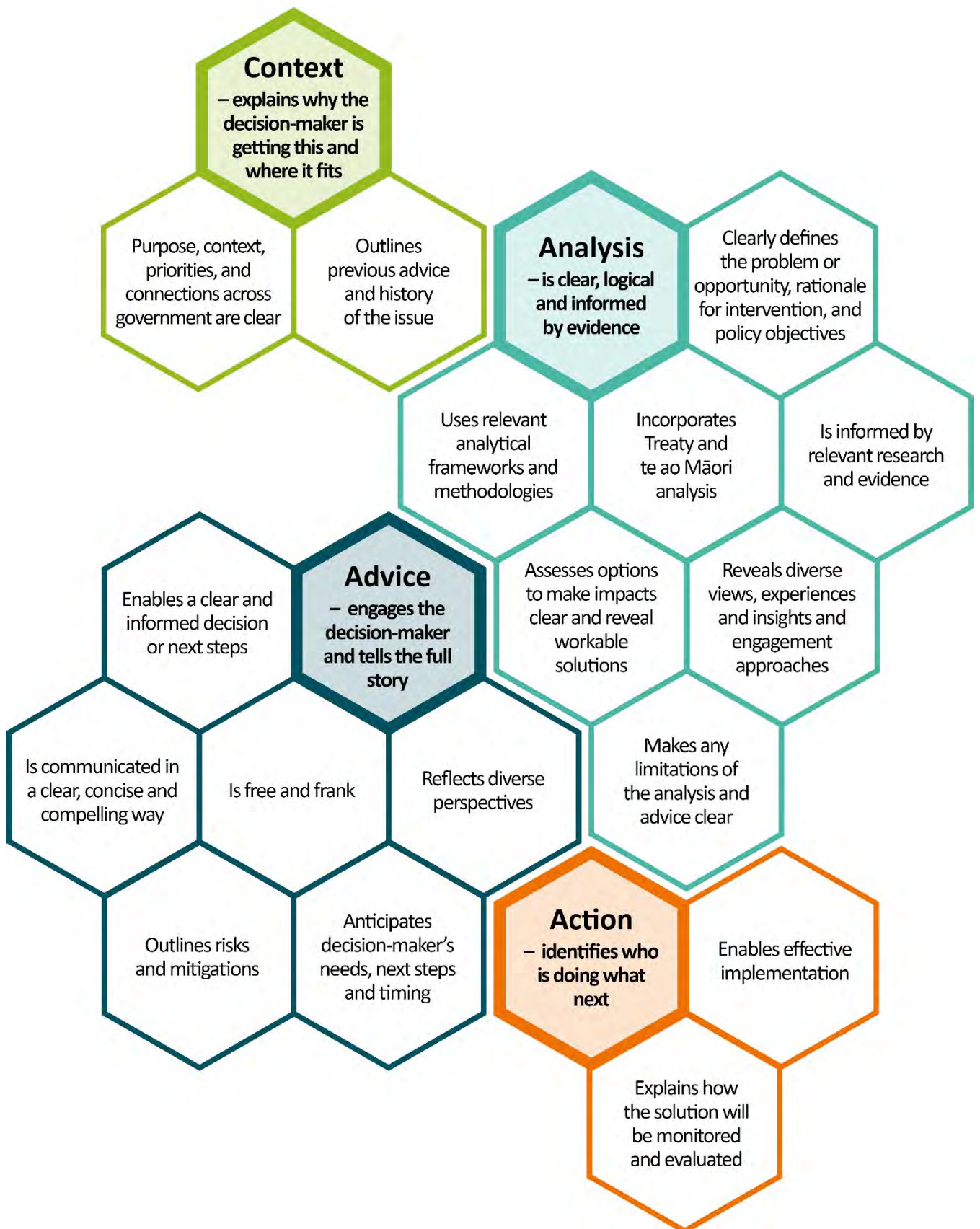We recommend that the Policy Profession Board:

1. **review** the attached slide pack, while keeping the contents of this paper in mind

2. **discuss** the matters identified for discussion in the slide pack, particularly in:

    a. Slide 6: Policy quality performance – next steps with metrics

    b. Slide 8: Performance insights – lifting the floor

    c. Slide 10: Performance insights – drawing connections across the PQF

3. **provide feedback** on which areas of further policy performance data analysis outlined in Slide 14 the Policy Project should prioritise

4. **provide feedback** on the possible future changes outlined in Slide 15 to the Policy Project focus in supporting agencies to improve policy performance (based on insights from the data analysis)

5. **note** that the results in the attached slide pack will be shared with the Tier 2 Policy Leaders Network at their 2 March 2021 meeting

6. **note** that these results may also be considered by those undertaking the interim evaluation of the Policy Project that is commencing shortly

7. **agree** that the Policy Project undertake a similar information request in 2021/22, so analysis of comparable data over time:

    a. enables policy system performance changes over time to be well understood, and

    b. informs Policy Project priorities and ways of working appropriately.

Appendix 1: Table of Public Service Agencies

Agency name	Abbreviation
Department of Conservation	DOC
Department of Corrections	Corrections
Department of Internal Affairs	DIA
Department of the Prime Minister and Cabinet	DPMC
Inland Revenue Department	IRD
Land Information New Zealand	LINZ
Ministry for Pacific Peoples	MPP
Ministry for Primary Industries	MPI
Ministry for the Environment	MfE
Ministry for Women	MfW
Ministry of Business, Innovation, and Employment	MBIE
Ministry of Culture and Heritage	MCH
Ministry of Defence	MOD
Ministry of Education	MOE
Ministry of Foreign Affairs and Trade	MFAT
Ministry of Health	MOH
Ministry of Housing and Urban Development	HUD
Ministry of Justice	MOJ
Ministry of Social Development	MSD
Ministry of Transport	MOT
New Zealand Customs Service	Customs
New Zealand Police	Police
Oranga Tamariki—Ministry for Children	OT
Statistics New Zealand	STATS
Te Arawhiti – Office for Māori Crown Relations	Te Arawhiti
Te Kawa Mataaho – PSC	PSC
Te Puni Kōkiri – Ministry of Māori Development	TPK
The Treasury	TSY

Appendix 2: One page summary version of the Policy Quality Framework
Quality Standards for written policy and other advice

Context – explains why the decision-maker is getting this and where it fits

Purpose, context, priorities, and connections across government are clear

Outlines previous advice and history of the issue

Analysis – is clear, logical and informed by evidence

Clearly defines the problem or opportunity, rationale for intervention, and policy objectives

Uses relevant analytical frameworks and methodologies

Incorporates Treaty and te ao Māori analysis

Is informed by relevant research and evidence

Advice – engages the decision-maker and tells the full story

Enables a clear and informed decision or next steps

Assesses options to make impacts clear and reveal workable solutions

Reveals diverse views, experiences and insights and engagement approaches

Is communicated in a clear, concise and compelling way

Is free and frank

Reflects diverse perspectives

Makes any limitations of the analysis and advice clear

Outlines risks and mitigations

Anticipates decision-maker's needs, next steps and timing

Action – identifies who is doing what next

Enables effective implementation

Explains how the solution will be monitored and evaluated

Appendix 3: The Policy Quality Framework scoring scale and template

The following table (from *Using the Policy Quality Framework to assess papers – a guide to panels and processes)* outlines the scoring scale to be used in awarding a paper a score out of 5.

Half points can be awarded where a paper falls between two points on the scale.

Score	Meaning	Description
1	Unacceptable	Does not meet the relevant quality standards in fundamental ways • Lacks basic information and analysis • Creates serious risk of poor decision-making • Should not have been signed out • Needed fundamental rework
2	Poor	Does not meet the relevant quality standards in material ways • Explains the basic issue but seriously lacking in several important areas • Creates risk of poor decision-making • Should not have been signed out • Needed substantial improvement in important areas
3	Acceptable	Meets the relevant quality standards overall, but with some shortfalls • Provides most of the analysis and information needed • Could be used for decision-making • Was sufficiently fit-for-purpose for sign-out • Could have been improved in several areas
4	Good	Meets all the relevant quality standards • Represents good practice • Provides a solid basis for decision-making • Could have been signed out with confidence • Minor changes would have added polish
5	Outstanding	Meets all the relevant quality standards and adds something extra • Represents exemplary practice • First-rate advice that provides a sound basis for confident decision-making • Could have been signed out with great confidence • A polished product

The following pages from the paper scoring template demonstrate how the Policy Quality Framework elements link to the above scale. Each panel member completes one template for each paper being assessed before a panel meeting, with the (pre-moderation) paper score being decided by consensus during the meeting.

Context – explains why the decision-maker is getting this and where it fits

The paper is clear about the:	Yes	No	N/A
• purpose	○	○	○
• context	○	○	○
• priorities	○	○	○
• connections across government.	○	○	○
The paper outlines previous advice and history of the issue.	○	○	○

Comments: What are the paper's strengths? How could it have been improved?

Analysis – is clear, logical and informed by evidence

The analysis clearly defines the:	Yes	No	N/A
• problem or opportunity	○	○	○
• rationale for intervention	○	○	○
• policy objectives.	○	○	○
The analysis uses relevant analytical frameworks and methodologies.	○	○	○
The analysis incorporates Treaty and te ao Māori analysis.	○	○	○
The analysis draws on relevant research and evidence.	○	○	○
The analysis assesses options to make impacts clear and reveal workable solutions.	○	○	○
The analysis is clear about any strengths and limitations.	○	○	○
The analysis reveals diverse views, experiences and insights, and engagement approaches.	○	○	○

Comments: What are the paper's strengths? How could it have been improved?

Advice – engages the decision-maker and tells the full story

	Yes	No	N/A
The advice enables a clear and informed decision or next steps.	○	○	○
The advice is communicated in a clear, concise and compelling way.	○	○	○
The advice is free and frank.	○	○	○
The advice reflects diverse sector perspectives.	○	○	○
The advice outlines risks and mitigations.	○	○	○
The advice anticipates the decision-maker's needs, next steps, and is timely.	○	○	○

Comments: What are the paper's strengths? How could it have been improved?

Action – identifies who is doing what next

	Yes	No	N/A
The actions enable effective implementation.	○	○	○
The actions explain how the policy solution will be monitored and evaluated.	○	○	○

Comments: What are the paper's strengths? How could it have been improved?

Overall panel rating for paper

Based on consideration of the ratings above and panel discussion of the strengths and weaknesses of the paper, the panel should collectively assign an overall score between 1 and 5.

/ 5

If this paper is an exemplar, retain it for future reference
and make it available for others as an example of best practice.

Appendix 4: The Ministerial Policy Satisfaction Survey

Ministerial Policy Satisfaction Survey

Purpose

This survey asks for your feedback on the policy advice, both written and oral, you have received in the last [ENTER NUMBER] months from [ENTER AGENCY OR BUSINESS UNIT NAME] on [ENTER NAME OF POLICY IF SEEKING FEEDBACK ON A SPECIFIC PROJECT. IF NOT, DELETE].

It will take about 5 to 10 minutes to complete.

How the results will be used

Your feedback will help [ENTER AGENCY OR BUSINESS UNIT NAME] better serve you in the future.

Your responses will also be used to calculate a ministerial satisfaction score to be included in the [ENTER AGENCY OR BUSINESS UNIT NAME]'s information for the Estimates and Annual Report.

This is part of [ENTER AGENCY OR BUSINESS UNIT NAME]'s commitment to be transparent and accountable for our policy performance.

What is quality policy advice?

The policy advice you receive should support you to make a fully informed decision, or take the next steps, by answering the following questions:

Context-setting: Why am I getting this and where does it fit?

Analysis: Is the analysis clear, logical and informed by evidence?

Advice: Does it engage me and tell the full story?

Action: Is it clear who is doing what next?

Thank you for your feedback

General Satisfaction

1 Thinking about your interactions with [ENTER AGENCY OR BUSINESS UNIT NAME] over the last [ENTER NUMBER] months on [ENTER NAME OF POLICY IF SEEKING FEEDBACK ON SPECIFIC PROJECT. IF NOT, DELETE], how often did each of the following occur?

	Never	Some of the time	About half the time	Most of the time	Always
I was engaged early enough in the policy process	◯	◯	◯	◯	◯
I was engaged in a way that reflects how I like to work	◯	◯	◯	◯	◯
My feedback was taken on board	◯	◯	◯	◯	◯
I was able to access relevant expertise on the issues	◯	◯	◯	◯	◯
I received advice that demonstrates an understanding of my priorities and context	◯	◯	◯	◯	◯
I received advice within the agreed timeframes	◯	◯	◯	◯	◯

2 Please add any comments or suggestions on how [ENTER AGENCY OR BUSINESS UNIT NAME] could improve the support provided to you as Minister, or examples of what has worked well that should be continued.

_____

_____

_____

_____

2

Quality of policy advice

3 Thinking about the policy advice you have received from [ENTER AGENCY OR BUSINESS UNIT NAME] over the last [ENTER NUMBER] months on [ENTER NAME OF POLICY IF SEEKING FEEDBACK ON SPECIFIC PROJECT. IF NOT, DELETE], how often did the advice:

	Never	Some of the time	About half the time	Most of the time	Always
Clearly explain the problem or opportunity	◯	◯	◯	◯	◯
Make relevant connections with other portfolios and address any issues that arise	◯	◯	◯	◯	◯
Clearly explain the rationale for intervention	◯	◯	◯	◯	◯
Demonstrate that appropriate stakeholder engagement strategies were used	◯	◯	◯	◯	◯
Demonstrate how relevant research, evidence, and insights informed the advice	◯	◯	◯	◯	◯
Provide the advice you need to hear and not only what you want to hear	◯	◯	◯	◯	◯
Enable a clear and informed decision to be made, or next steps to be taken	◯	◯	◯	◯	◯
Consider the longer term implications for New Zealand as applicable	◯	◯	◯	◯	◯
Communicate in a clear, concise and coherent manner	◯	◯	◯	◯	◯
Adequately consider how the policy will be implemented and will work in practice	◯	◯	◯	◯	◯
Explain how the policy will be monitored or evaluated to check that it works	◯	◯	◯	◯	◯

3

4   Please add any comments or suggestions on how [ENTER AGENCY OR BUSINESS UNIT NAME] could improve the quality of its policy advice, or examples of what has worked well that should be continued.

_____

_____

_____

_____

Overall performance

5   I have confidence in the policy advice provided by [ENTER AGENCY OR BUSINESS UNIT NAME]:

Never	Some of the time	About half the time	Most of the time	Always
◯	◯	◯	◯	◯
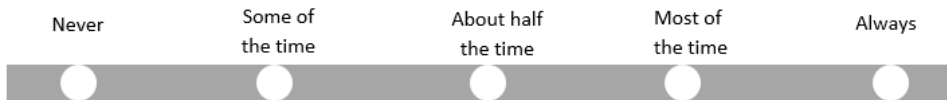
6   I have trust in the officials I engage with from [ENTER AGENCY OR BUSINESS UNIT NAME]:

Never	Some of the time	About half the time	Most of the time	Always
◯	◯	◯	◯	◯
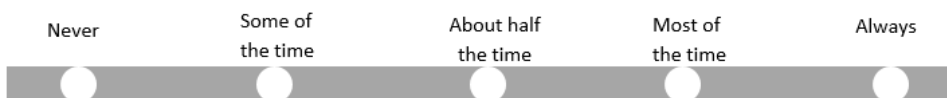
7   What is one thing that [ENTER AGENCY OR BUSINESS UNIT NAME] could do better?

_____

_____

_____

_____

Thank you for your feedback

4

Policy Performance 2019/20

Tier 2 Policy Leaders' Network
2 March 2021

Overview

Analysis of information request data – on the quality of policy advice and on ministerial satisfaction with policy services

In September 2020, the Policy Project asked all 28 public service agencies with a policy function to provide information on:

- a range of quality of policy advice indicators from policy quality review panels and ministers
- their use of the revised Policy Quality Framework (PQF) and how to improve it, and the Ministerial Policy Satisfaction Survey.

This is the third information request the Policy Project has undertaken since 2017 – and the first since all agencies adopted the PQF standardised scoring system for quality of advice assessment. In 2019/20 the revised PQF was also incorporated in the revised Ministerial Policy Satisfaction Survey. This year's results can therefore act as a baseline for assessing policy performance trends in subsequent years. They are also of interest in their own right.

The Policy Quality Framework

Objectives

➢ The objectives of the Policy Performance 2019/20 exercise are to analyse information received from public service agencies on policy quality assessments and ministerial satisfaction, to enable the Policy Project to:

- Better understand the performance of the policy system.

- Identify areas of common development needs – so the Policy Project can design and provide agencies with practical support to address these needs.

- Understand how to improve the Policy Quality Framework and Ministerial Policy Satisfaction Survey resources that the Policy Project has provided for the policy system.

- Identify where we most need to support policy practitioners in using the Policy Quality Framework (and related Policy Skills and Capability Frameworks).

2

What this presentation covers

This presentation provides a high-level overview of policy performance results and key insights from the data received through the 2019/20 and the previous information request. It encompasses:

1. **Policy quality performance**: comparative policy quality scores for all public service agencies in 2019/20 on three policy quality metrics, changes in mean scores since 2017/18, and next steps

2. **Panel types and frequency of quality of advice reviews:** information about the types and frequency of panel reviews of the quality of agencies' policy advice

3. **Performance insights:** implications of the quantitative policy quality metrics and thematic analysis of qualitative data from review reports for future Policy Project focus areas

4. **How agencies are using the PQF:** utilisation patterns and agency perceptions of usefulness

5. **Ministerial Policy Satisfaction Survey:** quantitative analysis of performance scores in responses to the surveys, and thematic analysis of ministers' comments about what matters to them

6. **Next steps for analysing the performance data**: seeking feedback on future data analysis priorities

7. **Implications for the Policy Project:** seeking feedback on what insights from this data analysis means for the Policy Project's future focus in supporting agencies to improve policy performance.

Caveats about the data

Agencies provided their own data – and there may be variations in how the PQF is applied

While the PQF provides a consistent framework, there are likely to be some variations in how it is used that need to be kept in mind when comparing different agencies. These are:

- variable sample sizes – raising questions about how well some sample results may represent quality of the whole population of papers produced for decision-makers in a year

- possible variations in how widely 'policy paper' is interpreted when deciding what should be in, or out, of the population of papers the sample is chosen from

- variations in the proportion of substantive policy advice papers different Ministers ask the policy teams in their agency to produce, compared with more procedural papers

- variations within and between panels, and over time in how they interpret and apply the PQF scoring scale

- variations in the degree of time pressure that papers had to be produced in, and the complexity of policy issues – which may vary over time, or affect some agencies differently to other agencies.

The Ministerial Policy Satisfaction Survey asked ministers for feedback on a range of measures, and also provided space for additional open-ended comments. The approach ministers take to scoring their agencies will inevitably vary not only according to their perceptions of the agency but also to the personal style of the minister and timing of the receipt of the questionnaire.

Methodologies used

➤ The quantitative methodology we have applied to the quality of policy advice scores data in the agency reports involves assessing each agency on three common metrics – mean score of papers, and percentage scoring at least 3 and at least 4 (see slide 5).

➤ We have not considered the different targets that individual agencies have selected. We did not attempt to adjust for different sample sizes or provide any moderation.

➤ The qualitative methodology we applied to the narrative parts of agency's quality report was based on a thematic analysis – considering which areas for development were most common.

➤ The qualitative methodology applied to the open-ended comments in the Ministerial Policy Satisfaction Survey involved grouping different comments according to common themes to identify aspects that ministers particularly value.

➤ The Policy Project is interested in drawing insights from these results for the policy system as a whole – individual agencies may draw different insights for improving their own performance based on the detail of their policy quality reviews and ministerial feedback.

1a. Policy quality performance – assessed by review panels

Mean quality score

Agency	Score
Customs	4.10
MFAT	4.00
Police	3.97
DIA	3.84
MfW	3.83
MPI	3.79
Te Arawhiti	3.77
MOT	3.76
MSD	3.73
Corrections	3.70
HUD	3.70
OT	3.64
MOH	3.63
MPP	3.60
IRD	3.54
DPMC	3.50
PSC	3.50
MfE	3.46
TPK	3.43
MOD	3.42
STATS	3.35
DOC	3.33
MOE	3.31
TSY	3.24
MBIE	3.20
MOJ	3.13
LINZ	3.06
MCH	2.80

Percent of papers over quality thresholds

- Papers scoring over 3
- Papers scoring over 4

Notes on the data

- All 28 agencies provided data on their mean policy quality score for a sample of papers assessed using the PQF.

- Papers scoring 1 or 2 in the PQF scale are assessed as below minimum standards. 3 is acceptable, 4 good, and 5 outstanding.

- The mean score ranges from 2.8 to 4.1. The public service mean of all agency mean scores is 3.55, with a standard deviation of 0.297.

- The clear majority of agencies (20 of 28) are within one standard deviation of the mean, with three above and five below. That is, there are a large group of agencies clustered around the mean, with few outliers.

5

1b. Policy quality performance – tracking over time

28 agencies provided data on their mean policy quality in the 2020 year. Of these, 21 had previously provided us with data for 2017/18 that could be compared with their scores using the revised PQF for 2019/20. The public service-wide mean score in 2019/20 was 3.55.

Analysis of the differences in mean quality scores between 2017/18 and 2019/20 for the 22 agencies for which reasonably comparable data was achievable reveals the following trends. The overwhelming majority of agencies – 16 out of 21 (76%) – showed no significant shift in performance between 2017/18 and 2019/20 (which we have defined as having a difference between mean scores for these years of less than half a point). Three agencies had a drop in performance of more than half a point (LINZ, MCH, MOJ). Two agencies improved by more than half a point (MPP and the Treasury).

Shifts in quality of policy advice performance 2018/2020

Category	Value
Comparison data unavailable	7
Significant drop in performance	3
No significant performance change	16
Significant improvement in performance	2

Notes on the data

➢ 2017/18 mean scores for 21 agencies were converted to revised PQF scores using a methodology developed by NZIER in 2020 and reviewed by the Policy Project.

➢ We haven't converted scores for agencies that had their panel use bespoke scoring systems in 2018, as this wouldn't be sufficiently accurate for a reliable comparison (MfE, MBIE, MoE, IRD Oranga Tamariki)

➢ In addition, the two public service departments that didn't exist in 2017/18 – HUD and Te Arawhiti – couldn't be included in the comparison.

6

1c. Policy quality performance – next steps with metrics

- ➢ The first Policy quality performance slide assessed systemic quality of policy advice performance using three performance measures. Each measure provides useful information for assessing a different aspect of performance – and any one of them alone will not give a full picture. Other metrics could also be used (e.g. the percentage of papers scoring less than 3).

- ➢ The analysis revealed that an agency's performance on one metric is a relatively poor predictor of performance on another. For example, two agencies with similar mean scores may have a very different spread of paper scores, with one agency having several low scoring papers and several high scoring papers, while another agency has most papers scoring near the middle of the range.

- ➢ Under current Treasury reporting requirements, agencies must use the refreshed Policy Quality Framework from 2019/20 to assess the quality of their policy advice papers, and must report on this in Annual Reports – and select their own policy quality measures and targets.

- ➢ Current guidance from the Policy Project on performance metrics in Annual Reports recommends that agencies report on both:
  - the mean score of their assessed sample, and
  - the percentage of papers that meet a performance target set by the agency (for example, 50% of papers scoring 4 or over).

For discussion

Should guidance be revised to recommend a more consistent approach to reporting against benchmarks?

- ➢ Should Policy Project guidance enable a fuller picture and easier comparison between agencies in future, by recommending that all agencies:
  - a. report on:
    - the mean score
    - the percentage of papers scoring over 3 ('adequate')
    - the percentage scoring over 4 ('good')
  - b. while also emphasising that each agency still sets its own targets for each measure?

- ➢ Note: policy performance reviews for future years could examine whether trends start to emerge between different performance metrics over time.
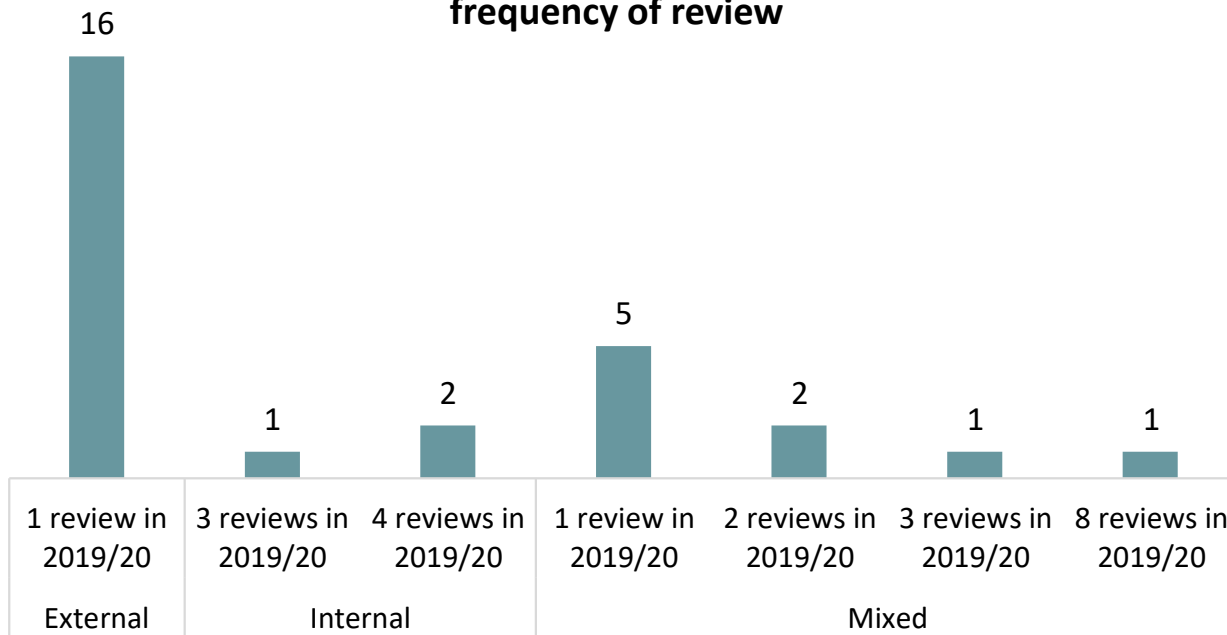
Most agencies use NZIER and do one annual review

Agencies are free to choose who does their policy quality assessment and how frequently. In 2019/20, 16 out of 28 agencies (57%) used NZIER to undertake policy quality assessments in a single annual review. We refer to this type of policy quality assessment as undertaken by an 'external' panel.

Nine agencies used panels composed of staff from the agency and other external expert members. We refer to this type of review as being by a 'mixed' panel. The remaining four agencies used panels drawing only from internal staff – referred to as 'internal'. Two of the internal panels also reported to us that they used external moderators.

Number of agencies using each type and frequency of review

| | 16 | | | 1 | 2 | | 5 | 2 | 1 | 1 |

1 review in 2019/20 — **External**
3 reviews in 2019/20 — 4 reviews in 2019/20 — **Internal**
1 review in 2019/20 — 2 reviews in 2019/20 — 3 reviews in 2019/20 — 8 reviews in 2019/20 — **Mixed**

Commentary and next steps

➢ Internal and mixed panels can potentially take a more strategic and nuanced view than external panels (e.g. considering the overall PQF score across a set of related papers, or average scores by subject matter area or group). Using an internal or mixed panel also helps departments grow internal capability in quality assessment of policy advice papers, with internal panel members becoming champions for quality policy advice in their teams.

➢ However, external panels and external members on panels (whether NZIER or otherwise) can provide a new perspective and draw on insights from reviewing a wide range of papers across government. Both have value.

➢ Agencies that are only doing an internal review may wish to consider using an external review at periodic intervals as a benchmarking exercise, and/or consider how to use external moderation (subject to capacity and demand, the Policy Project could have a role in this – having already undertaken this once for MoJ).

➢ Agencies that are only doing an external review may wish to consider how to develop internal quality assurance processes that draw on the PQF and provide for quicker feedback than can be provided though a single, external, yearly review.

➢ The Policy Project can provide cross-government training based on our *A guide on panels and processes for assessing policy advice papers*, and tailored support for agencies wanting to establish internal review processes.

3a. Performance insights – lifting the floor

Focus Area 1

Lifting the floor – addressing the lowest performing papers

Key points

- **Low performance is relatively common**: All agencies reported on the minimum quality score in their sample of papers. 22 of the 28 agencies (79%) had some papers scoring below a 3 (the standard for 'acceptable' under the PQF). Of these, 8 had a minimum score of 2.5, 13 had a minimum score of 2, and one had a minimum score of 1.5.

- **Papers scoring under 3 should not be signed out:** The rating scale of the PQF builds in the concept that papers scoring below 3 should not have been signed out. Under the PQF, a paper scoring 1 "creates serious risk of poor decision making", and needs "fundamental rework". A paper scoring a 2 "creates a risk of poor decision making" and needs "substantial improvement in important areas".

- **Low scoring papers create risks:** Low-scoring papers can have a particularly detrimental effect on a minister's confidence in an agency's policy capability. If a poor quality paper is the first in a series of papers on a related topic, there is a risk of poor decision making that has flow-on effects for a whole stream of advice.

Possible next steps (for discussion):

➢ The Policy Project: could include discussion of the incidence of low scoring papers and how to reduce this through more agency use of ex-ante quality assurance and other means on the agenda at the forthcoming Policy Quality Review Panel Forum.

➢ Agencies with low-scoring papers: could undertake further work on the contexts in which they are signed out, to identify and then plan how to address underlying factors. This could include identifying the types of papers that are most likely to be low scoring, and whether there are agency-specific or system-wide improvements that could lift the performance of these papers (for example, improved peer review processes).

➢ The Head of the Policy Profession: could ask Tier 2 Policy Leaders to reiterate the importance of policy staff using the PQF checklist before papers are signed out. This could be done when they're considering the system-wide report on Policy Performance 2019/20 at their 27 April meeting.

3b. Performance insights – policy quality improvement areas

Treemap of policy quality improvement areas:
- Armchair ride — 9
- Assess options to make impacts clear and reveal workable solutions — 9
- Proofreading/formatting — 8
- Make intervention logic clear — 6
- Incorporate Treaty and Te Ao Māori analysis — 5
- Is communicated in a clear, concise and compelling way — 5
- Improve QA systems and processes — 4
- Outline risks and mitigations — 4
- Use relevant analytical frameworks and methodologies — 4
- Anticipate decision maker's needs, next steps and timing — 2

Notes on the data

➢ Thematic analysis of the narrative sections of all 28 agencies 2019/20 policy quality reports revealed the common areas for development. Narrative statements have been thematically linked and summarised as consistently as possible. The larger the box, the more agencies identified that area needed development.

➢ The majority of areas for improvement are from the 'analysis' section of the PQF, followed by the 'advice' section.

➢ The need to give ministers more of an 'Armchair ride' was regularly identified by NZIER in the 16/28 agency quality review reports they wrote. This means providing a minister with all the support the minister may need. The PQF captures three different dimension of this:
   • Purpose, context, priorities and connections across government are clear.
   • Enables a clear and informed decision on next steps.
   • Anticipates decision maker's needs, next steps, and timing.

➢ The theme of 'Make intervention logic clear' arising mostly in NZIER reports, means developing a coherent narrative across the problem definition, options analysis, and recommended action.

Focus Area 2

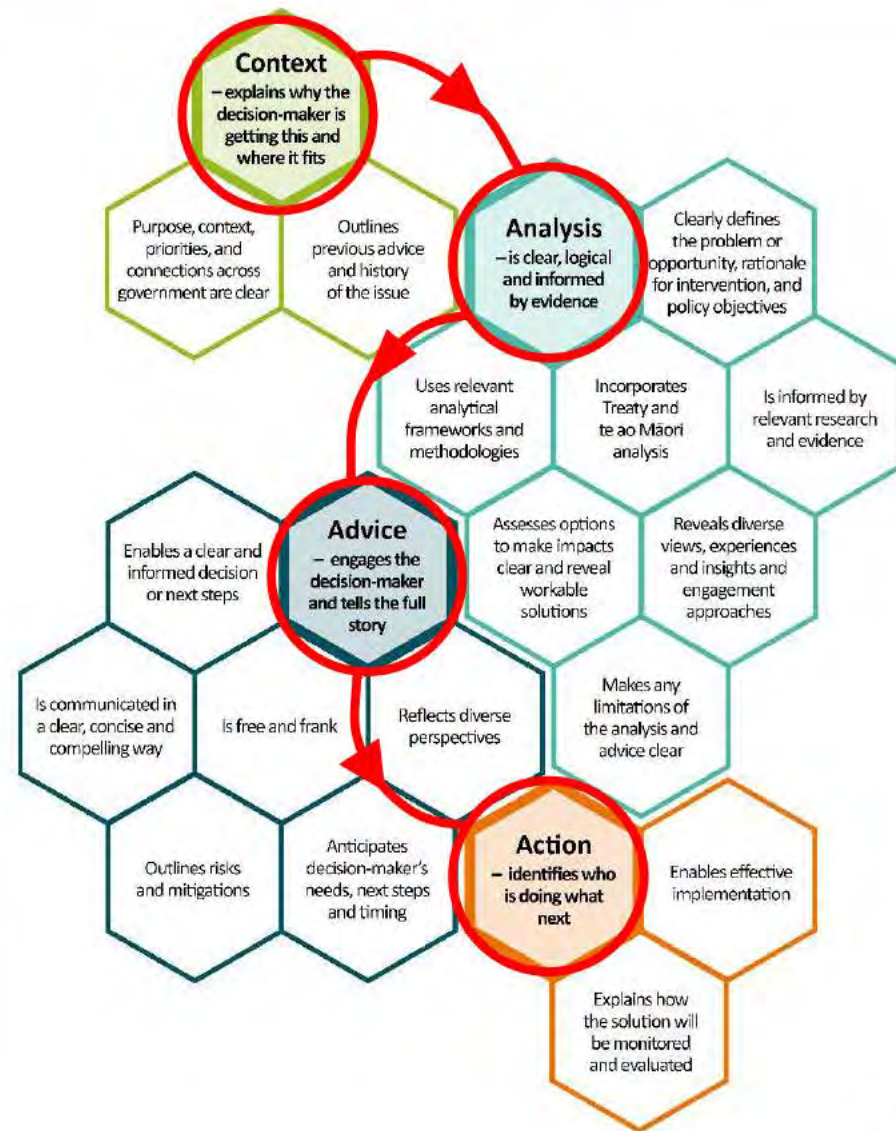### The whole is greater than the sum of the parts: drawing connections across the PQF

Key points

- Review panels consistently identified that the best papers create a logical chain from the context, through the analysis, to advice and action. This goes to the heart of creating robust policy advice.

- Some internal panels scored papers against each element of the PQF separately, and analysis of these assessments shows that a paper with a weak problem definition is likely to remain weak throughout. Similarly, if the analysis is poor, the advice will not score well.

- While each element of the PQF is an essential component for good policy advice, it can also be understood as a sequential building of robust advice from the outset.

Possible next steps (for discussion):

The Policy Project could:

- identify and disseminate annotated examples of best practice, that demonstrate how elements of the advice chain link together

- encourage agencies to spot issues in the problem definition and options analysis internally, ideally before final decision papers are written up

- include discussion of the policy quality chain of logic in training materials for new graduates and other policy professionals.

Context
– explains why the decision-maker is getting this and where it fits

Purpose, context, priorities, and connections across government are clear

Outlines previous advice and history of the issue

Analysis
– is clear, logical and informed by evidence

Clearly defines the problem or opportunity, rationale for intervention, and policy objectives

Uses relevant analytical frameworks and methodologies

Incorporates Treaty and te ao Māori analysis

Is informed by relevant research and evidence

Advice
– engages the decision-maker and tells the full story

Enables a clear and informed decision or next steps

Assesses options to make impacts clear and reveal workable solutions

Reveals diverse views, experiences and insights and engagement approaches

Is communicated in a clear, concise and compelling way

Is free and frank

Reflects diverse perspectives

Makes any limitations of the analysis and advice clear

Outlines risks and mitigations

Anticipates decision-maker's needs, next steps and timing

Action
– identifies who is doing what next

Enables effective implementation

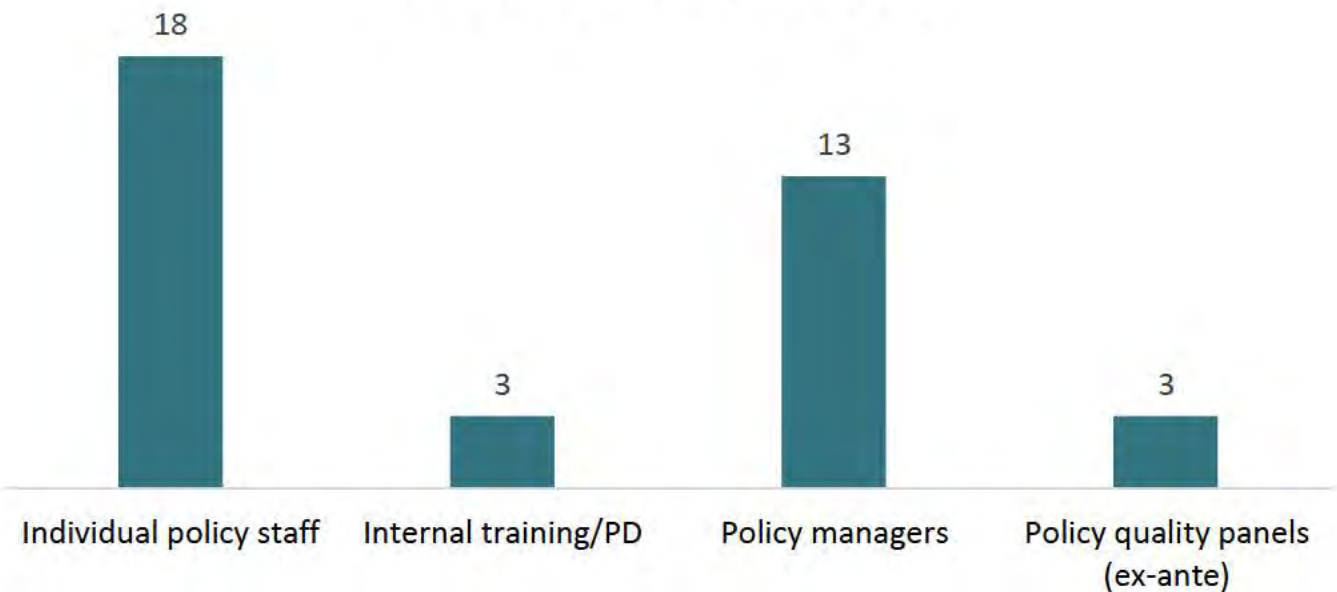Explains how the solution will be monitored and evaluated

4. Using the Policy Quality Framework

Agencies are using the PQF and rate it highly

As a minimum, all agencies are required to use the PQF to conduct ex-post reviews of the quality of policy advice for Estimates reporting – and all 28 reported that their panel does. Other ways that agencies can use it to support them in developing quality papers include: when developing them, when peer reviewing them, when assessing them before signoff to the minister (ex-ante), etc. 27 of 28 agencies are also using the PQF for some other purposes. Individual policy staff are using the PQF to develop and review papers in 18 of 28 agencies (64%), and managers are using the PQF in assessing papers in 13 out of 28 agencies (46%).

How are agencies using the PQF?

PQF Assessment

Question	Responses (out of 28)
PQF was easy to use	
Strongly agree	3
Agree	19
Neither agree or disagree	3
Disagree	1
No response	2
PQF includes all standards agency needs	
Strongly agree	2
Agree	21
Neither agree or disagree	3
Disagree	1
Strongly disagree	1
PQF could be applied to a range of papers	
Strongly agree	4
Agree	14
Neither agree or disagree	7
Disagree	3

Agencies were also asked how helpful they find the resource *Using the Policy Quality Framework to assess papers: A guide on panels and processes.* 79% of agencies rated it positively, with 15 out of 28 saying it was either "extremely helpful" or "very helpful".
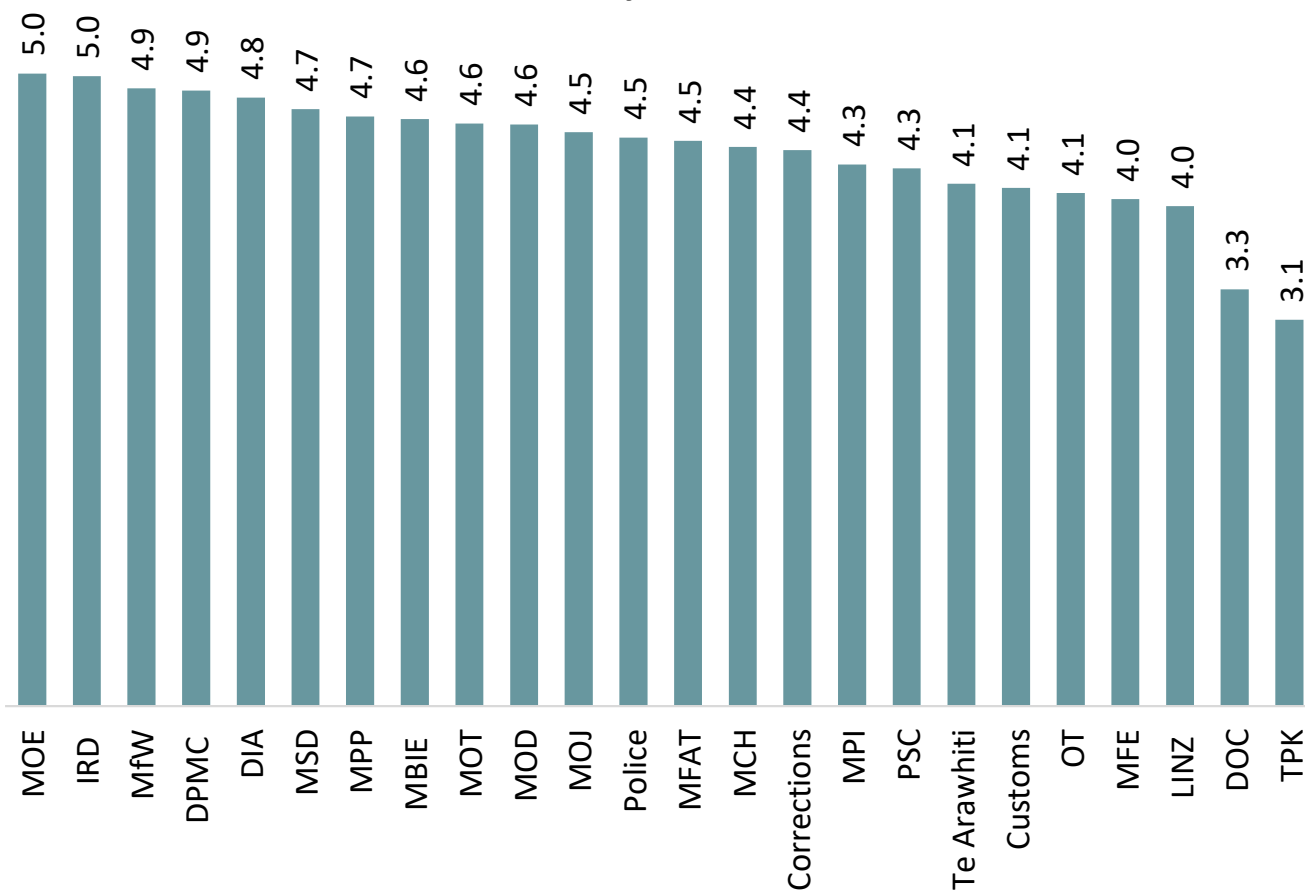
12

5a. Ministerial Policy Satisfaction scores

Average results, by agency

The Ministerial Policy Satisfaction Survey scores are reported in Annual Reports. They are calculated by averaging the scores each minister gave to subsets of 18 different aspects of their agency's performance in delivering policy services. The public service as a whole demonstrated high levels of overall ministerial satisfaction in 2019/20 – with a mean of 4.4, and a range of 3.1 to 5, out of 5.

Ministerial Policy Satisfaction Scores

Agency	Score
MOE	5.0
IRD	5.0
MfW	4.9
DPMC	4.9
DIA	4.8
MSD	4.7
MPP	4.7
MBIE	4.6
MOT	4.6
MOD	4.6
MOJ	4.5
Police	4.5
MFAT	4.5
MCH	4.4
Corrections	4.4
MPI	4.3
PSC	4.3
Te Arawhiti	4.1
Customs	4.1
OT	4.1
MFE	4.0
LINZ	4.0
DOC	3.3
TPK	3.1

Notes on the data

➢ Ministers can elect not to complete a survey. We received 54 surveys across 24 agencies (out of 28 agencies). For agencies with more than one minister completing the survey, we averaged the scores of their ministers.

➢ Three agencies asked ministers additional questions relevant to their work programme – the scores for those questions were not included in the averages reported opposite.

➢ We compared how agencies reported the results of ministerial surveys and panel's policy quality assessments in their Annual Reports. This varies widely in both cases.

➢ No clear relationship is evident between these ministerial survey scores and the ex-post policy quality scores of a sample of policy papers in slide 3. Agencies whose review panels assessed their policy advice papers as scoring highly or poorly on policy quality were not consistently scored highly or poorly by their minister(s).

➢ As both are measuring somewhat different but related aspects of performance, it's important that agencies focus on what they can learn from the quantitative and qualitative feedback both include.

➢ The written comments from ministers contain valuable insights on areas where agencies can improve policy performance (see next slide).

5b. Ministerial Policy Satisfaction – thematic comments analysis

Maintain open communication on any new issues, agency priorities, and timeframes
9

Discuss policy early and explicitly with minister to capture minister's priorities.
7

Focus on prompt delivery of outcomes – remembering policy is a step towards operational or legislative changes and ministers have timeframes in mind for achieving those changes
5

Engage with other agencies to improve whole of government delivery on priorities
4

Early communication to minister of any potential stakeholder issues
3

Maintaining agency capacity to respond quickly to new and emerging issues
3

Concise, clear writing that identifies the key issues
7

Recognising Treaty of Waitangi responsibilities and engaging with Māori
4

Whole of project thinking – planning in advance for smooth handover from policy to operations, effective implementation, and public outreach/education/communication (where relevant) at the conclusion of the policy phase
3

Cultural competency and diversity
2

Ensure advice is free and frank, and allow for full engagement with minister on difficult or contentious issues
2

Identify the appropriate level of detail needed – manage agency's workload and minister's time well
2

Timely response to requests from private secretaries and ministerial advisors
2

Notes on the data

➢ Thematic analysis of the open-ended survey responses has identified areas ministers value – whether they were mentioned as a strength or an area for improvement. Narrative statements have been summarised as consistently as possible, using terminology in the Policy Quality Framework and Policy Capability Framework.

Insights for improvement

➢ The ministerial survey provides particularly useful detail for policy advisers on how to give better effect to the quality indicator of "anticipates decision-makers needs, next steps, and timings". The Policy Project could consider drawing on this to create additional resources on this aspect.

➢ It's clear that ministers see agencies as full service providers and expect close working relationships between policy and operational service delivery, and engagement across government.

➢ There's close alignment between matters ministers value, and matters identified in the Policy Quality and Policy Capability frameworks.

6. Further analysis of 2019/20 policy performance data?

> **Which options below (or what else) do you consider top priority next steps, and why?**

- ➢ **Possible areas for further data analysis by the Policy Project in the coming months**:

    a) Detailed analysis of Ministerial Policy Satisfaction Survey scores, focused on how ministers score agencies on PQF elements – and how those scores compare with agency's panel scores.

    b) Triangulating the results with Cabinet Paper qualitative assessments by the CE and DCE (Policy), DPMC.

    c) Analysing patterns in agency targets for policy quality scores, and whether they are achieved.

    d) Assessing the impacts of COVID-19 on quality score patterns.

    e) Identifying aspects of the PQF that agencies consider need changing, including further analysis of open-ended responses from agencies on the PQF.

    f) Analysing agency feedback on Ministerial Policy Satisfaction Survey aspects needing change.

    g) Detailed analysis of different uses of internal or mixed panels and external moderation.

    h) Considering the implications of any of the above for whether refinements are needed to the *Policy Quality Framework: A guide on panels and processes for assessing policy advice papers.*

    i) Analysis of agency differences in policy pay scales and remuneration policy (from information request) – and in actual remuneration (drawn from PSC's Human Resource Capability Survey.

7. How can the Policy Project help improve performance?

Discussion – key questions and future directions

➤ Discussion of the implications of this analysis of policy performance in 2019/20 for the Policy Project's 2020/21 work programme is necessarily preliminary – as that also needs to be informed by the results of the interim evaluation. The evaluation will help us assess whether we are collecting the right data about policy performance in the right way, and what to change.

➤ **We would appreciate your feedback on the following questions**:

a) Do you agree that in supporting agencies to improve the quality of policy advice, the Policy Project should increase its focus on encouraging them to:

- lift the floor – identify and address factors that result in low scoring papers?

- ensure their policy advisers and managers understand the importance of drawing connections across the PQF and that the whole is greater than the sum of the parts?

b) What further steps could the Policy Project take to support agencies to enhance their quality assurance processes:

- to embed ongoing improvement into their organisation's approach to producing policy advice and

- focus on processes to identify and respond to gaps in performance as these emerge?