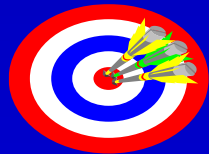
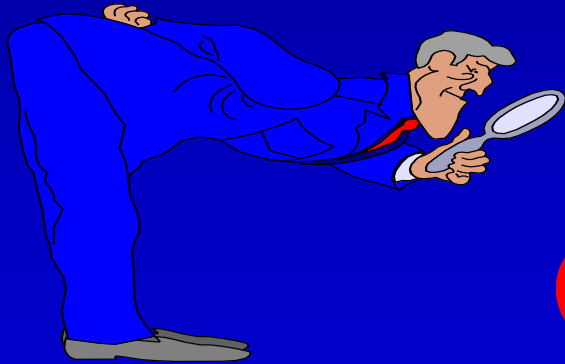


# Help!

I'm doing an impact evaluation,  
what evidence do I need?



Scott Bayley

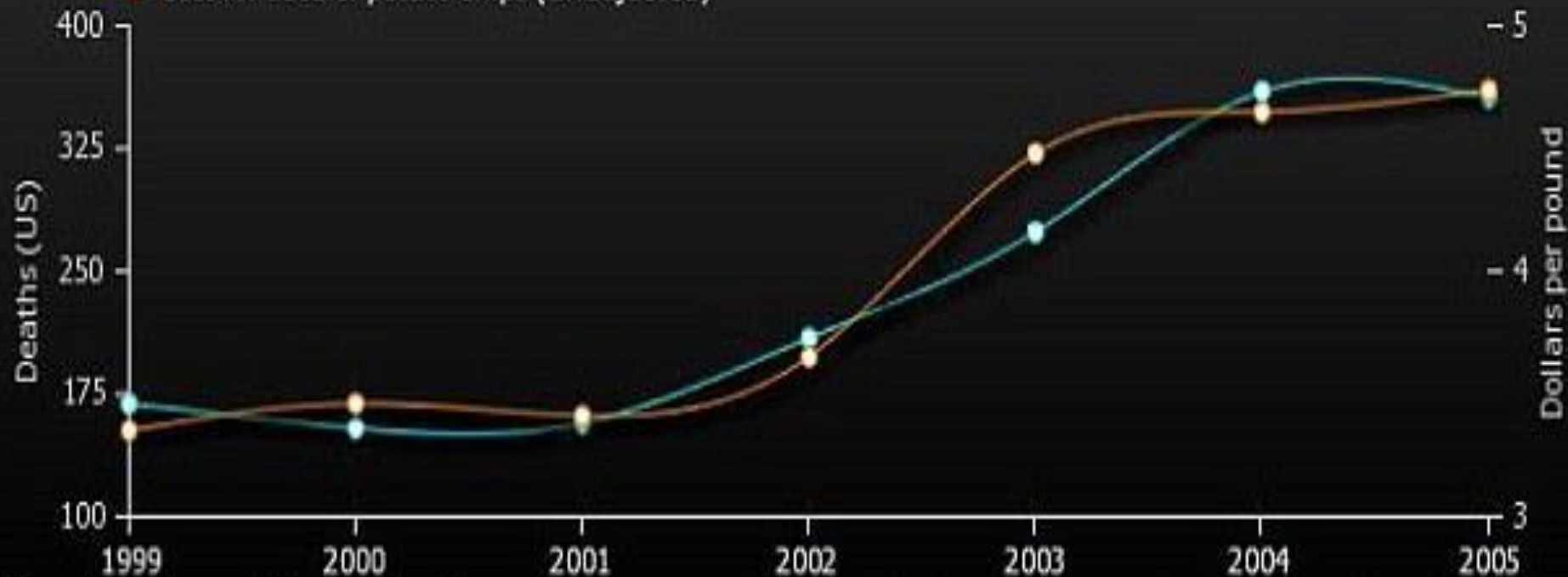
29 September 2023

# Overview

1. Examples: *What is going on here?*
2. Philosophy and evaluation models
3. Evidence requirements for IE
4. Lets practice
5. Summary
6. Quiz question & prize, discussion

# People who died falling out of a wheelchair correlates with the costs of potato chips

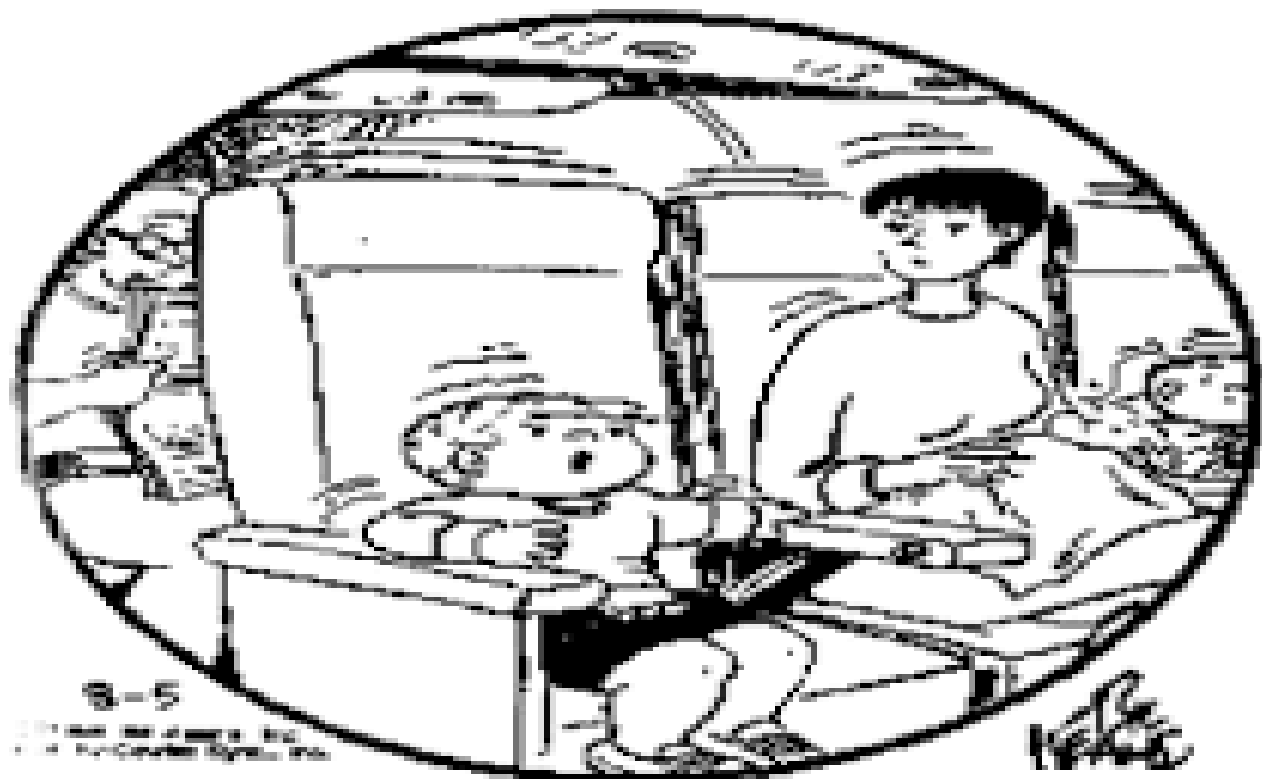
- People who died by falling out of their wheelchair
- Cost for 16oz of potato chips (unadjusted)



Student:

*I will study hard next week in order to get a good grade in today's exam*

# THE FAMILY CIRCUS



8-5

© 1988 by Disney, Inc.  
All Rights Reserved.

*Handwritten signature*

"I wish they didn't turn on that seatbelt sign so much! Every time they do, it gets bumpy."

# Philosophical perspectives



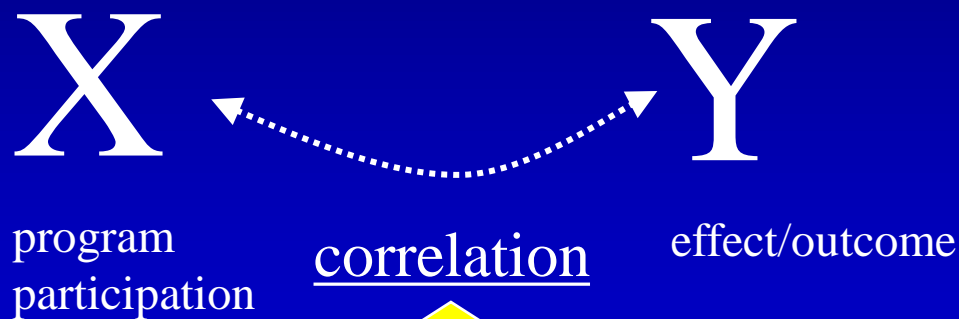
# Evidence requirements for IE

In order to conclude that program (X) causes outcome (Y), three criteria must be satisfied. These are:

1. Demonstrate an association between participating in program X and outcome Y;
2. Establish time order ( $X < Y$  in time); and
3. Rule out alternative explanations (Z).

①

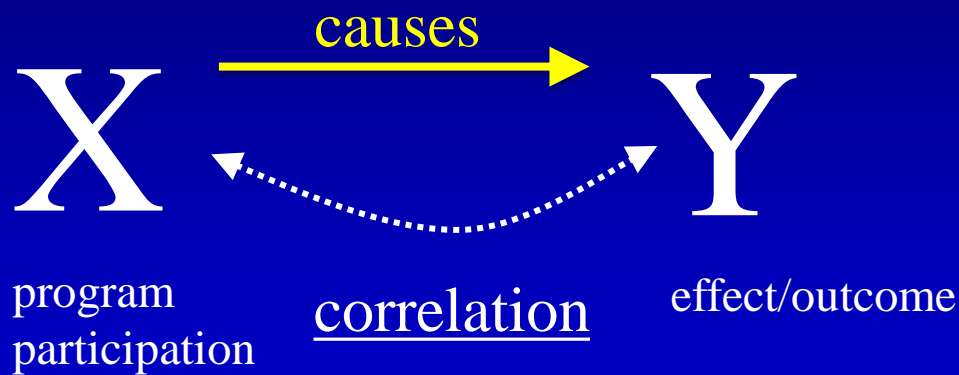
# Four possible explanations:



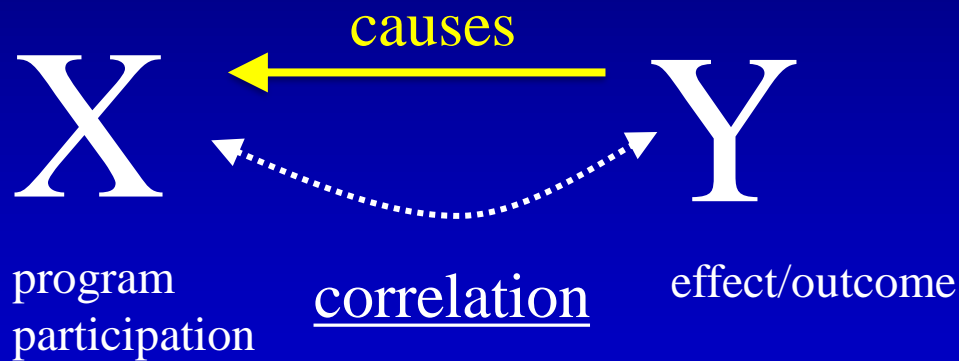
*pure chance/methodological error*

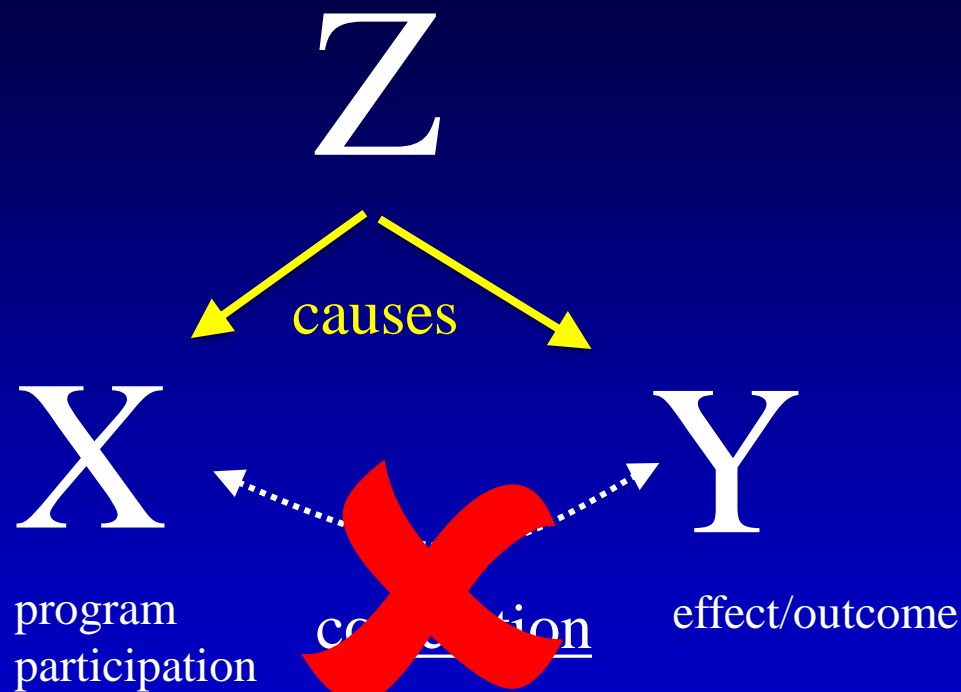


②



③

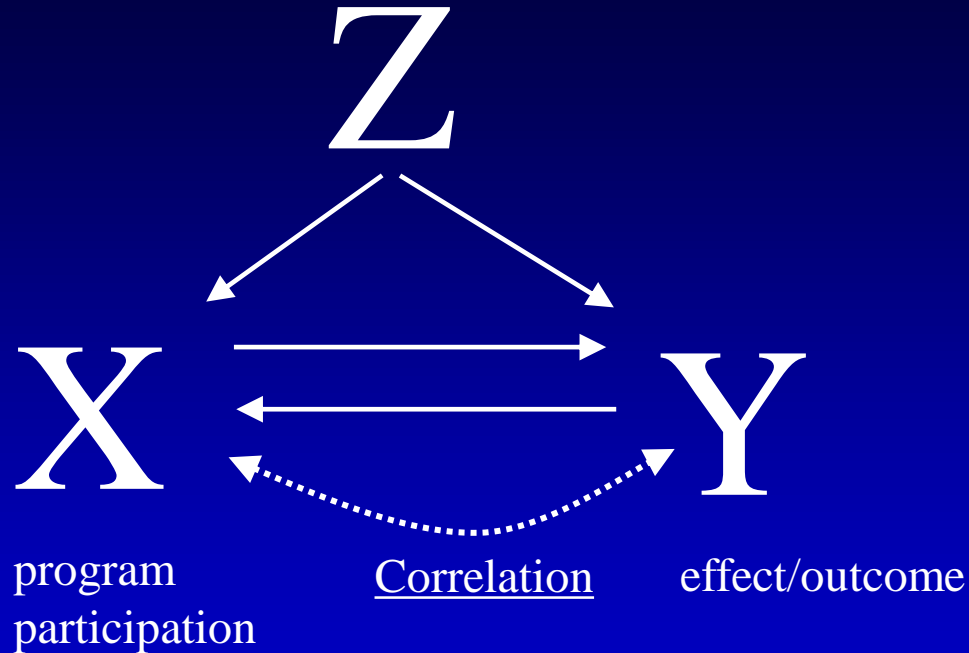




To be trustworthy and credible, an IE must:

1. Demonstrate an association between participating in program X and outcome Y;
2. Establish time order ( $X < Y$  in time); and
3. Rule out alternative explanations (Z).

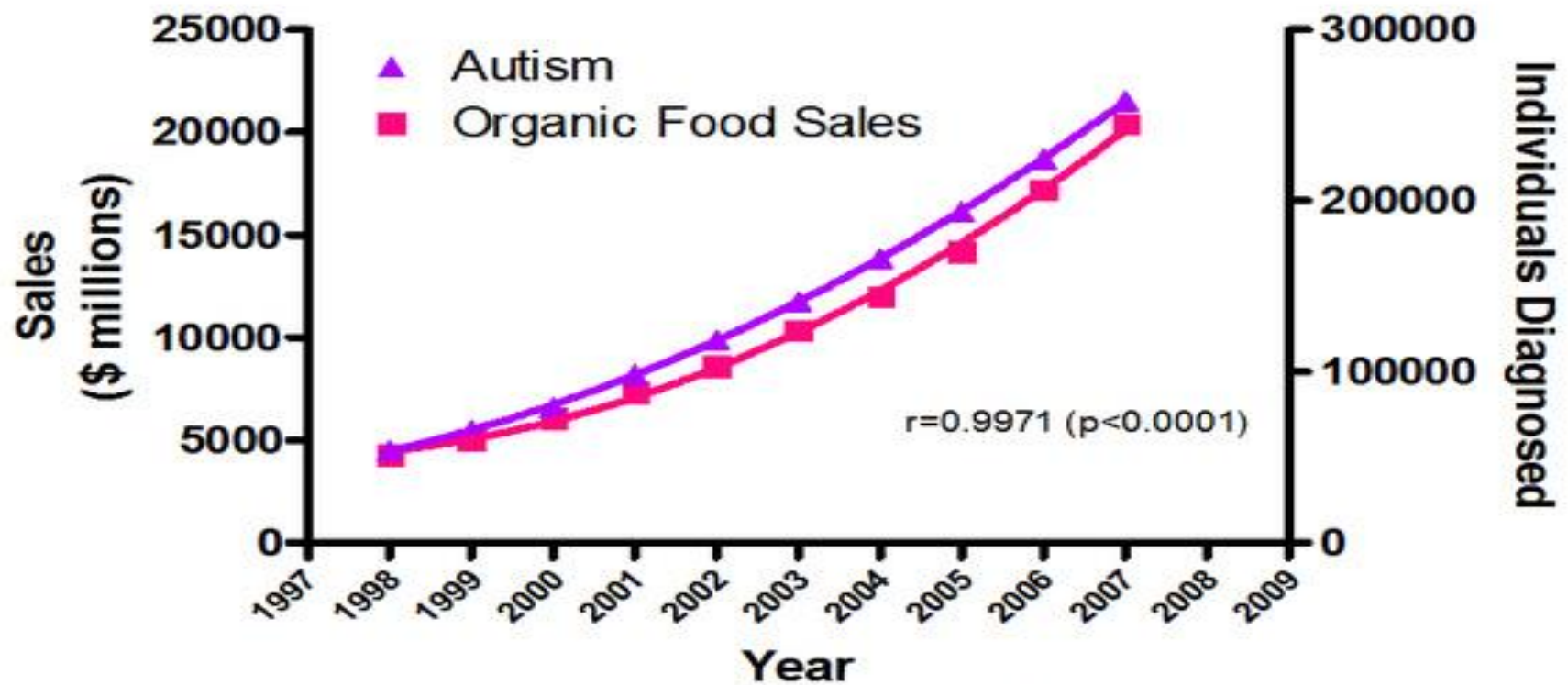
# Lets Practice



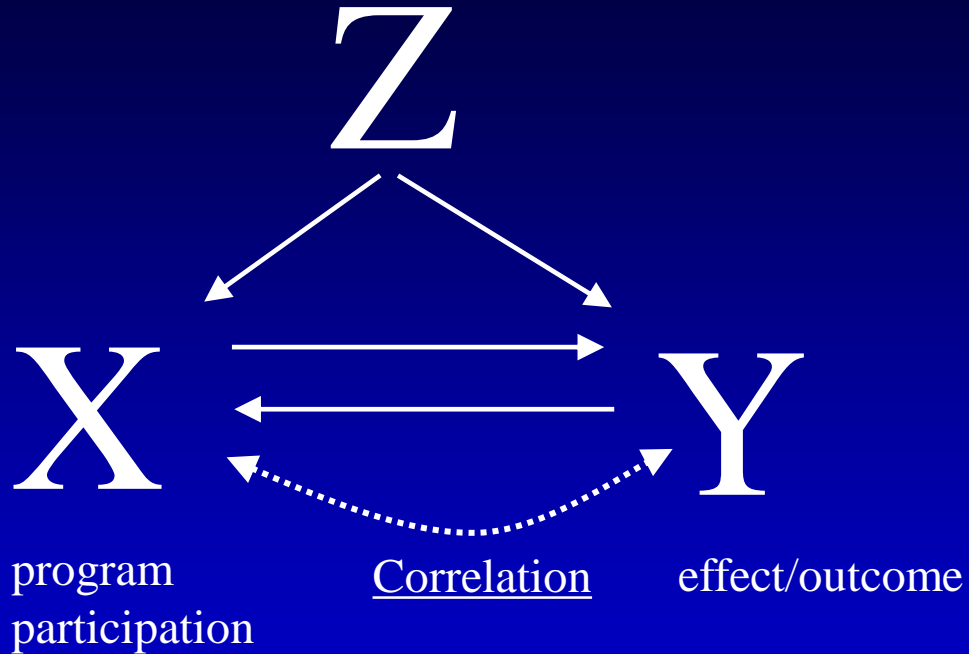
## Examples:

- iPhone sales and # deaths from falling down stairs
- The more firemen are sent to a fire, the more damage is done

## The real cause of increasing autism prevalence?



Sources: Organic Trade Association, 2011 Organic Industry Survey; U.S. Department of Education, Office of Special Education Programs, Data Analysis System (DANS), OMB# 1820-0043: "Children with Disabilities Receiving Special Education Under Part B of the Individuals with Disabilities Education Act"



## Examples:

- participating in job training and employment outcomes
- high levels of patient/doctor satisfaction with surgery for ulcers

# Juvenile Offender's Program

1983

440 juvenile offences

1984

New offenders program commences. Recreation activities, voluntary work, street worker employed

1985

350 juvenile offences

420 juvenile offences

\* 440 offences down to 350 offences (a 20% reduction) \*

➔ *Can we say that the program has been effective in reducing juvenile offending?*



440 offences down to 350 offences - 20% reduction:  
*(before vs after comparison)*

Gross  
Observed  
Change = net  
program  
impact + effects of  
other events  
& processes + methodological  
error

- the actual causal effect of the program in question
- these effects have direction, magnitude and duration

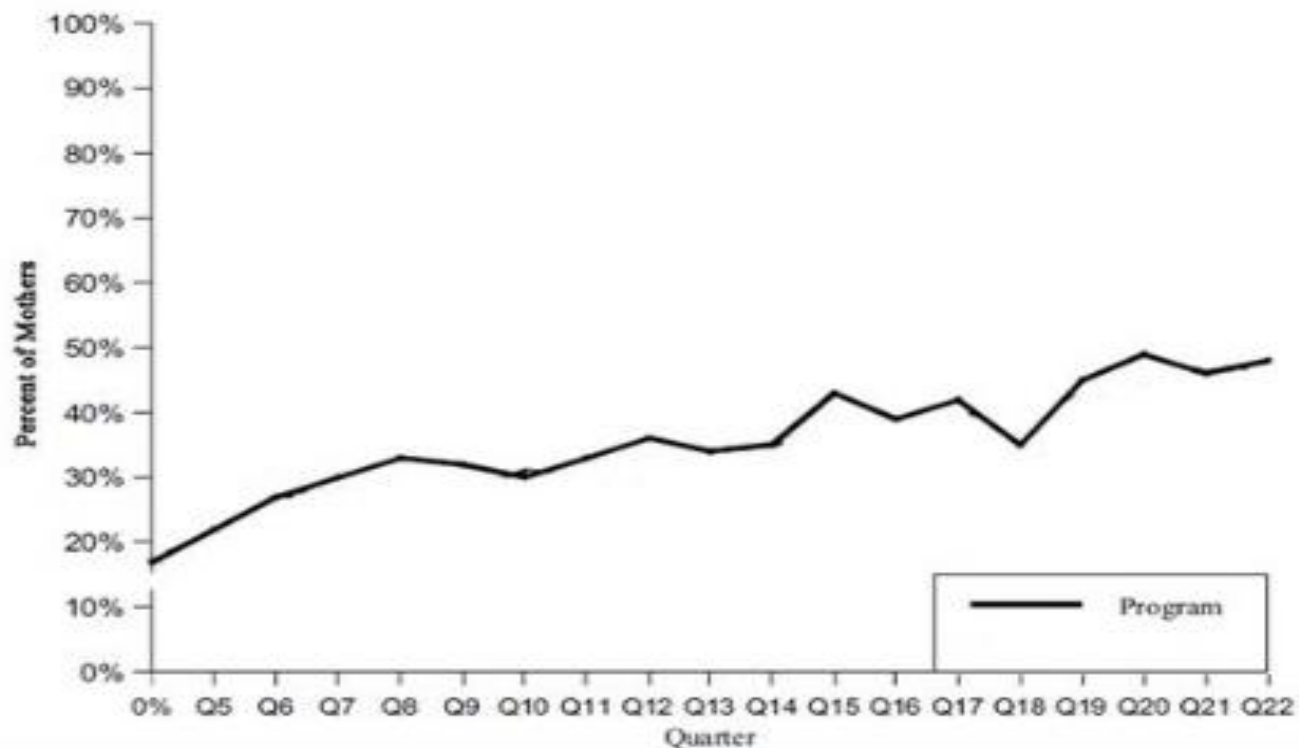
- i.e. extraneous influences/effects, also known as threats to internal validity
- impact of other programs/activities/events
  - general trends
  - maturation of clients
  - motivation of clients
  - differential selection
  - statistical regression

- measurement error
- sampling error
- analysis errors
- etc

Control for extraneous influences by: randomization, use of a control or comparison group, identifying and measuring them, logical elimination, or simply ignore them!

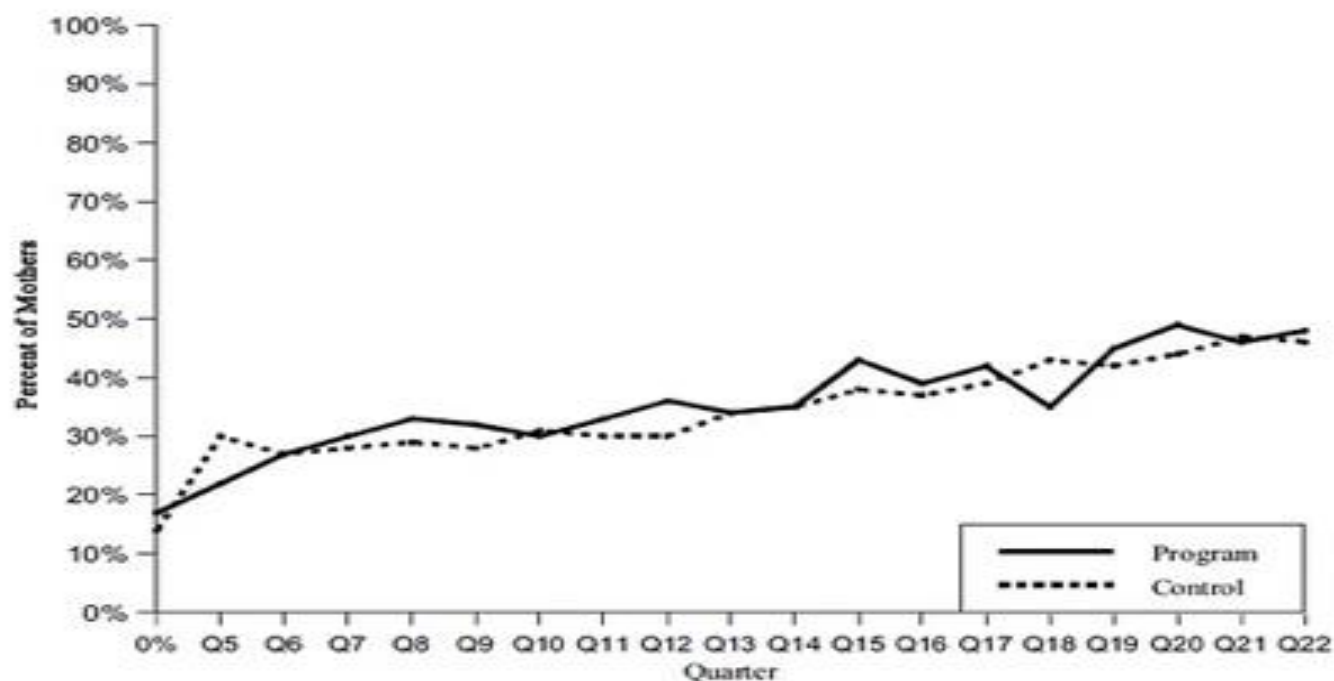
# Percent of Mothers Employed Over 5 years

---



# Percent of Mothers Employed Over 5 years

---



# Summary

1. Policy makers often want to know:  
Does it work? At what cost? What do stakeholders think? (but interactions are the norm)
2. In order to conclude that program (X) causes outcome (Y), three criteria must be satisfied:
  - Demonstrate an association between X & Y;
  - Establish the time order of X & Y; and
  - Rule out alternative explanations (Z).

## 2 Tools:

3.            Z  
                 ?  
         X            Y

Pearl & Mackenzie, 2019, The Book of Why: The New Science of Cause and Effect.

4.

**Observed Change = net program impact + effects of other events & processes + methodological error**

- the actual causal effect of the program in question
- these effects have direction, magnitude and duration

- i.e. extraneous influences/effects, also known as threats to internal validity
- impact of other programs/activities/events
  - general trends
  - maturation of clients
  - motivation of clients
  - differential selection
  - statistical regression

- measurement error
- sampling error
- analysis errors
- etc

5. Select the strongest possible research design. Add additional research designs and measures until you have adequately covered off on the 3 criteria. This is the key!

Example: Reynolds & West, 1987, 'A multiplist strategy for strengthening nonequivalent control group designs', Evaluation Review, 11, 6, 691-714. *(an excellent example of how to fix up a weak research design by adding additional features thereby improving your evaluation of the program's impact)*

6. Conducting an IE assumes a consistent intervention plus fidelity of implementation

7. Let's be careful not to confuse:

- theory building
- seeking evidence to support/confirm a theory,
- portraying the views and experiences of program participants/staff, or
- the opinions of experts

with impact evaluation (causal inference).

**These are very different things!**



## 8. Potential outcome comparisons include:

- program vs control/comparison groups
- intervention A vs intervention B vs intervention C ....
- dose response patterns
- outcome trends/trajectories over time
- outcomes across different locations, & across different target sub-groups



## 10. Stronger non-experimental research designs for IE include:

- Regression discontinuity design (modeling a known selection process)
- Interrupted time series with comparison group
- Multiple baseline designs
- Cohort designs
- Alternating treatment / removed treatment designs
- Comparative case studies or dose response models based upon a well developed understanding of the program and pattern matching (but be careful)

## Guiding principles:

Aim to:

- use high quality outcome measure(s), match local groups (program and comparison) on pre-treatment outcomes, control for self-selection (differences in motivation and incentives), measure outcomes several times - before, during, and after program participation (& allow for interactions).

# Impact Evaluation Approaches

## Method led

- case studies
- econometrics
- experiments
- process tracing



## Program theory led

- theories of change
- Mayne's contribution analysis  
(see Cook 2000)



## Driven by evidentiary criteria

- various philosophical schools (CM's 3 criteria)



# Quiz question and prize

In 2006 a World Bank study found that better M&E is associated with improved outcomes of development projects. Does this mean that M&E is an effective way to improve development results?

Z  
?  
X Y



# Questions & Discussion

**Scott Bayley**

Email: scottbayley\_evaluationervices@outlook.com

Website: <https://www.scottbayleyevaluationervices.com.au/>