Standards of Evidence:
A common language for decision making, implementation and evaluation in schools

Matt Deeble

Director

Social Ventures Australia

Dr Deb Kember, Deb Fisher,

State Schools Division

Department of Education, Queensland

Queensland
Government

While you are waiting……

To receive session resources in a digital pack:

- add your email address to one of the recording sheets

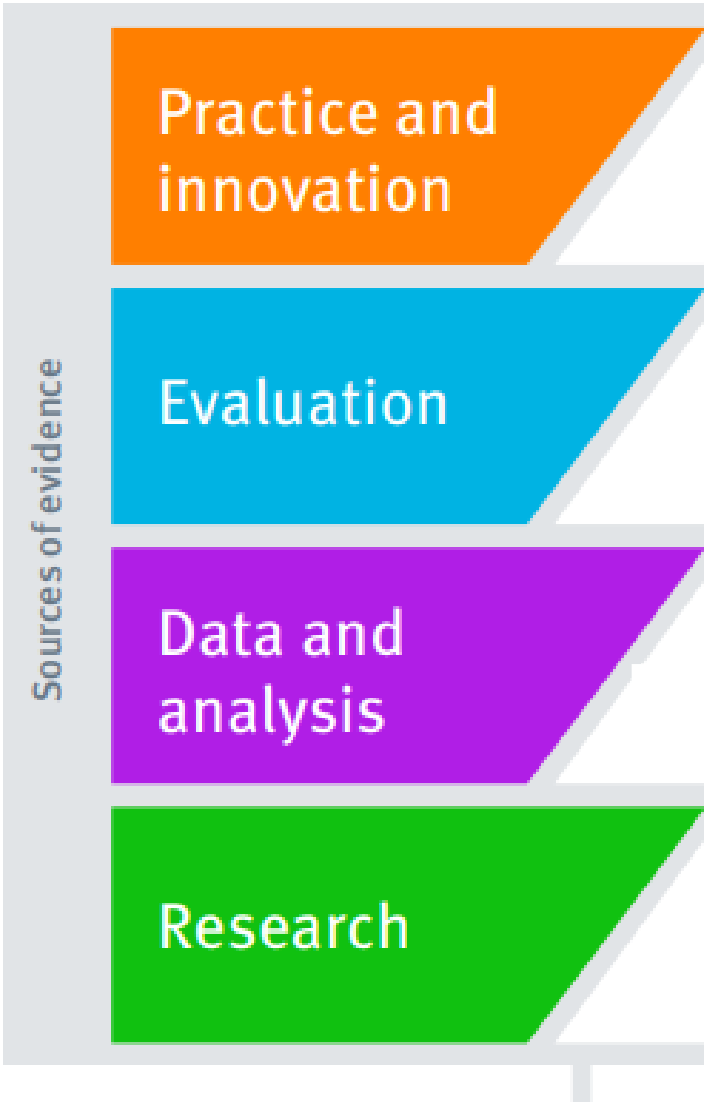- resources will be emailed to you before the end of the conference.

Queensland's state schools

- 552,512 fulltime students

- 1240 state schools

- 7 educational regions

- Queensland is home to the fastest improving educational system in Australia

Evidence Framework

Sources of evidence

- **Practice and innovation**
- **Evaluation**
- **Data and analysis**
- **Research**

- can generate reliable evidence of what works

- plays in the design and implementation of the department's initiatives

- quality, robust data informs our initiatives and helps us measure our impact

- rigorous research helps ensure our work is informed by the best available evidence

The Standards of Evidence

Design
Quality of attribution

Impact
Measurable change

Scalability
Potential to implement

Investment
Creating value

Strength of evidence

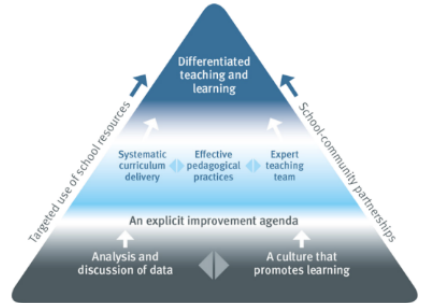| Design — Quality of attribution	Impact — Measurable change	Scalability — Potential to implement	Investment — Creating value
Verified attribution An independent source or sources verifies the demonstrated measurable change can be attributed to the initiative	**Very high** Very large measured improvement	**Larger scale implementation** An independent source or sources verifies the initiative has been implemented by others on a larger scale	**Cost-effectiveness known** Information on cost-effectiveness exists
Attribution Evidence from a high-quality study or studies establishes the measurable change can be attributed to the initiative	**High** Large measured improvement	**Implemented elsewhere** The initiative can be implemented in more than one setting and/or with different groups with at least the same impact. Documentation supports implementation	**Outcomes costed** Information on the costs per unit of outcome exists
Possible attribution A reasonable explanation supports attribution of the measurable change to the initiative	**Moderate** Medium measured improvement	**Well documented** Documentation exists about how the initiative is implemented and is comprehensive enough to enable implementation with different groups or in other settings	**Outputs costed** Information on the costs per unit of output exists
Attribution unclear A logical plan describes the initiative and its objectives, target and comparison groups, inputs, outputs, outcomes and timelines. It includes the measures and methods for collecting, storing and analysing data	**Low** Small measured improvement that can be reasonably linked to the initiative	**Could work** Documentation describing the implementation of the initiative in one setting exists	**Costs known** The costs of the initiative (for example, financial, human resources, infrastructure, program licensing, implementation etc.) are known
No attribution possible The plan or design of the study and the quality of the data or monitoring are not sufficiently robust to support attribution	**Unknown** Impact cannot be measured or unintended impact is identified	**Unknown** Insufficient documentation exists on the initiative	**Unknown** There is no or limited information on the cost

Our common language

Guiding questions

1. How are our students doing in their learning?
2. How do we know?
3. What are we doing to improve students' learning?
4. How do we know it is working?
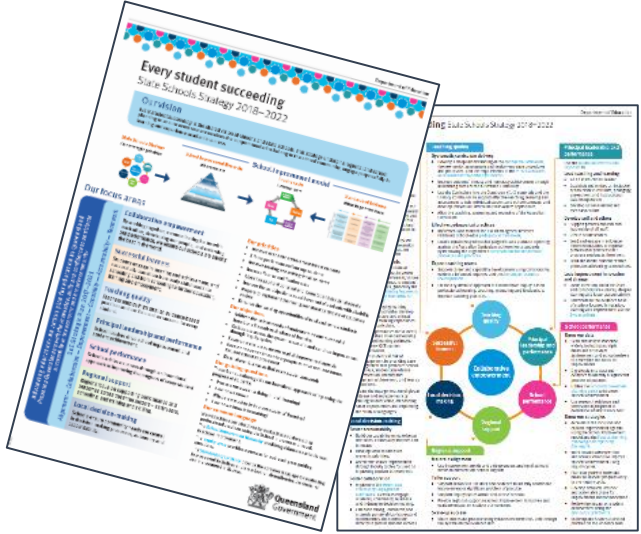
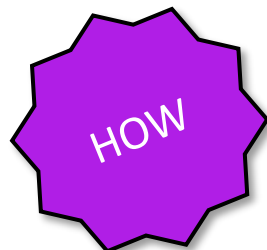# Queensland state schools improvement agenda

WHY

Every student succeeding

HOW

- ✓ Engage in inquiry
- ✓ Apply Standards of Evidence
- ✓ Access local evidence
- ✓ Draw on research
- ✓ Discuss evidence

WHAT

Evidence-informed continuous improvement

Dimensions in detail

Design
Quality of attribution

The **Design** dimension describes the extent to which any measurable change can be attributed to an initiative.

Impact
Measurable change

The **Impact** dimension is concerned with the strength of measurable change.

Scalability
Potential to implement

The **Scalability** dimension indicates the extent to which an initiative could be/has been reproduced with similar or greater success across other settings or groups.

Investment
Creating value

The **Investment** dimension is concerned with return on investment based on the evidence of cost-effectiveness or value for money.

The Standards of Evidence

Design

Quality of attribution

Verified attribution

An independent source or sources verifies the demonstrated measurable change can be attributed to the initiative

Attribution

Evidence from a high-quality study or studies establishes the measurable change can be attributed to the initiative

Possible attribution

A reasonable explanation supports attribution of the measurable change to the initiative

Attribution unclear

A logical plan describes the initiative and its objectives, target and comparison groups, inputs, outputs, outcomes and timelines. It includes the measures and methods for collecting, storing and analysing data

No attribution possible

The plan or design of the study and the quality of the data or monitoring are not sufficiently robust to support attribution

- What was the research question/focus? How is this relevant to our priority or problem of practice (and context)?

- (Precisely) How was it implemented? Who was involved and what were their attributes? What were the inputs – structures, processes, resources?

- What were the measures of success – implementation and outcomes? How did they monitor implementation and change?

- Was the methodology robust? (sample size, comparison data, analyses) How was research selected/rejected for meta analysis?

Impact
Measurable change

Very high
Very large measured improvement

High
Large measured improvement

Moderate
Medium measured improvement

Low
Small measured improvement that can be reasonably linked to the initiative

Unknown
Impact cannot be measured or unintended impact is identified

Did the research/review/meta analysis report a measurable difference to (targeted) student outcomes?

Did they supply baseline and endpoint data? Can we check their analyses, interpretations and conclusions?

How did they describe their claims of strength of impact?

What assumptions and predictions (bias) can we detect? Can we accept them?

How will we confirm/triangulate across multiple measures?

How will we determine the strength of impact?

Scalability
Potential to implement

Larger scale implementation

An independent source or sources verifies the initiative has been implemented by others on a larger scale

Implemented elsewhere

The initiative can be implemented in more than one setting and/or with different groups with at least the same impact. Documentation supports implementation

Well documented

Documentation exists about how the initiative is implemented and is comprehensive enough to enable implementation with different groups or in other settings

Could work

Documentation describing the implementation of the initiative in one setting exists

Unknown

Insufficient documentation exists on the initiative

- Was this conducted in one or more settings?

- How much information is available? Enough to implement (with 'fidelity')?

- Was there any independent verification of implementation at scale, with impact?

$Investment

Creating value

Cost-effectiveness known

Information on cost-effectiveness exists

Outcomes costed

Information on the costs per unit of outcome exists

Outputs costed

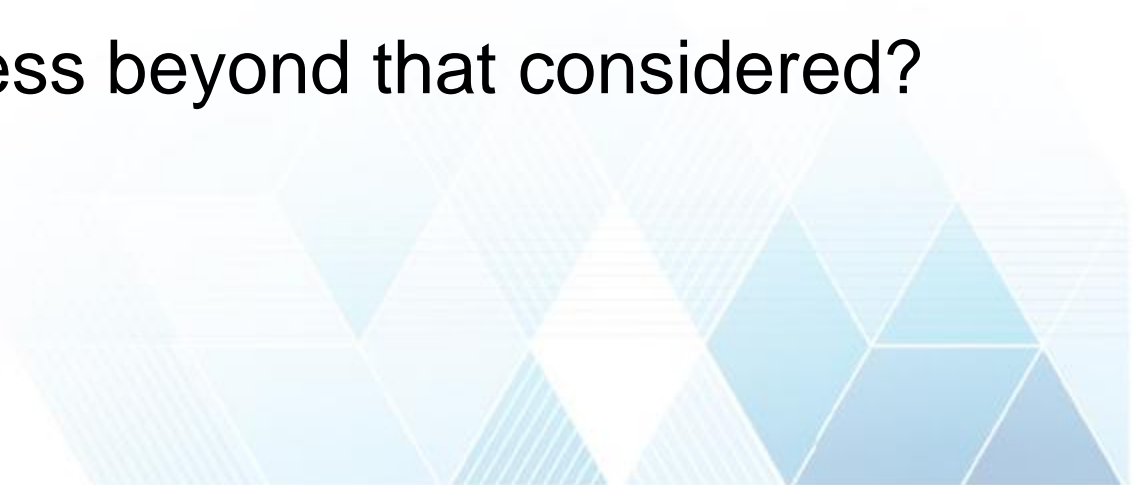Information on the costs per unit of output exists

Costs known

The costs of the initiative (for example, financial, human resources, infrastructure, program licensing, implementation etc.) are known

Unknown

There is no or limited information on the cost

- Were the resources described/listed/detailed?

- Were they costed?

- Was cost per student determined?

- Was cost-effectiveness beyond that considered?

Design

Quality of attribution

Verified attribution

An independent source or sources verifies the demonstrated measurable change can be attributed to the initiative

Attribution

Evidence from a high-quality study or studies establishes the measurable change can be attributed to the initiative

Possible attribution

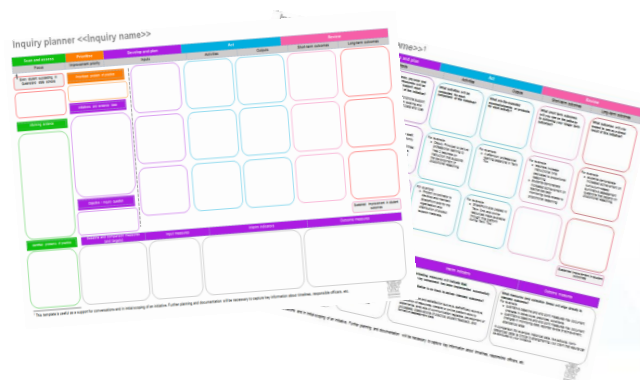A reasonable explanation supports attribution of the measurable change to the initiative

Attribution unclear

A logical plan describes the initiative and its objectives, target and comparison groups, inputs, outputs, outcomes and timelines. It includes the measures and methods for collecting, storing and analysing data

No attribution possible

The plan or design of the study and the quality of the data or monitoring are not sufficiently robust to support attribution

- What is the **problem of practice** being addressed? What was the **evidence** that led this being prioritised?

- What is your **improvement initiative**? What is its objective? How will be implemented? Who and what will be involved?

- What are the **measures of success**? How will the initiative be **monitored** to gauge progress? What were the interim measures? Was there a logical chain of evidence?

- Is there **alignment** between the problem of practice, the improvement initiative and the measures of success?

Inquiry planner <<inquiry name>>

Scan and Assess

Review

Prioritise

Act

Develop and Plan

Impact
Measurable change

Very high
Very large measured improvement

High
Large measured improvement

Moderate
Medium measured improvement

Low
Small measured improvement that can be reasonably linked to the initiative

Unknown
Impact cannot be measured or unintended impact is identified

How will we know if the initiative making a difference to **student outcomes**?

Do we have **baseline** and **endpoint data**? How will we **analyse** (and interpret and represent) data?

What assumptions and predictions (bias) do we have?

How will we confirm/triangulate across **multiple measures**?

How will we determine the **strength** of impact?

Scan and Assess

Review

Prioritise

Act

Plan

Scalability

Potential to implement

Larger scale implementation

An independent source or sources verifies the initiative has been implemented by others on a larger scale

Implemented elsewhere

The initiative can be implemented in more than one setting and/or with different groups with at least the same impact. Documentation supports implementation

Well documented

Documentation exists about how the initiative is implemented and is comprehensive enough to enable implementation with different groups or in other settings

Could work

Documentation describing the implementation of the initiative in one setting exists

Unknown

Insufficient documentation exists on the initiative

- What will we need to do/document to ensure sustainability, allow scaling (or inform adjustments)?

- What could we consider – RCTs, clusters, partners, external evaluators?

- How do we verify? How do we share?

- How do we support others to 'replicate' in context?

Investment

Creating value

Cost-effectiveness known

Information on cost-effectiveness exists

Outcomes costed

Information on the costs per unit of outcome exists

Outputs costed

Information on the costs per unit of output exists

Costs known

The costs of the initiative (for example, financial, human resources, infrastructure, program licensing, implementation etc.) are known

Unknown

There is no or limited information on the cost

- What **resources** (HR, financial, IT, other?) will be required to implement the improvement initiative?

- Is there an understanding of the **total cost** of this improvement initiative?

- How will we keep track of the costs directly related to this initiative?

- How could we calculate the cost of this improvement initiative **per student or teacher** involved?

- What more would we need to do to determine the cost per outcome?

Our invitation to you

- Positive, potential, promise

- Concerns, constraints, cautions

- Interesting, blind spots, improvements, applications

Documenting our learning
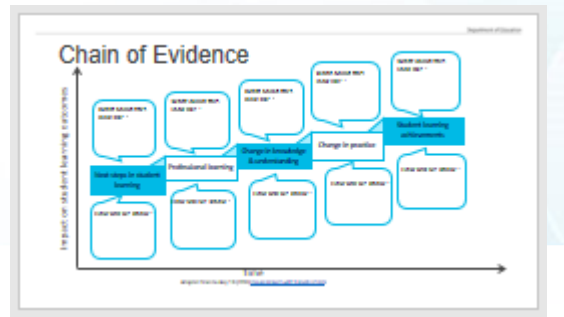
Inquiry planners

Actionable playbooks

The SIM GPS

Generated local evidence

Conversation & calibration tools

Evidence chains
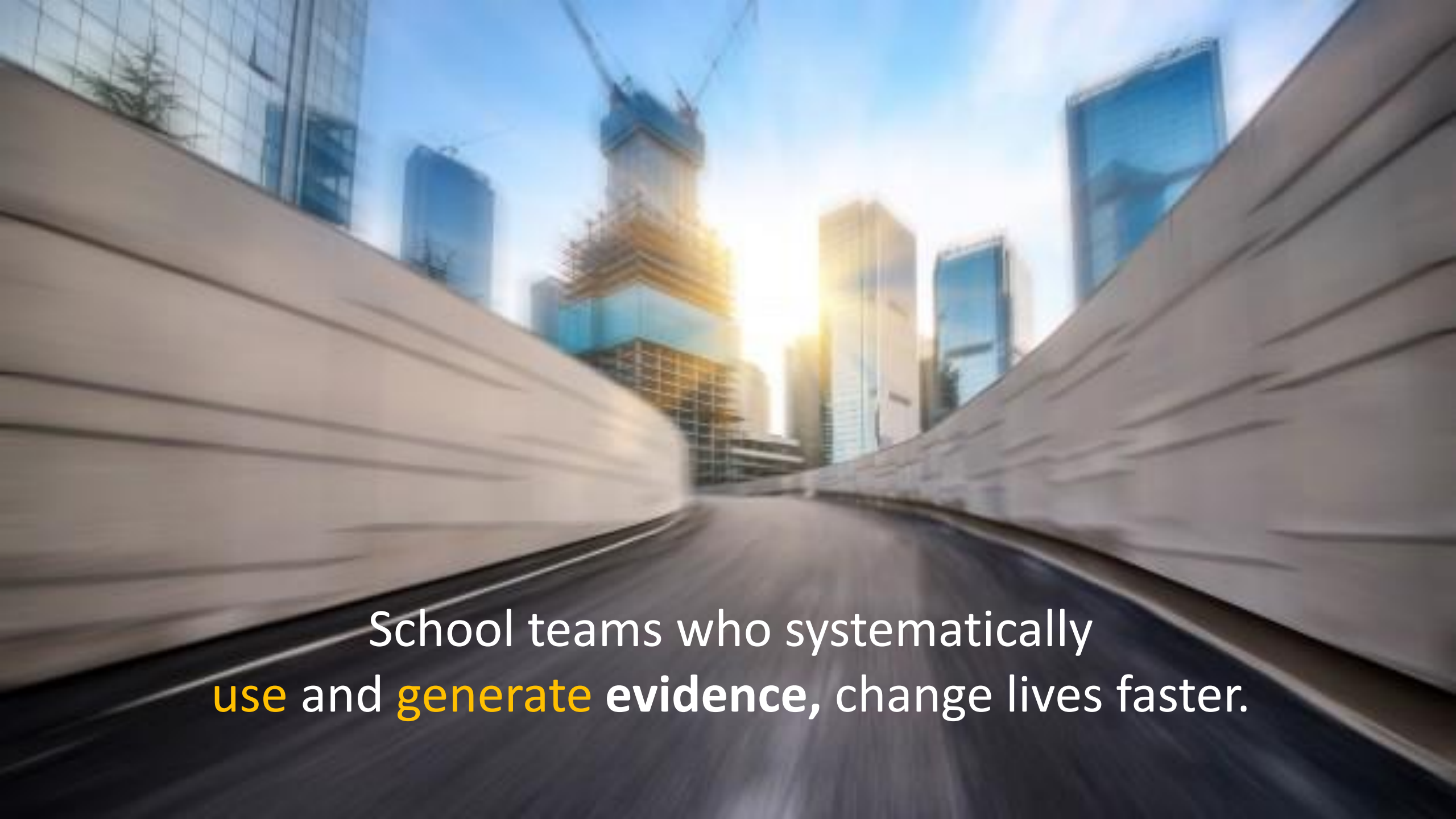
Dr Deb Kember, Director, School Improvement

Deb.kember@qed.qld.gov.au

Deb Fisher Director School Improvement

Deb.fisher@qed.qld.gov.au

State Schools Division

Department of Education, Queensland

School teams who systematically
use and generate **evidence,** change lives faster.