

V1 2019 USCOTS Workshop 1

Statistics Literacy For Decision Makers

Statistical Literacy Details Chapter 2

by
Milo Schield

USCOTS Workshop May 16, 2019
www.StatLit.org/pdf/2019-Schild-USCOTS-Slides2.pdf

V1 2019 USCOTS Workshop 2

Take CARE: Details Chapter 2 Outline

Associations: Comparison and Co-Variation

- Comparisons: Ordered and Arithmetic
- Comparisons: Kinds of Arithmetic

Take CARE: Solutions

- Confounder control: effect size, study design
- Assembly:
- Randomness: Test for statistical significance
- Error/Bias: Single & Double blind.

V1 2019 USCOTS Workshop 3

Stat Literacy studies Stats as Evidence in Arguments

The Point or the Target

The more disputable the point,
the stronger the evidence must be.

Statistic As Evidence

"All Statistics are Socially Constructed"
So, "Take CARE"!!

Statistics may be influenced by:

C	A	R	E
Context	Assembly	Randomness	Error

V1 2019 USCOTS Workshop 4

Associations: Two Kinds

Two-group comparisons:

- Men are taller than women
- Women live longer than men

Two-factor Covariation

- As height increases, weight increases
- The more height, the more weight

V1 2019 USCOTS Workshop 5

Comparisons: Two Kinds

Ordinal (Order): Women live longer than men

Arithmetic:

- Men shave six days more/week *than* women
6% is one percentage point more *than* the 5%
- Men shave seven times **as much as** women.
- Men save 600% **more** often *than* women.
6% is 20% **more** than 5%.
Men shave six **times more** often *than* women.
Women shave 7 **times less** often *than* men

V1 2019 USCOTS Workshop 6

Prevalence of Comparisons Google Ngrams

times as

times more

by a factor of

1900 1920 1940 1960 1980 2000

V1 2019 USCOTS Workshop 7

Confounding

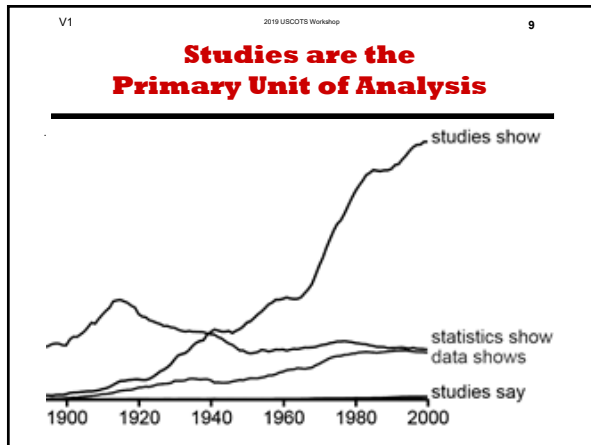
What things block or negate confounders?

1. Large effect size; large arithmetic comparison
2. Study design
3. Ratios
4. Comparison of ratios.
5. Selection and stratification
6. Standardizing

V1 2019 USCOTS Workshop 8

#1 Effect Size

1. *Does the association involve an effect size?*
If not, then no reason to think it is large
2. *Is the effect size material?* For example, a factor of 10 increase in 1 chance in 10,000.
3. *Is the effect size statistically significant?*
4. *Is the effect size large enough to ward off confounders?* A: RR>4, B: RR > 3, C: RR>2, D: RR > 1.5. Schield (2018, ICOTS).

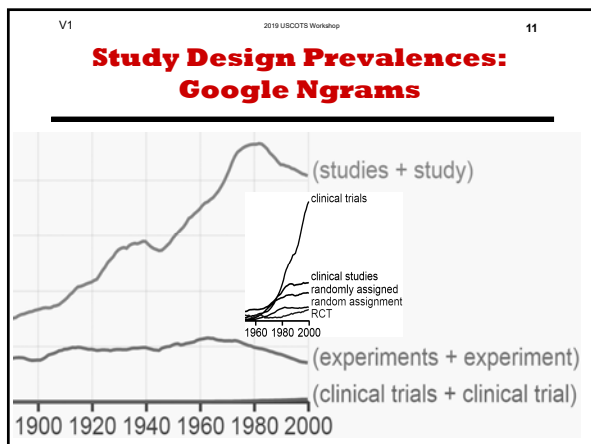


V1 2019 USCOTS Workshop 10

Six Basic Study Designs

Experiment	Observational
Researcher assigns/intervenes	Researcher is passive/observes
Repeatable: Scientific Exp.	Movie: Longitudinal Study
Randomized: Clinical Trial	Snapshot: Cross-sectional S.
Other: Quasi-Experiment	Someone says: Anecdotal

There are distinctions within these, but these six are enough to get started.



V1 2019 USCOTS Workshop 12

Random Assignment Nullifies Prior Confounding

Randomized controlled trials (RCT) are a major contribution of statistics to human knowledge.

By doing the impossible—controlling for all variations (known and unknown) — randomized trials can be considered a “statistical miracle.”

Experiments Gold std. RCT Silver std.

V1 2019 USCOTS Workshop 13

Random Assignment Examples

- 1747. Lind tests sailors with scurvy.
- 1935 Fisher: The Lady Tasting Tea.
- 1961 Perry Pre-School Project.
- 1974 RAND Health Insurance Experiment
- 1980s First AIDs trial video

V1 2019 USCOTS Workshop 14

Placebo Effect

Placebo Effect: Clinical trials where placebo group did as well as treatment group.

See migraine prophylaxis, positive response: Placebo meds, 22%. placebo acupuncture 38%. placebo surgery, 58%.

Note; Clinical studies, clinically proven, medical trials, medically proven, medical studies and controlled trials don't require randomization.

V1 2019 USCOTS Workshop 15

Study Designs

Quasi (Queasy)-Experiment
Nature or humans intervene on pre-existing groups

<i>Nature intervenes</i>	<i>Humans intervene</i>
Epidemics Plagues, outbreaks	Wars/Politics Change laws & policies
Natural disasters Earthquakes, tornadoes	Business/Education Change pricing/teaching

562 BC. Jews in Babylon test meat vs vegetarian diet.
 1796 Jenner administers cowpox to patient with smallpox
 1898 Lease of Hong Kong to the British for 99 years.
 1919-1933: US prohibits production/consumption of alcohol.

V1 2019 USCOTS Workshop 16

Quasi-Experiments: More Examples

1920 Watson's "Little Albert" study of social conditioning.
 1945 Post-WWII division of Germany into East and West.
 1945/48 Korea partition: North (USSR) and South (USA).
 1951 Asch Conformity Exp. 74% agreed w peers' falsehood.
 1954 Salk polio vaccine*. Biggest public health experiment.
 1968 Bystander Effect. Less likely to act if in a group.
 1987-2014: US states allow concealed carry of weapons (CCW)

* Salk: Second graders were treatment group; 1st and 3rd graders were control.
www.medicine.mcgill.ca/epidemiology/hanley/c622/salk_trial.pdf

V1 2019 USCOTS Workshop 17

Longitudinal Studies: Examples

Retrospective longitudinal studies : subjects recall past events. Cheap, quick.
 Prospective longitudinal studies: follow subjects through time.
 Expensive, time-consuming. Minimizes recall bias and sampling bias.
 Cross-sectional results are more reliable.

Prospective studies:

- 1921 Terman (Stanford) study of the gifted
- 1948 Framingham Study: Follow all inhabitants of Framingham MA
- 1951 British Doctors Survey
- 1976 Harvard Nurses Study
- 1979 Bouchard study of twins raised apart
- 1979 National Longitudinal Study of Youth (NLSY)

V1 2019 USCOTS Workshop 18

Cross Sectional Associations: Examples

- 1948 Framingham Study: Cross-sectional data associated heart attacks with high blood pressure, high cholesterol and smoking.
- 1951 British Doctors Survey. Cross-sectional data strongly associated lung-cancer deaths with smoking.
- 1979 Bouchard study of twins raised apart. Similarities between twins are due more to genes, less to environment.
- 1979 National Longitudinal Study of Youth. Cross-sectional data showed that social outcomes more strongly associated with individual IQ than with parents' socio-economic status. See *The Bell Curve* (1994) by Herrnstein and Murray.

V1 2019 USCOTS Workshop 19

Evaluating Study Designs Grades are Starting Points

CONTROL OF CONFOUNDERS	
Physical Control (Grade = Quality)	
Experiment	Observational Study
A+ Scientific	C Longitudinal
A- Random Assign	D Cross-sectional
B Quasi-Exper	F Anecdotal story

Which are cheapest?
Which are most common in the media?
Examples of uncontrolled quasi-experiments?

V1 2019 USCOTS Workshop 20

From Association to Causation

Association is not causation vs
Association is often evidence of causation.

Don't cross in the middle of the block vs.
look both ways before you do.

Sex is not love (Danny Kaplan) vs.
sex and love can be closely related.

V1 2019 USCOTS Workshop 21

Chance: Law of Very Large Numbers

The unlikely is almost certain given enough tries

Math: Suppose there is one chance in N for a given rare event on the next try.
The chance of having **at least*** one such event in N tries is over 50%—it is expected.

* Chance of having just one event < 50%.

V1 2019 USCOTS Workshop 22

Chance: Statistical Significance

Consider matched statistics from two groups.
If their 95% intervals don't overlap, then their difference is statistically significant. Otherwise, the difference may be statistically insignificant.

Suppose 70% of gals dream in color (40% of guys) and the 95% margin of error is 10 points.
The associated 95% confidence intervals are 60 to 80% for gals (30 to 50% for guys).
The 30 point difference is statistically significant.

V1 2019 USCOTS Workshop 23

Case Study: The Prontosil Experiment

Before 1936, as many as one in three expectant moms died from puerperal fever following birth.

Gerhard Domagk, a German doctor, developed Prontosil to fight against streptococcal infections.

In 1936, Prontosil was administered to 38 newly delivered mothers, all suffering from puerperal fever. Three died and thirty-five survived.

V1 2019 USCOTS Workshop 24

Case Study: The Prontosil Experiment

When Prontosil was administered earlier in the course of the infection, no mother died.

In 1936, Prontosil was used to treat Franklin D. Roosevelt, Jr., the President's son.

This was the moment when the world realized that drugs were potent alternatives to surgery.

V1 2019 USCOTS Workshop 25

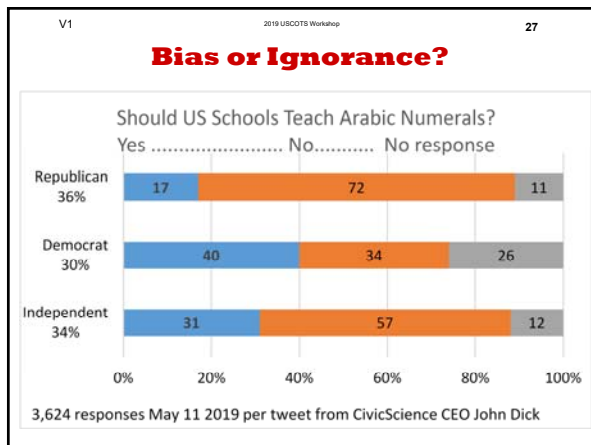
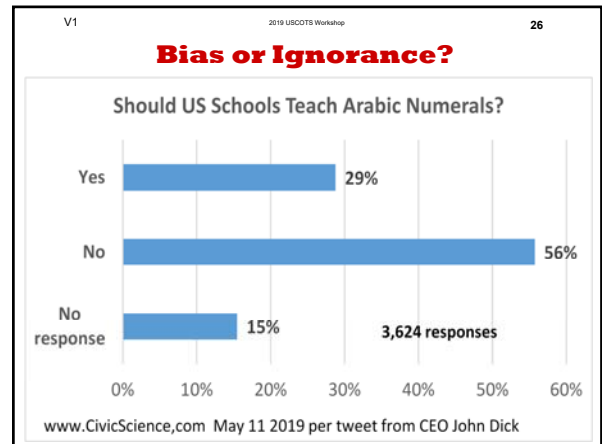
Case Study

Do Magnets Reduce Pain?

Fifty subjects having pain associated with post-polio syndrome were randomly assigned. The treatment group received concentric magnets; the control group received inert placebo magnets. A major decrease in pain was reported by 75% in the treatment group 19% in the control group.

- Natural Health, August, 1998. Page 52.

Effect size. Study design.
Hypothetical thinking using Take CARE.



Statistics Literacy For Decision Makers

Statistical Literacy Details Chapter 2

**by
Milo Schield**

USCOTS Workshop May 16, 2019

www.StatLit.org/pdf/2019-Schild-USCOTS-Slides2.pdf

Take CARE: Details

Chapter 2 Outline

Associations: Comparison and Co-Variation

- Comparisons: Ordered and Arithmetic
- Comparisons: Kinds of Arithmetic

Take CARE: Solutions

- Confounder control: effect size, study design
- Assembly:
- Randomness: Test for statistical significance
- Error/Bias: Single & Double blind.

Stat Literacy studies Stats as Evidence in Arguments

The Point or the Target

The more disputable the point,
the stronger the evidence must be.

Statistic As Evidence

“All Statistics are Socially Constructed”

So, “Take CARE”!!

Statistics may be influenced by:

C	A	R	E
Context	Assembly	Randomness	Error

Associations: Two Kinds

Two-group comparisons:

- Men are taller than women
- Women live longer than men

Two-factor Covariation

- As height increases, weight increases
- The more height, the more weight

Comparisons: Two Kinds

Ordinal (Order): Women live longer than men

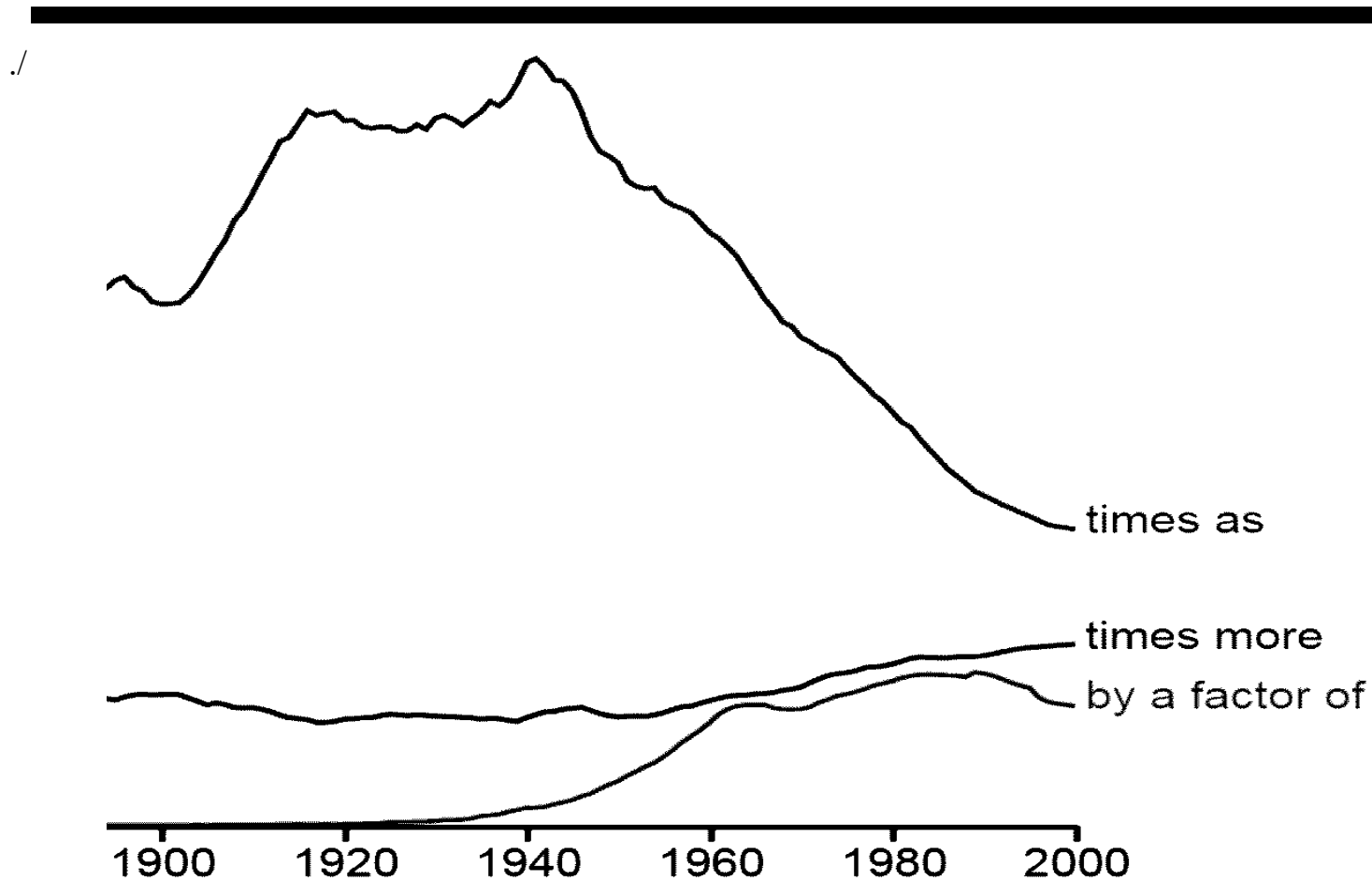
Arithmetic:

- Men shave six days more/week *than* women
6% is one percentage point more *than* the 5%
- Men shave seven times **as much as** women.
- Men shave 600% **more** often *than* women.
6% is 20% **more** than 5%.

Men shave six **times more** often *than* women.

Women shave 7 **times less** often *than* men

Prevalence of Comparisons Google Ngrams



Confounding

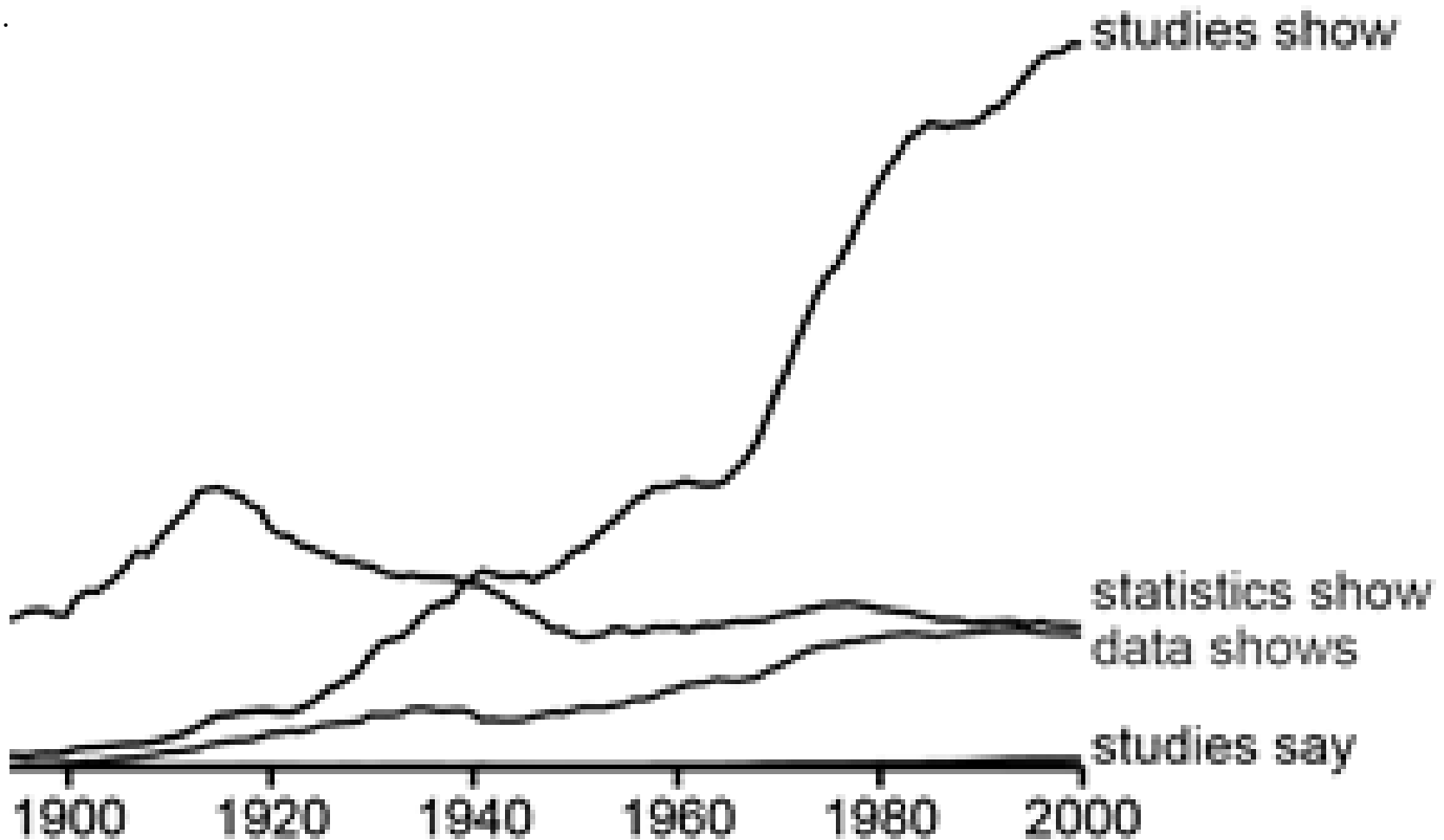
What things block or negate confounders?

1. Large effect size; large arithmetic comparison
2. Study design
3. Ratios
4. Comparison of ratios.
5. Selection and stratification
6. Standardizing

#1 Effect Size

1. *Does the association involve an effect size?*
If not, then no reason to think it is large
2. *Is the effect size material?* For example, a factor of 10 increase in 1 chance in 10,000.
3. *Is the effect size statistically significant?*
4. *Is the effect size large enough to ward off confounders?* A: $RR > 4$, B: $RR > 3$, C: $RR > 2$, D: $RR > 1.5$. Schield (2018, ICOTS).

Studies are the Primary Unit of Analysis



Six Basic Study Designs

Experiment

Researcher assigns/intervenes

Repeatable: Scientific Exp.

Randomized: Clinical Trial

Other: Quasi-Experiment

Observational

Researcher is passive/observes

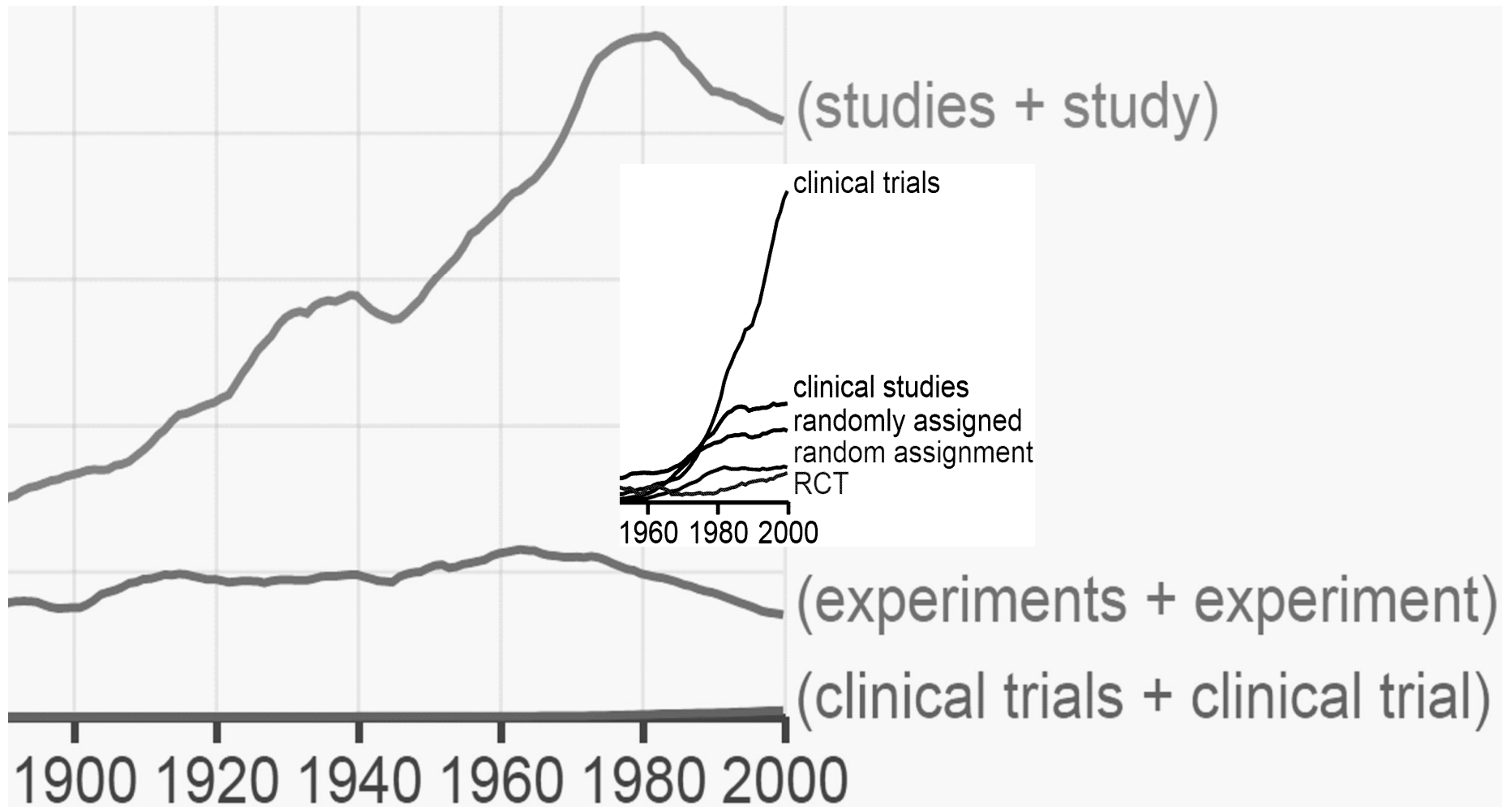
Movie: Longitudinal Study

Snapshot: Cross-sectional S.

Someone says: Anecdotal

There are distinctions within these, but these six are enough to get started.

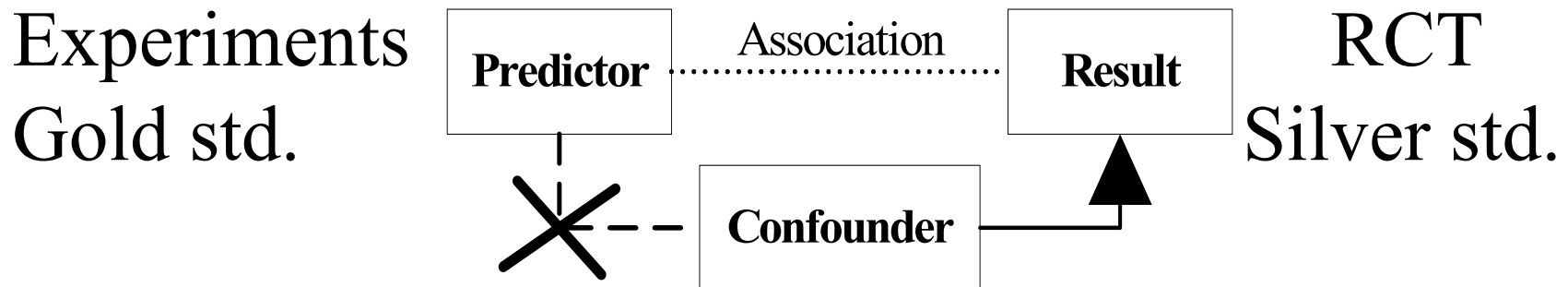
Study Design Prevalences: Google Ngrams



Random Assignment Nullifies Prior Confounding

Randomized controlled trials (RCT) are a major contribution of statistics to human knowledge.

By doing the impossible—controlling for all variations (known and unknown) — randomized trials can be considered a “statistical miracle.”



Random Assignment Examples

- 1747. Lind tests sailors with scurvy.
- 1935 Fisher: The Lady Tasting Tea.
- 1961 Perry Pre-School Project.
- 1974 RAND Health Insurance Experiment
- 1980s First AIDs trial video

Placebo Effect

Placebo Effect: Clinical trials where placebo group did as well as treatment group.

See migraine prophylaxis, positive response:
Placebo meds, 22%. placebo acupuncture 38%.
placebo surgery, 58%.

Note; Clinical studies, clinically proven, medical trials, medically proven, medical studies and controlled trials don't require randomization.

Study Designs

Quasi (Queasy)-Experiment

Nature or humans intervene on pre-existing groups

Nature intervenes

Epidemics

Plagues, outbreaks

Natural disasters

Earthquakes, tornadoes

Humans intervene

Wars/Politics

Change laws & policies

Business/Education

Change pricing/teaching

562 BC. Jews in Babylon test meat vs vegetarian diet.

1796 Jenner administers cowpox to patient with smallpox

1898 Lease of Hong Kong to the British for 99 years.

1919-1933: US prohibits production/consumption of alcohol.

Quasi-Experiments: More Examples

1920 Watson's "Little Albert" study of social conditioning.

1945 Post-WWII division of Germany into East and West.

1945/48 Korea partition: North (USSR) and South (USA).

1951 Asch Conformity Exp. 74% agreed w peers' falsehood.

1954 Salk polio vaccine*. Biggest public health experiment.

1968 Bystander Effect. Less likely to act if in a group.

1987-2014: US states allow concealed carry of weapons (CCW)

* Salk: Second graders were treatment group; 1st and 3rd graders were control.
www.medicine.mcgill.ca/epidemiology/hanley/c622/salk_trial.pdf

Longitudinal Studies: Examples

Retrospective longitudinal studies : subjects recall past events. Cheap, quick.

Prospective longitudinal studies: follow subjects through time.

Expensive, time-consuming. Minimizes recall bias and sampling bias.

Cross-sectional results are more reliable.

Prospective studies:

- 1921 Terman (Stanford) study of the gifted
- 1948 Framingham Study: Follow all inhabitants of Framingham MA
- 1951 British Doctors Survey
- 1976 Harvard Nurses Study
- 1979 Bouchard study of twins raised apart
- 1979 National Longitudinal Study of Youth (NLSY)

Cross Sectional Associations: Examples

- 1948 Framingham Study: Cross-sectional data associated heart attacks with high blood pressure, high cholesterol and smoking.
- 1951 British Doctors Survey. Cross-sectional data strongly associated lung-cancer deaths with smoking.
- 1979 Bouchard study of twins raised apart. Similarities between twins are due more to genes, less to environment.
- 1979 National Longitudinal Study of Youth. Cross-sectional data showed that social outcomes more strongly associated with individual IQ than with parents' socio-economic status. See *The Bell Curve* (1994) by Herrnstein and Murray.

Evaluating Study Designs

Grades are Starting Points

CONTROL OF CONFOUNDERS			
Physical Control (Grade = Quality)			
Experiment		Observational Study	
A+	Scientific	C	Longitudinal
A-	Random Assign	D	Cross-sectional
B	Quasi-Exper	F	Anecdotal story

Which are cheapest?

Which are most common in the media?

Examples of uncontrolled quasi-experiments?

From Association to Causation

Association is not causation vs
Association is often evidence of causation.

Don't cross in the middle of the block vs.
look both ways before you do.

Sex is not love (Danny Kaplan) vs.
sex and love can be closely related.

Chance: Law of Very Large Numbers

The unlikely is almost certain given enough tries

Math: Suppose there is one chance in N for a given rare event on the next try.

The chance of having **at least*** one such event in N tries is over 50%—it is expected.

* Chance of having just one event $< 50\%$.

Chance: Statistical Significance

Consider matched statistics from two groups. If their 95% intervals don't overlap, then their difference is statistically significant. Otherwise, the difference may be statistically insignificant.

Suppose 70% of gals dream in color (40% of guys) and the 95% margin of error is 10 points.

The associated 95% confidence intervals are 60 to 80% for gals (30 to 50% for guys).

The 30 point difference is statistically significant.

Case Study: The Prontosil Experiment

Before 1936, as many as one in three expectant moms died from puerperal fever following birth.

Gerhard Domagk, a German doctor, developed Prontosil to fight against streptococcal infections.

In 1936, Prontosil was administered to 38 newly delivered mothers, all suffering from puerperal fever. Three died and thirty-five survived.

Case Study: The Prontosil Experiment

When Prontosil was administered earlier in the course of the infection, no mother died.

In 1936, Prontosil was used to treat Franklin D. Roosevelt, Jr., the President's son.

This was the moment when the world realized that drugs were potent alternatives to surgery.

Case Study

Do Magnets Reduce Pain?

Fifty subjects having pain associated with post-polio syndrome were randomly assigned.

The treatment group received concentric magnets; the control group received inert placebo magnets.

A major decrease in pain was reported by 75% in the treatment group 19% in the control group.

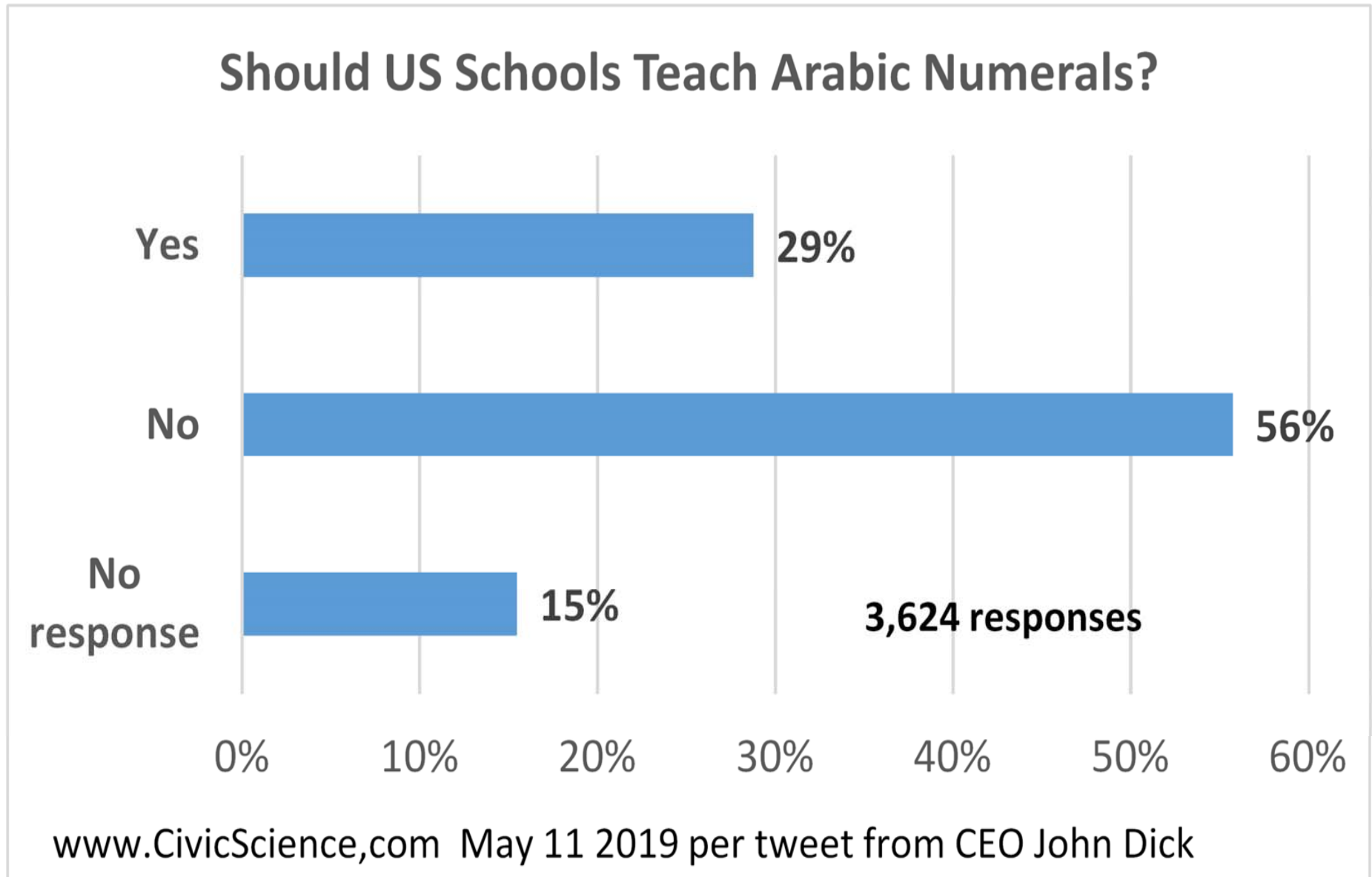
- Natural Health, August, 1998. Page 52.

Effect size. Study design.

Hypothetical thinking using Take CARE.

/

Bias or Ignorance?



Bias or Ignorance?

