

Evidence based policy: Why is progress so slow and what can be done about it?

APO 15th
Anniversary

28th Nov, 2017

Nicholas Gruen

ngruen@lateraleconomics.com.au

@ngruen1



Lateral Economics

Outline

1. Introduction
2. Arteries and capillaries
3. Thick and thin problems
4. Evidence based programs
5. Evaluation
6. Program logic
7. Accountability
8. Institutionalising evidence-based policy

1 Introduction



Evidence-based policy

New
Perfor
Imple
progr
NSW T
Depart
Depart
Depart
Depart

Conclusion

The NSW Government's program evaluation initiative is largely ineffective, as it is not providing sufficient information to government decision makers on the performance of programs.

No information is provided on the performance of programs that have been evaluated. The information that is provided is limited to a list of programs being evaluated in the upcoming financial year, with little assurance that the right programs are on the list. NSW Treasury and DPC are not using evaluation outcomes to analyse agency funding proposals in their advisory role to the NSW Government.

For program evaluation to be effective, agencies should demonstrate they are evaluating the right programs, and the outcomes from completed evaluations should inform advice to the NSW Government on investment decisions.

dit
ce
H WALES



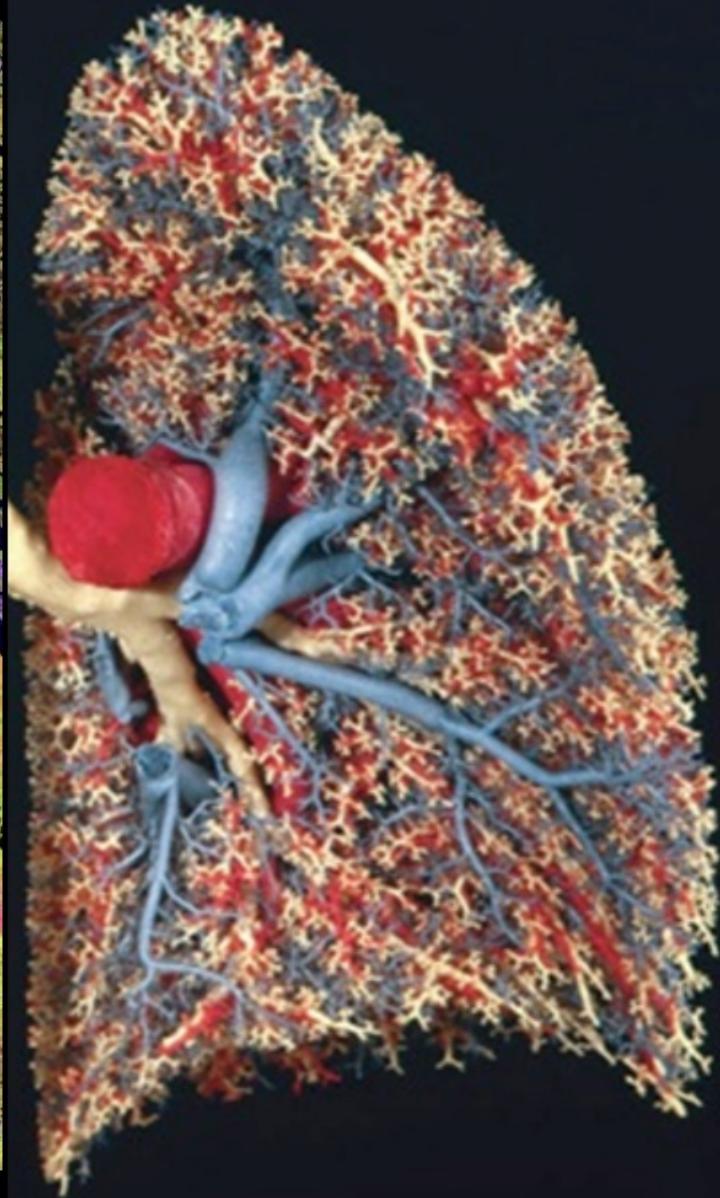
@ngruen1



Lateral Economics⁵

2 Arteries and capillaries

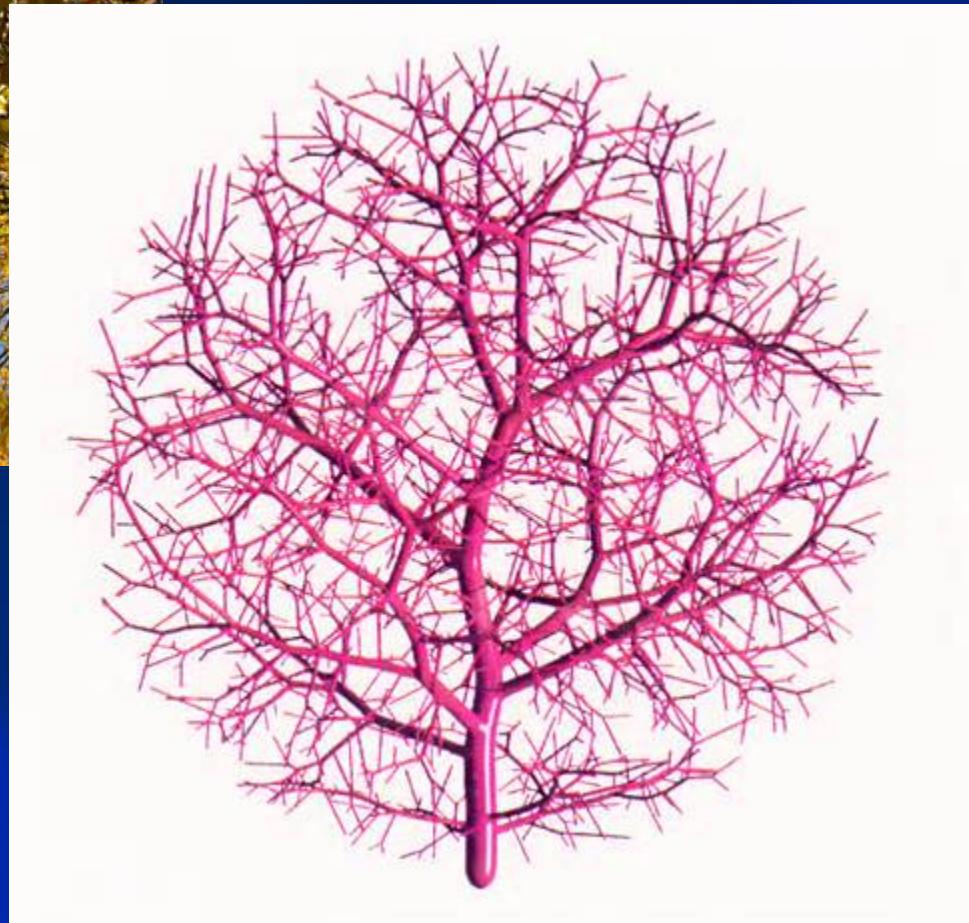




@ngruen1



Lateral Economics



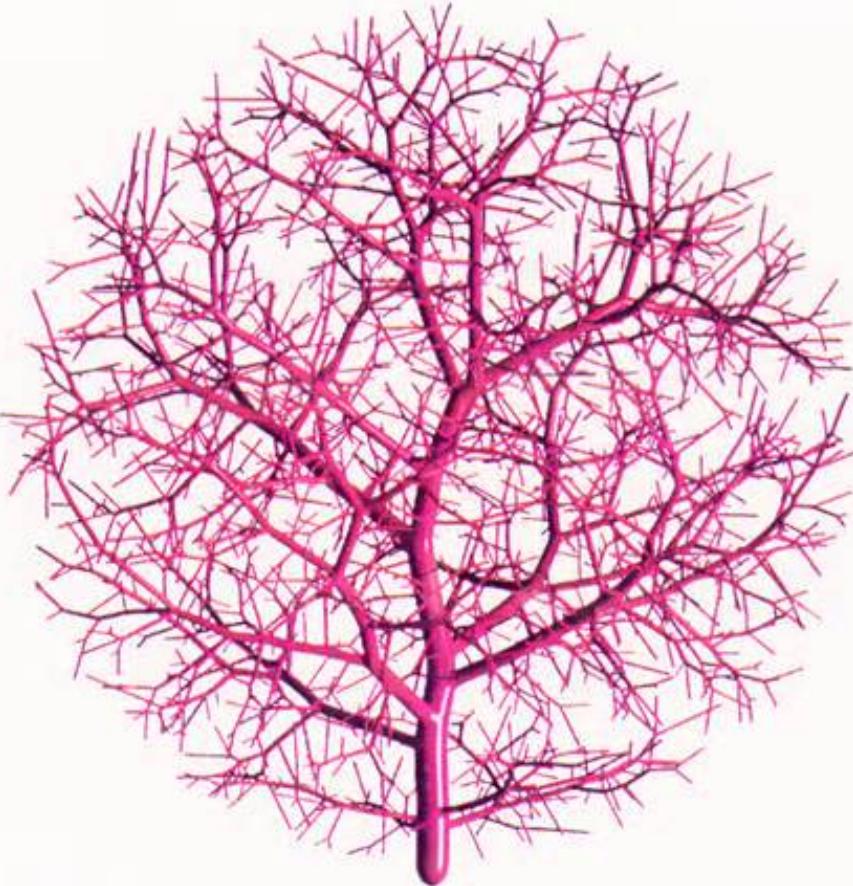
@ngruen1



Lateral Economics

The human world is ...

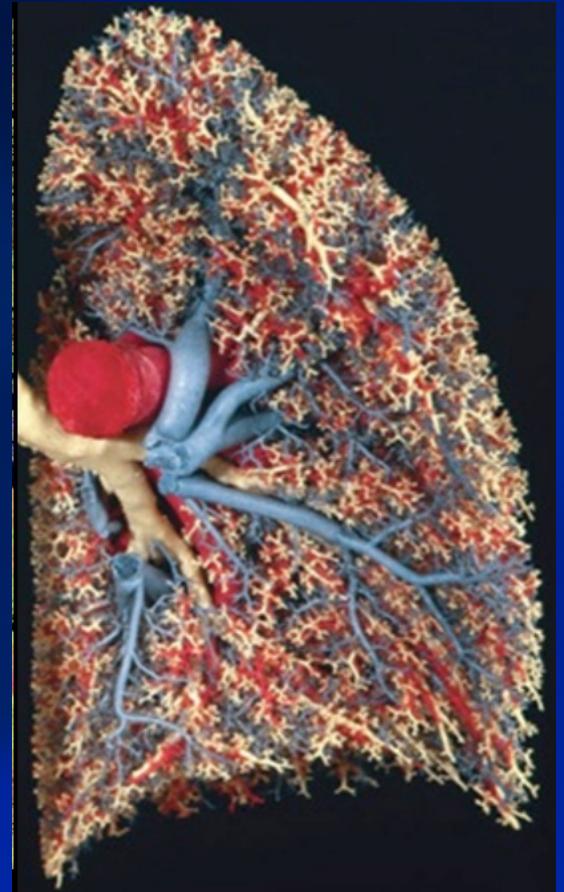
A hierarchy of trunks and branches
Arteries and capillaries



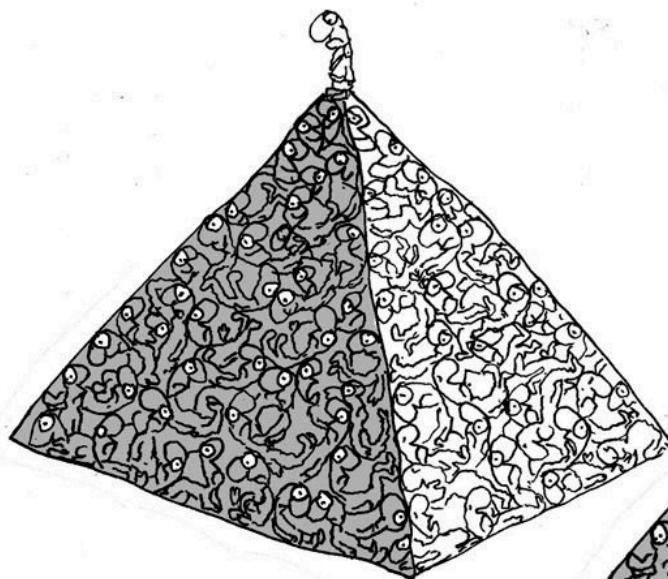
This could be a

- Software program
- Profession or discipline
- Industry
- Encyclopedia
- Catalogue
- Org chart
- Corporate accounts
- Corporate KPIs

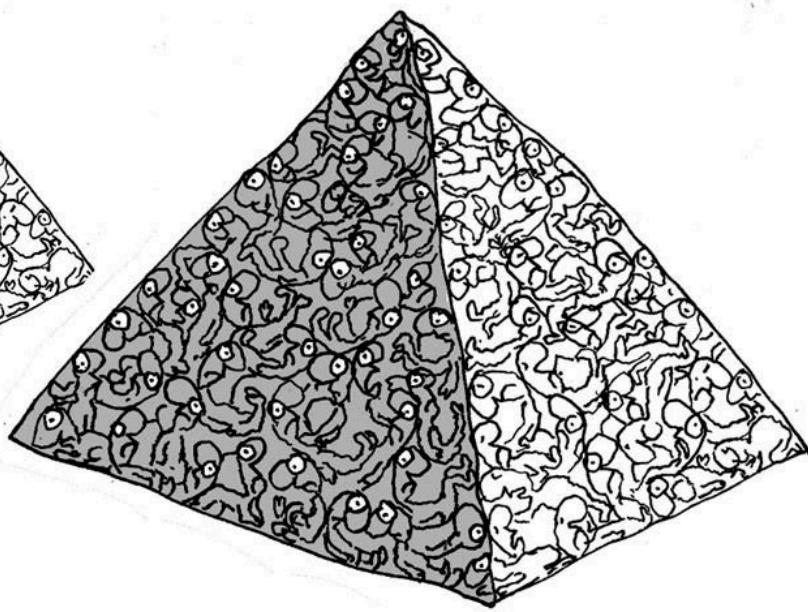
Democracies are hierarchies



THE ANCIENT PYRAMIDS



WITH DICTATOR.



DICTATOR REMOVED.

Leunig

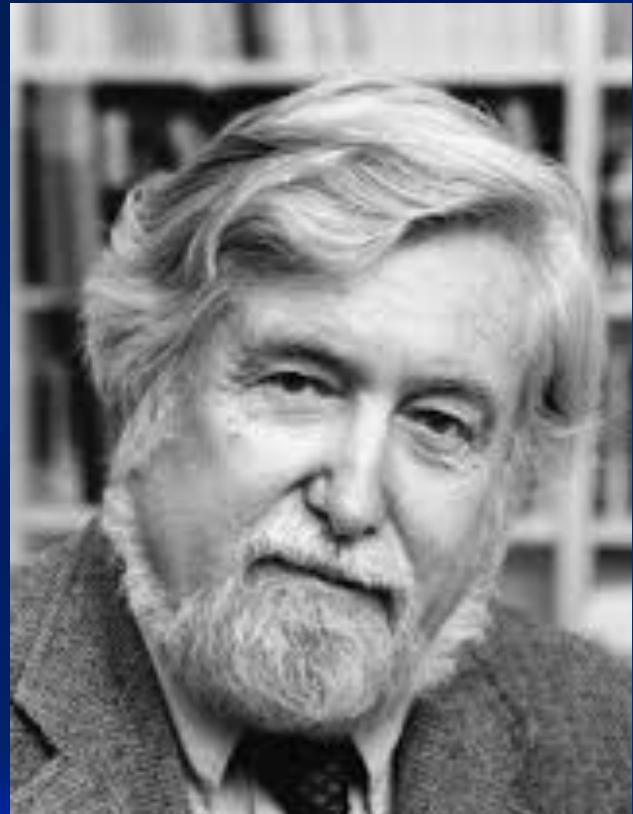
3: Thick, thin: understanding of context and motives



Thick and thin



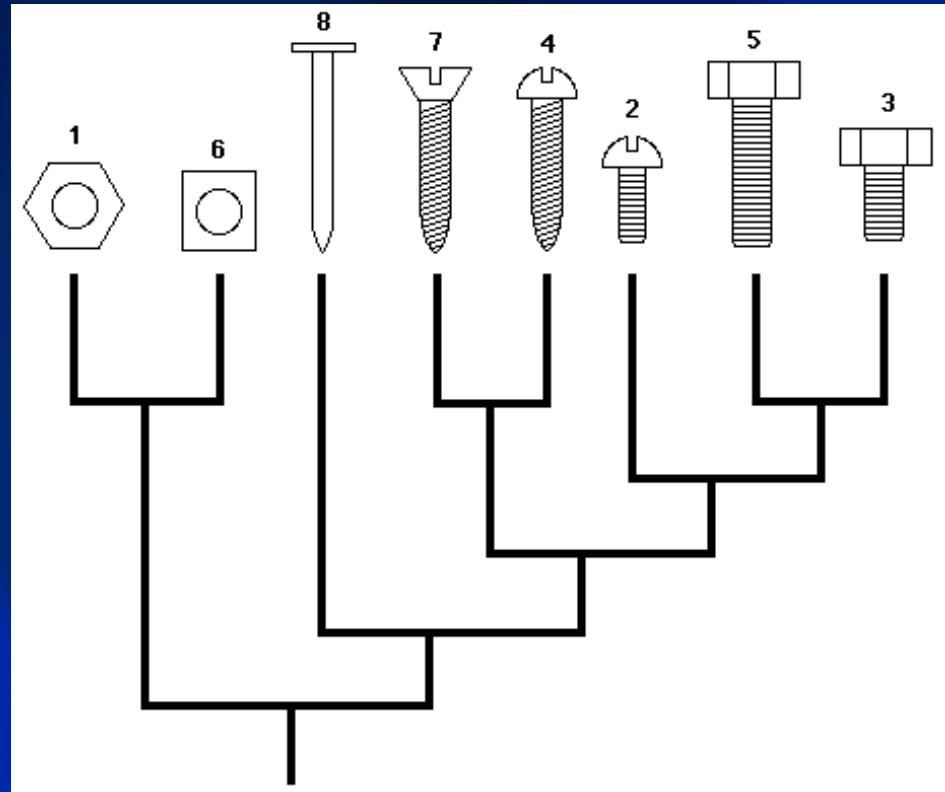
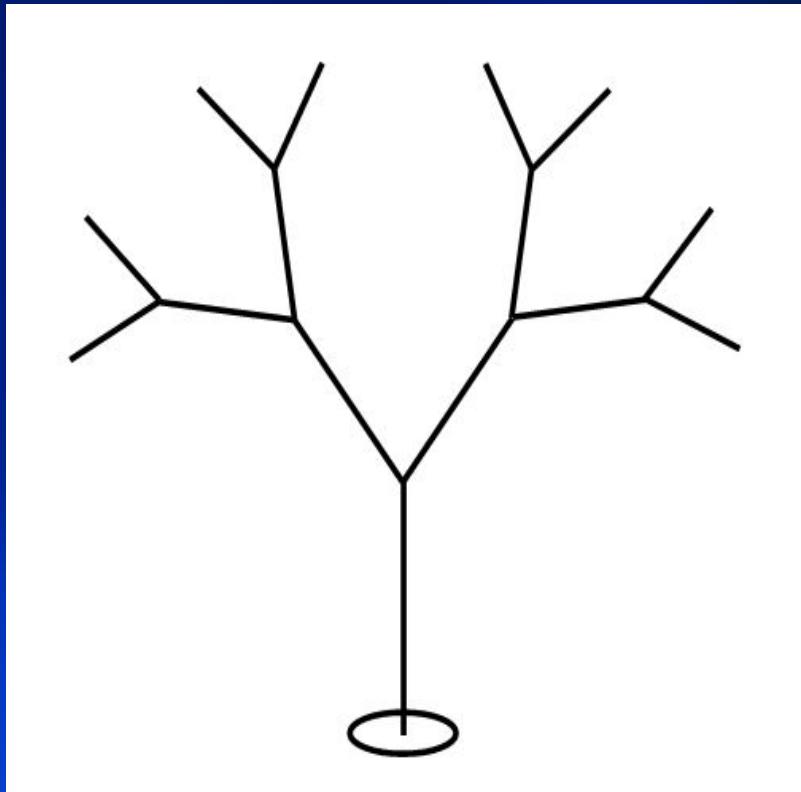
Gilbert Ryle



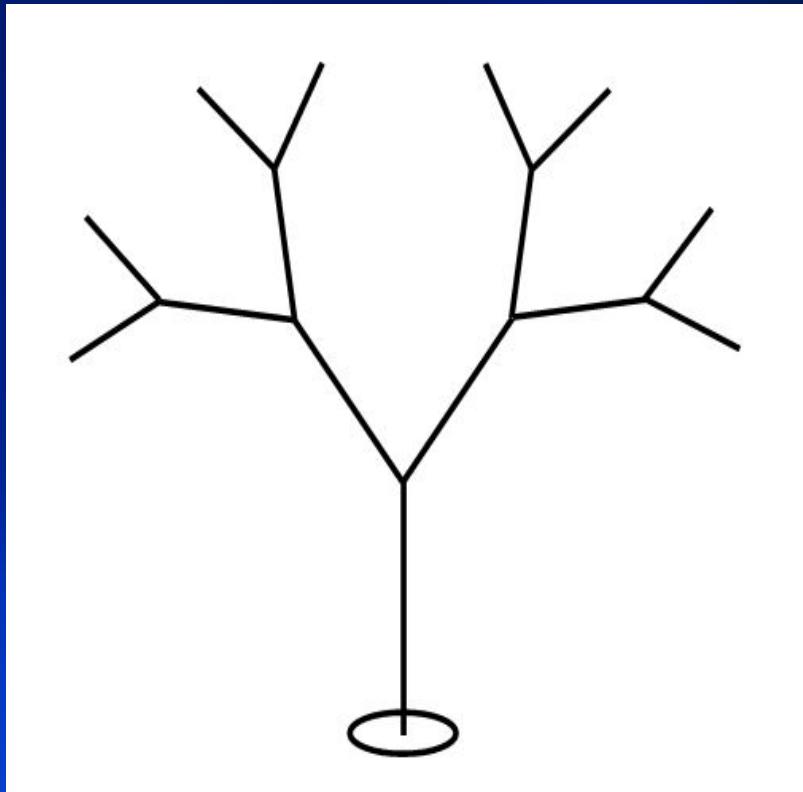
Clifford Geertz

Thin to thick
From what to why

Thin problems are mechanical



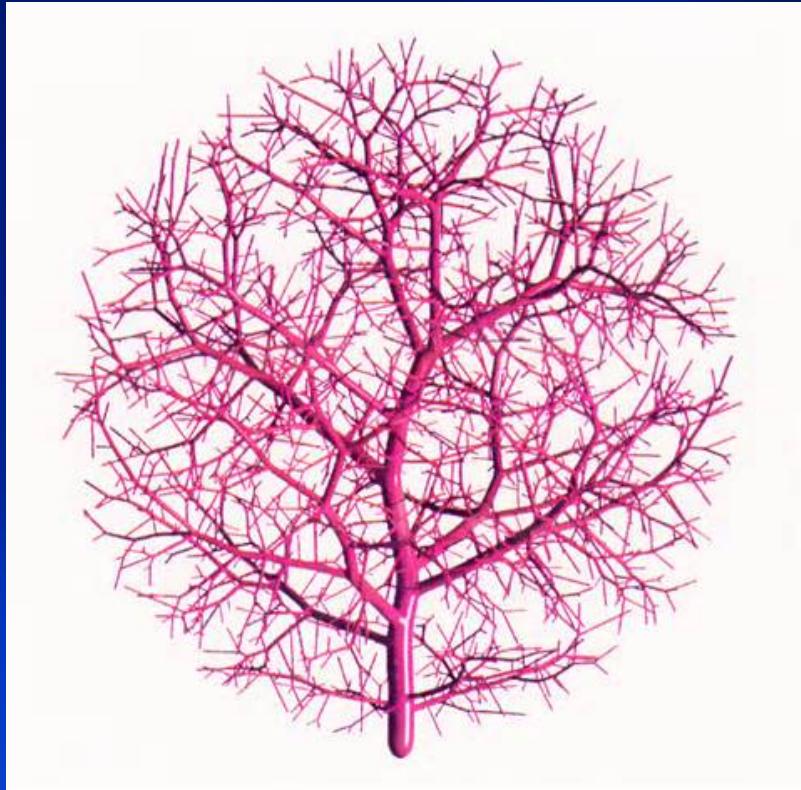
Thin policy delivery



Top-down policy can work

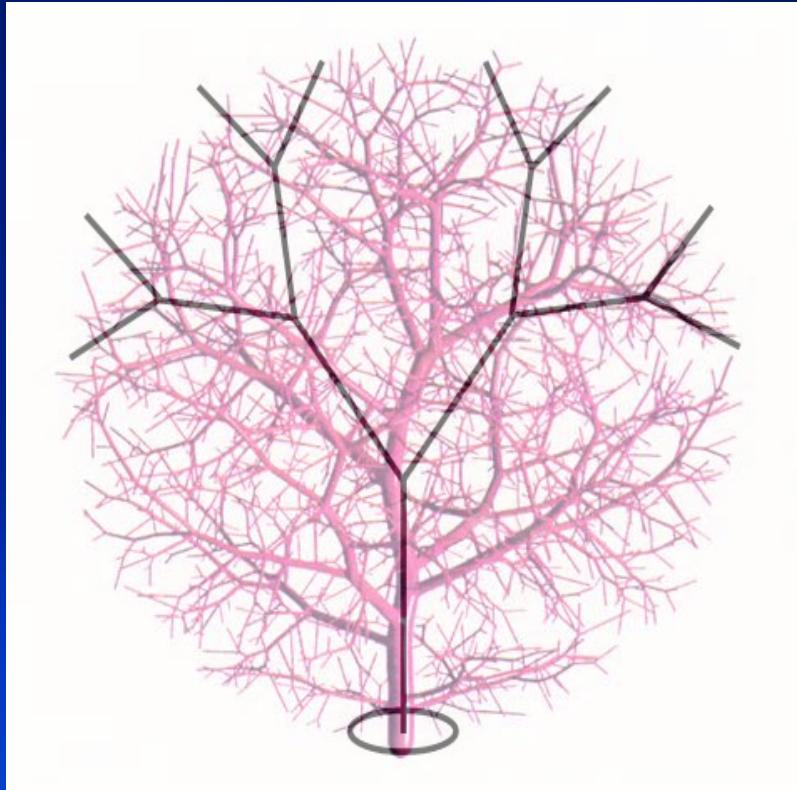
- Tax and family benefits changes
- Student loans
- Child Support Agency
- Stroke of the pen deregulation
 - shopping hours, two airline policy

Thick policy delivery



- Regulation
- IT
- Education
- Health
- Defence
- Transport
- Employment services
- Social support (Indigenous and other)

Ideologies are thin, issues are thick



Income management

User Choice

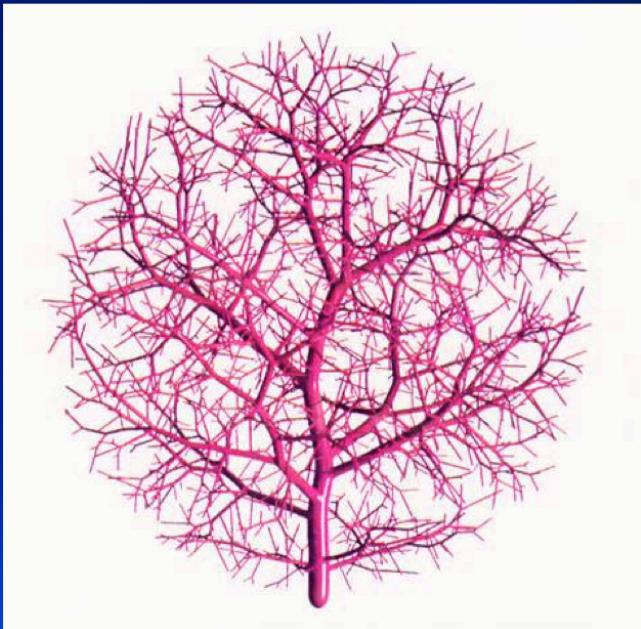
Diversity

Core values

Individual responsibility

Collective responsibility

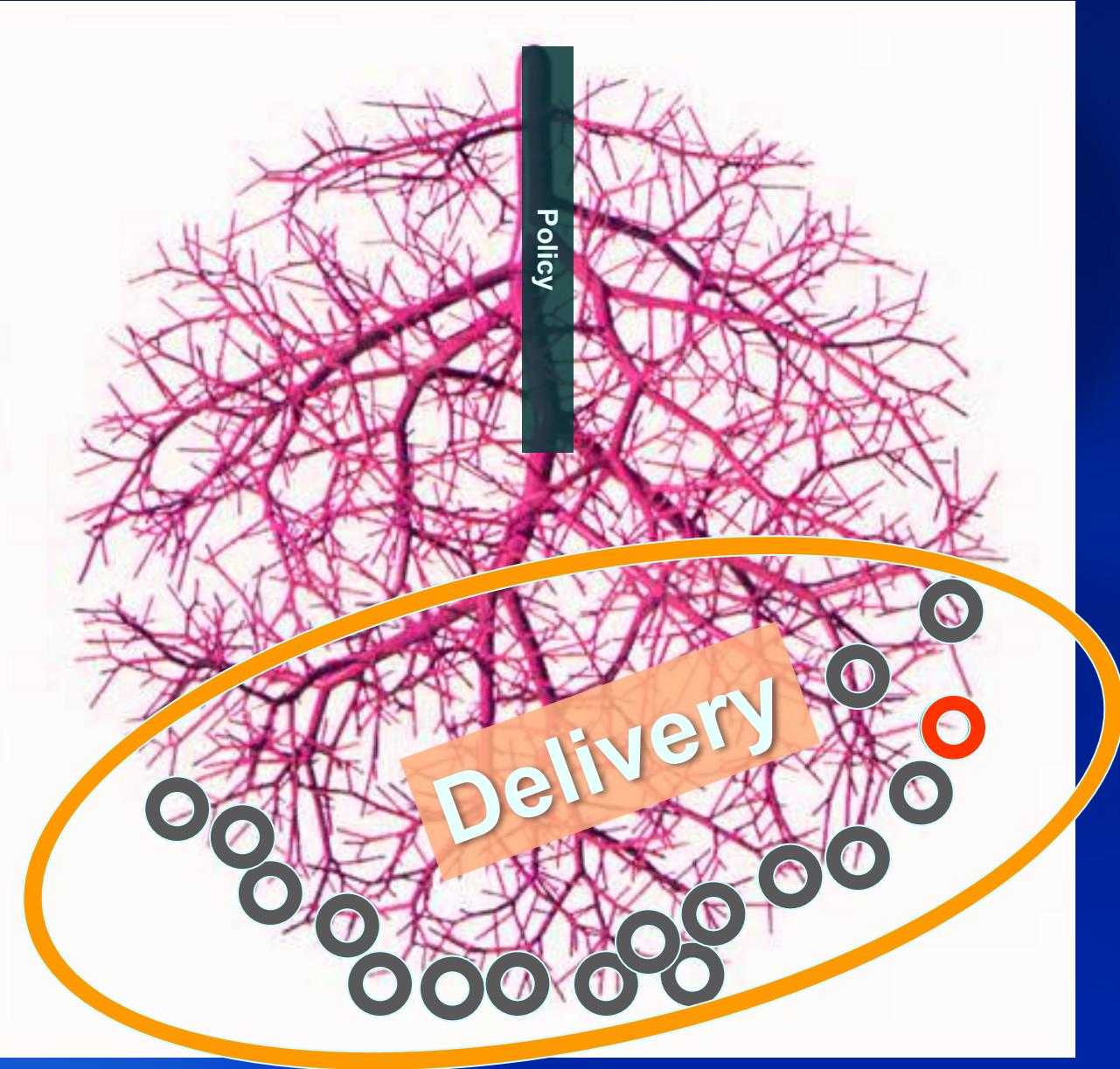
Arteries, capillaries and status



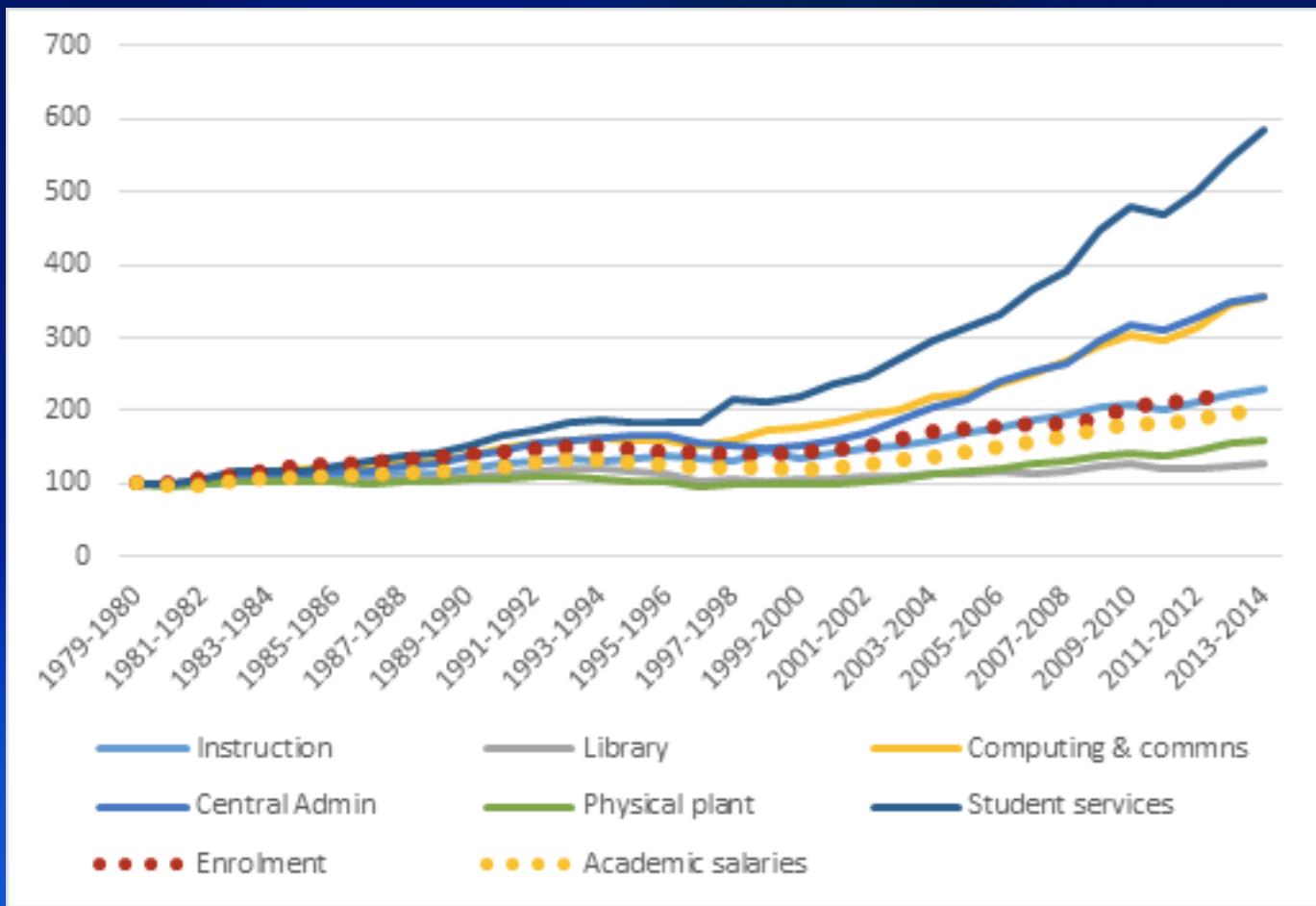
The ambitious know full well that the road to the top is through policy, generating ideas, managing the blame game, being visible in Ottawa circles, and central agencies, not through program management.

Donald J. Savoie “What Is Government Good At? A Canadian Answer” 2015

<= Learning goes upward



Academia



The arteries are willing, but the capillaries are weak

NSW Audit Office on Reg Review

Regulatory burden increased

Over the life of the ‘one-on, two-off’ initiative overall net legislative regulatory burden increased by \$16.1 million.

The numeric test was met with 237 instruments repealed and 54 introduced — an overall ratio of roughly four repeals for every new instrument. However, most of these repeals related to redundant legislation with little or no regulatory burden.

Legislative complexity increased

The stock of legislative regulation increased.

By 1.4% per year compared with 1.1% falls previously.

IT

Endless policy cycles and revisions accrue.

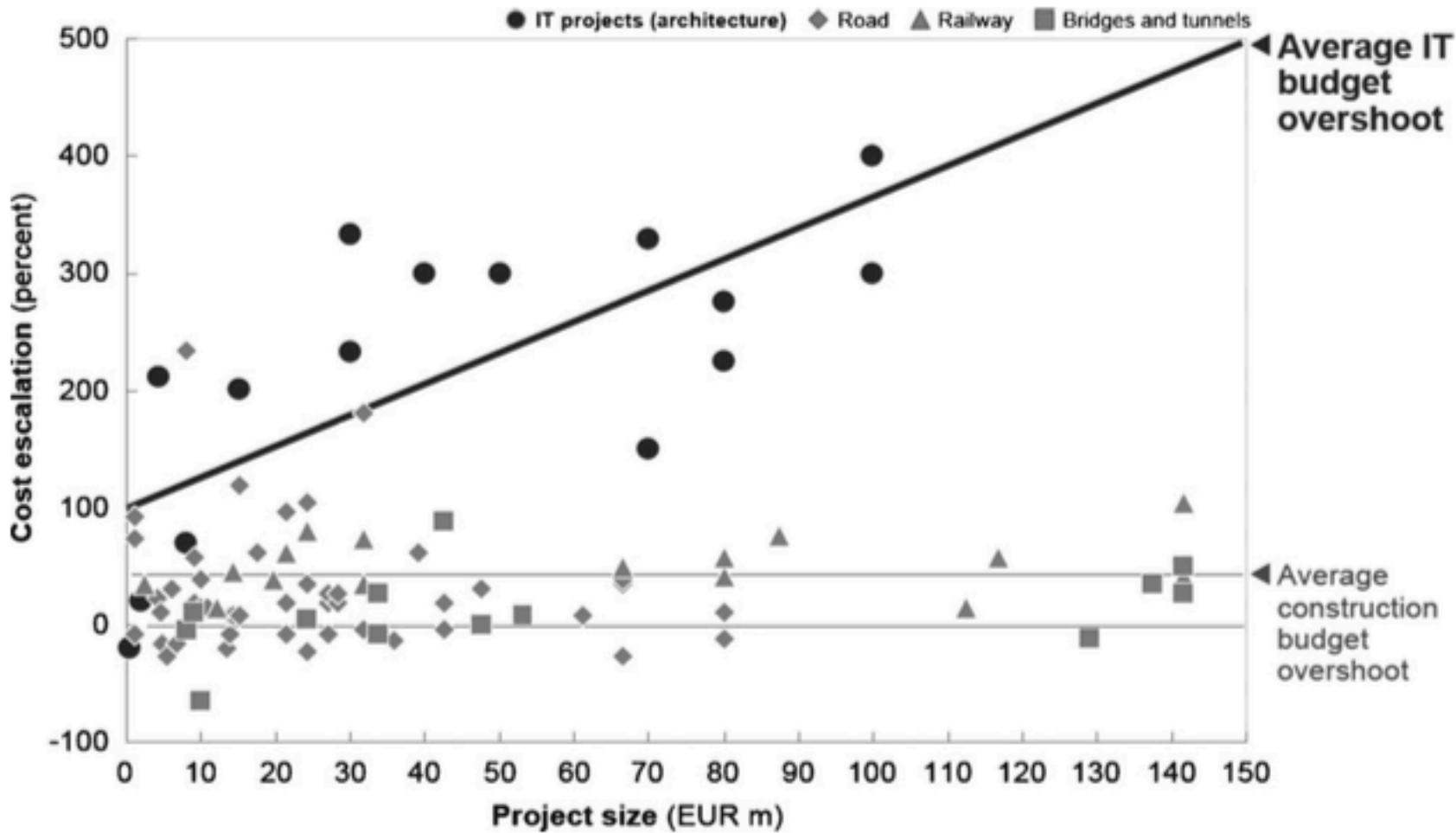
Subs to Ministers, private office communications, correspondence across departments and occasional harvesting of consultation feedback.

Rarely ... does user need get a look-in except internal users.

How the departmental needs can so often trump the needs of public users is beyond me.

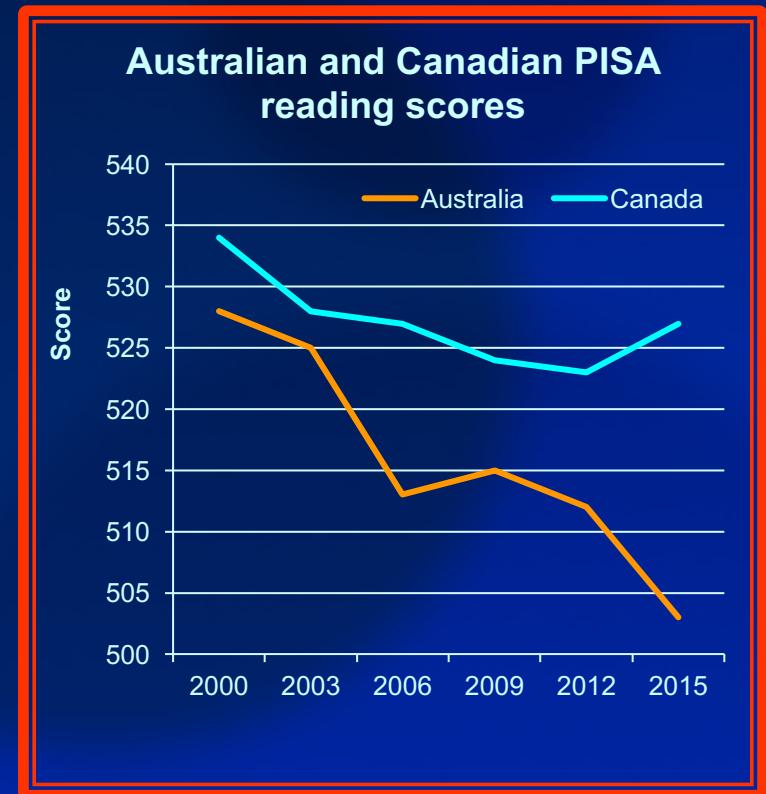
Mike Bracken, The strategy is delivery, UK

Figure 4: Cost overrun in construction projects and IT projects compared.



Schools

Cost of divergence
between Australia and
Canada according to
the HALE index of
wellbeing is
\$17 bil of human
capital each year



Health: micro-detail as a thicket

In our report, you can read how hospitals are required to sign up to IP restrictions preventing data transfer between wards.

Or how cancer researchers use foreign data sets because local ones are more restricted.

Or how a nationally-funded research project into vaccination is nearly 7 years into a saga to be allowed access to Commonwealth and States' data sets. It expects to be finally allowed full access in another year or so.

These are pretty disgraceful events.

They are the tip of the iceberg.

Peter Harris, Chairman, Productivity Commission, 2017



The cult of announceables

@ngruen1



Lateral Economics²⁹

Further Education



48 Secretaries of State

Responsible for Further Education since 1981

13,000 qualifications

In the Further Education space, as noted by Philip Hammond in Budget 2017

Regional government



17 years

Government Offices, 1994-2011

12 years

Regional Assemblies, 1998-2010

1 year

Local Authority Leaders' Boards, 2009-10

1992

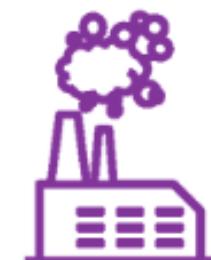
'New' unitary authorities created

2009

'New' unitary authorities created

Source: Phil Swann, 2016

Industrial strategy



2008

Mandelson's 'new activism'

2012

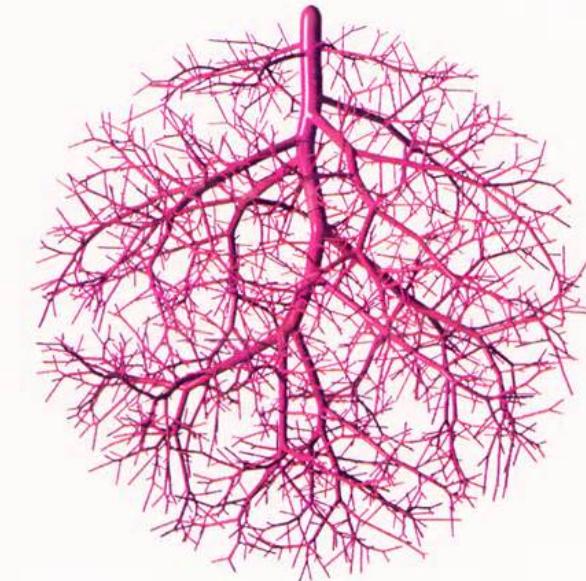
Cable's industrial strategy

2015

Javid's 'industrial approach'

2016

May's industrial strategy

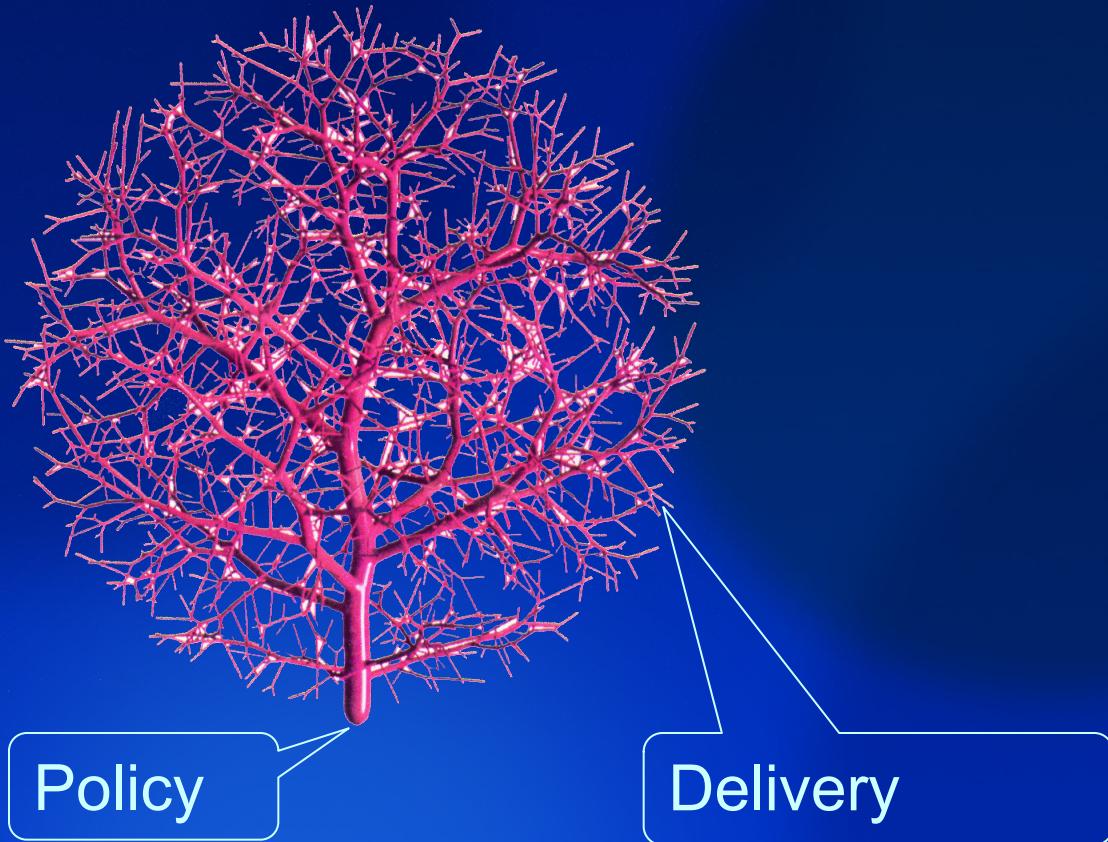




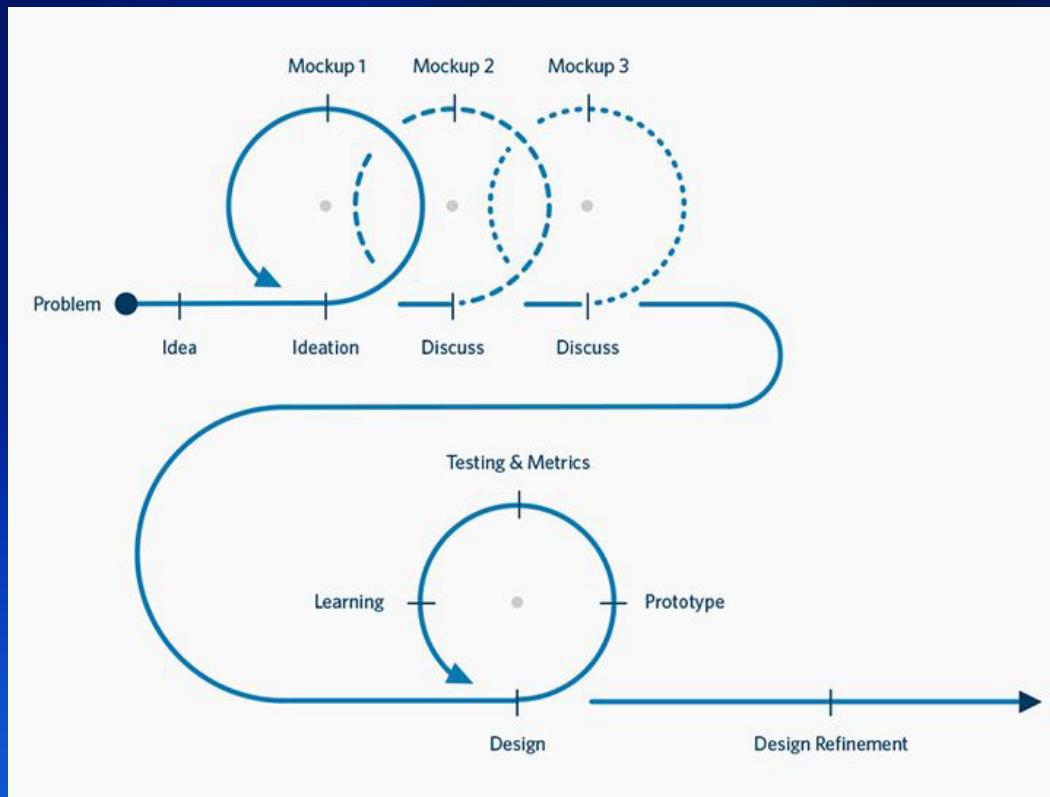
In 50 years, Commonwealth administration of Indigenous Affairs has cycled through 21 different ministers, and 11 different structures under them. Ten of the 11 structures have occurred in the last 30 years.

4 Evidence based programs

Evidence based *delivery*



Design: evidence based delivery



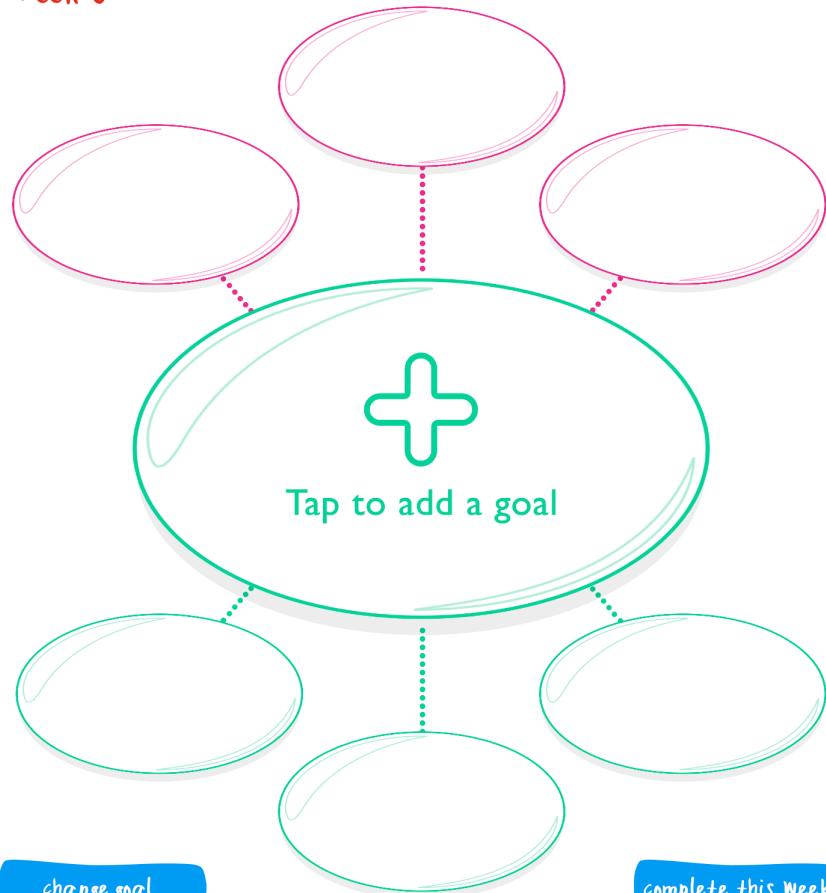
Empathic bond





my bubbles
week 0

The Tester Family



I've worked with that family for 3 years and I just learnt more about them in 2 hours.

Case worker

Families commented:
“you’re the only one who has ever asked what would work for my family”.

Family coach



THE AUSTRALIAN CENTRE FOR SOCIAL INNOVATION
BOLD IDEAS. BETTER LIVES.

PRESENTS

@ngruen1



Lateral Economics³⁷



@ngruen1



Lateral Economics ³⁸

5 What evidence?

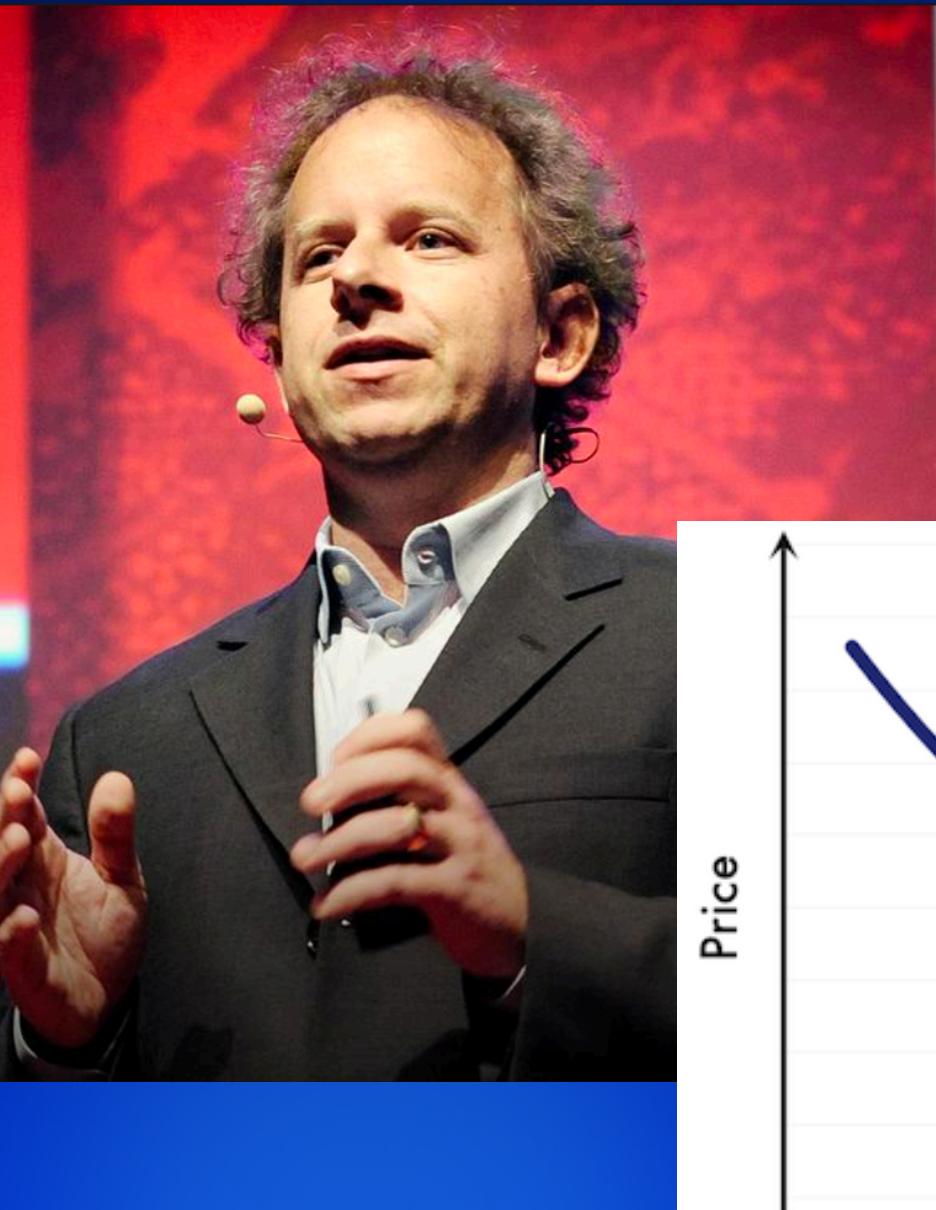
Program efficacy depends on causality

Generalisability:

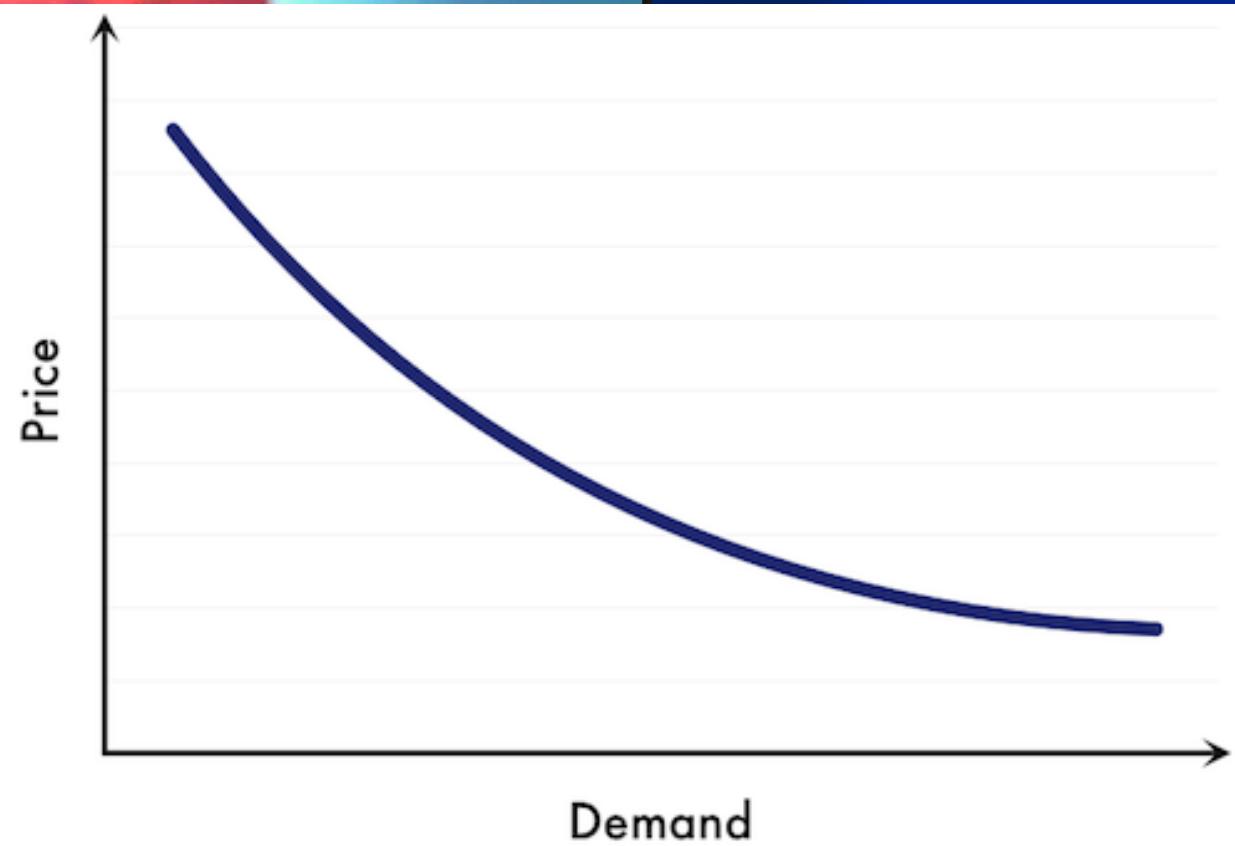
- It works somewhere – from an impact evaluation
- It works in general – from synthesis of a range of impact evaluations
- It will work for us – is a question of judgement; the potential that it will work for us depends on the context in which it is implemented and the quality of implementation.

Family by Family – 6 week scoping in new suburbs

- To optimise efficacy and
- Test for validity



What kind of
data do we
need?



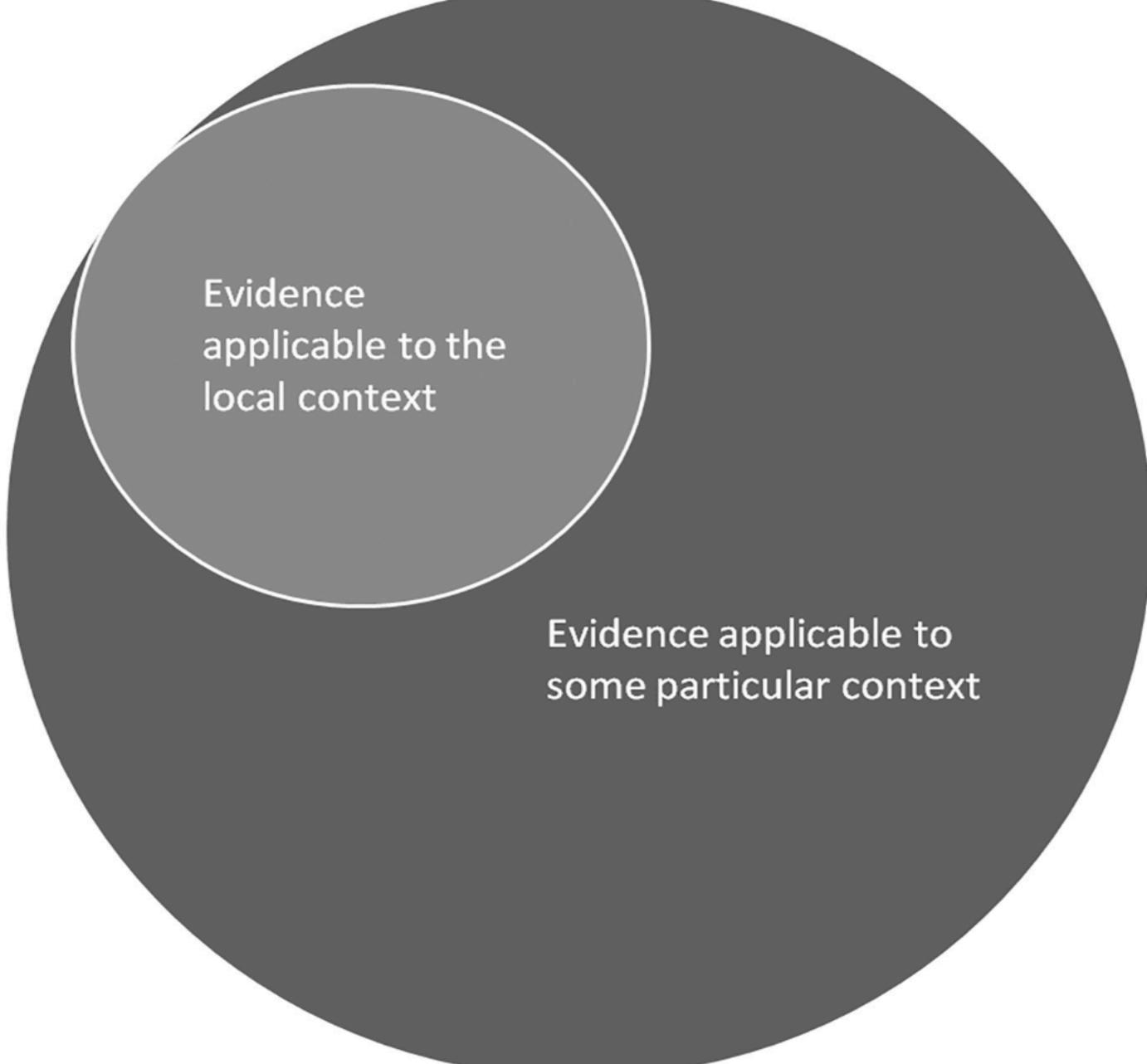
Data on what *causes* what

“Our success at Amazon is a function of how many experiments we do per year, per month, per week, per day....”

Jeff Bezos

“Last year at Google the search team ran about 6,000 experiments and implemented around 500 improvements based on those experiments. The ad side of the business did about the same. Any time you use Google, you are in many treatment and control groups. The learning from those experiments is fed back into production and the system continuously improves.”

Hal Varian, chief economist at Google



Evidence
applicable to the
local context

Evidence applicable to
some particular context

e
o the good

Figure 6.3 Evidence may be more or less applicable in the local policy context. als.

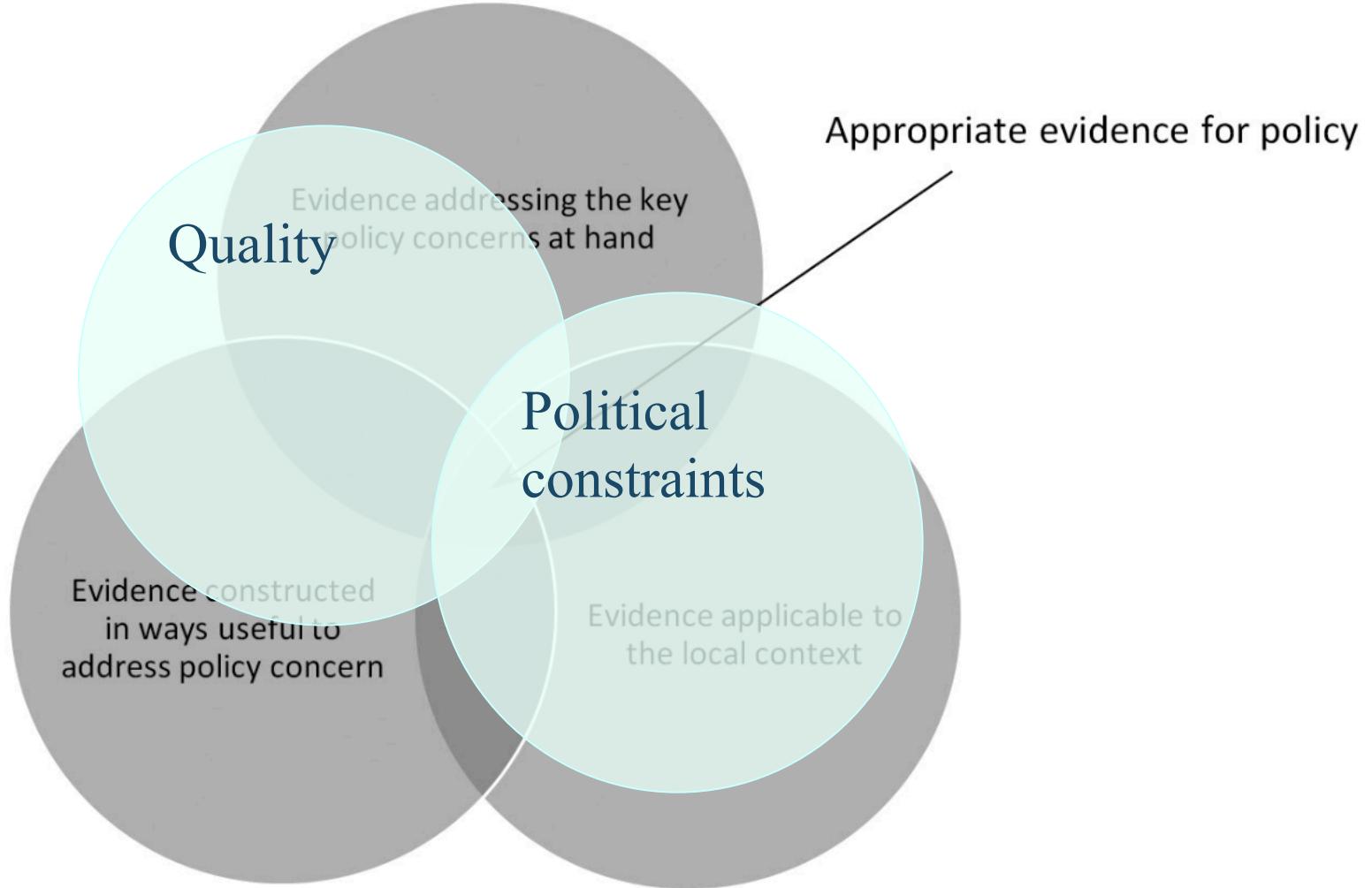


Figure 6.4 Appropriate evidence for policy context.

6 Program logic

Parkinson: Indigenous funding has weak evidence base

by
**David
Donaldson**
26.08.2016

[1 Comment](#)

PEOPLE

[Martin Parkinson](#)

DEPARTMENTS

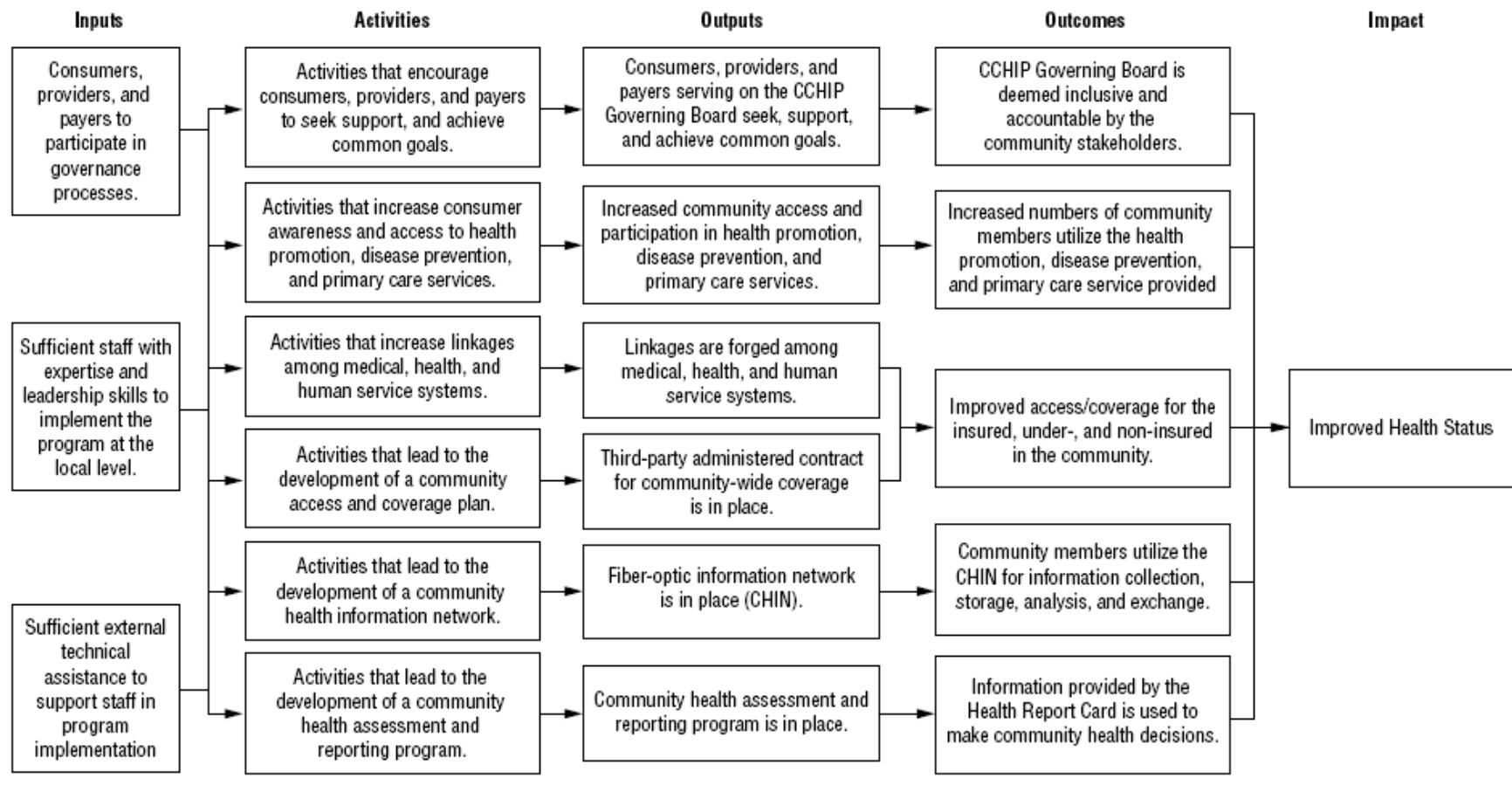
[Department of the
Prime Minister and
Cabinet](#)

TAGS

[Indigenous](#), [Martin
Parkinson](#), [Program
evaluation](#), [Evaluation](#)



Evaluation and partnerships are both in short supply in the Indigenous Affairs portfolio, warns the nation's top bureaucrat. Public servants need to learn to operate in the worlds of Indigenous people, not the other way around.

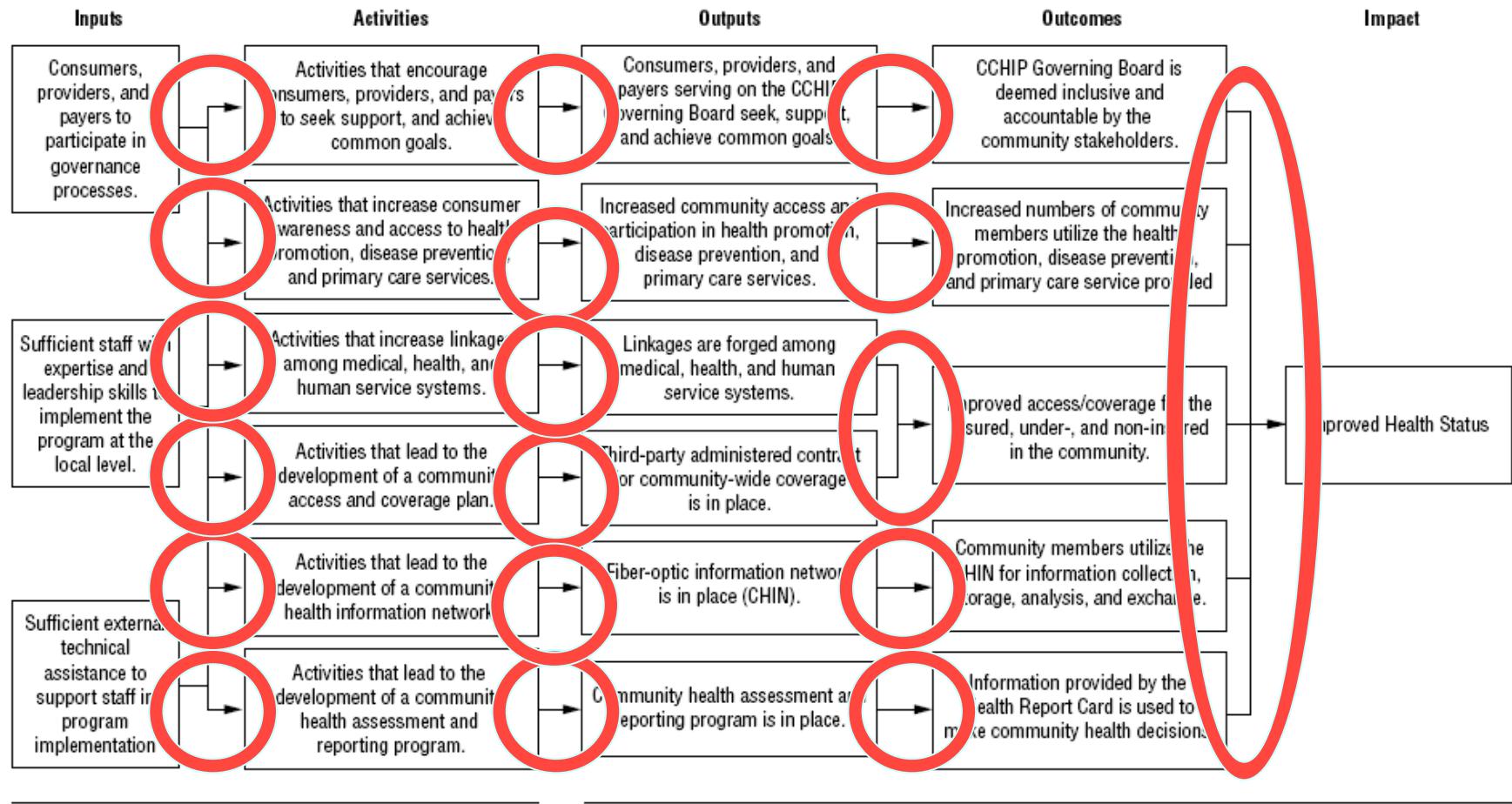


Your Planned Work



Your Intended Results

Example of an Outcome Approach model (example drawn from the Calhoun County Health Improvement Program, funded under the Comprehensive Community Health Models of Michigan initiative).



Your Planned Work



Your Intended Results

Example of an Outcome Approach model (example drawn from the Calhoun County Health Improvement Program, funded under the Comprehensive Community Health Models of Michigan initiative).

ROUTLEDGE STUDIES IN GOVERNANCE AND PUBLIC POLICY

The Politics of Evidence

From evidence-based policy to the good governance of evidence

Justin Parkhurst



One of the main conceptual holdovers from the world of evidence-based medicine has been the widespread, and often uncritical, embrace of so-called 'hierarchies' of evidence

Justin Parkhurst, *The Politics of Evidence: From evidence-based policy to the good governance of evidence* 2017

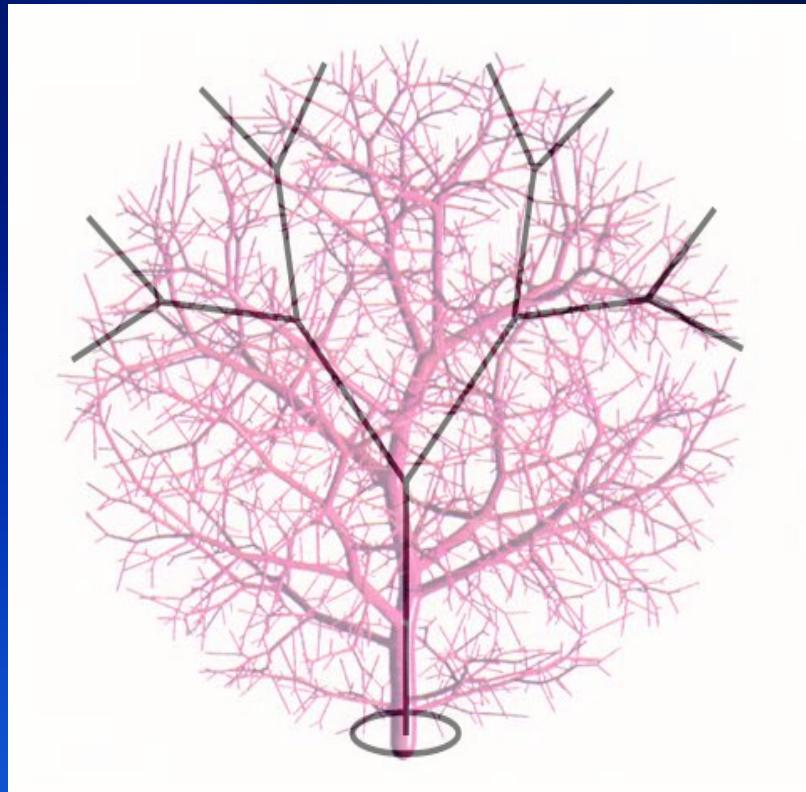


Lateral Economics⁴⁹

RCTs are one of many tools

- Angus Deaton
- William Easterly
- Dani Rodrik
- Sanjay Reddy
- “Randomization is a metaphor and not a gold standard,” James Heckman
- And “Student’s” collaborator, the experimental maltster and barley farmer, Edwin S. Beaven.

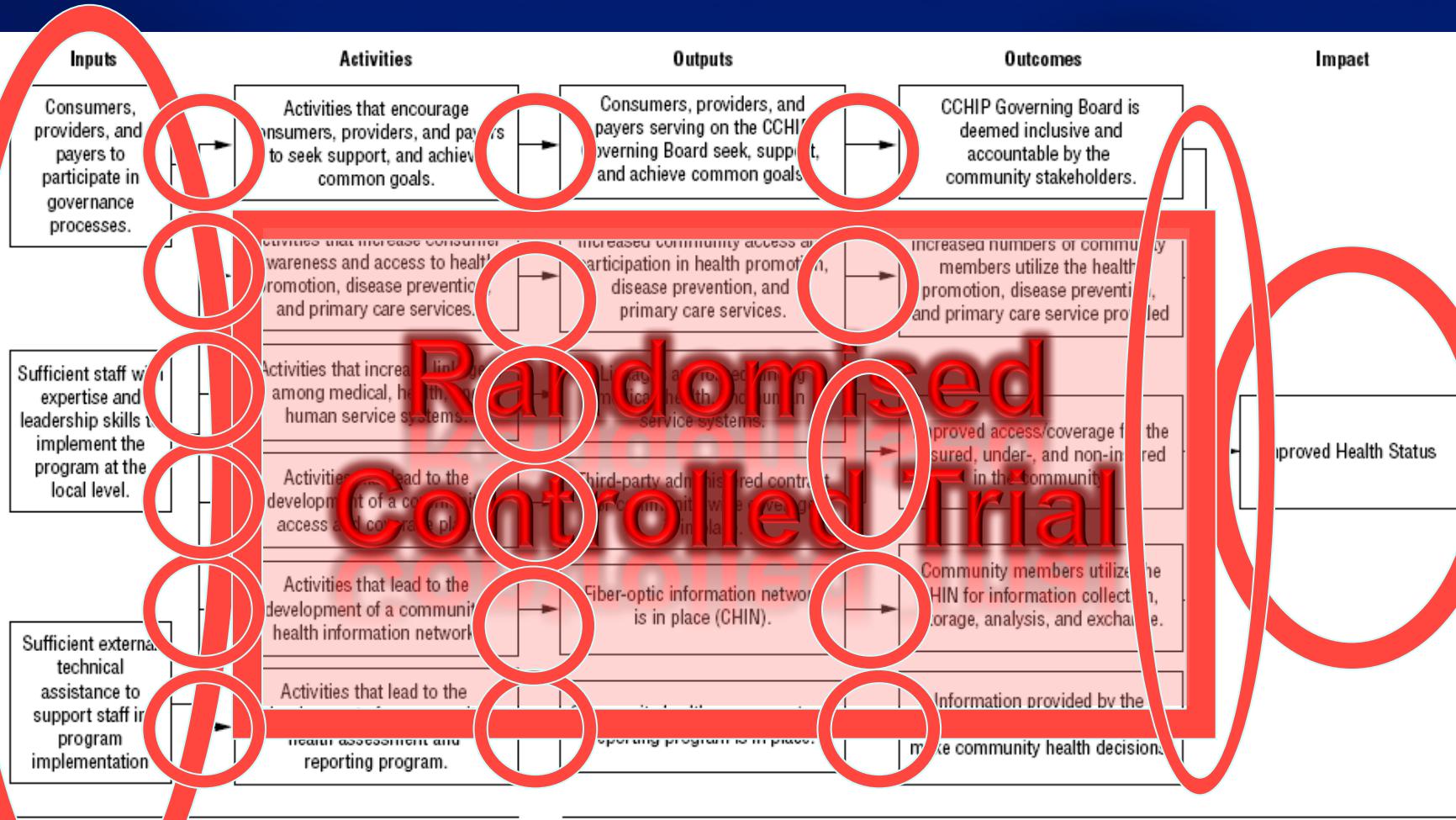
RCTs can be important but they're thin



Hayek on scientism: 1942

In the hundred and twenty years or so during which this ambition to imitate Science in its methods rather than its spirit has now dominated social studies, it has contributed scarcely anything to our understanding of social phenomena...

Demands for further attempts in this direction are still presented to us as the latest revolutionary innovations which, if adopted, will secure rapid undreamed of progress.



Example of an Outcome Approach model (example drawn from the Calhoun County Health Improvement Program, funded under the Comprehensive Community Health Models of Michigan initiative).

Deaton and Cartwright on RCTs

RCTs are valuable. Yet some enthusiasm for them seems based on misunderstandings. That:

- randomization allows a precise estimate of the treatment effect
- that randomization is required to solve selection problems
- lack of blinding does little to compromise inference
- statistical inference in RCTs is straightforward because it requires only the comparison of two means.

None of these statements is true.

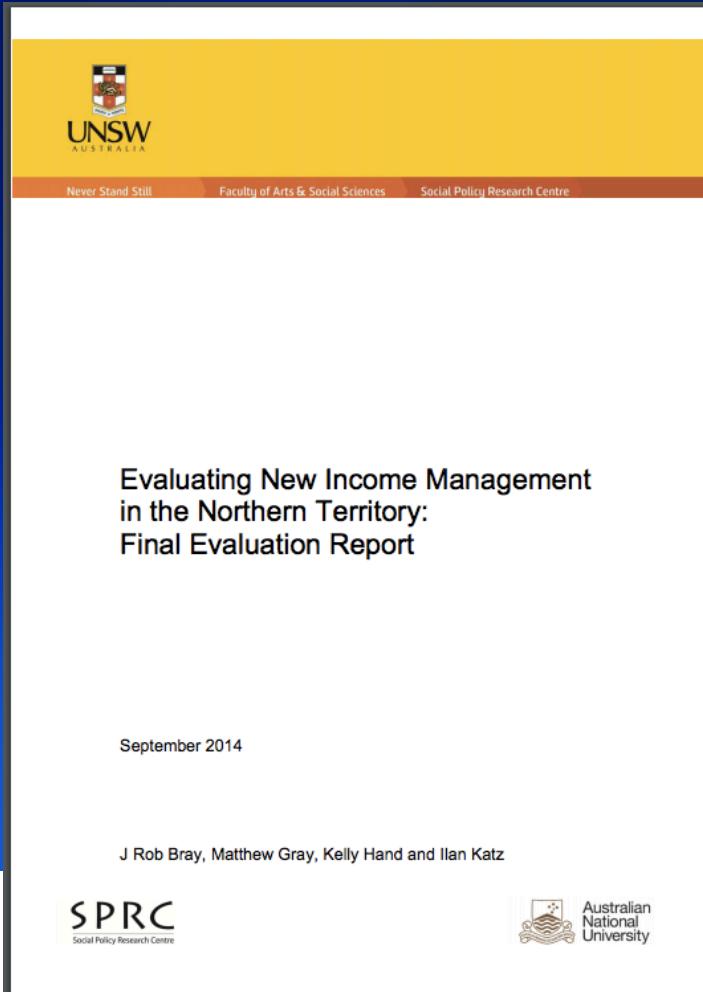
RCTs require minimal assumptions and little prior knowledge, an advantage when persuading distrusting audiences, but disadvantage for scientific progress.

The lack of connection between RCTs and other scientific knowledge makes it hard to use them to increase the scope of the exact context in which they are conducted.

They can be valuable in building knowledge, provided they are combined with other methods, to discover not “what works,” but why things work.

RCT McNuggets: RCTs as scientific trinkets

A very small part of evidence based policy ...



The image shows the cover page of a report. At the top, there is a yellow bar with the UNSW Australia logo and the text 'Never Stand Still', 'Faculty of Arts & Social Sciences', and 'Social Policy Research Centre'. The main title 'Evaluating New Income Management in the Northern Territory: Final Evaluation Report' is centered in a large, serif font. Below the title, the date 'September 2014' is printed. At the bottom, the authors' names 'J Rob Bray, Matthew Gray, Kelly Hand and Ilan Katz' are listed, along with the logos for SPRC (Social Policy Research Centre) and the Australian National University.

UNSW
AUSTRALIA

Never Stand Still Faculty of Arts & Social Sciences Social Policy Research Centre

Evaluating New Income Management
in the Northern Territory:
Final Evaluation Report

September 2014

J Rob Bray, Matthew Gray, Kelly Hand and Ilan Katz

SPRC Social Policy Research Centre

Australian National University

Data on what *causes* what

“Our success at Amazon is a function of how many experiments we do per year, per month, per week, per day....”

Jeff Bezos

“Last year at Google the search team ran about 6,000 experiments and implemented around 500 improvements based on those experiments. The ad side of the business did about the same. Any time you use Google, you are in many treatment and control groups. The learning from those experiments is fed back into production and the system continuously improves.”

Hal Varian, chief economist at Google

Evidence based *policy*



Policy

Delivery





@ngruen1



Lateral Economics⁵⁸

Evidence based Policy: Good things to have

Expertise

- To understand difficulties and maximise the chance to learn

Built round the practical needs of field workers to improve

Causal data (A/B testing)

- To help us learn and improve

Openness

- To build a community of practice and collaborative problem solving

Independence

- To keep us honest, externally, internally, up, down

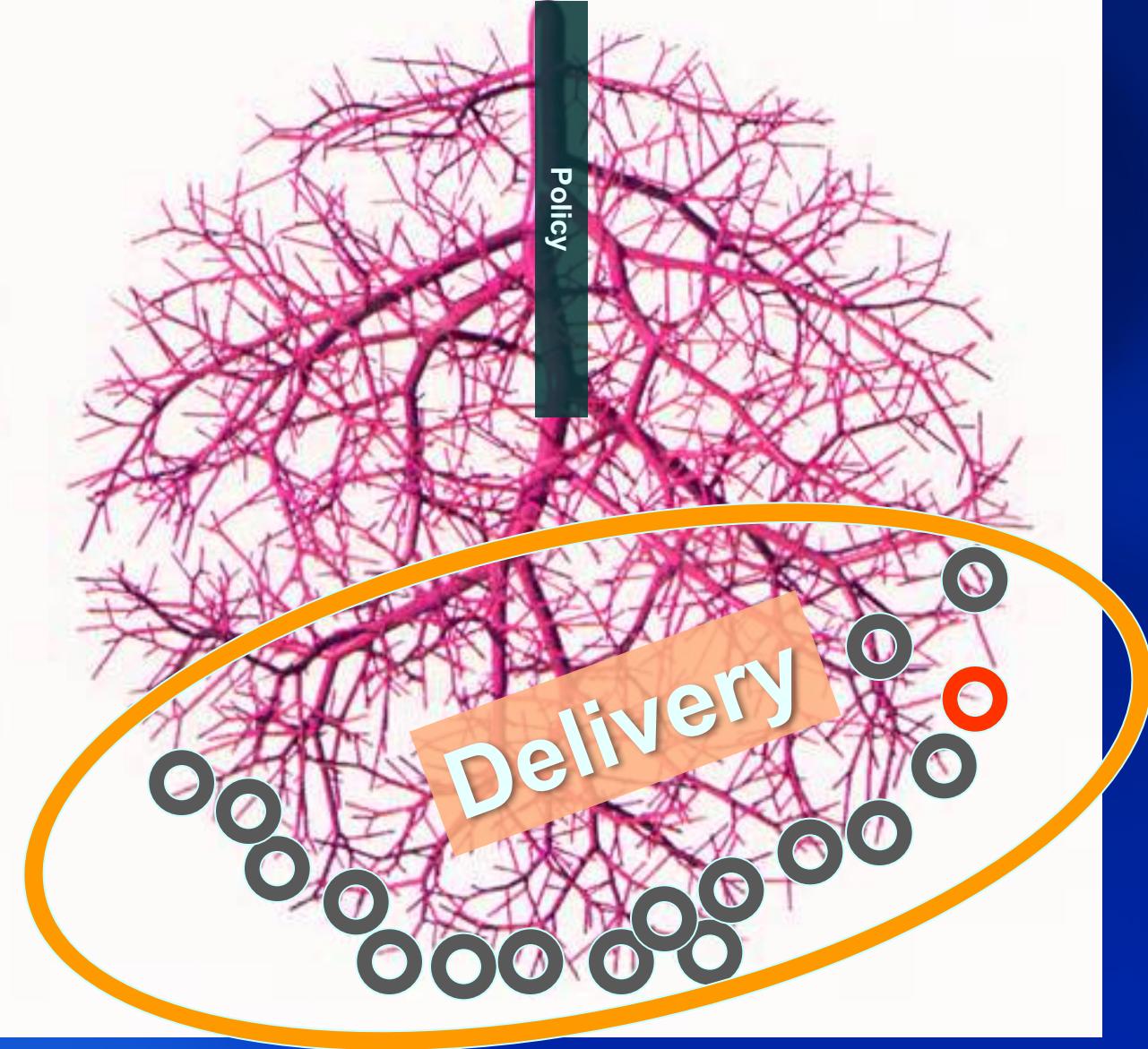
Incentive compatibility

- To keep us trying

Evaluation planned and built in. Not retrospective and bolt on

- So it's efficacious

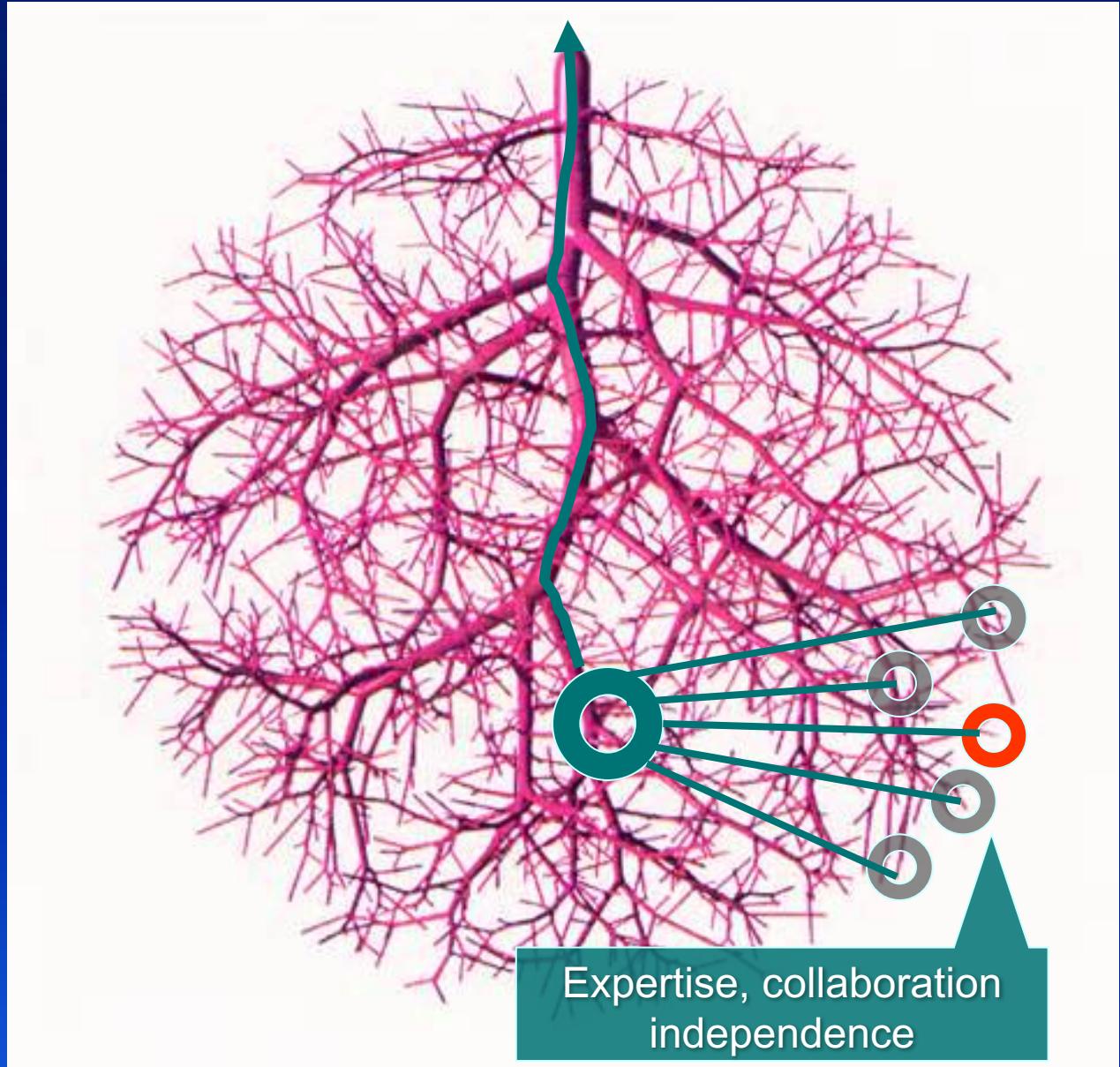
<= Learning goes upward



8 Institutionalising evidence-based policy



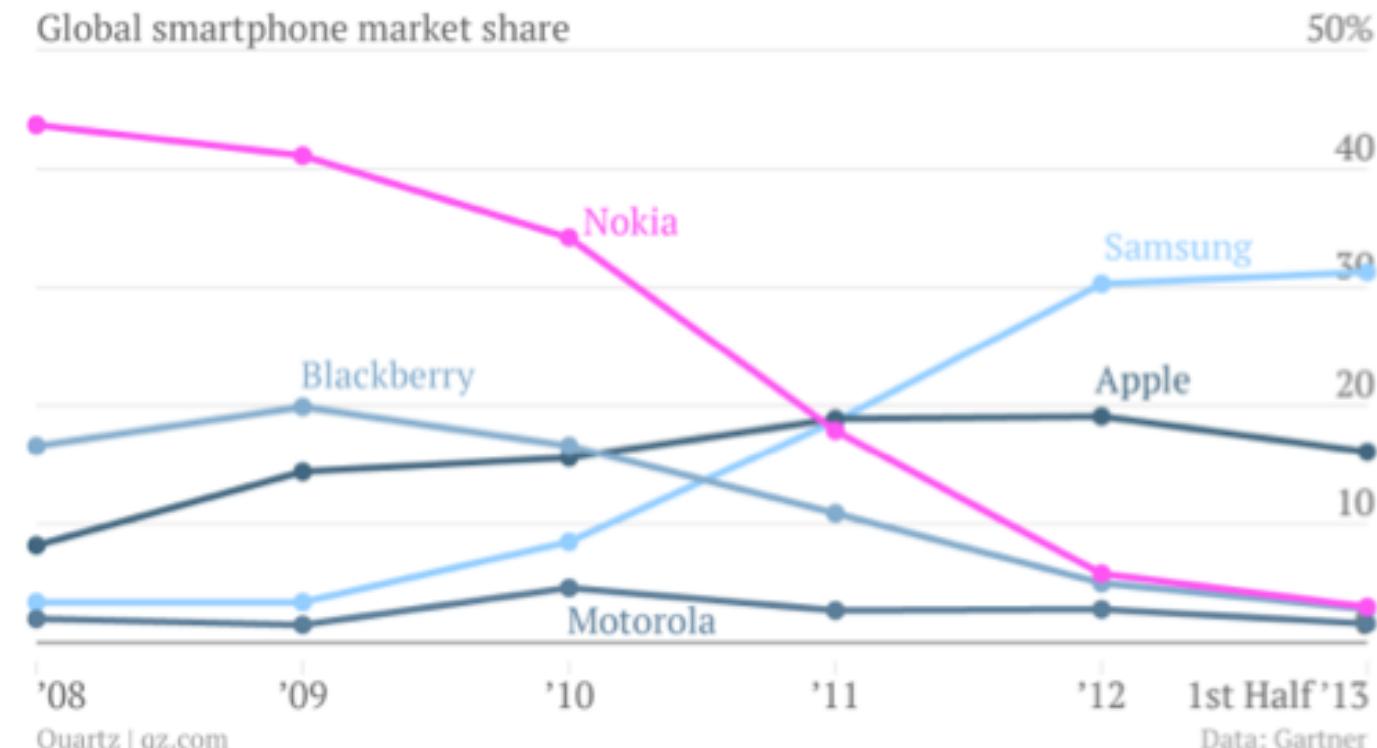
Continual
testing against
the facts of the
life world



The institutional imperative



Figure 5.2: Markets adjust spontaneously to improvements in technology



Source: <http://tinyurl.com/y825v9au>

Too much innovation remains at the margin of public administration. Opportunities are only half-seized; new modes of service delivery begin and end their working lives as ‘demonstration projects’ or ‘pilots’; and creative solutions become progressively undermined by risk aversion and a plethora of bureaucratic guidelines.

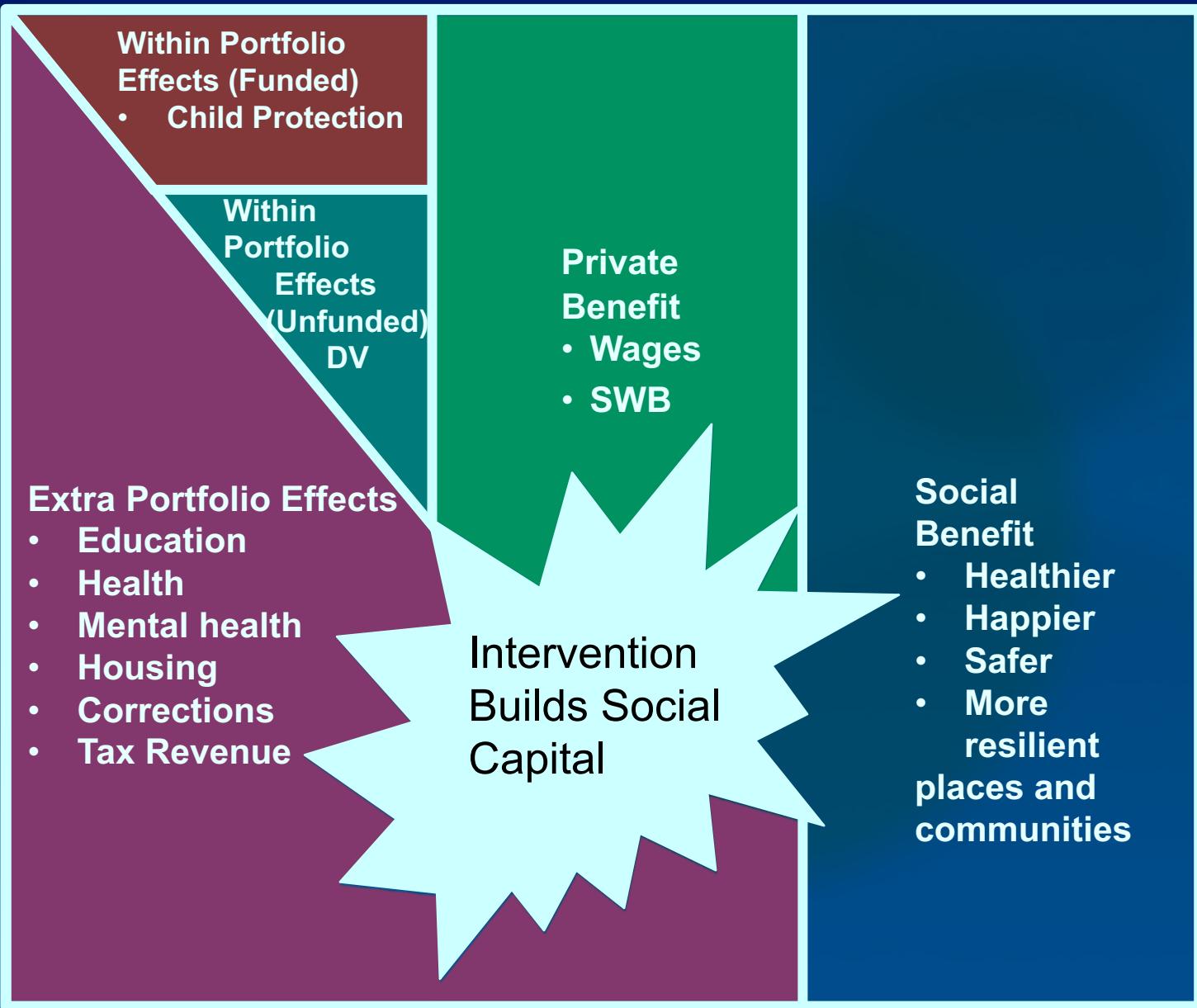
Peter Shergold, 2013.

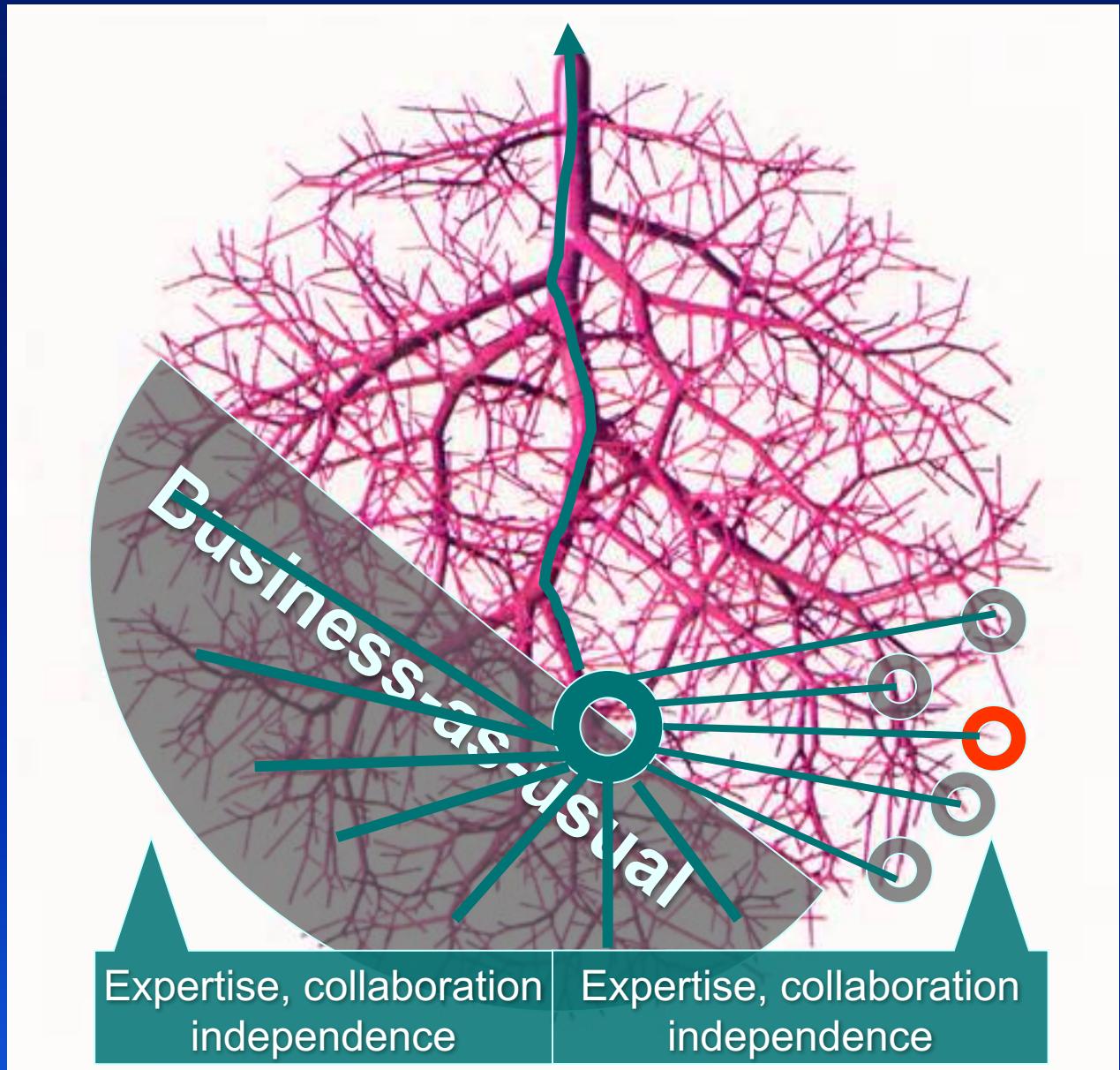


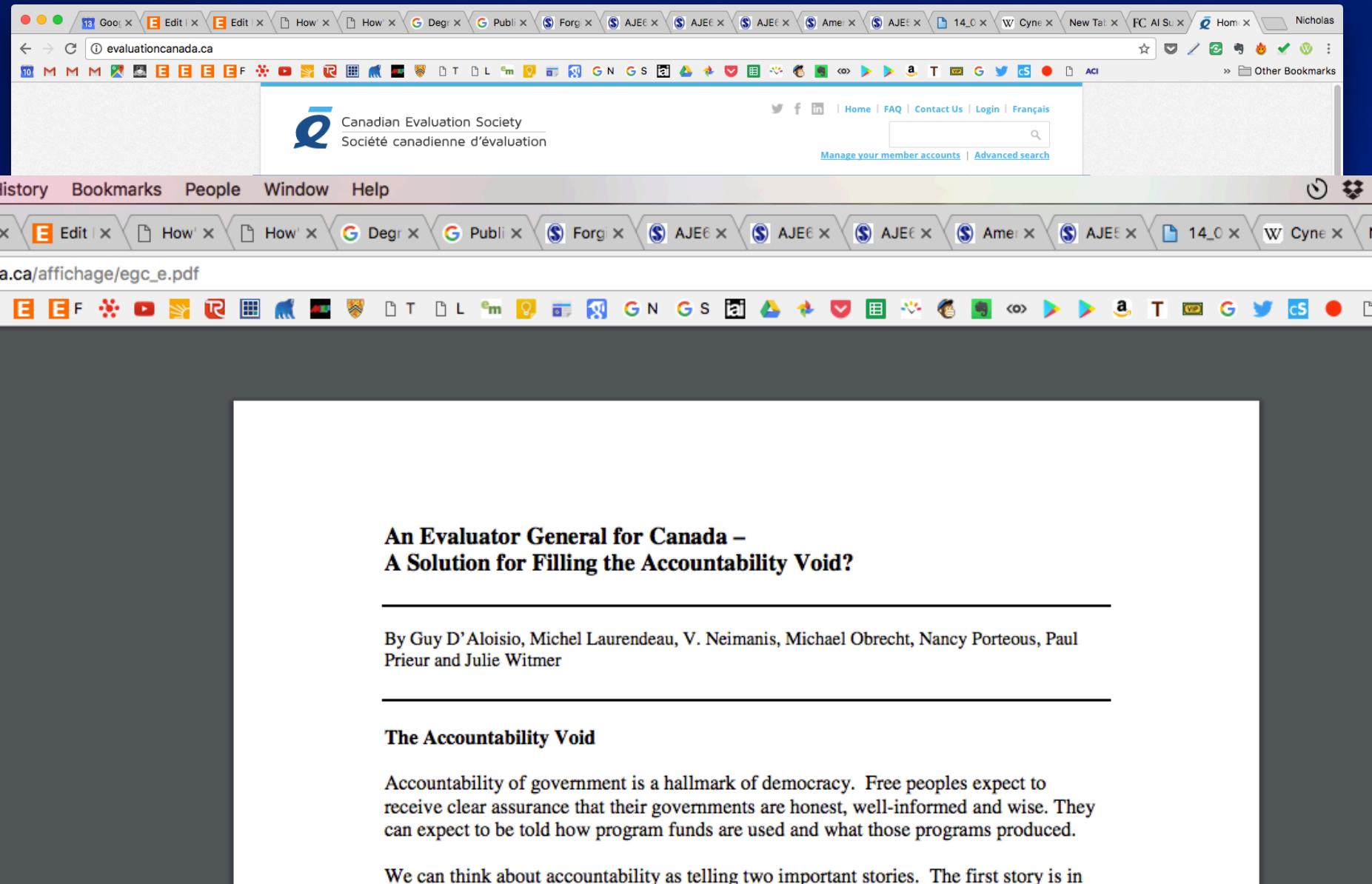
@ngruen1



Lateral Economics⁶⁷







@ngruen1



Lateral Economics⁷⁰

Figure 1 - Current Situation: An Accountability Void

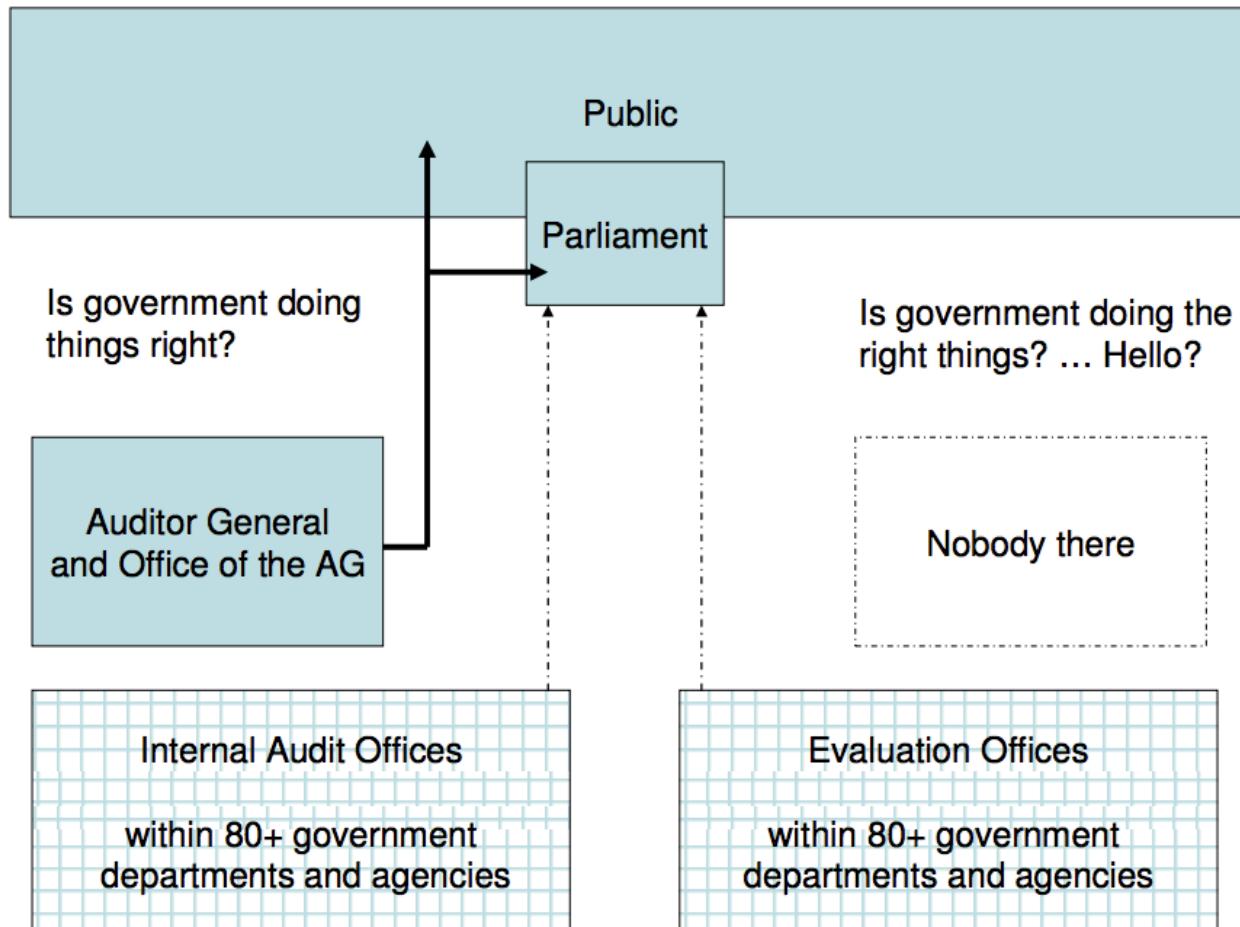
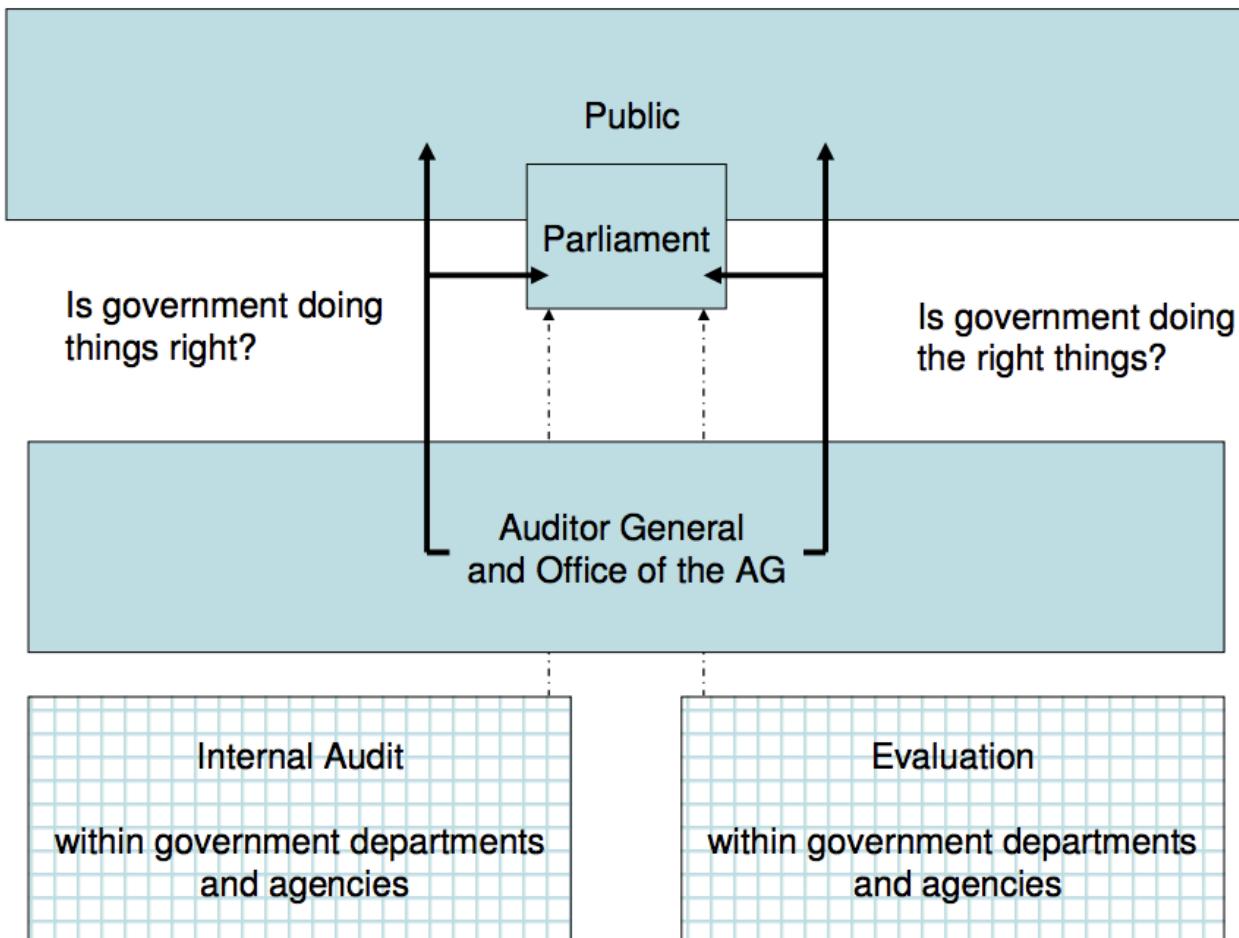
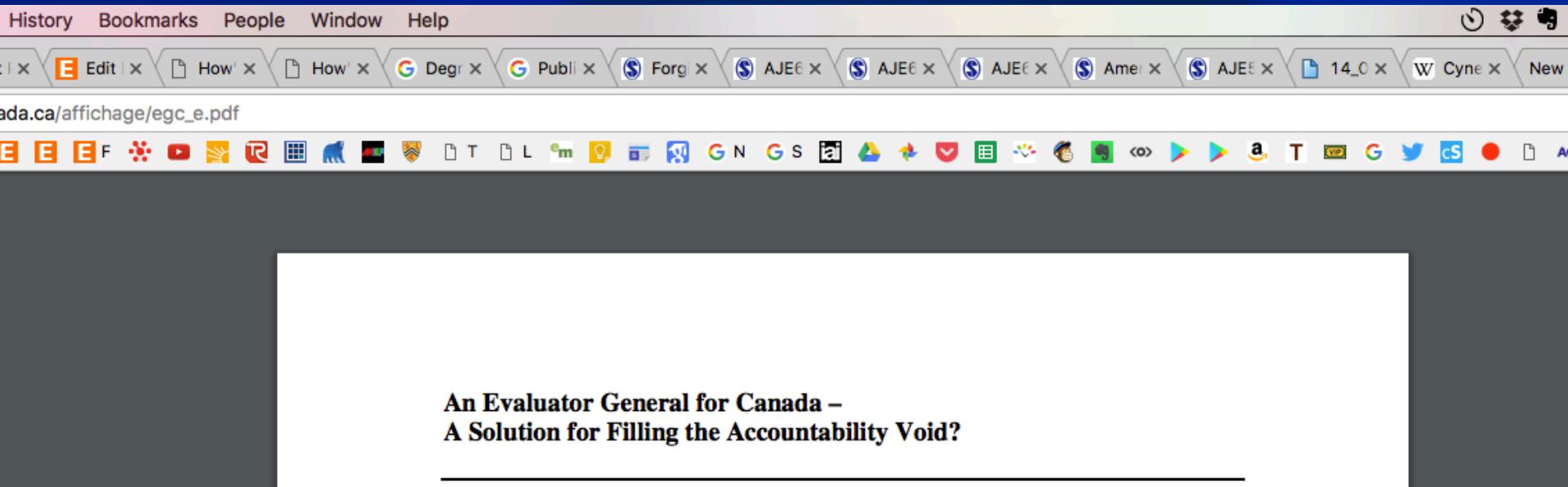


Figure 2 - A Possible Solution: The Auditor General Fills the Void



But it would be a mistake to expect that one person, the Auditor General, would be able to function in both paradigms simultaneously, just as it would be erroneous to imagine that an audit official grounded in chartered accountancy could work effectively in the area of effectiveness assessment where no absolute bottom line can ever be reckoned.



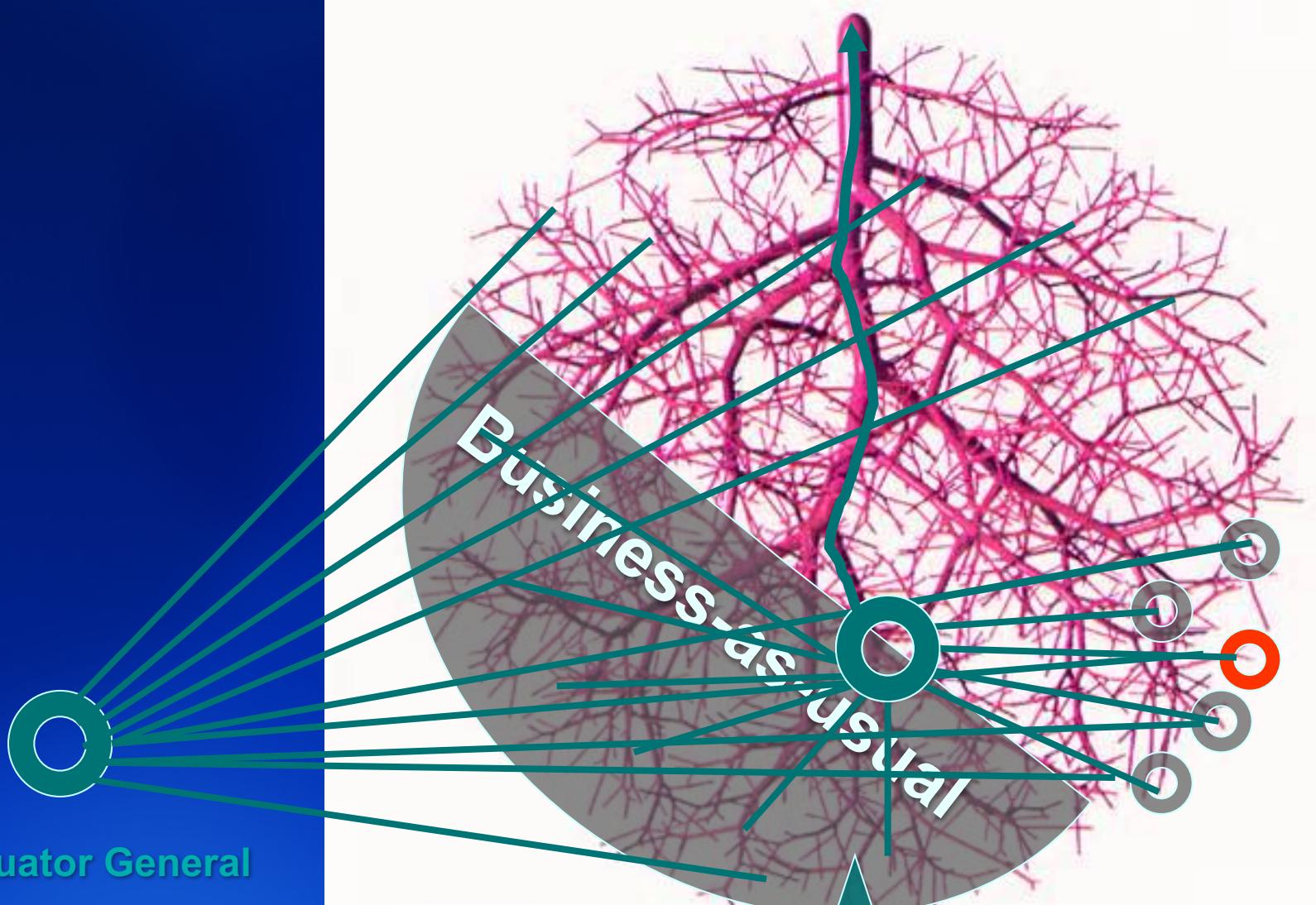
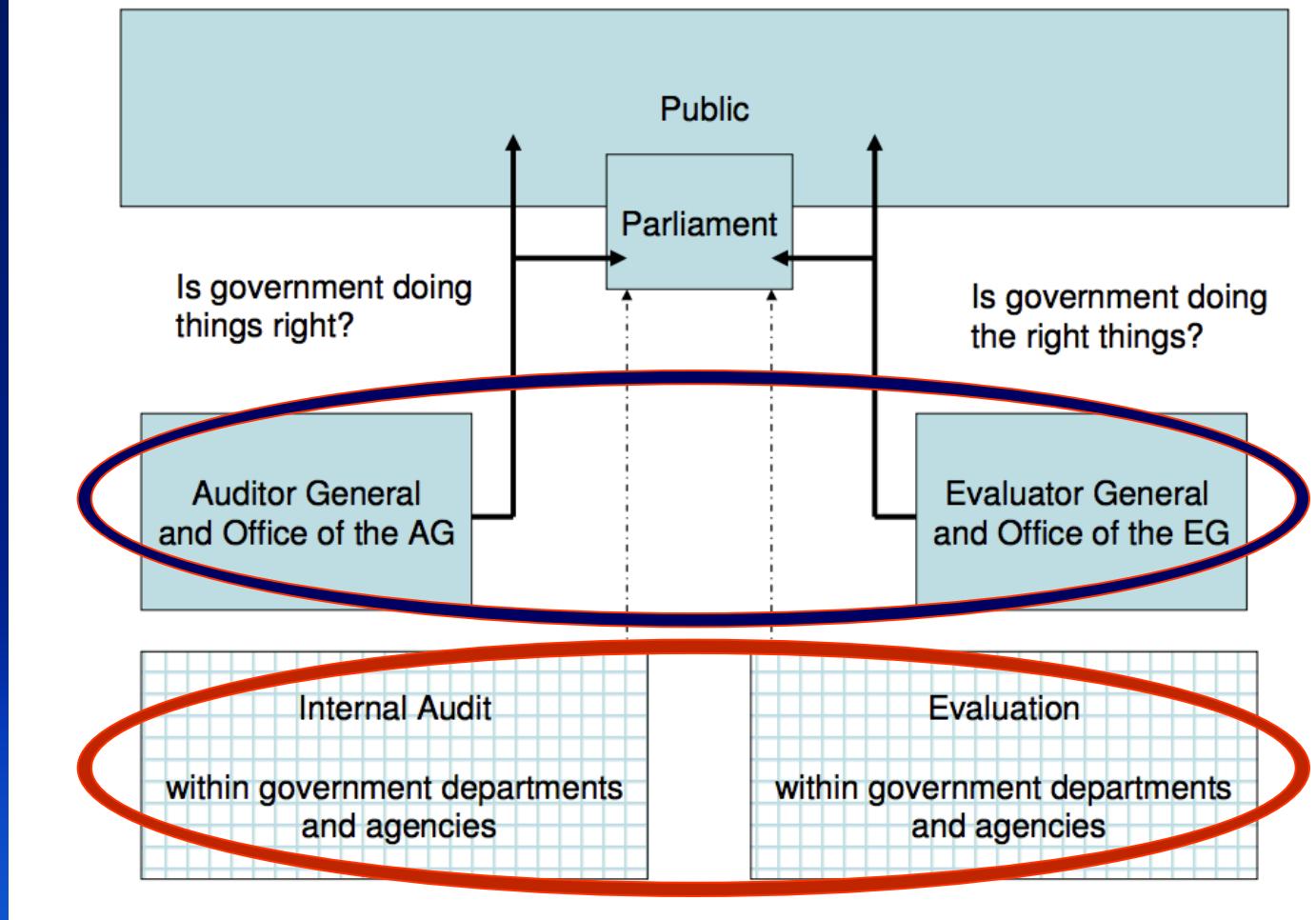


Figure 3 - Preferred Approach: An Evaluator General for Canada



Two systems

Competitive provision

- Delivering services
 - Departments of state
- Making choices

Direct provision

- Integrity
 - Auditor Gen'l, Ombsmn
- Information
- Knowing what we're doing
- Understanding policy choices
 - PC
 - PBO

Accountability

It is not necessary to abandon the notion or being accountable for what has to be done but to return to the meaning and focus on systems of accountability that both justify [it can be justified] and explain what has been done. This requires careful consideration of *who* is being held **accountable, to whom, for what, how, and with what consequences**.

More thoughtful and comprehensive approaches to accountability should demonstrably support good performance and encourage responsibility.

Patricia Rogers, ANZSOG

Evidence based delivery: Good things to have

Expertise

- To understand difficulties and maximise the chance to learn

Built round the practical needs of field workers to improve

Causal data (A/B testing)

- To help us learn and improve

Openness

- To build a community of practice and collaborative problem solving

Independence

- To keep us honest, externally, internally, up, down

Incentive compatibility

- To keep us trying

Evaluation planned and built in. Not retrospective and bolt on

- So it's efficacious

Starting out

Grow expertise and independence of evaluation

- Can start very small
 - within government agencies

Identify some priority sectors (and/or regions)

- Indigenous policy (Aus)
- Child protection (Aus)
- Loneliness (UK)
- Wellbeing (UK)

Set some system targets with independent reporting on them (from Auditor General)



@ngruen1



Lateral Economics⁸⁰



ngruen@lateraleconomics.com.au

The end

@ngruen1



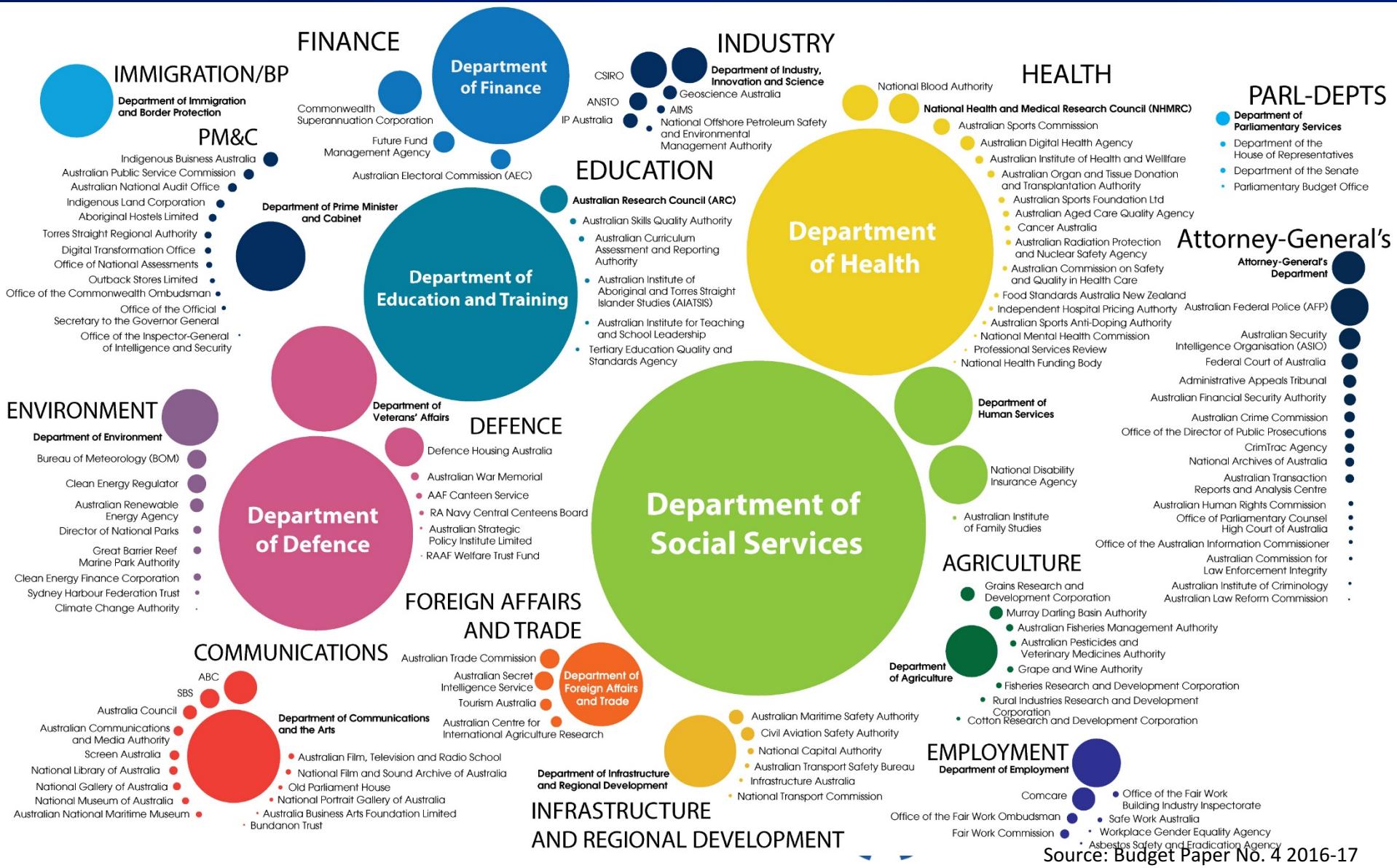
Lateral Economics⁸¹

@ngruen1



Lateral Economics⁸²

2016–17 Federal Budget—\$450.6 billion



2016–17 Federal Budget—R&D \$10.1 billion

