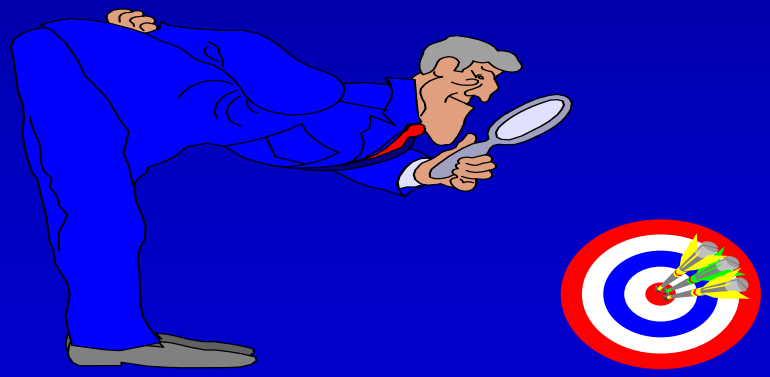
I’m doing an impact evaluation:

What evidence do I need?



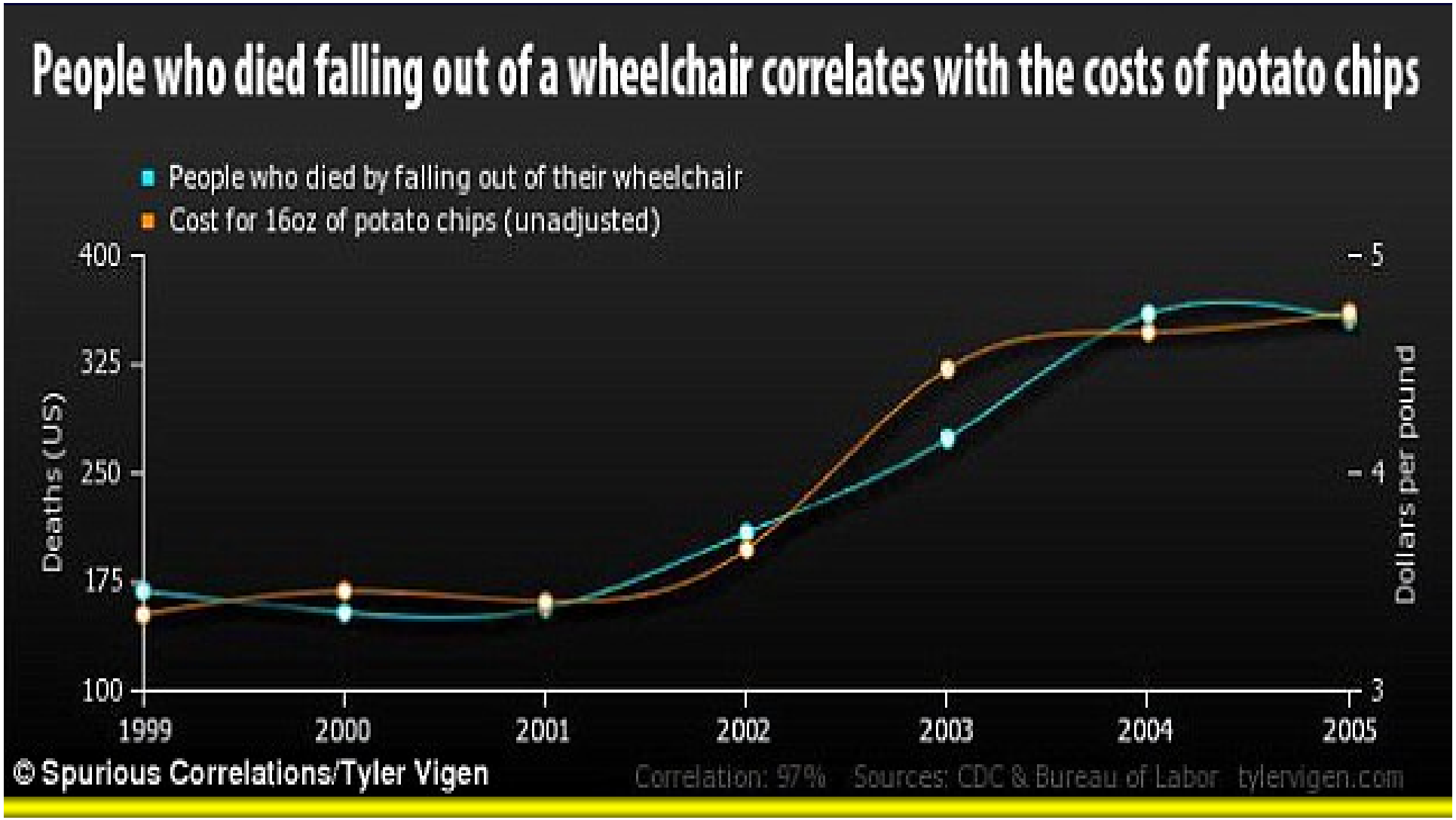
Scott Bayley

AES conference, Canberra Sept 2017

Overview

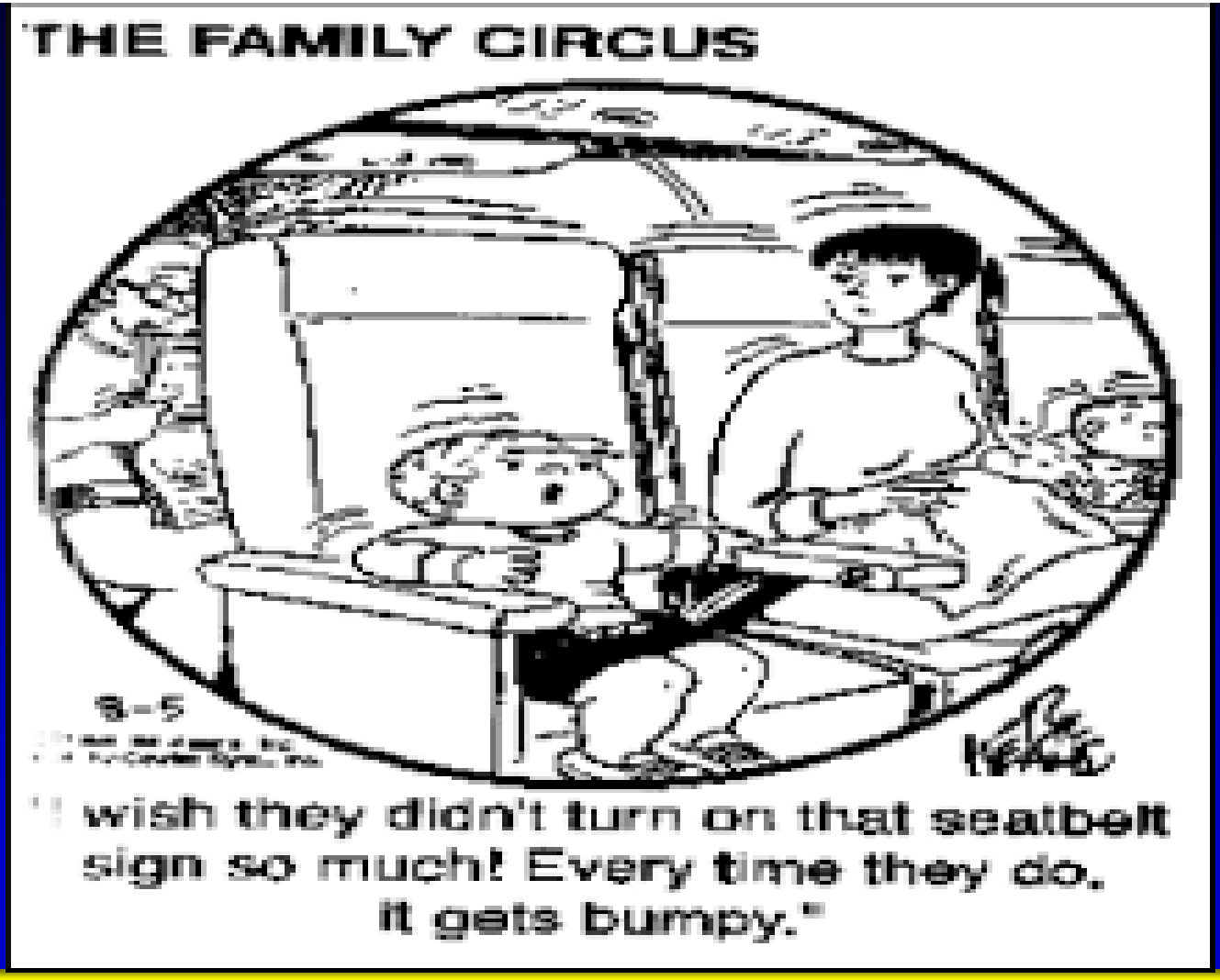
1. Examples: *What is going on here?*
2. Philosophy and evaluation models
3. Evidence requirements for IE
4. Lets practice
5. Summary

Quiz & prize


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)
Deaths (US)
400
325
250
175
100
Dollars per pound
5
4
3
1999
2000
2001
2002
2003
2004
2005
© Spurious Correlations/Tyler Vigen
Correlation: 97%
Sources: CDC & Bureau of Labor
tylervigen.com

Student:

I will study hard tomorrow in order to get a good grade in today's exam


THE FAMILY CIRCUS
8-5
"I wish they didn't turn on that seatbelt sign so much! Every time they do, it gets bumpy."

Philosophical perspectives

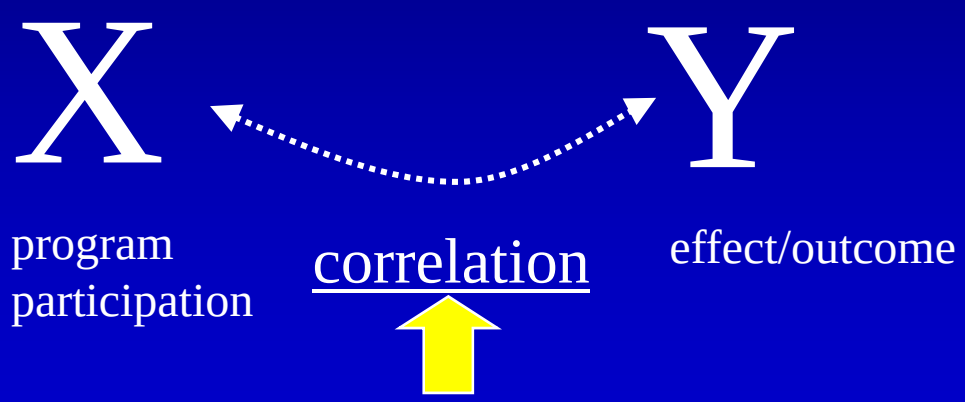
Evidence requirements for IE

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

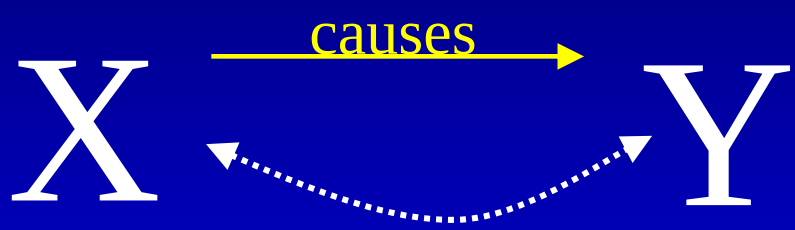
1. Demonstrate an association between participating in program X and outcome Y;
2. Establish time order (X<Y in time); and
3. Rule out alternative explanations (Z).

Four possible explanations:

①



②



program participation

correlation

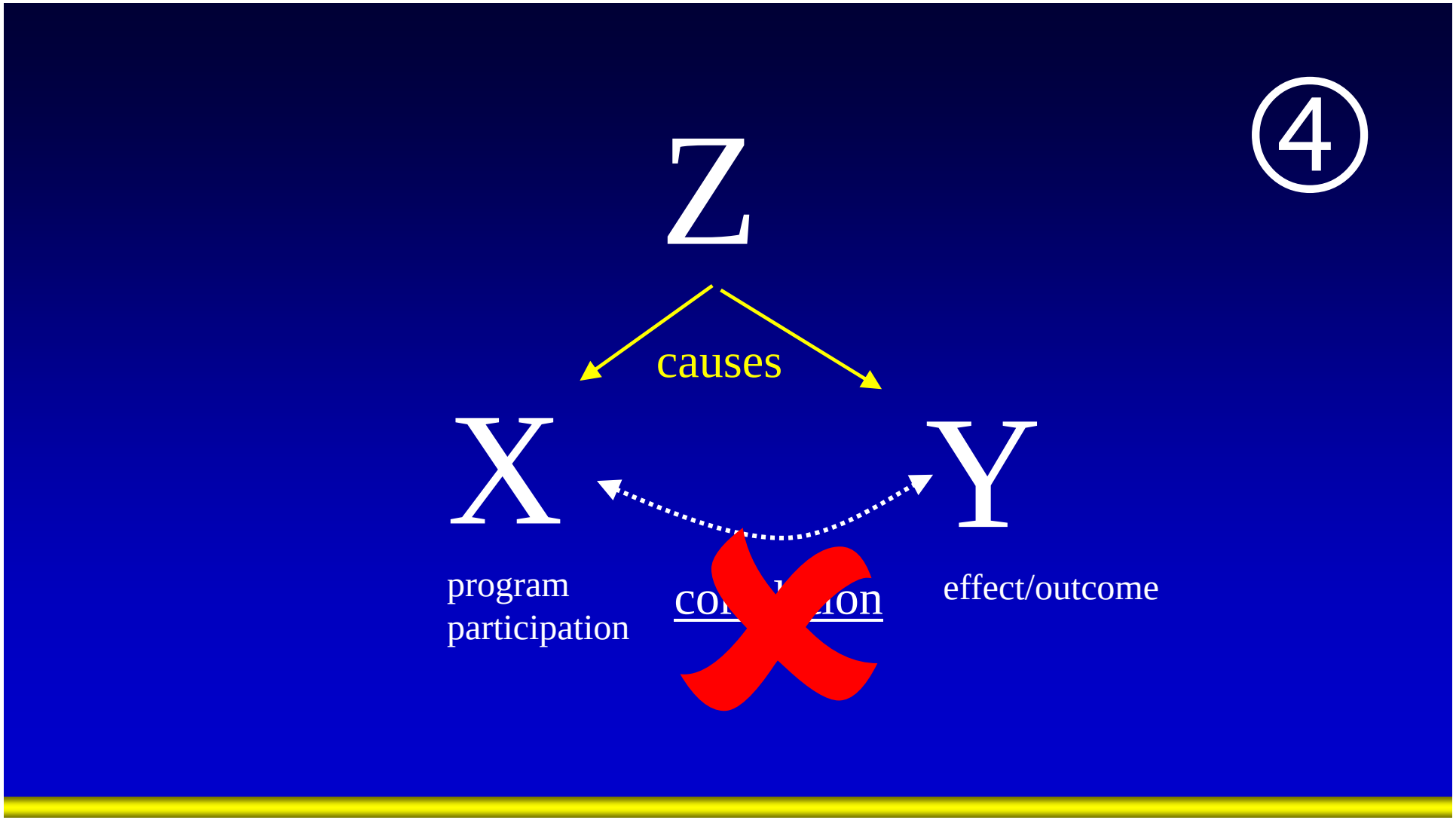
effect/outcome

③

program participation

correlation

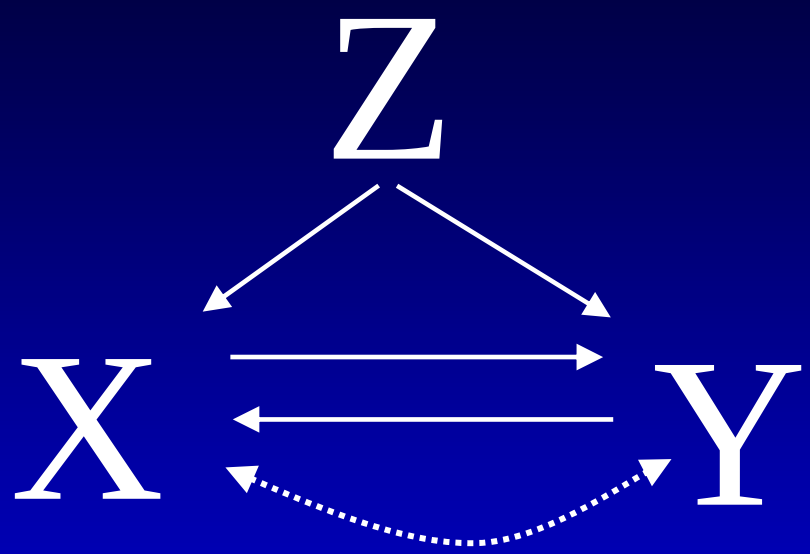
effect/outcome


④
Z
causes
X
Y
program
participation
effect/outcome

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



program participation

Correlation

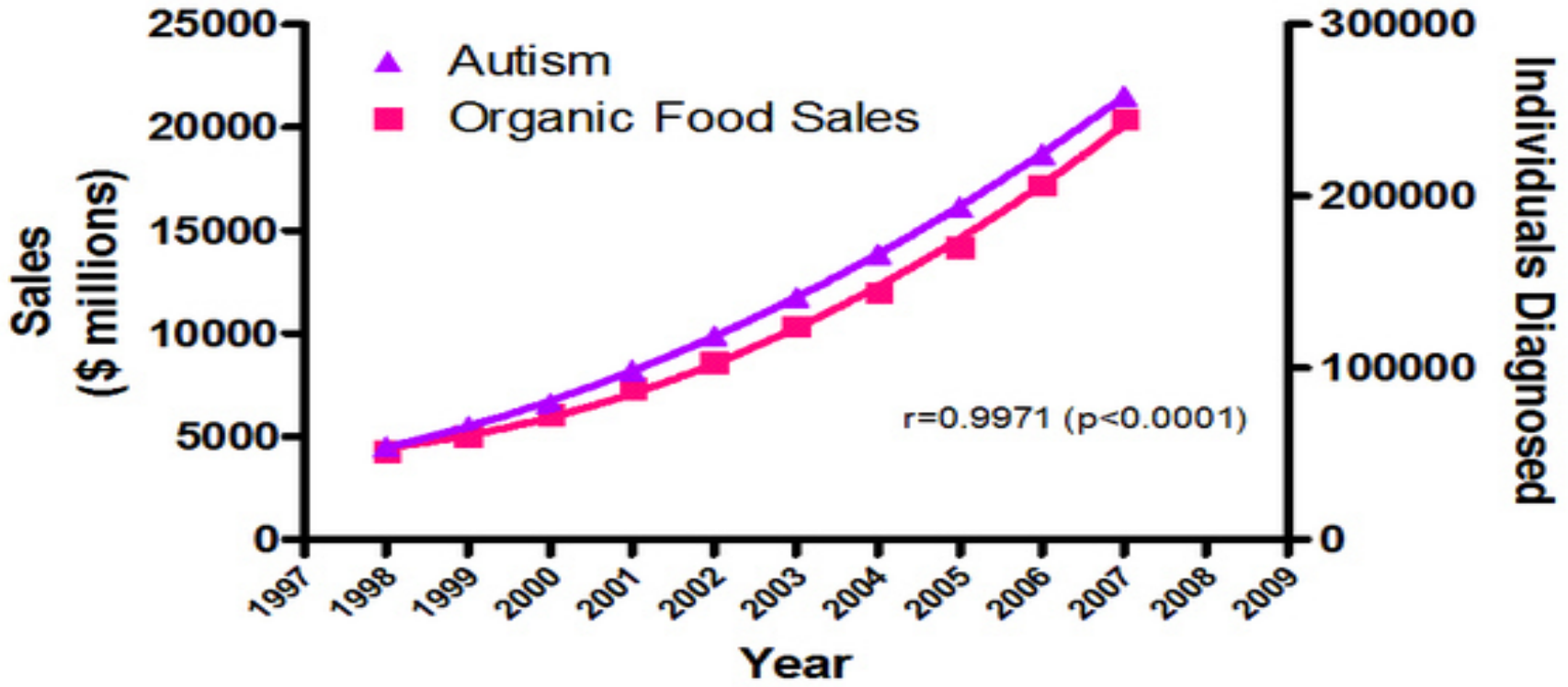
effect/outcome

Examples:

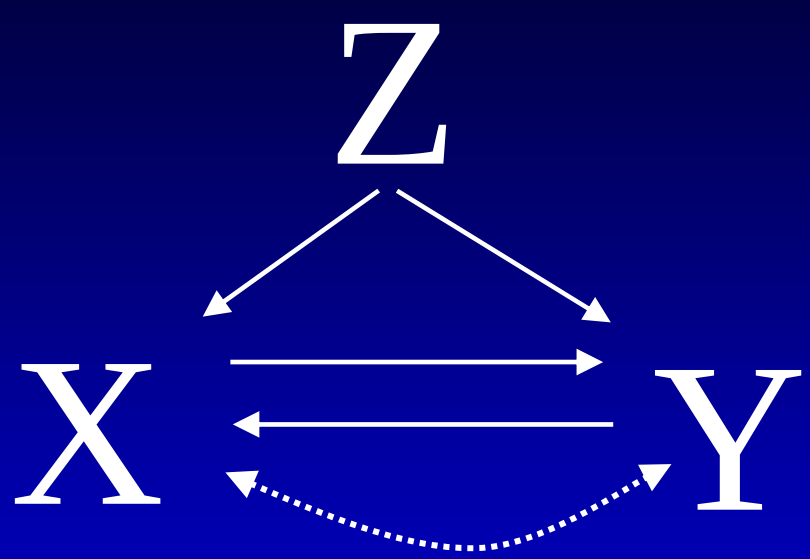
-# fire engines at location and the value of fire damage

-teenagers sleeping <5 hrs night and mental illness

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:

-attending 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

420 juvenile offences

1985
350 juvenile offences

- 440 offences down to 350 offences (20% reduction)

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

440 offences down to 350 offences (20% reduction):

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, identifying and measuring them, logical elimination, or try simply ignoring them!

Summary

1. For policy making we want to know:
   What works? For who? In what context? How? At what cost?

2. In order to conclude that program (X) causes outcome (Y), three criterion must be satisfied:

- Demonstrate an association between X & Y;
- Establish the time order of X & Y; and
- Rule out rival explanations (Z).

3.

Z

?

X Y

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 criterion. This is the key!

6. 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)

7. 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 well developed program logic models and pattern matching
- Match local geographically proximate groups on pre-treatment outcome measures while controlling for self selection

Today's quiz question

If people with higher levels of education earn larger salaries, does this mean that education is an effective way to reduce poverty?

Z

?

X Y