



*World-class  
evaluation &  
organisational  
consulting*

# Causal inference: Nuts and bolts

A Mini Workshop for the  
anzea Wellington branch

E. Jane Davidson, Ph.D.  
May 21, 2009

# Agenda

---

- Causation, outcomes and coincidences
- Two applied examples:
  - Leadership development programme
  - Performance appraisal & bonus system
- 8 strategies for inferring causation
- Hands-on practice
- Certainty about causation – how certain do we need to be?
- Guidelines for inferring causation
- Summary and further references

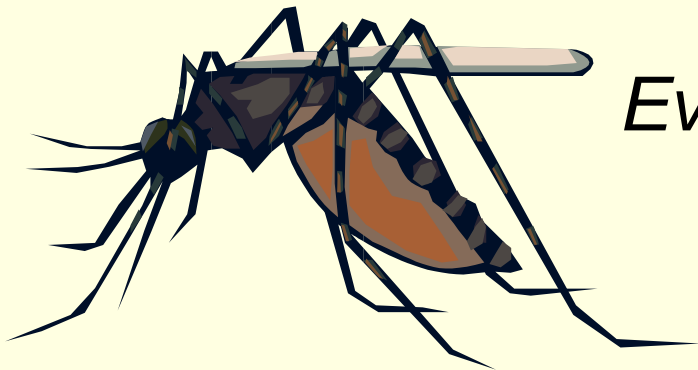
# What is causation?

---

- “Causation. The relation between mosquitos and mosquito bites. Easily understood by both parties but never satisfactorily defined by philosophers or scientists.”

-- Michael Scriven

*Evaluation Thesaurus* (4<sup>th</sup> ed.)



# From one evaluation report ...

---

- “This is an *outcome* evaluation, in which actual changes are measured and documented, but it is not able to measure whether or not it is the programme itself which *causes* the changes, sometimes called the *impact* of the programme.”

# Outcome or coincidence?

---

- If an “outcome” is not *caused* by the intervention, it is NOT an outcome; it’s merely a coincidence
- Coincidences cannot be documented as intervention outcomes
- Therefore, causal inference is a *crucial* part of linking inputs to outcomes
- You do NOT necessarily need a randomised experimental design to infer causation! (although they can be a good option)

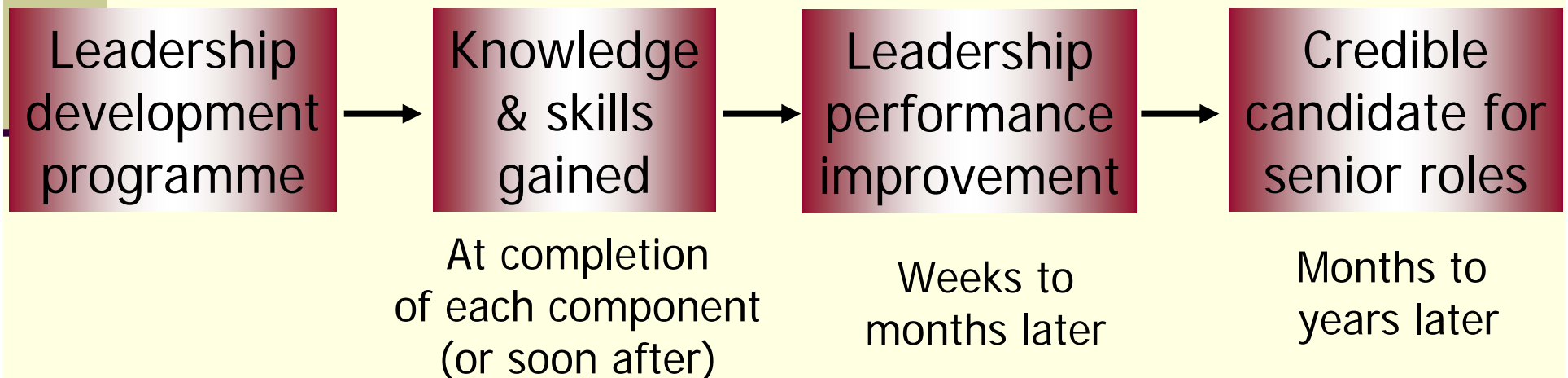
# Ex#1: Leadership development programme

---

- In the evaluation of a [central government-wide] leadership development programme, could recent promotions and career advances (e.g. appointments to chief executive and other senior positions) be attributed to the programme?
- Level of certainty required: about 70-80%
- Most important counterexplanation:  
Pure coincidence, would have happened anyway as part of the individual's career progression

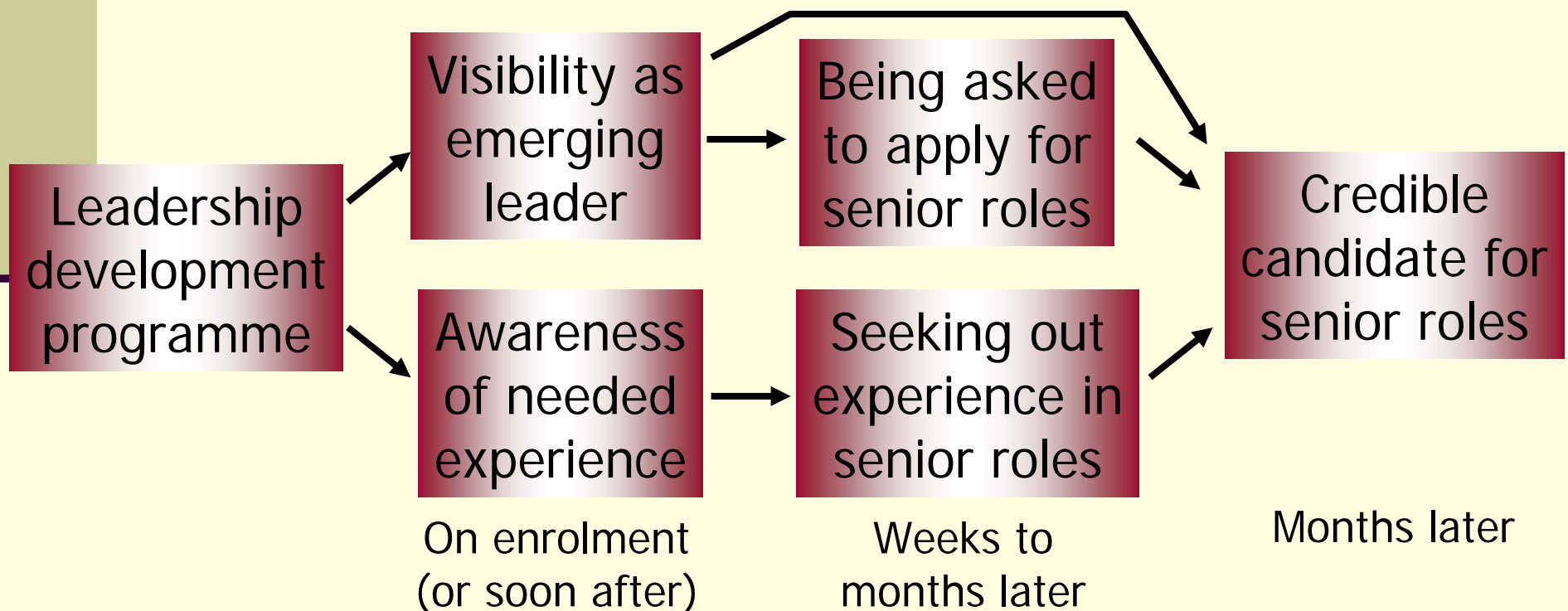
## (a) Logical timing of outcomes

- Map the intervention to a logical chain of outcomes
  - Distal outcomes should not occur before proximal outcomes
  - The timing of outcomes should make sense
- So, was it too soon for this chain of events to have occurred?



## (a) Logical timing of outcomes

- For some who had relatively recently joined the programme:
  - Too early to see major promotions due to knowledge and skills gained
  - So, maybe there was another causal path or two?

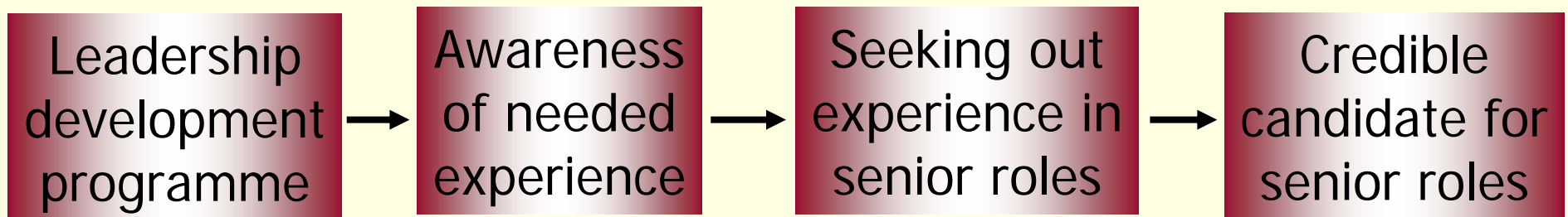




## (b) Match content to outcomes

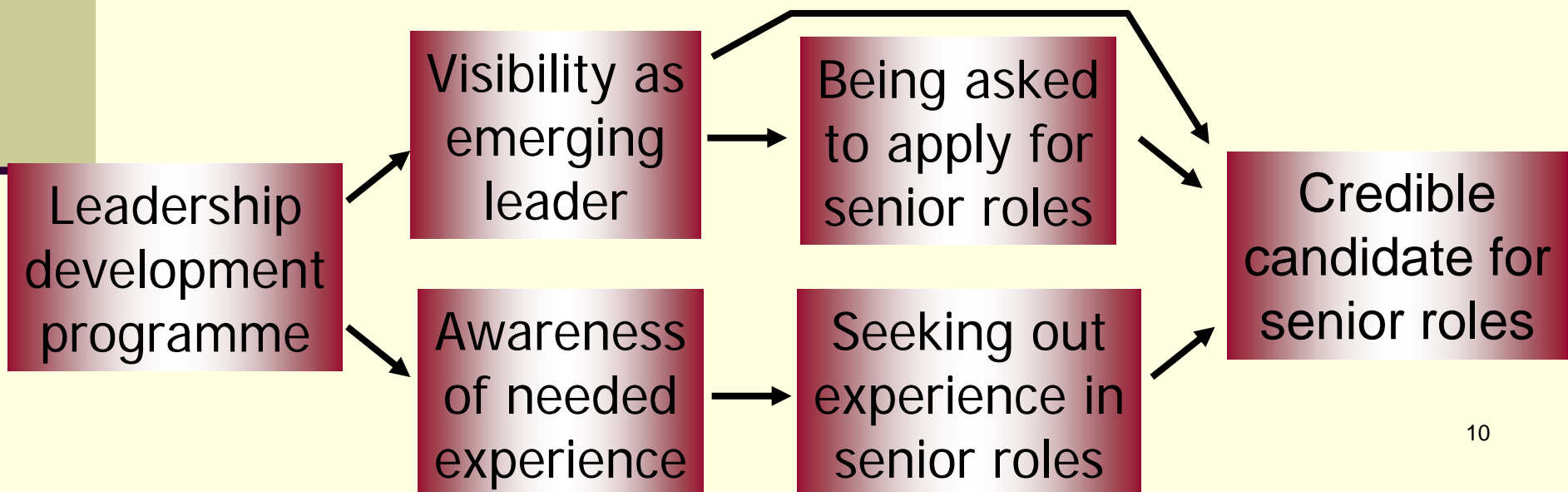
---

- Was the leadership development programme a cause of the participants' promotions/career advancement?
  - Did those who took more senior roles end up in positions that were consistent with the development needs that were identified in their initial assessment and individual development plan?



## (c) Check causal mechanisms

- Causal claims are a lot more convincing if backed by a logical explanation – especially if empirically tested:
  - How are ELPers seen by those who conduct searches for top-level positions?
  - Do candidates actively seek out roles to develop them in areas identified as needed in the programme?



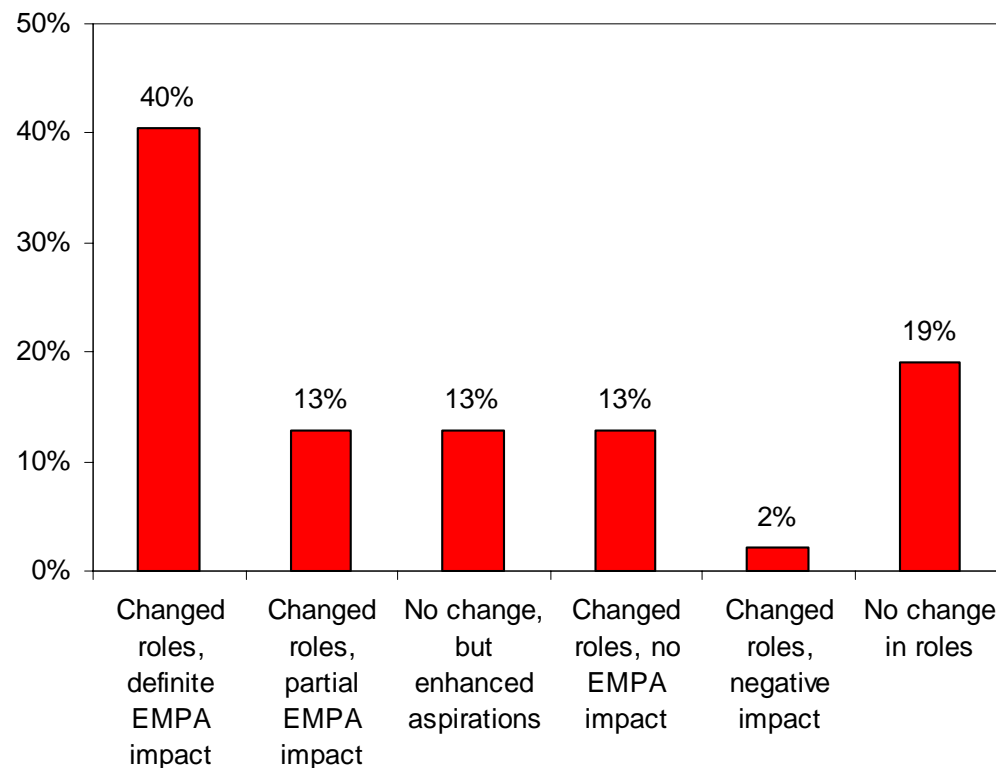
## (d) Ask observers



- Causation is often directly experienced or observed:
  - By intervention participants themselves
  - By managers, peers, and other stakeholders
  - By intervention implementers (e.g., trainers)
  
- In this case:
  - Ask participants whether, to what extent, and why they attribute their new appointment to the program
  - Ask the participant's previous chief executive, senior manager, or mentor whether, to what extent, and why they attribute the participant's new appointment to the program
  - Ask the selection committee whether the participant's enrolment in the program influenced their decision

# “Just ask people”!

**Since enrolling in the Executive Master of Public Administration, have you changed job roles or responsibilities at all? If so, would you consider this change a career advancement? To what extent did the EMPA contribute to your getting the new role or responsibilities? How?**



*“Yes- I've been promoted. It was a significant career advancement. EMPA provided me with the breadth of learning experiences that I could bring to the role - firstly actually being accepted for the course is rated highly, secondly I understand that my interview went well because of my ability to relate EMPA learnings to real life issues.”*

## (e) Modus operandi method

- The detective metaphor:
  - Just as criminals have a “modus operandi” (patterns of behavior), so too do some interventions create distinctive/characteristic patterns of effects



- In this case:
  - The program’s individual development plan identified needed experiences → candidate sought these out → obtained more senior position to gain that experience
  - Another “signature trace” we might expect if this is true is that the individual continues to push outside his/her comfort zone and seek out stretch experiences within the new job that will address other gaps/needs

# Inferring causation in Example #1

---

- Causal inference strategies used in this case:
  - a) Check whether the timing of outcomes makes sense
  - b) Check if content of intervention matches nature of outcome
  - c) Identify and check the causal mechanism(s)
  - d) Ask those who have observed or experienced the causation first-hand
  - e) Look for distinctive effect patterns (modus operandi method)
- Recall that the standard of proof needed to be about 70-80% certainty
- Conclusion: Enough evidence of a causal link in several (but not all) cases, even if other causes were in play

## Ex#2: Performance appraisal system

---

- In the evaluation of a performance appraisal and bonus system, did it cause the detected improvements in motivation performance?
- Level of certainty required:  
again, about 70-80%
- Most important counterexplanation:  
The quality of management caused improved motivation and performance

## (a) Use a comparison or control group

---

- Compare with a “control” or “comparison” group
  - Random assignment (a true experiment) is often possible because full, simultaneous rollout is frequently impossible
  - If not possible, comparisons with the best possible control group match is a very good alternative
  - You need to have sufficient numbers to make this method really useful
- In this case:
  - Some managers hadn't implemented the system at all, so we had a 'natural' control group!



## (b) Examine the dose-response link

---

- Patchy implementation of interventions
  - It's actually an evaluation opportunity, not a disaster!
  - Look for the relationship between “dose” (extent and quality of implementation) and “response” (magnitude of outcomes)
  
- In this case:
  - Using a survey, we included a question about extent and quality of implementation of the initiative
  - Simple option: compare outcomes for people who experienced thorough vs. weak implementation
  - More complicated: partial correlation between index of implementation and posttest outcome, controlling for pretest outcome *[Note that change scores are too unreliable to use as a dependent variable (DV), hence the use of this analysis strategy.]*

## (c) Control for extraneous variables

---

- Even with random assignment or a good comparison group, worries often remain about rival explanations
  - => use statistical techniques (e.g., regression) to control for extraneous influences or variables
  
- Recall that the most plausible alternative explanation was that improvements in motivation and performance were due to quality of management
  - Included a 'quality of management' scale in the organizational survey
  - Perform regression analysis with:
    - DV = posttest outcome (motivation or performance)
    - IVs: step 1 = pretest outcome, step 2 = quality of management, step 3 = index of implementation
    - *[Note that change scores are too unreliable to use as a DV,<sup>18</sup> hence the use of the above analysis strategy.]*

## (d) Ask observers



- Causation is often directly experienced or observed:
  - By intervention participants themselves
  - By managers, peers, and other stakeholders
  - By intervention implementers (e.g., trainers)
- In this case:
  - Ask participants whether, to what extent, and why they felt the new performance appraisal and bonus system had motivated them or helped them perform better
  - Ask open-ended questions about what else made a large difference to how motivated they felt in their work and the extent to which they felt they could really add value
  - Ask managers and HR about what they had observed where the system had been implemented well vs. poorly

# Inferring causation in Example #2

---

- Causal inference strategies used:

- a) Use a control or comparison group
- b) Look at the relationship between “dose” and “response”
- c) Control statistically for extraneous variables
- d) Ask those who have observed or experienced the causation first-hand

- Conclusions:

- No improvement was detected due to the performance appraisal and bonus system
- Motivation increases were determined primarily by effective management
- Performance was harder to pin down

# 8 strategies for inferring causation

---

1. Ask those who have observed or experienced the causation first-hand
2. Check if the content of the intervention matches the nature of the outcome
3. Look for distinctive effect patterns (modus operandi method)
4. Check whether the timing of outcomes makes sense
5. Look at the relationship between “dose” and “response”
6. Use a comparison or control (experimental or quasi-experimental design)
7. Control statistically for extraneous variables
8. Identify and check the causal mechanisms

# Hands-on practice

---

1. Choose an example of a programme or policy one or more of your group is familiar with
2. Identify one of the most important possible outcomes
3. Identify the most compelling counterexplanation
4. Choose two or three causal inference strategies that are likely to be most convincing – explain how and why

# Certainty About Causation

---

- Academic training:
  - Teaches us to be terribly cautious about our conclusions: “The evidence appears to suggest ...”
  - Pushes for at least 95% certainty ( $p < .05$ ) – and even then we don’t call it “proof”
- This language is often incredibly frustrating for clients
- In many contexts, decisions are made based on much less certainty ( $\approx$  60-70% or less?)
- Need to match methods – and the way we talk about the certainty of our conclusions – with decision-maker needs (not with academic conventions)

# Inferring Causation

---

## ■ Basic Principles

- Look for evidence in favor of *and against* the intervention as cause
- Look for evidence in favor of and against the most likely alternative causes
- Use the most cost-effective causal inference strategies successively until required level of certainty is reached

## ■ Use a mix of the eight strategies

1. Ask observers
2. Match content to outcomes
3. Modus operandi
4. Logical timing
5. Dose-response link
6. Comparisons
7. Control variables
8. Causal mechanisms



# Causal inference guidelines

- Determine what level of certainty you need about causation
    - Beyond reasonable doubt?
    - Balance of evidence?
  - Identify the [next] biggest threat (or counterexplanation) to your causal claim
  - Use two or three (or add more) complementary strategies to test the intervention vs. the counterexplanation as cause
  - Have you ruled out the counterexplanation to the required level of certainty? ↘ YES
  - Are there any other serious counterexplanations?
  - Finished! ↙ NO
- 
- ```
graph TD; Q1[■ Determine what level of certainty you need about causation] --> Q2[■ Identify the [next] biggest threat (or counterexplanation) to your causal claim]; Q2 --> Q3[■ Use two or three (or add more) complementary strategies to test the intervention vs. the counterexplanation as cause]; Q3 -- NO --> Q2; Q3 -- YES --> Q4[■ Have you ruled out the counterexplanation to the required level of certainty?]; Q4 -- YES --> Q1; Q4 -- NO --> Q5[■ Are there any other serious counterexplanations?]; Q5 -- NO --> Q4; Q5 -- YES --> Q6[■ Finished!];
```

# Summary of key points

---

- An outcome is not an outcome unless there is a demonstrated causal link
- The causal link doesn't have to be demonstrated to 100% certainty; we need to match the level of certainty to the decision making context (not throw in the methodological kitchen sink)
- You DON'T necessarily need (although they are often useful):
  - Randomised experimental designs
  - Quantitative methods
- But you DO always need (for validity and credibility):
  - A mix of strategies for inferring causation
  - At least one qualitative source of causal evidence

# References

---

- Davidson, E. J. (2004). *Evaluation methodology basics: The nuts and bolts of sound evaluation*. [Sage Publications]
- Scriven, M. (1974). Maximizing the power of causal investigations: The modus operandi method. In W. J. Popham (Ed.), *Evaluation in education: Current applications* (pp. 68-84). Berkeley, CA: McCutchan Publishing.
- Davidson, E. J. (2003). Linking organizational learning to the bottom line: Methodological issues, challenges, and suggestions. In T. E. Kramer (Ed.), Linking organizational learning to the bottom line [Special issue]. *The Psychologist-Manager Journal*, 6(1), 54-67.

