



Social Investment Agency
Toi Hau Tāngata

Impact Measurement Standards

June 2026



New Zealand Government
Te Kāwanatanga o Aotearoa



Impact Measurement Standards

Purpose of the standards

The Impact Measurement Standards have been designed to enable a consistent approach for measuring and using evidence on outcomes and impact when evaluating the effectiveness of the government's social spending.

The standards offer a structured framework for using evidence in the design of programmes and services, assessing the quality of evidence, and ensuring that appropriate outcome measures are in place from the outset. They also support decision-making, continuous learning, and improvement by helping agencies understand what works, for whom, and in what contexts.

These standards are one tool agencies can use as they move toward embedding a social investment approach within their portfolios.

Who do the standards apply to?

These standards apply to government agencies delivering or commissioning programmes or services to support improved social outcomes. This includes agencies for child wellbeing and protection, education, health, housing, justice, policing, social development and affiliated services. They will be less applicable to services contracted to deliver a specified product or service for a discrete purpose (e.g. immunisation services).

Senior leadership, policy, research and evaluation, commissioning, and operational functions within agencies will need to work together to apply the standards effectively.

How should you use the standards?

Agencies should start applying the standards across their portfolios of programmes and services to:

- Use evidence on outcomes and impact to guide decision-making, prioritisation, and resource allocation.
- Identify and address gaps in evidence and support continuous learning and adaptive improvement to refine programme design and delivery over time.
- Review and report evidence of outcomes and impact.
- Work with service delivery organisations and partners to agree how outcomes and impact will be measured, including appropriate data collection and sharing.
- Design and implement new programme or service proposals so that outcomes and impact can be meaningfully measured from the outset.

The Social Investment Agency will use the standards when working with other agencies to apply the social investment approach and to understand the impact of current social sector spend. We are also working to embed this standard in wider funding and commissioning processes and requirements.

Impact Measurement Standards

Level 1: The Theory of Change is based on evidence

All new programmes and services should meet Level 1

- Describes how the programme or service's (short-term, medium-term and long-term) inputs, activities and outputs will result in the expected outcomes and impacts.
- Works with partners, communities and other stakeholders to identify and list expected outcomes and impacts, and how they will be measured including using Kaupapa Māori approaches and partnering with Māori where relevant, recognising these as important approaches to understanding outcomes and impact.
- Documents evidence, assumptions and limitations, and identifies areas of uncertainty that may require further testing.

Level 2: Evidence indicates changes in outcomes or impacts

All programmes and services should strive to meet Level 2 once implemented

- Meets all Level 1 standards.
- Reports and measures changes in short-term, medium-term and long-term outcomes and impacts for those receiving the programme or service.
- Captures information at a level that enables outcomes to be meaningfully measured, often at an individual or whānau level where appropriate.
- Shows changes in outcomes and impact and provides evidence of a programme's contribution to these changes, while recognising that other factors may also have influenced the results.
- Considers how outcomes and impacts vary for different population groups.
- Considers caveats and addresses limitations and states them clearly.

Level 3: Evidence demonstrates if a programme or service has contributed to, or caused, observed outcomes or impacts

All programmes and services should consider the feasibility and appropriateness of reaching Level 3, recognising this may not be suitable for all types of interventions

- Meets Levels 1 and 2 standards.
- Uses research methods that allow robust conclusions about the effect of the programme or service on outcomes and impacts. This may involve comparing outcomes with a similar group who did not receive the programme or service, where appropriate and feasible.
- May use (qualitative) methods that do not involve numeric data but allow for robust tests of whether the intervention does or does not lead to the outcomes.
- Captures information at an individual or whānau level.
- Ensures that the analysis recognises potential sources of bias and appropriately identifies and addresses them.
- Considers how outcomes and impacts vary for different population groups.
- Documents the methodology in detail so that it can be replicated.

Explanation

While the strength of evidence can increase across all levels, higher levels do not necessarily indicate more appropriate or more useful evidence in all contexts. The choice of approach should be guided by the nature of the programme or service, the evaluation questions, and what is feasible and appropriate in practice.

For this reason, it is also important to consider the level of confidence in the evidence when interpreting findings. Confidence in evidence depends on factors such as data quality, methodological robustness, consistency of findings, and the plausibility of the Theory of Change. Clearly articulating confidence helps support transparent and well-informed decision-making.