Australian Government
The Treasury

Australian Centre for Evaluation
ace

Reflections on embedding a culture of evaluation across the Commonwealth: from principles to practice

Suzanne Butler & Harry Greenwell
Australian Centre for Evaluation

evaluation.treasury.gov.au

27 September 2023

Overview

1. Context – Independent Review of the Australian Public Service (APS) | 2019

2. Reflections – Implementing the Commonwealth Evaluation Policy (released December 2021)

3. Next steps – About the Australian Centre for Evaluation (from July 2023)

4. Championing high-quality impact evaluation
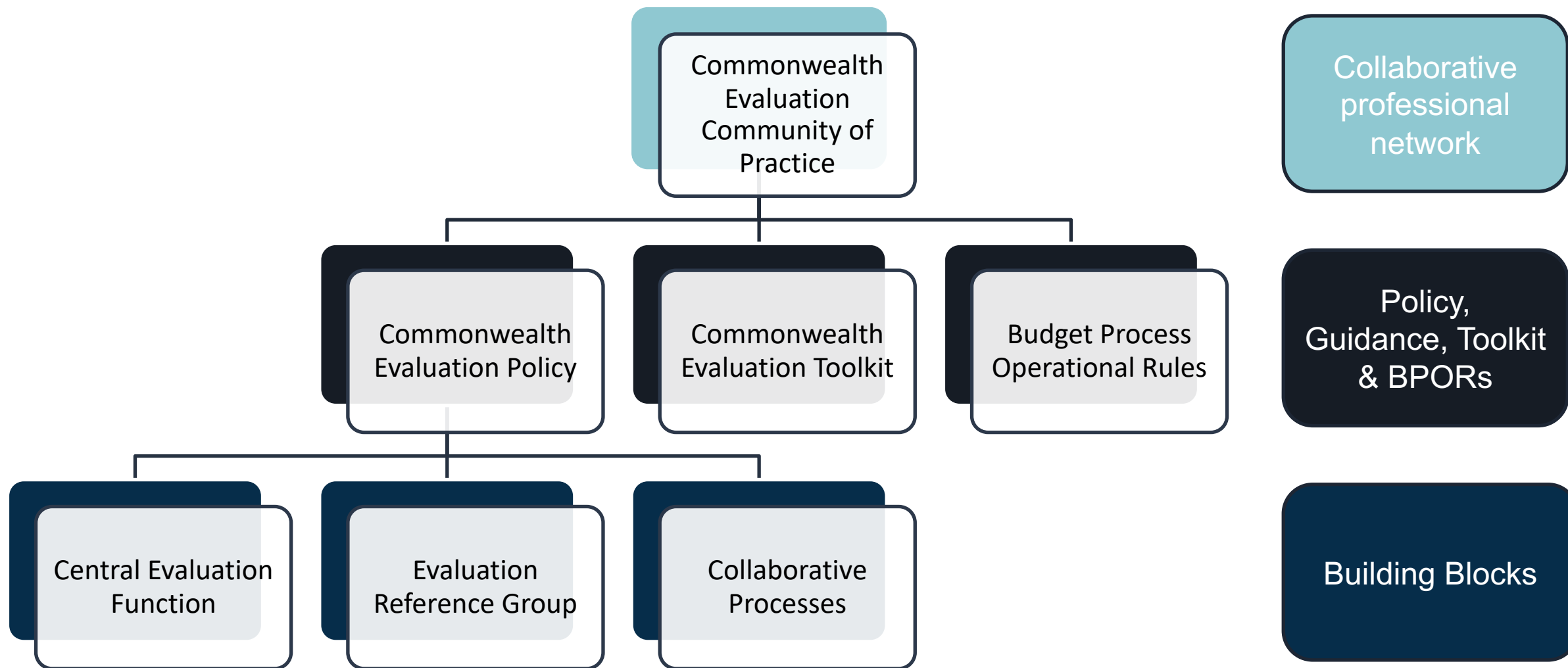
5. Open discussion

Context
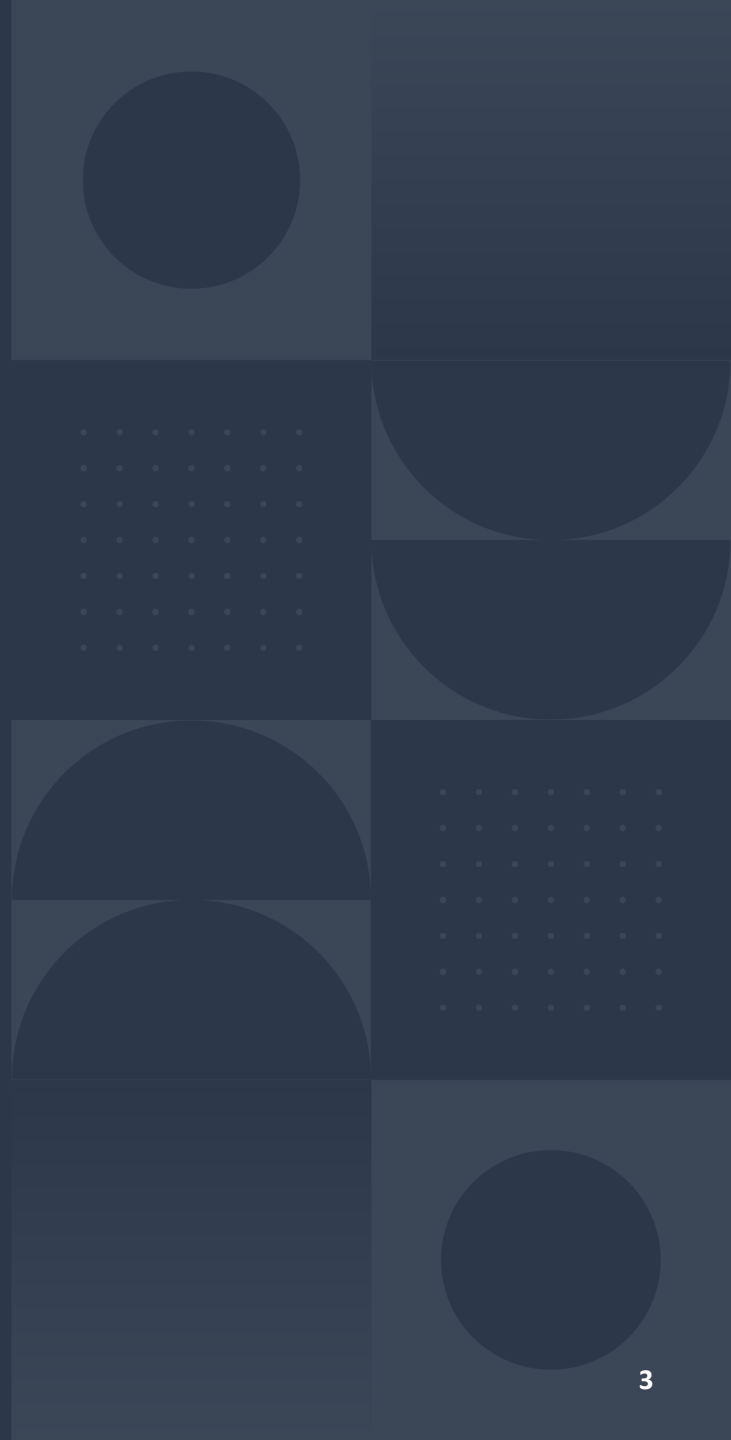
Context – Independent Review of the APS

OUR PUBLIC SERVICE OUR FUTURE

INDEPENDENT REVIEW OF THE AUSTRALIAN PUBLIC SERVICE

Embed a culture of evaluation and learning from experience to underpin evidence-based policy and delivery.

- Finance to develop, for Secretaries Board agreement, an APS-wide approach to build evaluation capability and ensure systematic evaluation of programs and policies.
- Finance to establish a central enabling evaluation function to support APS evaluation practices and expertise.
- Agencies to establish in-house evaluation functions and annual plans, and publish evaluations, unless exempt by the Cabinet.
- Government to amend Cabinet and budget requirements to establish a systematic approach for formal evaluations.

Establishing strong foundations

```
                    ┌─────────────────────────┐
                    │     Commonwealth        │
                    │     Evaluation          │         ┌──────────────────────┐
                    │   Community of          │         │    Collaborative     │
                    │     Practice            │         │    professional      │
                    └────────────┬────────────┘         │     network          │
             ┌───────────────────┼───────────────────┐  └──────────────────────┘
    ┌────────┴────────┐ ┌────────┴────────┐ ┌─────────┴─────────┐
    │  Commonwealth   │ │  Commonwealth   │ │  Budget Process   │  ┌──────────────────────┐
    │  Evaluation     │ │  Evaluation     │ │  Operational      │  │     Policy,          │
    │  Policy         │ │  Toolkit        │ │  Rules            │  │  Guidance, Toolkit   │
    └────────┬────────┘ └─────────────────┘ └───────────────────┘  │     & BPORs          │
  ┌──────────┼──────────┐                                          └──────────────────────┘
┌─┴─────────┐┌─┴───────┐┌─┴─────────┐
│ Central   ││ Evaluat-││ Collabor- │                              ┌──────────────────────┐
│ Evaluation││ ion     ││ ative     │                              │                      │
│ Function  ││ Reference││ Processes │                              │   Building Blocks    │
│           ││ Group   ││           │                              │                      │
└───────────┘└─────────┘└───────────┘                              └──────────────────────┘
```

Commonwealth Evaluation Policy

- Principles-based policy to support fit-for-purpose evaluative approaches in line with the public sector resourcing, reporting and performance frameworks established under the *Public Governance, Performance and Accountability (PGPA) Act 2013*

- The Policy is designed to:
  - embed a culture of evaluation and learning from experience across the Commonwealth to underpin evidence-based policy and delivery
  - improve the way entities assess implementation, measure the impact of government programs and activities, and frame policy decisions on revised or new programs
  - improve the quality of performance reporting for the purposes of accountability, continuous improvement and decision-making

Commonwealth Evaluation Toolkit

Reflections
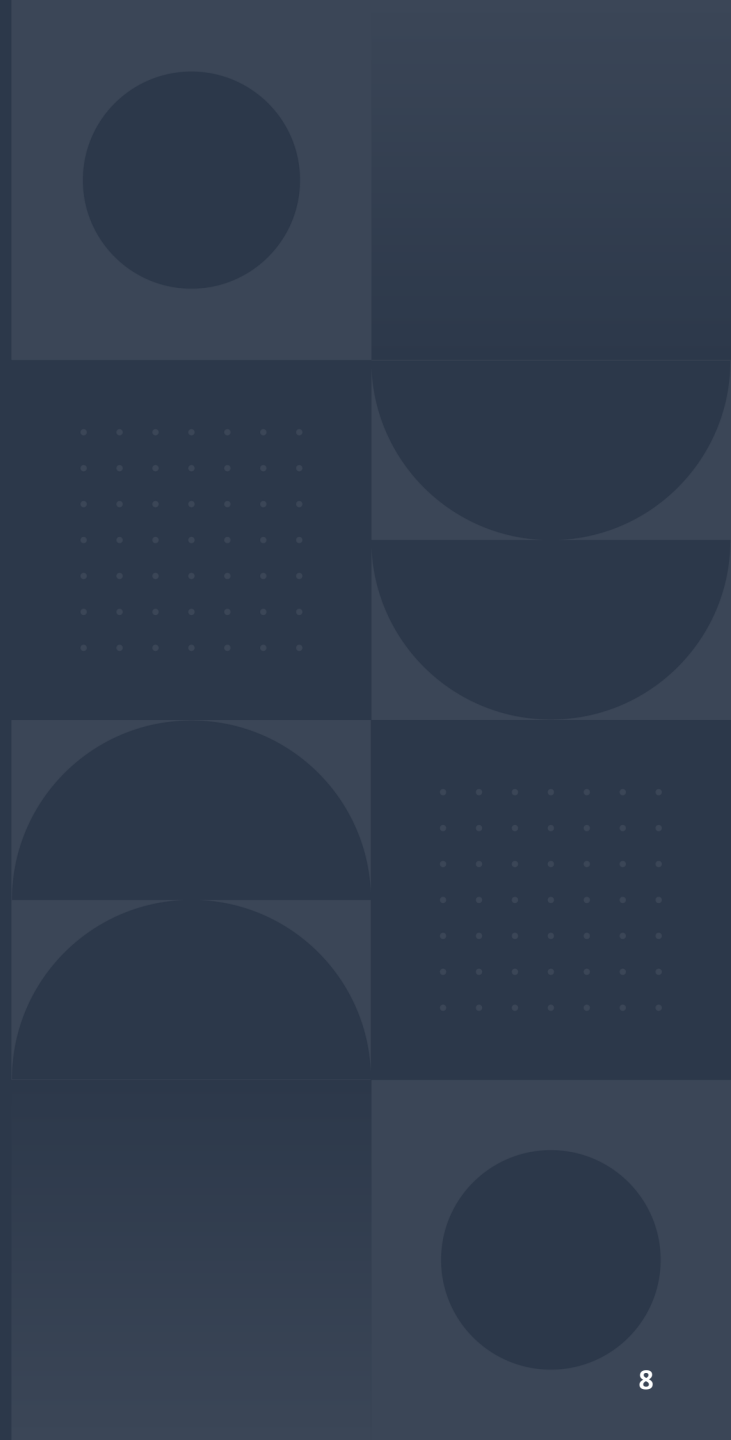
Key issues

- Conceptual issues related to *"embedding a culture of evaluation"*

- Interdependencies related to improving the use of evaluation, data and evidence

- Critical importance of understanding the operating context

- Stakeholder feedback on how to embed *"enduring cultural change"*

Conceptual issues – "evaluation" vs "Evaluation"

A "culture of evaluation" involves a broad spectrum of activities and practices, ranging from:

- "evaluative thinking" or the use of an "evaluative mindset" in everyday business activities

- "evaluative-like" activities that occur in various related business processes (e.g. budget, risk, program and/or policy management; performance monitoring; corporate performance planning and reporting; compliance monitoring; statutory and/or strategic reviews etc)

- "formal evaluations" commissioned to generate robust evidence to support decision-making and deliver better outcomes at various stages of the policy cycle.

Interdependent APS Reform Initiatives

"Better use of data to support policy decisions, improve and tailor services, facilitate risk-based regulation and conduct targeted evaluations will be one of the most significant changes to effect the APS in coming decades"

(p173, Independent Review of the Australian Public Sector)

Understanding the Commonwealth operating context

- **189** Commonwealth entities and companies

- **204** Outcomes funded through appropriations

- **505** Budget-funded programs (many of which include multiple sub-programs)

Stakeholder feedback – enduring change requires:

1. A flexible, principles-based evaluation policy approach

2. Guidance that targets people with no evaluation experience

3. Practical tools and support to help people use evaluation in their daily work

4. Staged, incremental reforms designed to support enduring cultural change

About the Australian Centre for Evaluation (ACE)

2023–24 Budget Decision

In the context of *An Ambitious and Enduring APS Reform Plan*, the Albanese Government provided $10 million over 4 years from 2023–24 (and $2.1 million per year ongoing) to establish a central evaluation function within Treasury to provide leadership and support to improve evaluation capability across Government and leading a small number of flagship evaluations each year.

Budget Paper No. 2: Budget Measures (p 213)
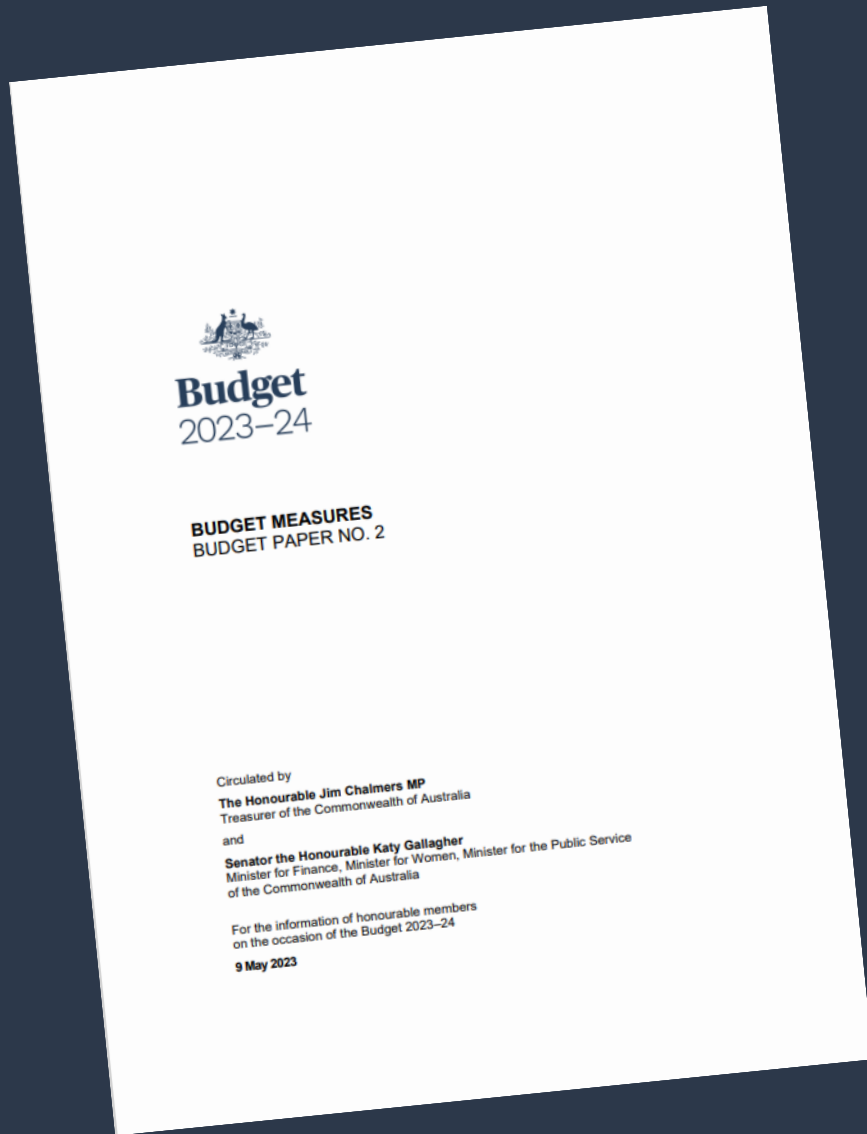
Structure of the ACE

ACE Branch Head

Evaluation Leadership, Policy & Capability Unit

- *Leadership*: Provide APS-wide leadership on evaluation through new governance arrangements and an annual State of Evaluation report.

- *Evaluation policy*: Support agencies' to embed the Policy, with initial focus on evaluation planning.

- *Capability*: Support APS-wide evaluation capability building efforts.

- *Budget process*: Oversee agencies' evaluation effort in the Budget process.

Impact Evaluation Unit

- *Selected impact evaluations*: Initiate 3-5 impact evaluations each year in partnership with agencies. Lead or contribute to research design, data collection, data analysis, ethics review and privacy assessment.

- *Agency-led impact evaluations*: Provide support and resources to agencies to plan, commission and undertake impact evaluations.

Work Streams

Noting the ACE's operating model is still under development, our key work streams include:

- embedding the Commonwealth Evaluation Policy and promoting the Commonwealth Evaluation Toolkit

- championing high-quality impact evaluations, including identifying opportunities to partner with agencies on selected evaluations

- promoting better evaluation planning and use in Budget and Cabinet processes

- supporting and overseeing evaluation capability building across the Australian Government.

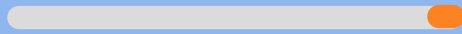# Australian Population: 25,422,788 people

TOTAL SEX DISTRIBUTION

49 % are male, 51% are female

0.4% WERE UNHOUSED

122,494 Australians were estimated to be unhoused

27.6% WERE BORN OVERSEAS

1 in 5 Australians were born overseas

0.3% LIVED WITH A SAME-SEX PARTNER
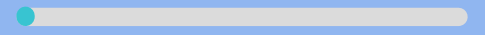
1 in 3 of these couples were married

5.8% HAVE A DISABILITY

5.8% of Australians require assistance with core activities.

3.2% ARE FIRST NATIONS
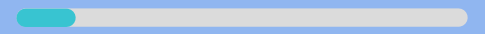
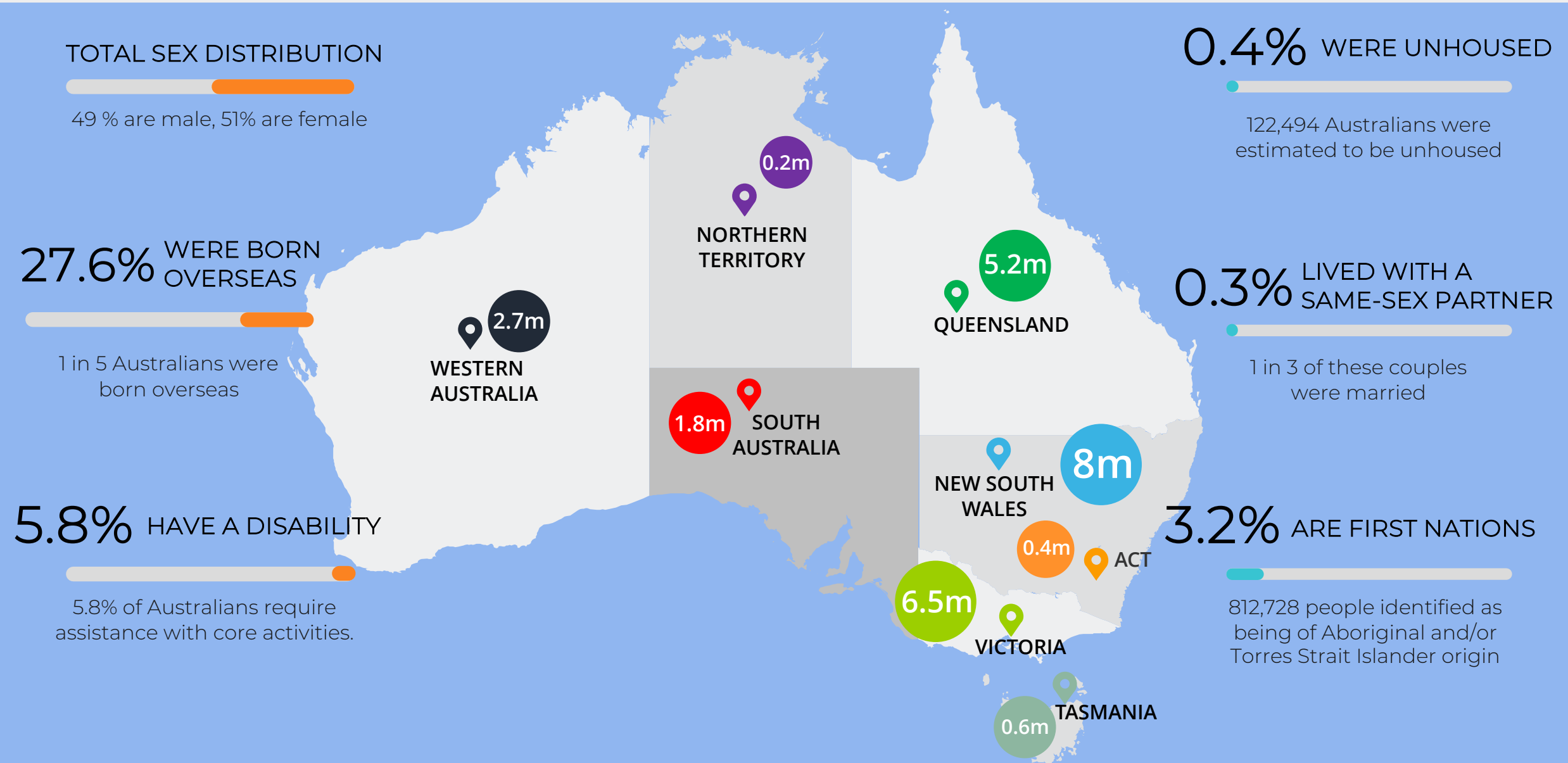812,728 people identified as being of Aboriginal and/or Torres Strait Islander origin

0.2m NORTHERN TERRITORY

5.2m QUEENSLAND

2.7m WESTERN AUSTRALIA

1.8m SOUTH AUSTRALIA

8m NEW SOUTH WALES

0.4m

ACT

6.5m VICTORIA

0.6m TASMANIA

Championing high-quality impact evaluation

Harry Greenwell

Director, Impact Evaluation Unit

Why focus on impact evaluations?

- Counterfactual-based impact evaluations, such as randomised control trials (RCTs), estimate average causal effects.

- These estimates are often important for decision makers.

- But estimates of average causal effects are often lacking – even where feasible – or they are heavily assumption driven.

Five maxims for how we champion RCTs

- How we champion impact evaluation matters
- The following maxims align with the principles in the Commonwealth Evaluation Policy.

1   2   3   4   5

1. RCTs should be ethical & culturally appropriate

- Ethical review

- Procedures to ensure RCTs are culturally appropriate

- Transparency – publication by default

2. RCTs should be robust (well-designed)

- We shouldn't be mesmerised by randomisation. A poorly designed RCT is just a poor evaluation.

- Important features of good design:

  - Pre-registration, with a detailed pre-analysis plan

  - Sufficient statistical power to produce estimates that are precise, as well as accurate

  - Avoid threats that undo the magic of randomisation such as 'spillovers' or missing data

3. RCTs should be useful

- We should only conduct RCTs if the average causal effect is likely to provide useful evidence.

- Even where RCTs are useful, they will often be more useful when part of mixed method designs.
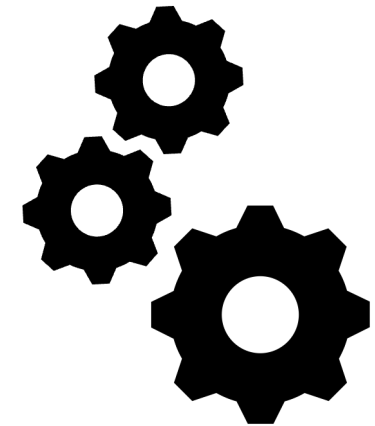
4. Aligned with the Government's priorities

- We will encourage RCTs that test the details of policy delivery and implementation.

- But we will *especially* champion RCTs that address problems aligned with the Government's priorities.

- And where RCTs are not possible, we will avoid devaluing other evaluation evidence that is brought to bear.
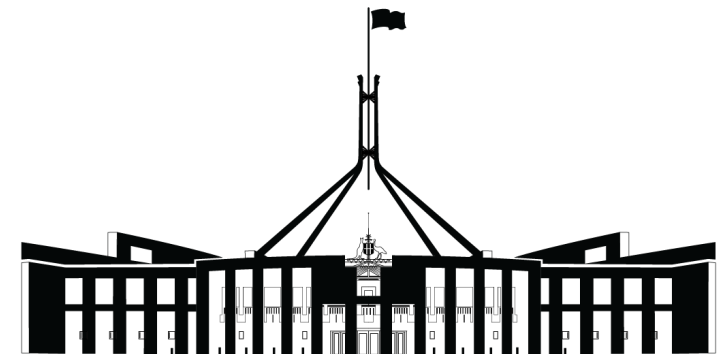
Source: Department of Parliamentary Services

5. RCT evidence should be used with care

- We will be careful about how we report *average* effects.

- We will be careful about over-generalising.

Conclusion

- Estimates of the average causal effects are often important for policy makers. But these estimates are often lacking, or heavily assumption driven.

- That is why we will champion RCTs, and other counterfactual-based impact evaluation methods, that are:

  – ethical and culturally appropriate

  – well-designed

  – useful

  – aligned with the Government's priorities

  – used with care

Questions & Answers

Contact us: evaluation@treasury.gov.au