

Try, Test and Learn Fund – Evaluation Activity Planning 2017

Stage / task (Overview)	Proposed evaluation-related activities	Indicative timing	Potential Evaluation Unit role and notes
TTL Fund evaluation This task involves an evaluation of the TTL Fund as a whole. This may take the form of a Post Implementation Review.	Scoping PIR	January	EU to lead and advise? TTL team to prepare documentation and provide support as appropriate
	Procuring evaluation provider to deliver PIR	February	EU to lead and advise? TTL to undertake administrative and documentation tasks as appropriate
	Managing evaluation provider contract	February onwards	EU responsible for contract management? TTL to provide input as needed
Idea assessment This stage involves the assessment of the TTL Fund idea submissions. It also involves the collation of a shortlist of recommended ideas to progress to co-development, to be provided to the Inter-Departmental Committee.	Evaluation content in assessment team training and supporting documentation (i.e. operationalising “potential for robust evaluation of outcomes” and other evaluation issues in assessment criterion 1, see footnote over page).	January	EU to provide input and advice? TTL team to prepare documentation.
	Short evaluation presentation in assessment team induction/training session	February	EU to prepare content and deliver? TTL team to provide TTL specific content as needed.
	Specialist evaluation input into assessments of ideas as required? (some possible approach include: evaluation team moderation of selection of ideas; small team of specialised assessors completing review against criteria 1 for all subs; more detailed evaluation readiness assessment of top scoring ideas?)	22 Feb – 10 March	EU to provide staffing assistance / input to assessment process as appropriate.
	Consideration of shortlist from evaluation perspective – i.e ensuring recommended ideas have a feasible evaluation workload and a mix of complex and more straightforward evaluation needs.	6 March – 10 March	EU to provide input to shortlisting as appropriate.
Co-development This stage involves collaborative activities including idea submitters and relevant experts to develop and refine shortlisted ideas to detailed project proposals. Project proposals will be reviewed by the Minister, with a selection proceeding to funding and delivery.	General planning of evaluation-focused components of the co-development approach	January – February	TTL Team will be planning co-development more broadly. EU can provide direct input or can advise through a workshop/meeting(s).
	Specific planning of evaluation-focused co-development activities for each shortlisted ideas	13 March – 17 March	This could be conducted by EU staff? Or by contracted evaluation consultants.
	Conducting evaluation-focused activities, such as program logic or theory of change workshops with co-development partners.	17 March – 14 April	This could be conducted by EU staff? Or by contracted evaluation consultants.
	Evaluation scoping and planning for each project under co-development	11 March – 14 April	EU to lead? Could include advice from Expert Advisory Panel members, and assistance from contracted evaluation consultants. May also include idea submitters if evaluation partners were included in idea submission team.
	(If required) procuring evaluation specialists to conduct evaluation planning for projects during co-development period.	13 March – 17 March	EU to lead and advise? TTL to undertake administrative and documentation tasks as appropriate
Funding	Procuring evaluation providers to deliver evaluations of funded projects	13 May to 26 May	EU to lead and advise? TTL to undertake administrative and documentation tasks as appropriate

This stage involves selecting a delivery partner to implement selected ideas. This will generally involve a direct or restricted grant process.			
Delivery and evaluation of projects This stage involves the delivery and evaluation of selected projects.	Managing evaluation contracts	June onwards	EU responsible for contract management? TTL to provide input as needed

Footnote (from Try, Test and Learn Fund Trance 1 Handbook)

Appropriateness for trial and evaluation

The idea will be assessed on its potential for a robust evaluation of its outcomes and how suitable it is to be run as a trial.

The Department will consider whether an idea offers a testable hypothesis that could be empirically evaluated, including whether it is suitable for best practice methods including randomised controlled trials.

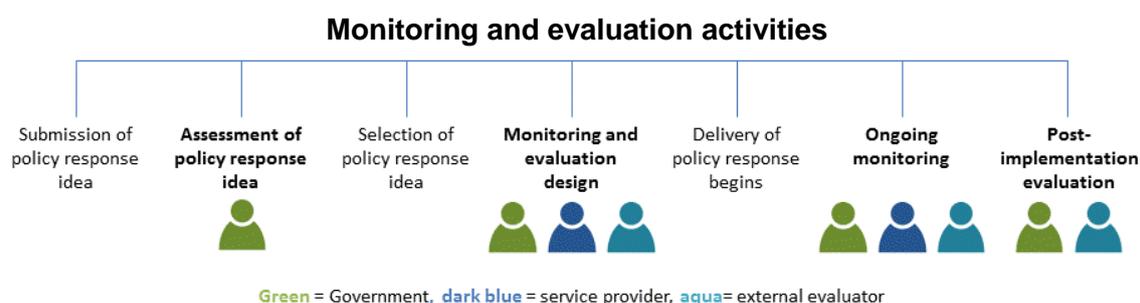
The Department will also consider whether the idea has the potential to be delivered to a sufficient number of participants for statistically significant analysis of program outcomes. The appropriate number of participants for statistically significant analysis will vary by program and will be explored with Proposers of shortlisted ideas during the co-development stage.

Ideas that are suitable for trial will need to have the potential to be scaled up or replicated and delivered to more individuals or communities if they are successful. They should also have the potential to be discontinued after an initial trial period if they are not.

The Department will also consider whether the idea would be overly complex, impractical or expensive to implement and evaluate.

Monitoring and Evaluation of Policy Responses

Monitoring and evaluation will be the key to testing and learning from the policy responses trialled under the Try, Test and Learn Fund. The diagram below sets out the implementation activities that will involve consideration of monitoring and evaluation issues, and which players will have a role.



All policy responses will be evaluated using a combination of analysis of Priority Investment Approach data and modelling, and supplementary data collection and evaluation. It is likely that external evaluators will be engaged to evaluate policy responses in liaison with service providers and the Government. The roles of these stakeholders are captured below.

Monitoring and evaluation stakeholder roles

Stakeholder	Roles
Government	<ul style="list-style-type: none"> Assessing policy responses for evaluation readiness Engaging external evaluators (for most policy responses) Liaising with external evaluators and service providers
External evaluator	<ul style="list-style-type: none"> Design monitoring and evaluation for relevant policy response Carry out ongoing monitoring and evaluation as appropriate Liaising with Government and service providers
Service provider	<ul style="list-style-type: none"> Collecting and sharing data on policy responses Liaising with Government and external evaluators

s 22

Assessment of policy responses

Evaluation readiness will be one of the factors considered in the assessment and selection of policy responses, which is discussed in the separate paper about the application process.

Monitoring and evaluation design

The monitoring and evaluation design for a policy response will define outcomes and timeframes, data and data collection, and analytical methods to be used. Where they are engaged, external evaluators will lead this design process.

Outcomes

The Try, Test and Learn Fund will support targeted trials of policy responses for defined groups of several thousand people. While these trials may have significant impacts on the target groups, it is unlikely that these impacts will be visible in the annual population-wide valuations. Instead, the outcomes will be measured through tailored analysis and evaluations conducted internally and by external, independent evaluators.

The outcomes used to measure success will be guided by the goals of the Priority Investment Approach.¹ The specific outcomes will differ for each policy response and the barriers to work that it seeks to address. They will involve questions such as:

- Are priority groups participating in work or education at increased levels?
- Are priority groups less reliant on welfare (e.g. receiving lower rates of payment)?
- Are priority groups finding work and exiting the welfare system sooner?
- Are the average future lifetime costs for priority groups increasing?

Sources and methods

The Priority Investment Approach model and dataset will be the first data source considered in the monitoring and evaluation design for a policy response. The questions above can be tracked using indicators in this model and dataset, such as:

- Number of people on payment
- Payment transitions
- Earnings from work
- Payment rate (full rate or part rate due to income from work)
- Participation in education
- Average future lifetime costs (noting that the influence of policy changes on these costs may not be reflected in the model for two or more years)

Outcomes will be measured by comparing indicators for people participating in the policy response with a similar group of people not participating. This approach allows us to measure the impact that the policy is having. The methods used to achieve this comparison may include:

- Randomised controlled trials
- Quasi-experimental methods
- Difference-in-difference analysis

Importantly, data sources beyond the Priority Investment Approach model and dataset will be used to ensure a complete picture of what's working and how. Surveys, post-participation monitoring, qualitative data and customer focus groups or interviews will be utilised where appropriate to complement econometric, statistical and other quantitative techniques.

Outcomes that may be captured with this additional analysis include priority groups' capacity for work, increasing work skills and knowledge, wellbeing, social inclusion, financial stress and family functioning.

¹ Improve the lifetime wellbeing of people and families in Australia through an Investment Approach that: increases the capacity of people to live independently of welfare; decreases the Commonwealth's long-term social security liabilities; and reduces the propensity for intergenerational welfare dependence.

Monitoring and evaluation

In most cases, the external evaluator will carry out the bulk of the monitoring and evaluation activities for a given policy response. Monitoring will involve continuous assessment that generates detailed information on the progress of the policy response. Evaluation will involve a systematic study of the effectiveness, efficiency and impact of the policy response. Both of these activities will be used to inform decisions to continue, amend or conclude the implementation of the policy response. Evaluation will also be used to gather lessons to inform future policy design.

DRAFT

We are keen to brainstorm your key questions and issues about evaluation, data and the Priority Investment Approach for use in our issues register

Evaluation and the Priority Investment Approach

- **Evaluation using the Priority Investment Approach:** using the tools and administrative data of the Approach to help to evaluate programs
- **Evaluation of the Try, Test and Learn (TTL) Fund:**
 - Evaluating the TTL Fund as an overall measure
 - Evaluating the separate interventions funded under the TTL Fund

Your questions and issues

- What **questions** about the Priority Investment Approach need to be resolved to help your team to move forward?
- What are the chief **issues and risks** facing evaluation using the Priority Investment Approach, and evaluation of the TTL Fund and its interventions?
- Who will be the key **stakeholders** to resolve these issues, risks and questions?

Our discussion questions

- **Stakeholders:** what role will DEX play in evaluation under the Priority Investment Approach? How will other stakeholders be involved?
- **Evaluation processes:** what will we need to make evaluation of TTL interventions and evaluation using the Priority Investment Approach a success?
- **Access to data:** what will be the processes for access to the data and tools of the Priority Investment Approach? What data governance / management issues arise?

Try, Test and Learn Fund

s 22

The Government announced the \$96 million Try, Test and Learn (TTL) Fund in the 2016-17 Budget. The TTL Fund will finance new or innovative policy responses that seek to support the priority groups identified through the actuarial analysis, and the goals of the Priority Investment Approach. The TTL Fund will also seek to foster collaborative and data-driven policy making across government and with external stakeholders.

The TTL Fund will enable the development, implementation and tailoring of innovative policy responses aimed at addressing individuals' barriers to participation and supporting people with the capacity to work to do so. The policy responses will be evaluated to assess their efficacy, with the results being used to transform and inform our investment in existing programs or make the case for investment in new policy responses. More information on how policy responses will be evaluated is provided below.

s 22

Evaluation of Try, Test and Learn Policy Responses

Rigorous monitoring and evaluation of Try, Test and Learn Fund policy responses will be very important. It will allow informed decisions to be made about the continuation of policy responses financed by the Fund. It will also ensure that the Fund generates high quality

evidence on what works, how, and for whom. This evidence can then be used to change and adapt existing programs and to make the case for new policy responses.

Try, Test and Learn Fund policy responses will be evaluated using a range of evaluation approaches depending on the nature of each response. The first step for each policy response will be working out how the Priority Investment Approach model and datasets can contribute to its evaluation. Other evaluation methods will then be incorporated to support the Priority Investment Approach analysis and create a complete picture of the effectiveness of a particular policy response, and how and why it is effective or ineffective.

Priority Investment Approach model and datasets

The Priority Investment Approach broadly entails two tools which can be used in the evaluation of policy responses:

- The actuarial model that generates the Priority Investment Approach annual valuations. This model simulates people's future life trajectories and interactions with the payment system.
- The longitudinal administrative dataset which contributed to the development of the actuarial model. This is a quarterly longitudinal social security dataset based on administrative data from the welfare payment system, and currently spans the period 2001 to 2015.

The capacity of these tools to evaluate the impacts of Try, Test and Learn Fund policy responses will continue to increase as they are developed over the coming years.

Longitudinal administrative dataset

The longitudinal dataset of social security administrative data has the potential to be used to evaluate policy responses in several ways. Any of the indicators in the dataset which were relevant to the impacts of a given policy response could be analysed to support the response's evaluation. The indicators in the dataset include, for example, payment receipt, payment rate, earnings, birth of children, caring responsibilities, experience of domestic violence and disability or mental health conditions.

A range of methods could be used to analyse the indicators described above, including:

- Quasi-experimental studies of participant impacts: longitudinal analysis of participants or trial locations compared to analysis of control groups or locations. Control groups or locations would be matched to the participant group or location.
- Randomised control trials: random assignment of individuals or communities to participation in a policy response. Longitudinal analysis of these participating individuals or communities with control groups. Control groups would be matched to the participant group or locations.

Actuarial model

The other Priority Investment Approach tool available to support the evaluation of policy responses is the actuarial model. The model can be used to understand the long term

impact of decisions made today and in the future at a fairly high 'system' level. The baseline valuation model provides a good foundation for accurately projecting future lifetime costs at the whole population level and for particular subgroups.

However, like all statistical models, the ability of the actuarial model to differentiate average future lifetime costs between different people is limited by the extent of the modelling variables included in the model. That is, the groups considered by the model can only be as specific as the characteristics provided by the modelling variables. As more modelling variables are included, the model will be able to provide future lifetime costs for more refined groups.

The baseline model used a limited number of modelling variables to determine people's future lifetime costs and pathways. These modelling variables included welfare class, age, gender, partner status, number of dependent children, highest level of education attained, duration in the welfare system, age entered welfare system, and so on. The 2016 valuation will include additional modelling variables, such as earnings indicator and DSP medical condition type.

Other evaluation methods

Other evaluation methods and data sources will support evaluation with the actuarial model and dataset to ensure a complete picture of both the impact of policy responses, and also how and why they are effective or ineffective. Supporting analysis will use both existing data sources (e.g. Data Exchange, program or service data) and customised data collection such as surveys, interviews or action research methods.

Using a rich, mixed-methods approach to evaluation will ensure that we generate detailed understanding of the impact Try, Test and Learn Fund responses are having on individuals and communities and have robust evidence on which to base future policy decisions.

Measuring success –Try, Test and Learn Fund outcomes and evaluation

The Try, Test and Learn Fund will support targeted trials of policy responses for defined priority groups of several thousand people. While these trials may have significant impacts on the target groups, it is unlikely these will be visible in the annual population wide valuations.

Instead, the outcomes of Try, Test and Learn Fund policy responses will be measured in tailored analysis and evaluations conducted internally and by external independent evaluators.

The measures of success will differ for each individual Try, Test and Learn Fund response and the barriers to work they are addressing. Further, the approach for the evaluation of policy responses will be designed with stakeholders as part of the overall consultative design process for the Fund.

Overall, the success of the Try, Test and Learn Fund will be tied to achieving the goals of the Priority Investment Approach, which entail improving the lifetime wellbeing of people and families in Australia by:

- Increasing the capacity of people to live independently of welfare
- Decreasing the Commonwealth's long-term social security liabilities
- Reducing the propensity for intergenerational welfare dependence.

The success of the Try, Test and Learn Fund will be measured against the following outcomes:

- Are priority groups participating in work or education at increased levels?
- Are priority groups less reliant on welfare, for example are they receiving part rather than full rates of payments?
- Are priority groups finding work and exiting the welfare system sooner?
- Are the average future lifetime costs for priority groups decreasing?

Indicators of these outcomes can be tracked in the Priority Investment Approach model and dataset. The key indicators will be:

- Number of people on payment
- Payment transitions
- Earnings from work
- Payment rate – full rate or part rate due to income from work
- Participation in education
- Average future lifetime costs (the model may take two or more years to reflect policy changes in future costs).

For individual priority groups and responses there will be additional specific outcomes.

For example, for young carers a key outcome will be the proportions of young carers who are participating in training or education, and who leave the system once their caring role ceases rather than transitioning onto an unemployment payment.

In the long-term, we will also be able to track the outcomes of people's children to measure whether the risk of intergenerational welfare dependence has changed.

The Priority Investment Approach model and dataset will be the first port of call for measurement of the outcomes.

We will measure outcomes by comparing indicators for people participating in a new policy or support to a similar group of people not participating. This allows us to see and measure the impact the policy is having.

In some cases we will use the Priority Investment Approach data to conduct Randomised Control Trial designs. In other cases we will find control groups to match the group or location that has already been select to participate in a policy trial.

Importantly, we will also use other data sources and evaluation methods to ensure a complete picture of what's working and how. Surveys, post-participation monitoring and qualitative data and customer focus groups or interviews will be utilised where appropriate to complement econometric, statistical and other quantitative techniques.

Outcomes that may be captured with additional analysis include priority groups' capacity for work, increasing work skills and knowledge, wellbeing, social inclusion, financial stress and family functioning.

This will support the Priority Investment Approach's important goal of improving people's lives, not just reducing payment recipient numbers or costs.

Monitoring and evaluation design

Evaluation across the TTL Fund policy cycle

Evaluation of TTL Fund policy responses will involve a systematic study of the effectiveness, efficiency and impact of the policy response. As outlined in the sections above, evaluation will be considered at all stages of the TTL Fund process as summarised below:

- during the **idea selection process**, evaluation readiness will be used as an assessment criterion to ensure all ideas that are selected have the potential to be effectively and rigorously evaluated
- during the **pre-market co-development stage**, adjustments may be made to an initial idea to increase the efficiency and effectiveness of evaluation or allow for sophisticated methodologies such as randomised controlled trials. The general approach to evaluation would be established at this time
- during the **selection process** service providers will be sought who have the capacity to support evaluation during the implementation of a program or project

⁶ It is important to note that while the agility of funding is a key strength of the TTL Fund, it will be critical to give service providers security to deliver initiatives, particularly larger initiatives, confidently. Public messaging will consider this balance between agility and security carefully.

⁷ As outlined earlier in this paper, the suitability of a program being trialled and ceased would be considered in the assessment process. For example, a program that involved large investments up front (e.g. in funding, delivery partnerships/relationships, participants' time etc.) may be less suitable for the TTL Fund due to the difficulty in ceasing these programs if they are not working effectively.

- during the **co-development stage**, DSS will work with the service provider and any external evaluation providers to develop and plan for the evaluation and conduct any baseline data collection required
- during the **implementation stage**, data from monitoring and evaluation activities will be used to assess performance against outcomes and KPIs and inform decision making at stop-go points
- **after the implementation stage:**
 - where feasible, post-participation evaluation would be undertaken to ascertain the long-term effect of an initiative. This may also include supporting evaluation continuity if a trial is scaled up outside the TTL Fund
 - final evaluation results will be publically shared and added to a knowledge store of evidence gathered via the Try, Test and Learn Fund policy responses. Final evaluation results may also be used by Commonwealth agencies, service providers or other stakeholders to seek other funding sources to continue a policy trialled through the Try, Test and Learn Fund.

Overarching measures of success

The outcomes used to measure success will be guided by the goals of the Priority Investment Approach. The specific outcomes will differ for each policy response and the barriers to work that it seeks to address. They will involve questions such as:

- are priority groups participating in work or education at increased levels?
- are priority groups less reliant on welfare (e.g. receiving lower rates of payment)?
- are priority groups finding work and exiting the welfare system sooner?
- are the average future lifetime costs for priority groups decreasing?

Evaluation data sources and methods

While TTL Fund policy responses may have significant impacts on the target groups, it is unlikely that these impacts will be visible in annual Priority Investment Approach population-wide valuations. Instead, the outcomes will be measured through tailored analysis and evaluations conducted internally and by external, independent evaluators.

The Priority Investment Approach model and dataset will be the first data source considered in the monitoring and evaluation design for a policy response. The questions above may be able to be tracked using indicators in this model and dataset, such as the number of people on payment, payment transitions, or average future lifetime costs (noting that the influence of policy changes on these costs may not be reflected in the model for two or more years).

Outcomes will be measured by comparing indicators for people participating in the policy response with a similar group of people not participating. This approach allows us to measure the impact that the policy is having. The methods used to achieve this comparison may include:

- randomised controlled trials⁸
- quasi-experimental methods
- difference-in-difference analysis.

Importantly, data sources beyond the Priority Investment Approach model and dataset will be used to ensure a complete picture of what's working and how. Surveys, post-participation monitoring, qualitative data and customer focus groups or interviews will be utilised where appropriate to complement econometric, statistical and other quantitative techniques. Outcomes that may be captured with this additional analysis include priority groups' capacity for work, increasing work skills and knowledge, wellbeing, social inclusion, financial stress and family functioning.

The Priority Investment Approach Taskforce is considering the best approach to securing human research ethics clearance for evaluations conducted under the Try, Test and Learn Fund, and could consider using an existing ethics committee such as the AIFS Ethics Committee or the Australian Health Ethics Committee.

s 22