
Date 17 October 2018

THE MARSDEN FUND COMBINED PANEL TRIAL

A report delivered to the Marden Fund Council

Dr Jason Gush, Programme Manager Insights & Evaluation

Royal Society Te Apārangi

EXPLORE | DISCOVER | SHARE



Contents

Background	4
Agreed measures of the CPT	5
Expert judgements by panel and assessors.	5
Panel and Assessor EOI grading compared with Referee grading of Full Proposals.	5
Workload impact on Assessors and Panellists	7
Perceptions of the process by Applicants and Panellists.	8
Summary	11
Recommendations	11

Figures

Figure 1: Proportion of assessments made by panellists (i.e., both CPT and non-CPT panels) and assessors at each levels of expertise.	5
Figure 2: Relationship between EHB, SOC, and HUM panellists' post-meeting grades and the mean referee estimate for proposals at each level of self-declared expertise.	6
Figure 3: Relationship between EHB, SOC, and HUM panellists' post-meeting grades and the mean referee estimate for proposals at each level of self-declared expertise.	6
Figure 4: Distribution of mean time taken to assess an EOI by different types of reviewer.	7
Figure 5: Survey responses indicating agreement that the EOI assessment burden is too high.	8
Figure 6: Survey responses expressing level of agreement with the statement: "The Combined Panel Trial process is less likely to select excellent proposals than the Panel process".	8
Figure 7: Survey responses indicating broad belief that assessment would be/was of personal benefit to those assessing.	9
Figure 8: Survey responses expressing level of agreement with the statement: "I am comfortable that my EOI was seen by up to 10 of my NZ-based colleagues".	10

Background

The Marsden Fund's Review¹ placed an expectation that the Council would explore alternatives to the traditional Panel process in the belief that:

1. the Panel process would become unsustainable with the number of EOIs received by the Fund predicted to grow; and,
2. the Panel process may have difficulty in fairly assessing proposals that fall into disciplinary gaps between panels, particularly inter-disciplinary proposals.

In order to both meet this expectation and have an alternative agreed in time for the publication of the 2018 guidelines, the Council developed a process to address these perceived problems that could be operated in the limited timeframe available. This process became the Combined Panel Trial (CPT). Three panels were agreed for the pilot CPT: Economics and Human Behaviour (EHB); Humanities (HUM); and, Social Sciences (SOC). In order to address the disciplinary problems, the Fund would seek expert reviewers to assess proposals. As the only ready pool of individuals with known expertise available to the Fund would be its applicants, the Trial would borrow from an NSF experiment (the Mechanism Design merit review pilot²) and use applicants as assessors. This creates a process that would scale with the number of applications.

In its August meeting the Council agreed to assess the CPT in four key areas:

- What proportion of EOIs were assessed by one or more experts, and what proportion of assessments were made at differing levels of expertise?
- Is expertise a predictor of how well an assessor/panellist is able to assess an EOI? Specifically, are the EOI judgements of self-declared experts closer to the referee's quality estimate than non-experts?
- How much effort is required in conducting these assessments?
- What is the community perception of each assessment strategy?

¹ <https://www.mbie.govt.nz/assets/9c551e0eb9/marsden-fund-assessment.pdf>

² Report to the NSB on the NSF's Merit Review Process FY2015
<https://nsf.gov/nsb/publications/2016/nsb201611.pdf>

Agreed measures of the CPT

Expert judgements by panellists and assessors.

As reported in August, the CPT assessment process resulted in a much higher proportion of EOI's reviewed by self-declared experts than the traditional Panel process (82% cf. 65%), and by assessors with a higher confidence in their expertise than expressed by panellists (figure 1).

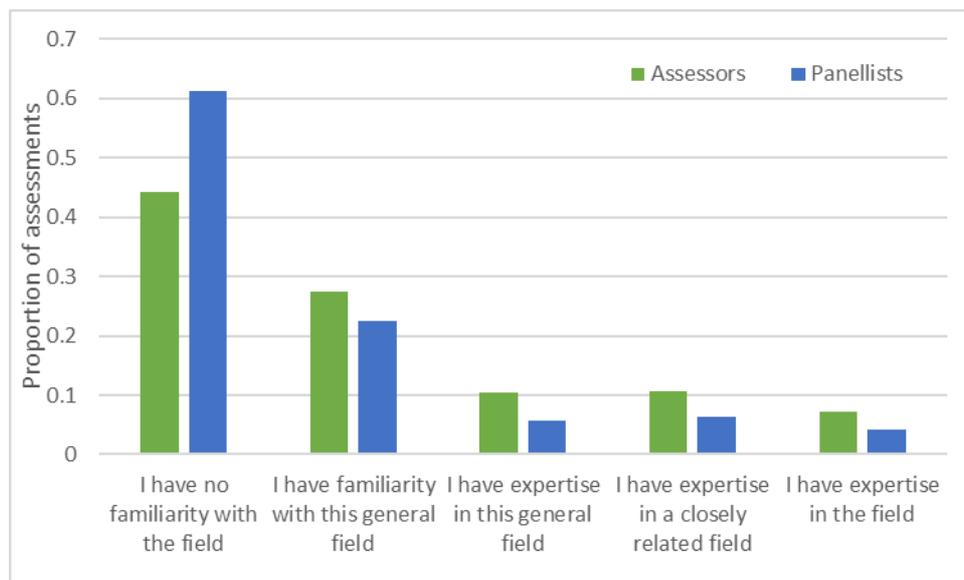


Figure 1: Proportion of assessments made by panellists (i.e., both CPT and non-CPT panels) and assessors at each levels of expertise.

Panel and Assessor EOI grading compared with Referee grading of Full Proposals.

A key belief behind the design of the trial was that experts will provide better judgements of a proposal than non-experts. If we accept that referee judgements are an unbiased estimate of proposal quality AND that full proposal quality can be properly anticipated by a review of the EOI; the hypothesis is that expert grades should have a stronger positive correlation with the mean referee grade than those with least familiarity. To a limited extent this hypothesis is supported, but the evidence is by no means strong (figures 2 and 3)

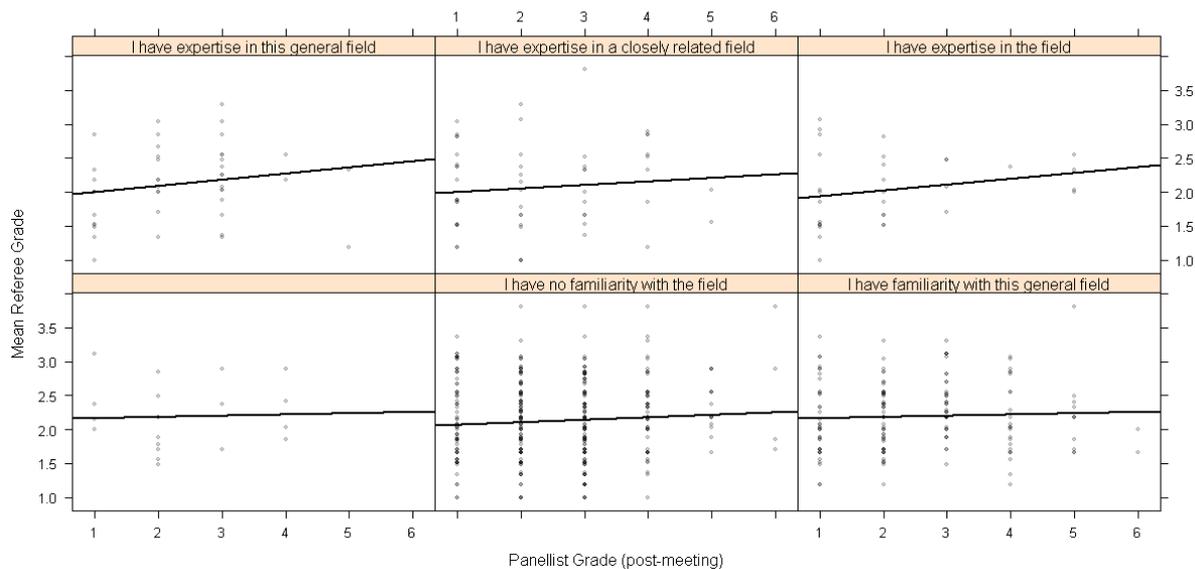


Figure 2: Relationship between EHB, SOC, and HUM panellists' post-meeting grades and the mean referee estimate for proposals at each level of self-declared expertise.

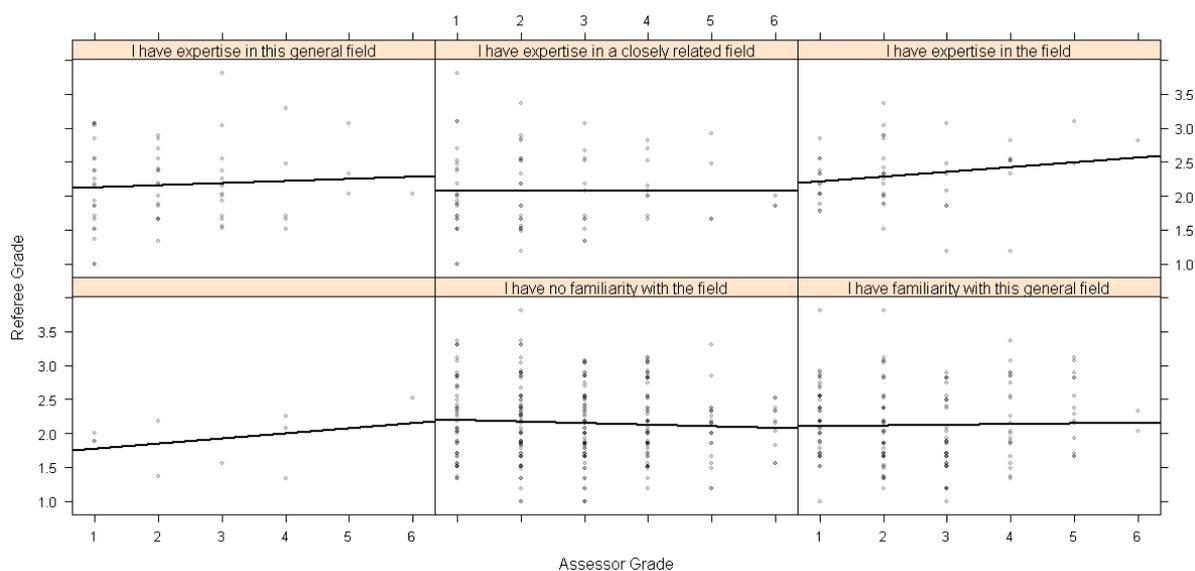


Figure 3: Relationship between EHB, SOC, and HUM panellists' post-meeting grades and the mean referee estimate for proposals at each level of self-declared expertise.

This lack of a robust and increasing relationship between assessment precision with increasing expertise is something of a disappointment; however there are at least three reasons why this expected relationship may not have been revealed:

- self-assessed expertise is an unreliable predictor of actual expertise;
- it is intrinsically difficult to anticipate a full proposal's quality from the parent EOI; and,
- the CPT process inadvertently prevented experts from fully expressing their expertise.

With regard to the last point, post-publication of the Funding guidelines, the Executive anticipated that asking assessors to fit their 26-27 EOI to the same distribution as was expected of panellists,

could lead to problems. This requirement was also the subject of some criticism in feedback provided by assessors. As this is the easiest and, in my opinion, most plausible explanation for the failure to demonstrate the expected relationship, the need for this requirement should be reviewed in any further assessment trials for the Fund.

Workload impact on Assessors and Panellists

As part of an attempt to measure the resource costs of the assessment process, the Executive undertook a survey of the time taken for panellists and assessors to review their EOIs. Response rates were high with 72% of CPT panellists, 75% of assessors and 86% of non-CPT panellists answering their survey. The mean time to assess an EOI was: 20.8 minutes for non-CPT panellists; 33.4 minutes for CPT panellists, and assessors falling in between at 28.8 minutes (figure 4).

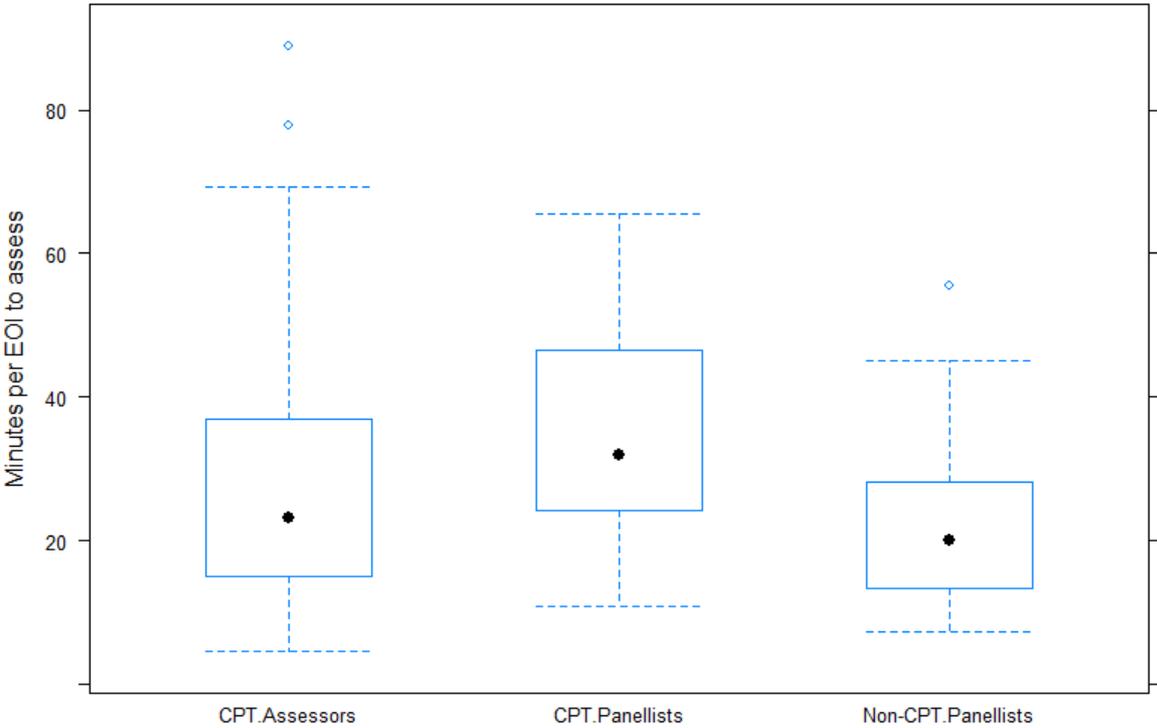


Figure 4: Distribution of mean time taken to assess an EOI by different types of reviewer.

In addition to requesting an estimate of the time taken, participants in the trial (i.e., all NZ-based standard investigators on applications to the HUM, EHB and SOC panels, and their panellists) were asked to give an estimate of the workload entailed in assessing the EOI round. The response rates for eligible uninvited applicants was still respectable at 43%, while those who had been invited to assess but not done so was rather lower at 26%. Given their differing circumstances this question was tailored to suit each group, e.g., panellists were asked for their agreement with the statement: “The workload required in assessing my Panel's EOIs was too high”, whereas eligible applicants in CPT panel proposals that had not been invited to assess were offered “The level of workload expected by the Combined Panel Trial (i.e. reviewing ~30 EOIs) would have been too high”.

Notably, those who participated in assessment tended to have a better opinion of the workload required; however, all groups leaned towards the sentiment that their assessment burden was/would be too high for the individual (figure 5).

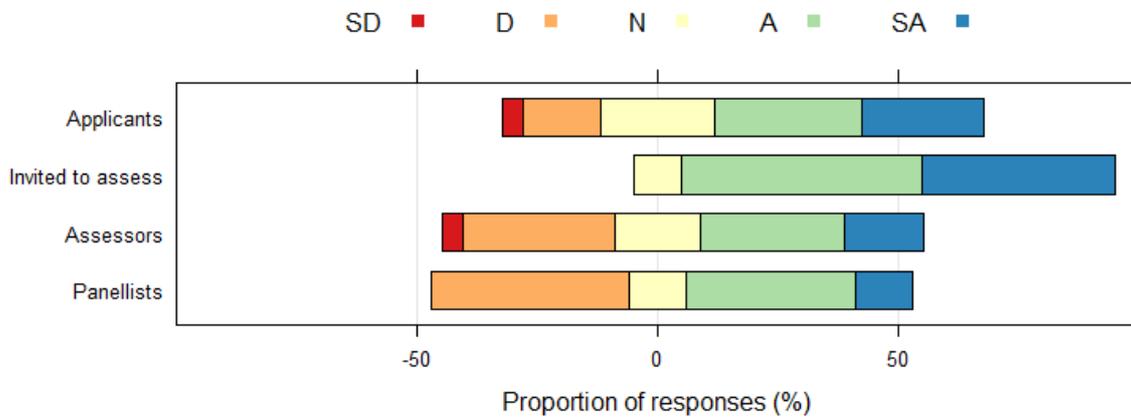


Figure 5: Survey responses indicating agreement that the EOI assessment burden is too high. Legend codes SD “Strongly Disagree”, D “Disagree”, N “No opinion”, A “Agree”, and SA “Strongly Agree”.

Perceptions of the process by Applicants and Panellists.

In terms of perception of the CPT process, the four participating groups were asked to give their opinions on three questions: could the CPT process provide judgements that were as good, or better, than the traditional panel process; is the process of assessing others’ EOIs likely to be of value to those assessing; and for applicants, whether they were comfortable with their EOI being seen by other applicants. The responses to each question are shown in the sections that follow.

Responses indicate weak preference for CPT process assessment amongst applicants, and weak preference for panel assessment amongst panellists.

All applicant groups tended towards disagreement that the CPT process would be inferior to the panel process; in contrast panellists tended to agree that the panel process would be superior (figure 6).

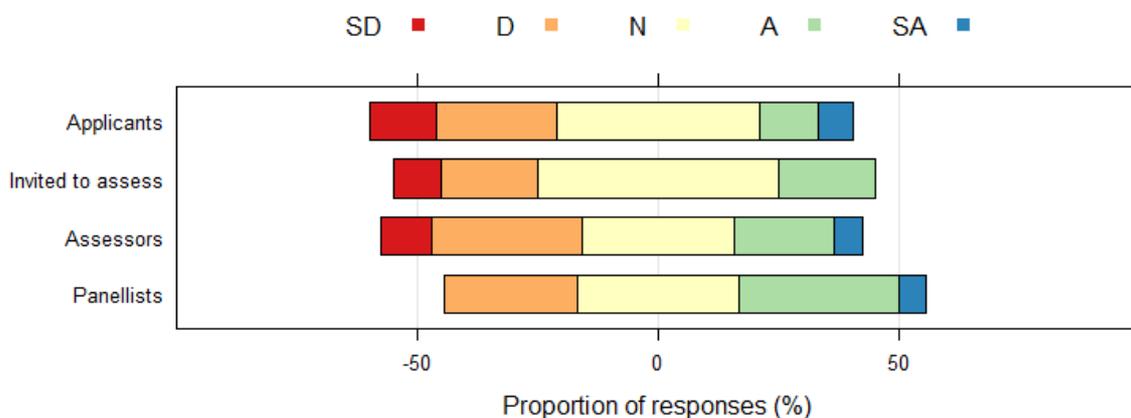


Figure 6: Survey responses expressing level of agreement with the statement: “The Combined Panel Trial process is less likely to select excellent proposals than the Panel process”.

All groups likely to express belief that assessing Fund proposals was/would be of personal benefit.

Given the differing circumstances the question was tailored to suit each group, e.g., panellists were asked for their agreement with the statement: “I gained some intellectual benefit from reviewing proposals to the Fund”, whereas invitees that had not assessed and those eligible applicants in CPT panel proposals that had not been invited to assess were offered “I would have gained some intellectual benefit from reviewing my colleagues’ EOIs”. All groups tended to be extremely positive that the experience of assessment would be of value, with even the most negative group non-scoring invitees agreeing at five times the frequency as disagreeing (figure 7).

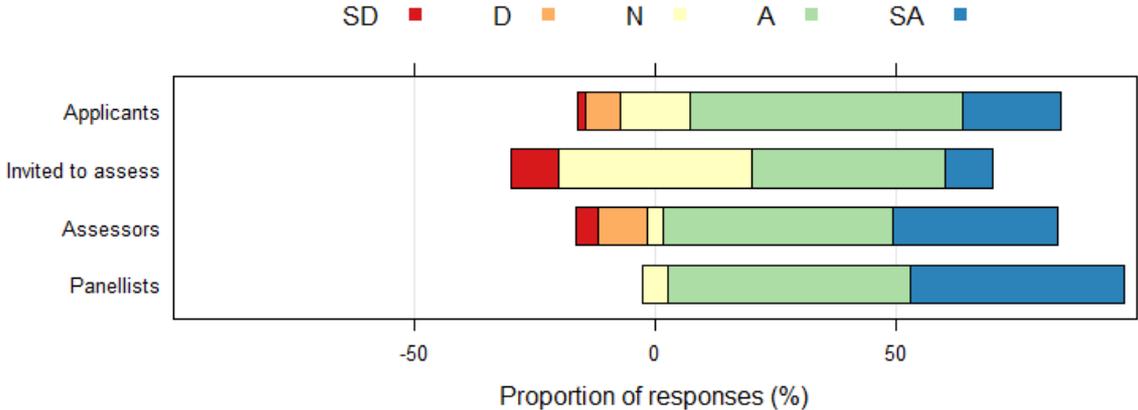


Figure 7: Survey responses indicating broad belief that assessment would be/was of personal benefit to those assessing. Legend codes SD “Strongly Disagree”, D “Disagree”, N “No opinion”, A “Agree”, and SA “Strongly Agree”.

Responses indicate broad comfort with CPT process amongst applicants with the exception of those who had been invited to assess but did not.

As a key part of the CPT process was peer review by co-applicants, we tested the comfort of applicants for having their application reviewed in this fashion. Responses were very positive with the exception of non-scoring invitees. Although the proportion of agreement and disagreement were equal for this group, this question resulted in the highest proportion of negative sentiment for any group and question (figure 8).

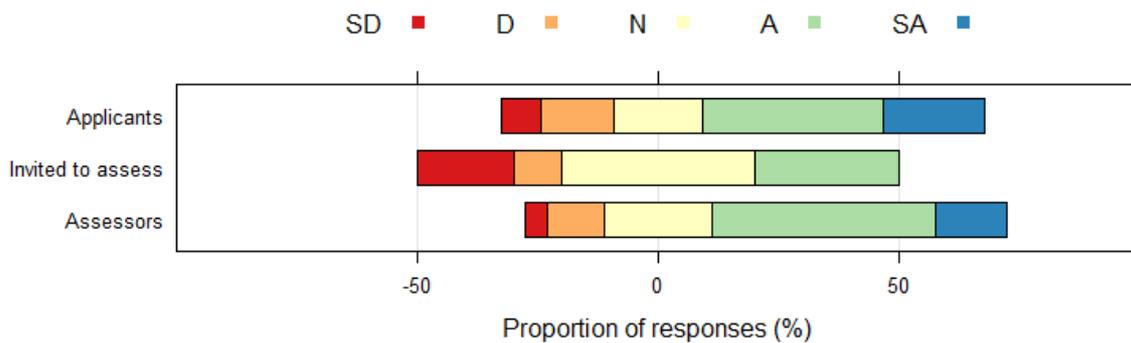


Figure 8: Survey responses expressing level of agreement with the statement: “I am comfortable that my EOI was seen by up to 10 of my NZ-based colleagues”. Legend codes SD “Strongly Disagree”, D “Disagree”, N “No opinion”, A “Agree”, and SA “Strongly Agree”.

Despite having been explained in the Fund’s application guidelines that each applicant had to affirm that they had read in order for their EOI to be submitted, this question appeared to spark the first realisation for some of what the CPT process entailed. That realisation seems to have been unpleasant given that the majority of comments expressing negative, and occasionally very negative, views on the trial despite the Likert-like questions suggesting modest overall support.

Summary

The pilot Combined Panel Trial has been run with a mixture of strong and weak success. The goal of a scalable process with a higher-rate of expert assessment was achieved, but without robust evidence that this process met or improved the existing panel process. In addition, with assessors taking similar amounts of time to panellists to assess each EOI, the CPT process does not appear to provide a way of reducing the overall burden of assessment to the research system; instead, it can be expected to spread that burden more widely across the sector.

Where the CPT process appears to have been weakest is in effective communications to the individuals that were participating. A report on this and an analysis of survey comments will be prepared in the near future. On a positive note, once made aware of the trial's existence via the survey the perceptions of participants tended to be supportive of the trial.

Recommendations

If the Council wishes to rerun/extend the CPT in anything like the current form, several interacting changes are suggested from the experiences of those conducting and participating in the trial.

- Remove the requirement to fit EOIs to a grading distribution, and instead provide robust guidance on what each grade entails;
- By removing the distribution, this would also allow for fewer EOI to be sent to each assessor reducing individual burden;
- Consider further reducing burden by omitting CVs from the EOI to review;
- With fewer EOIs being sent to more assessors, remove the requirement to send EOIs for assessment to those the Executive expects to be lacking in expertise;
- Explore alternative communication routes for making the community affected by any future trial aware that the trial will be taking place; and,
- Explore options for establishing non-applicant pools of assessors having known expertise