V1 2019 USCOTS Workshop 1

Teaching Statistical Literacy

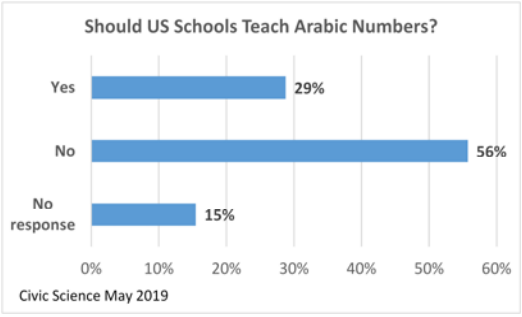
Chapter 1
by
Milo Schield

Half-Day Workshop
USCOTS May 16, 2019

www.StatLit.org/pdf/2019-Schield-USCOTS-slides1.pdf

V1 2019 USCOTS Workshop 2

First Sharia math, then Sharia law!!!



V1 2019 USCOTS Workshop 3

Working Moms; Better Kids

http://money.com/money/5272659/working-moms-better-kids/

V1 2019 USCOTS Workshop 4

Outline

Introduction:
A1. Who takes intro statistics
A2. SAT level of our students by college
A3. Math level of our students by major
Exp vs. Obs: What kinds are relevant?
A3. Kinds of influence on statistics
How common are these influences?
A4. Grammar: Association vs. causation

V1 2019 USCOTS Workshop 5

Goals of this Workshop

1. Present my view of statistical literacy
2. Expose you to lots of new ideas
3. Present a coherent structure for teaching
4. Show the importance of English grammar
5. Show simple ways of handling significance
6. Show simple ways of handling confounding
7. Show how confounding changes significance
8. Role-model analyzing studies

V1 2019 USCOTS Workshop 6

Fraction of 4-year Undergrads that take Intro Stats?

57%

Schield (2016, IASE)

V1 2019 USCOTS Workshop 7

Fraction of Course Gain that Stat Students Loose in 4 Months

50%

Tintle et al, 2013

V1 2019 USCOTS Workshop 8

Student Attitudes Toward Stats

Of those taking Stat I:

- less than 1% take *Stat II* (10-yrs @ U. St. Thomas)
- less than 0.2% major in statistics (nationwide).
- most see less value in statistics after the course than they did before. Schield and Schield (2008).
- too many say "Worst course I ever took" [anecdotal]

www.amstat.org/misc/StatsBachelors2003-2013.pdf 1,135 stat majors in 2013 at 32 colleges www.StatLit.org/pdf/2015-Schield-UST-Enroll-in-Statistics.pdf

V1 2019 USCOTS Workshop 9

What fraction of 4-Yr Intro Stat students are taught outside Math?

50%

Estimates by Schield (2015, Statchat)

V1 2019 USCOTS Workshop 10

Who takes Intro Statistics at Four-Year Colleges?

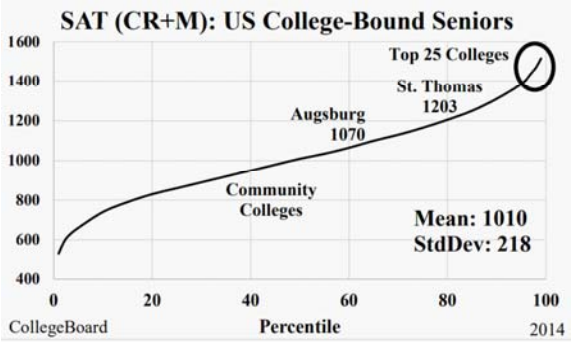
Table 1: Distribution of Majors in Stat 101

%	Major
38%	Business or Economics
19%	Social Science or History
13%	Health
10%	Psychology
9%	Engineering
9%	Biological Science
2%	Math or Statistics
100%	All students in these majors

Schield (2016, IASE). Inferred from data in 2012 US Statistical Abstract.

V1 2019 USCOTS Workshop 11

Where are your students?



Schield (2016, IASE)

V1 2019 USCOTS Workshop 12

SAT Math Percentile by Major

SAT MATH	PERCENTILE	MAJOR
613	80%	Math/Stats
585	72%	Physical Sciences
579	70%	Engineering
554	62%	Comp. Science
551	61%	Biological
550	61%	Social Sciences
522	51%	Business
522	51%	English Lang/Lit
506	46%	History
498	43%	Communication
489	40%	Psychology
482	38%	Education

Business Insider (2014), College Board (2015)

SAT Math Scores: Average by Student Major

Percentiles of all those taking the Math SAT

Schield (2016, IASE)

V1 2019 USCOTS Workshop 13

GAISE 2016 Update

The real world is complex and can't be described well by one or two variables.

If students do not have exposure to simple tools for disentangling complex relationships,

they may dismiss statistics as an old-school discipline only suitable for small sample inference of randomized studies.

V1 2019 USCOTS Workshop 14

GAISE 2016 Update

Multivariable thinking is critical to make sense of the observational data around us

- *learn to identify observational studies*
- *learn to consider potential confounding factors*
- *use … stratification … to show confounding*

This report recommends that students be introduced to multivariable thinking, preferably early in the introductory course and not as an afterthought at the end of the course.

V1 2019 USCOTS Workshop 15

Most Important Topics: Student Choices

Rank	The most important topics in Statistical Literacy for Managers
1	Take CARE: Confounding, Assembly, Randomness and Error/bias
2	Confounding
2	Hypothetical thinking: plausible confounders, plausible definitions
4	Statistics are more than numbers. They include the context
5	Association-causation (Luck-skill) including the grammar
5	Bias: Placebo, Single blind; double blind
5	Named Ratios and Ratio grammar; Percent, Percentages, Rates
5	Read tables and graphs

Schield (2016, ASA)

V1 2019 USCOTS Workshop 16

A-B-C Words: A = Association

Statistical association is not the same as Basketball Assoc.

Association words assert association explicitly or describe associations involving fixed conditions or unrepeatable events.

Association: Height is *associated* with age in children
Obesity is *correlated* with (related to) diabetes.

Prediction: Graduating from high school *predicts* success in life.

---

*Comparisons: People with degrees earn *more than* those without
Whites have a *higher* risk of suicide *than* blacks.

*Co-variation: *As children get older,* their weight *increases*.

* Manipulation is impossible, or treatment or outcome cannot be repeated.
Schield (2018, SL4DM)

V1 2019 USCOTS Workshop 17

A-B-C Words: C = Causation

Causation words assert causation, sufficiency or contra-factual

Causation: A bomb *caused* the fire. Insomnia is a side *effect*.
Lightning *resulted in* a fire. Spark results in a fire.

Sufficient: The more X you do, the more Y *you will get*.
Prevent, *stop*, *end*, *start*, *kill*, *produce*, *cure*, *avoid*, *ban*, *quit*, *block*, *ward off*, *stave off*, *cancel*, *hinder*, or *eliminate*.[6]

Contra-factual: Those who do X *will* get more Y *than if they had not done X.*

V1 2019 USCOTS Workshop 18

A-B-C Words: B = Between

Between words describe association but imply causation

Verbs: Red wine *cuts* cancer risk. TV *ups* kids' risk of flunking.
Gene X *increases* health risk. Smoking *raises* asthma risk.

Connectors: Nuts *linked to* cancer. Trauma *tied to* heart disease.

Contributor Diet *contributes* to diabetes. Age is *factor* in infertility

Nouns: Spinach is *asthma protector*. Bad water is a *killer*.

Logicals: Anxiety increases *due to* (*because of*) high stake testing

---

*Compare: People who take antidepressants have fewer migraines
Asthma attacks more likely for smokers *than* non-smokers.

*Covariation: *As* teacher pay *increases*, student scores increase.
The more hours worked, the *more likely* a promotion

*Manipulation is possible, and treatment and outcome are repeatable.

A-B-C Words: Distribution in Headlines

Of the 2,000 news headlines analyzed[6],
71% involved A, B or C.

Of those headlines involving A, B or C,

- **86% were "between" claims**,
- 11% sufficiency, 3% causation, 3% association.

6. Schield and Raymond (2009).

Association is not causation

This statement is ambiguous. It can mean:

1 Association is not sufficient to prove causation
2 Association provides no evidence for causation.

Teachers may intend #1; students often hear #2.

A better statement would be:
Association is evidence of causation somewhere.

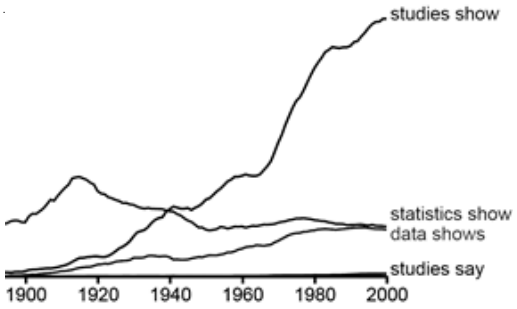
Association is not causation

No idea has stifled the growth of statistical literacy as much as the endless repetition of the words "correlation is not causation".

This phrase seems to be primarily used to suppress intellectual inquiry --
by encouraging the unspoken assumption that correlational knowledge is somehow an inferior form of knowledge.

John Myles White (2010):
www.johnmyleswhite.com/notebook/2010/10/01/three-quarter-truths-correlation-is-not-causation/

Studies are the Primary Unit of Analysis



Harvard Case Studies: Title or Abstract

#	INFERENTIAL		CONTROL/CONFOUND	
22	"clinical trial"	18	2,263	control
7	"statistical significance"		234	"control of" 200
4	"statistically significant"		113	"take (ing) into account"
3	"standard error"		30	"compensate (ing) for"
1	"sampling error"		19	"control (ed, ing) for"
1	"margin of error"		18	confound (er, ing)
1	"prediction interval"		17	"adjust(ed, ing) for"
1	p-value		3	"sampling bias"
0	"sampling distribution"		0	"alternate explanation"
0	"confidence interval"		0	"common cause"
0	"null hypothesis"		0	"effect modifier"
0	"reject the null"		0	"Simpson's paradox"
0	"random assignment"		0	"lurking variable"

Statistical Literacy : An Overview

Statistics are numbers in a context
Association is not causation

RANDOMNESS and CAUSATION	CONFOUNDING and CAUSATION
Chance, independence and sampling distributions	Comparisons, ratios, models and study designs
Margin of error, hyp tests & statistical significance	Epidemiological causation (Bradford-Hill)
Random assignment and causation (Fisher: RCT)	*Confounder conditions for nullification (Cornfield)*

V0.7 Conditional probability, medical tests and Bayesian reasoning
Coincidence, Simpson's Paradox and regression to the mean

V1 2019 USCOTS Workshop 25

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 26

Statistical Literacy : Assembly

Living with AIDs

All (1,000)	White (non-Hispanic)	Black (non-Hispanic)	Hispanic
434	150	186	78

Q1. Which group is largest?

Consolidate White (Non-Hispanic) with Hispanic.

Q2. Which group is largest?

V1 2019 USCOTS Workshop 27

Statistical Literacy : Randomness

Five non-quantitative Topics:

1. **Regression to the Mean**
   Sport Illustrated Cover
2. **Statistically significant**
3. **Chance-Related Mistakes:**
   Three Door problem; Birthday problem

- **Better than chance**
- **Unlikely to be chance**

V1 2019 USCOTS Workshop 28

Statistical Literacy : Error/Bias

Three kinds of error

1. **Subject/respondent error:**
2. **Researcher/measurement error:**
3. **Sampling error:**

V1 2019 USCOTS Workshop 29

Statistical Literacy : Assembly

Child Abuse Statistics

Each year, more than 7,000 children in Minnesota are confirmed to be victims of physical or sexual abuse, emotional maltreatment, or neglect.

V1 2019 USCOTS Workshop 30

Statistical Literacy : Recommendation

More college students (over half) take intro statistics than any other course (except English).

One-size fits all is no longer viable. Statistics education must support Stat 101 and 100/102.

Statistics education should (1) support different flavors for different majors, and (2) agree on the contributions of statistics to human knowledge.

V1 2019 USCOTS Workshop 31

Willful Ignorance

The past success of statistics has depended on vast, deliberate simplifications amounting to willful ignorance.

This very success now threatens future advances in medicine, the social sciences, and other fields.

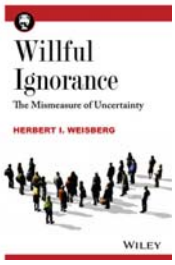
Limitations of existing methods result in frequent reversals of scientific findings/recommendations, to the consternation of scientists and the public.

Herbert I. Weisberg

V1 2019 USCOTS Workshop 32

Willful Ignorance Herbert Weisberg

The past success of statistics has depended on vast, deliberate simplifications amounting to willful ignorance.



Limitations of existing methods result in frequent reversals of scientific findings and recommendations, to the consternation of scientists and the lay public.

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-Schield-USCOTS-Slides2.pdf

V1 2019 USCOTS Workshop 2

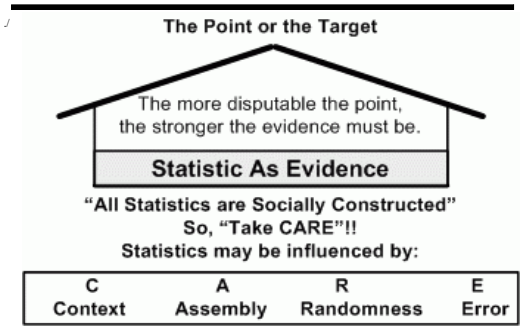
Take CARE: Details Chapter 2 Outline

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



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

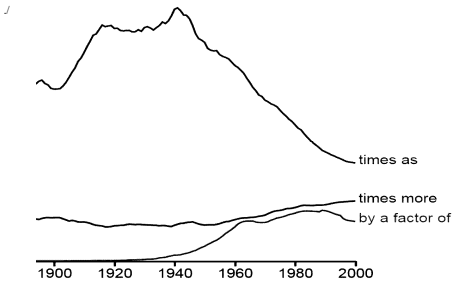
Comparisons: Two Kinds

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



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

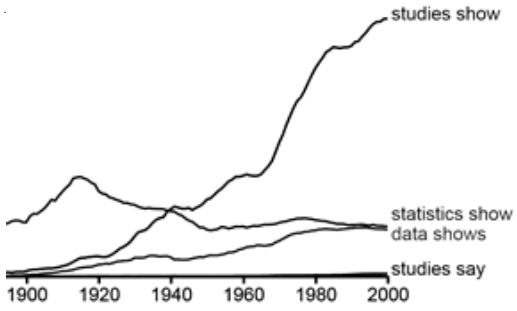
V1 2019 USCOTS Workshop 8

#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 9

Studies are the Primary Unit of Analysis



V1 2019 USCOTS Workshop 10

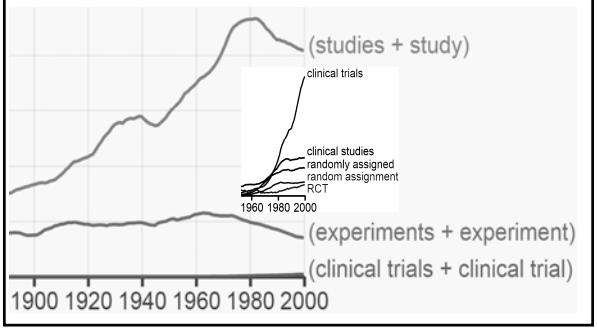
Six Basic Study Designs

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 11

Study Design Prevalences: Google Ngrams



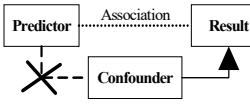
V1 2019 USCOTS Workshop 12

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

Nature intervenes	*Humans intervene*
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 Brouchard 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 Brouchard 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.

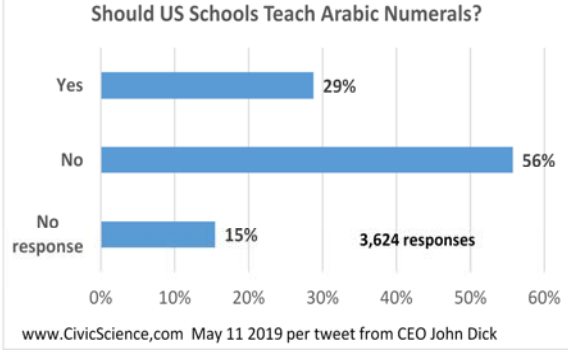
A major decrease in pain was reported by 75% in the treatment group 19% in the control group.

- Natural Health, August, 1998. Page 52.

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

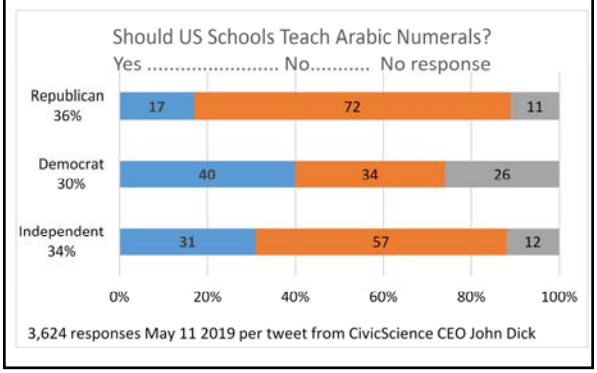
V1 2019 USCOTS Workshop 26

Bias or Ignorance?



V1 2019 USCOTS Workshop 27

Bias or Ignorance?



Statistics Literacy For Decision Makers

Chapter 3: Measurements

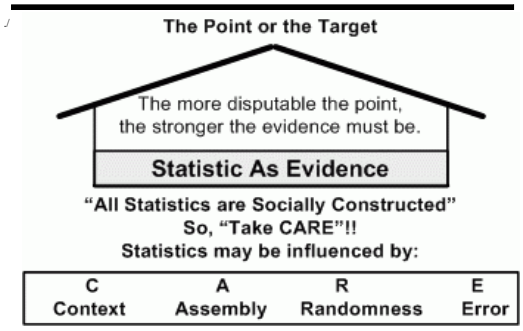
by
Milo Schield

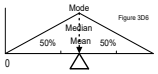
Half-Day Workshop
USCOTS May 16, 2019
www.StatLit.org/pdf/2019-Schield-USCOTS-Slides3.pdf

Measurements: Chapter 3 Outline

Distributions
Measures of center
Two-group comparisons of Means & Medians
Two-variable co-variation
Spread
Slope and simple regression

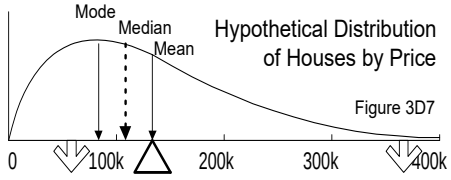
Stat Literacy: Study Statistics as Evidence in Arguments





Measures of Center

In an asymmetric distribution, mean, median and mode typically align alphabetically with mean most sensitive to extremes. Why?



Mean, median, mode: Alphabetically. Why?

Suppose that house prices in your town have a positive near-symmetric distribution

Suppose Bill and Melinda Gates move to your town. They built two Mac-Mansions.

How does that change the mode, median and mean of the original distribution?

Mode? Median? Mean?

Most relevant in the short run? In the long-run?

Issues:

1. Mean is more sensitive to outliers.
   Yet statisticians prefer the mean. Why?
2. Omit measure: *City1 income more than City2.*
3. Omit characteristic: *Midtown is a median city.*
4. Assume the mean exists. *1.8 kids per family.*
5. Ambiguity in specifying the group

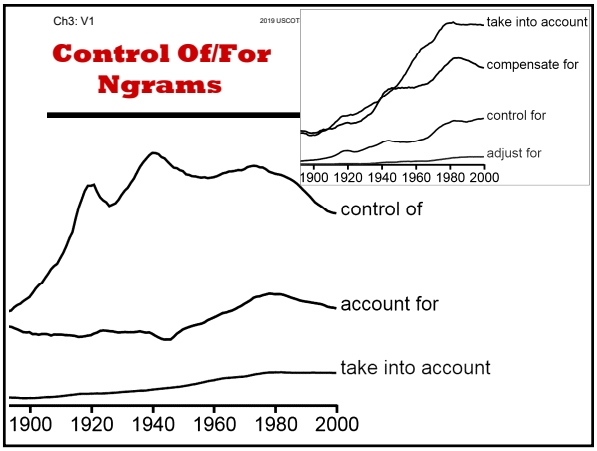
Controlling Confounding: Control Of

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

Controlling Confounding: Control For

CONTROLLING FOR CONFOUNDERS			
Take into account (mental)			
Can do by hand		*Calculator/Computer*	
1	Select/Stratify	4	Linear Regression
2	Form Ratios	5	Logistic Regression
3	Standardize	6	Multivariate Regress

Control Of/For Ngrams



Crude Associations

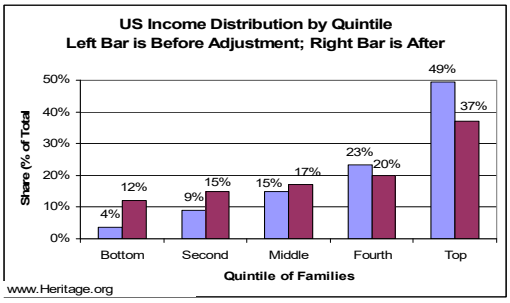
A **crude association** is an association in which nothing else has been taken into account.

Less likely to get pregnant:

- Short young adults than tall.
- Adults that shave daily than those that don't
- Adults with long hair than those with short.

What one takes into account is an assumption. Teachers should say, "Check your assumptions."

Crude Association versus an Adjusted Association



Prison Expense: Crude vs Adjusted Associations

State	Total	# Inmates	Per Inmate
CA	$2.9B	136K	$21,385
NY	$1.9B	69K	$28,426

Total ↑	Per Inmate ↓
50% more	25% less

State	Total	# Inmates	Per Inmate
MD	$481M	21,623	$22,245
KS	$159M	7,148	$22,245

Total ↑	Per Inmate ↔
3 times..	Same

State	Total	# Inmates	Per Inmate
MN	$184M	4,865	$37,825
ME	$48M	1,424	$33,711

Total ↑	Per Inmate ↑
260% more	12% more

Crude Ratio Associations
It's the Mix!!!

Ratio associations can be still be confounded. Averages are ratios.

NAEP Math 8	Internet Access at Home		
State	All	Yes	No
Virginia (VA)	▲ 275	282	258
Texas (TX)	273	▼ 285	▼ 260

NAEP Math 8	Internet Access at home		
State	All	Yes	No
Virginia (VA)	▲ 275 (100%)	282 (69%) ▲	258 (31%) ▼
Texas (TX)	273 (100%)	285 (53%)	260 (47%)

Simpson's Paradox: Time
It's the Mix!!

SAT Verbal flat, but every group improved.

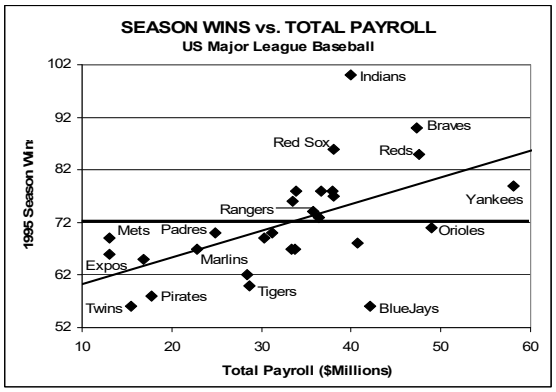
SAT-Verbal	---- Scores ----			---- Distribution ----		
Group	1981	2002*	Chg	1981	2002*	Points
White	519	527	+8	85%	65%	-20
Black	412	431	+19	9%	11%	+2
Asian	474	501	+27	3%	10%	+7
Mexican	438	446	+8	2%	4%	+2
Puerto Rican	437	455	+18	1%	3%	+2
Amer. Indian	471	479	+8	0%	1%	+1
ALL	504	504	0			

Will an Association Reverse?
The Cornfield Conditions

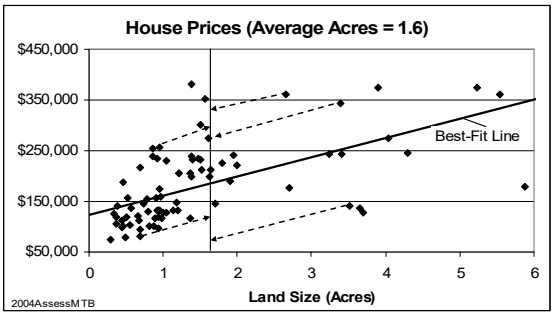
After learning about Simpson's Paradox, one student said, "I'll never trust another statistic." This is cynicism: not a good outcome.

Not all confounders can reverse an association. Jerome Cornfield proved that a confounder association must be "bigger" than the observed.

Cornfield's conditions are one of the three biggest contributions of statistics to human knowledge.



Regression Standardizes



Regression Standardizes
An Example:

The data shows that house prices increase by $39,000 per bedroom. This is a crude association.

$16,000 per bedroom if land is *controlled for*,

$9,000 per bedroom after *accounting for* land and house size,

$5,000 after *adjusting for* land, house size, and number of bathrooms.

TV for toddlers interferes with brain growth, says study:

Children under two should not be allowed to watch television because it increases their chances of suffering attention problems later in life, says an American study.

A study of 1,345 children found that each hour spent in front of the set every day increased the risks of attention deficit disorders by 10%.

U.S. journal, *Pediatrics*

Time to Double given Growth Rate

If a child's risk of Attention Deficit Disorder increases by 10% for every extra hour of watching TV, how many hours do they have to watch to double their risk?

Rule of 72*: Time to double = 72 / Rate

72 divided by 10% per hour = 7.2 hours

* Assuming compounding

How to Relate this to Math Colleagues

Don't talk about confounding or effect size.

Talk about assumptions.

- What one controls for is an assumption.
- What one fails to control for is an assumption.

AAU&C Quantitative Literacy VALUE rubric:

Assumptions: Ability to make and evaluate important assumptions in estimation, modeling, and data analysis.

AAC&U Quantitative Literacy VALUE Rubric

Interpretation, Representation, Calculation, Application, **Assumptions**, and Communication

Assumptions: Ability to **make and evaluate** important assumptions in estimation, modeling, and data analysis.

www.statlit.org/pdf/2009QuantitativeLiteracyRubricAACU.pdf
www.aacu.org/peerreivew/2014/summer/RealityCheck

Teaching Statistical Literacy

Chapter 4: Using and Describing Ratios

by
Milo Schield

Half-Day Workshop
USCOTS May 16, 2019
www.StatLit.org/pdf/2019-Schield-USCOTS-Slides4.pdf

Workshop Schedule

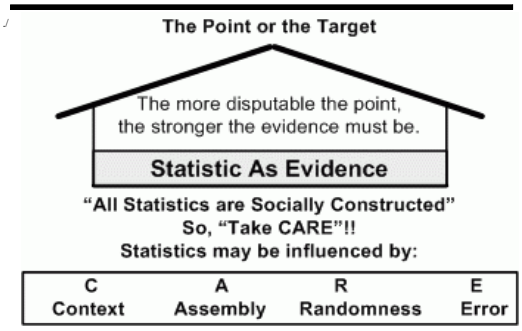
Start		Topic
1:00	1	Statistical Literacy Intro
1:30	2	StatLit Details
2:15	3	Measurements
2:45	4	Named Ratio Grammar
3:30	5	Comparing Count Ratios
4:00	6	Untangling Statistics

Ratios: Chapter 4 Outline

Per grammars:
- Percent grammar
- Percentage grammar
- Reading half tables and tables w/o margins
- Rate grammar

Ordinary Preposition grammars:
- Chance grammar
- Ratio grammar

Stat Literacy: Study Statistics as Evidence in Arguments



Evaluate these Using Just Assembly/Assumptions

1. One in five children face hunger [2019 billboard in St. Paul]
2. Two absences per month = Likely to fail a grade
3. Ninth-grade attendance better predicts graduation than 8th grade test score
4. Attendance alone explains 31% of the variance in performance
5. Budget cuts lead to deaths in Federal prisons
6. 22 million victims of human trafficking trapped worldwide.
7. The National Rifle Association is a terrorist organization.
8. Ban assault weapons
9. 2016 Memphis. 228 homicides. Down 500 police officers.

Forming Ratios

CONTROLLING FOR CONFOUNDERS			
Take into account (mental)			
	Can do by hand		*Calculator/Computer*
1	Select/Stratify	4	Linear Regression
2	Form Ratios	5	Logistic Regression
3	Standardize	6	Multivariate Regress

From Comparisons to Ratios: Using Prepositions

ARITHMETIC COMPARISONS
Using Conjunctions or 'Change -By'

Difference:	Ratio:	Relative Difference :
more (greater) than increase by #	*times [as much as] increase by a factor of*	*% (times) more than increase by X%*

RATIOS (Using Prepositions)

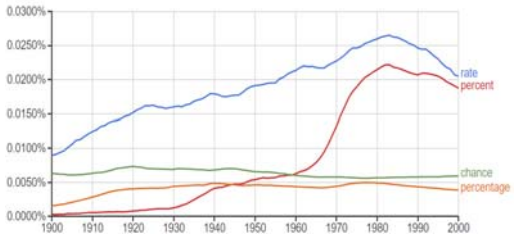
Common Prepositions :	**Per Grammar:**
Of, in, for. To [4 to 3; 4-3; 4:3] *4 out of [every] 5; cut in half*	*miles per gallon ; mph deaths per 1,000 men*

RATIOS (Using Prepositions)

Common Prepositions : *Of, in, for. To* [4 to 3; 4-3; 4:3] *4 out of [every] 5; cut in half*	**Per Grammar:** *miles per gallon ; mph deaths per 1,000 men*
Named-Ratios	Named-Ratios
Ratio Grammar: *ratio of women to men student-teacher ratio*	**Percent Grammar:** *85% of military personnel are men*
Chance Grammar: odds/risk/probability *chance of [our] winning; chance that we will win chance to win ; chance for a win*	**Percentage Grammar**: fraction/share *percentage of men who bet*
	Rate Grammar: prevalence, incidence *rate of n per d* *Men died at a rate of n per d*

Light-edge boxes need clause for part and whole (cannot compare ratios.
Dark-edge boxes have part and whole in phrases (can compare ratios)

Prevalence of Named Ratios



Two Kinds of Percents

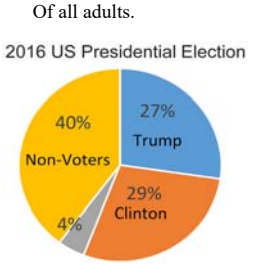
Which kind of percents are these: part-whole or percent compare?

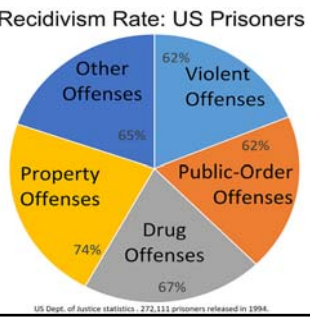
1. The youngest child's share of the candy.
2. Interest charged per year by the Mafia (criminals).
3. People live 100% longer on average in US than in Swaziland.
4. The advertisement said "40% off".

.

Part-Whole Using Pie Charts

Of all adults.





Four Different Grammars; Confusion of the Inverse

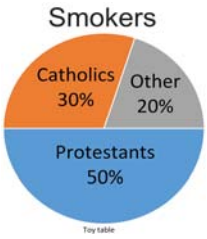
1. 40% of US adults did not vote for president in 2016.
2. The *percentage* of US adults who didn't vote was 40%
3. The non-voter *rate* among US adults in 2016 was 40%.
4. There was a 40% *chance* that an adult was a non-voter.

Confusion of the inverse exchanges part with whole.

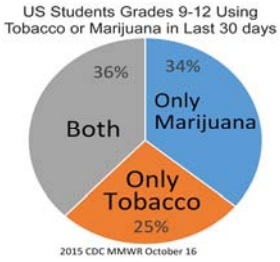
1. "The percentage of men who are in the military"
   .NE. "the percentage of the military who are men".
2. The percentage of smokers among women .NE.
   "the percentage of smokers who are women".

Use Percent Grammar <X% of Whole are Part>

Describe the 30%



Describe the 36%



Tables: Use Percent Grammar <X% of Whole are Part>

1. What percentage of men are art majors?
2. What percentage of art majors are men?
3. What percentage of students are male art majors?

Students	Men	Women	ALL
Humanities	28	72	100
Arts	4	36	40
Science	48	12	60
ALL	80	120	200

100% Tables: Percent Grammar <X% of Whole are Part>

Describe the 10%

Students	Men	Women	ALL
Humanities	28%	72%	100%
Arts	10%	90%	100%
Science	80%	20%	100%
ALL	40%	60%	100%

Describe the 5%

Students	Men	Women	All
Humanities	35%	60%	50%
Arts	5%	30%	20%
Science	60%	10%	30%
ALL	100%	100%	100%

Use Percent Grammar <X% of Whole are Part>

Table 33: World Population by Religion and Continent (1996)

(Millions)	Total	Asia	Europe	North Am	Other
Total	5,804	3,513	728	296	1,563
Christian	1,955	303	556	256	1,096
Muslim	1,126	778	32	5	316
Nonreligious	887	753	90	21	44
Hindus	793	787	2	1	4
Buddhists	325	322	2	1	1
Atheists	222	175	41	2	6
All Other	496	395	5	10	96

Table 1333. 1997 U.S. Statistical Abstract.

Percentage Grammar Four form

1. The percentage of seniors who smoke is 15%.
2. Among seniors, the percentage who smoke is 15%.
3. Among Seniors, the percentage of smokers is 20%.
4. Among men, the percentage of seniors who smoke is 20%

Numbers 3 and 4 are problems.
"Of" introduces whole in percent grammar.

Percentage Grammar Sports Grammar

Sports grammar is readily understood with a natural whole:

- *percentage of defective cans; percentage of tire failures*

Without a natural whole, sports grammar is ambiguous.

- *percentage of female smokers;*
- *percentage of working males*
- *percentage of infant deaths;*
- *percentage of single mothers*

Half Tables when Parts of 100% Table are Binary

Describe the circled 60%. Use percent grammar.

Class Last Year	Percentage who are Retained	Percentage who are Not Retained	All
Freshman	60%	40%	100%
Sophomore	75%	25%	100%
Junior	90%	10%	100%
Senior	10%	90%	100%
ALL	70%	30%	100%

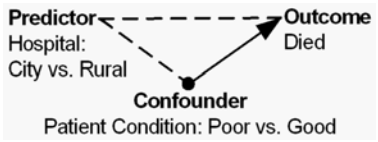
If 60% returned, what percentage did not return?

So, the right two columns are redundant.
Eliminating them will save space!

Confounding

Mortality by Hospital

Hospital	Total	Died	Death Rate
City	1,000	55	5.5%
Rural	1,000	35	3.5%
Both	2,000	90	4.5%



Statistics Literacy For Decision Makers

13: Confounding & Cornfield

by
Milo Schield

Half-Day Workshop
USCOTS May 16, 2019
www.StatLit.org/pdf/2019-Schield-USCOTS-Slides13.pdf

Workshop Schedule

1:00 Ch 1 Statistical Literacy – Introduction
1:30 Ch 2 Statistical Literacy – Details

2:15 Ch 3 Measurements
2:45 Ch 4 Ratios

3:30 Ch 13 Standardizing
4:00 Feedback

.

Confounding: Chapter 13 Outline

Cornfield-Fisher debate

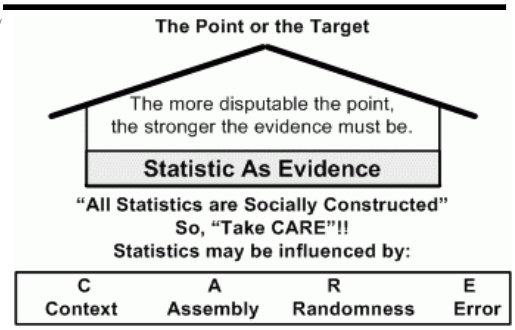
Cornfield conditions

Standardizing percentages, rates and averages

Standardizing percentage & number attributable

Statistical significance and confounding

Stat Literacy: Study Statistics as Evidence in Arguments



Cornfield-Fisher Debate

Doctors had noticed the strong association between smoking and lung cancer. Statisticians argued that this evidence strongly supported the claim that smoking was a cause of lung cancer.

Fisher, a smoker, noted that *association is not causation in observational studies*.

Fisher produced data. Identical twins were more likely to share a smoking preference than were fraternal twins. This statistic supported genetics as an alternate explanation for the association.

Cornfield-Fisher Debate

Now when the world's leading statistician says something that every statistician agrees is true, most reasonably-minded statisticians would back off.

And when the world's leading statistician produces data indicating a plausible confounder, it seems incredible that anyone would reply.

Jerome Cornfield did!

Cornfield Conditions

Cornfield **proved** that the relative risk of lung cancer had to be greater for a confounder (e.g., genetics) than for the predictor (e.g., smoking) in order to nullify or reverse the observed association.

Cornfield pointed out that smokers were about 10 times as likely to get lung cancer as non-smokers.

Fisher's data involved a factor of two.

Fisher never replied.

Contributions to Human Knowledge

"Cornfield's minimum effect size is as important to observational studies as is the use of randomized assignment to experimental studies.

No longer could one refute an ostensive causal association by simply asserting that some new factor (such as a genetic factor) might be the true cause.

Now one had to argue that the relative prevalence of this potentially confounding factor was greater than the relative risk for the ostensive cause."

Schield (1999). [This was written 20 years ago!]

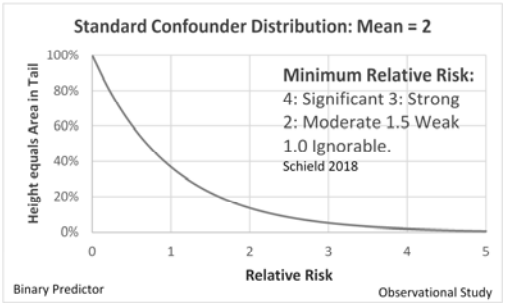
Confounder Distribution

Since confounders may be unknown, there is no way to derive or infer their distribution.

Schield (2018) argued that we needed a standard for confounder: a standard confounder distribution.

He proposed an exponential (one factor determined) with a mean relative risk of 2.
This applied if predictor and confounder are binary.

Confounder Distribution Unknown & Unknowable

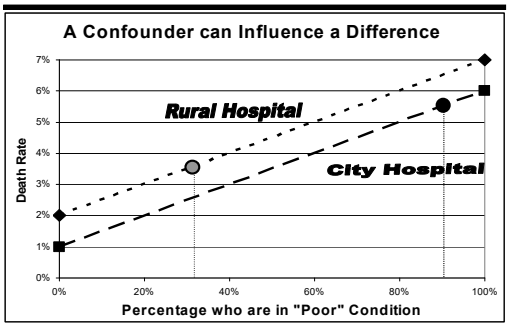


Controlling for a Confounder: Graphical Technique

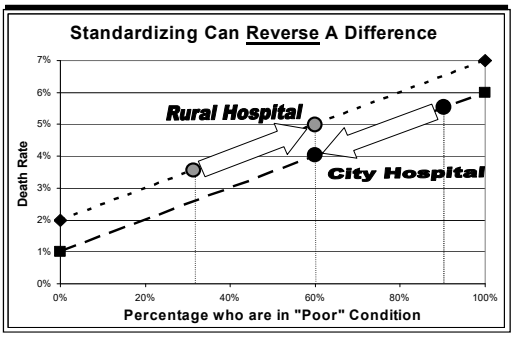
Wainer introduced a simple graphical technique that made the control of a binary confounder a relatively simple matter.

Schield (2006). Presenting Confounding Graphically Using Standardization, *STATS* magazine. www.statlit.org/pdf/2006SchieldSTATS.pdf

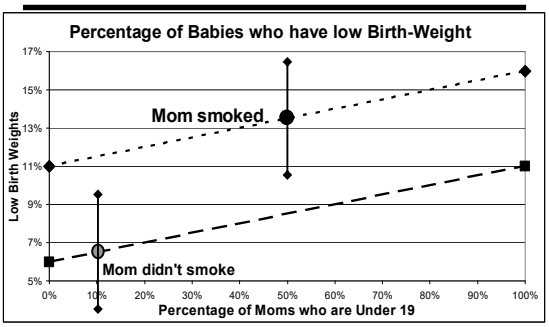
Crude Association: Death Rate: City > Rural



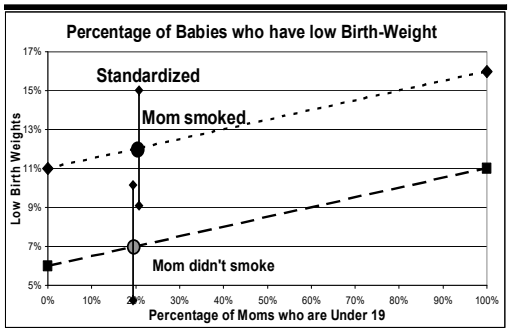
Controlling for a Confounder: Death Rate: City < Rural



Crude Association: Statistically Significant



Standardized Association: Statistically Insignificant



Confounder Effect on Statistical Significance

Controlling for a confounder can transform a statistically-significant association into an association that is statistically insignificant.

Although statistical educators are clearly aware of this, there is nothing in any introductory textbook that alerts students to this possibility.

The failure to show a significance reversal is *statistical negligence.*

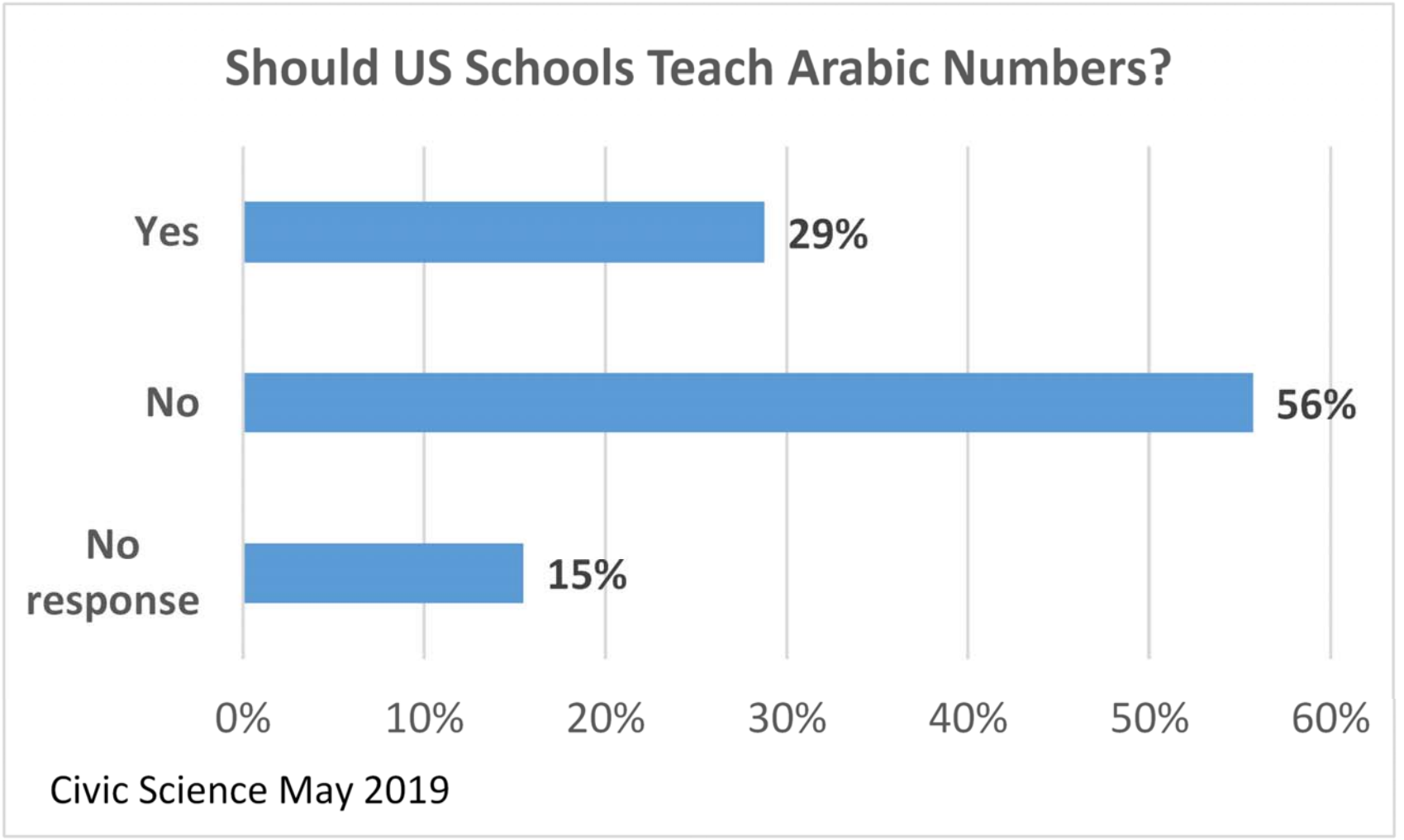
Teaching Statistical Literacy

Chapter 1
by
Milo Schield

Half-Day Workshop
USCOTS May 16, 2019

www.StatLit.org/pdf/2019-Schield-USCOTS-slides1.pdf

First Sharia math, then Sharia law!!!



Working Moms; Better Kids

http://money.com/money/5272659/working-moms-better-kids/

Outline

Introduction:

A1. Who takes intro statistics

A2. SAT level of our students by college

A3. Math level of our students by major

Exp vs. Obs: What kinds are relevant?

A3. Kinds of influence on statistics
How common are these influences?

A4. Grammar: Association vs. causation

Goals of this Workshop

1. Present my view of statistical literacy
2. Expose you to lots of new ideas
3. Present a coherent structure for teaching
4. Show the importance of English grammar
5. Show simple ways of handling significance
6. Show simple ways of handling confounding
7. Show how confounding changes significance
8. Role-model analyzing studies

Fraction of 4-year Undergrads that take Intro Stats?

57%

Schield (2016, IASE)

Fraction of Course Gain that Stat Students Loose in 4 Months

50%

Tintle et al, 2013

Student Attitudes Toward Stats

Of those taking Stat I:

- less than 1% take *Stat II* (10-yrs @ U. St. Thomas)
- less than 0.2% major in statistics (nationwide).
- most see less value in statistics after the course than they did before. Schield and Schield (2008).
- too many say "Worst course I ever took" [anecdotal]

www.amstat.org/misc/StatsBachelors2003-2013.pdf 1,135 stat majors in 2013 at 32 colleges www.StatLit.org/pdf/2015-Schield-UST-Enroll-in-Statistics.pdf

What fraction of 4-Yr Intro Stat students are taught outside Math?

50%

Estimates by Schield (2015, Statchat)

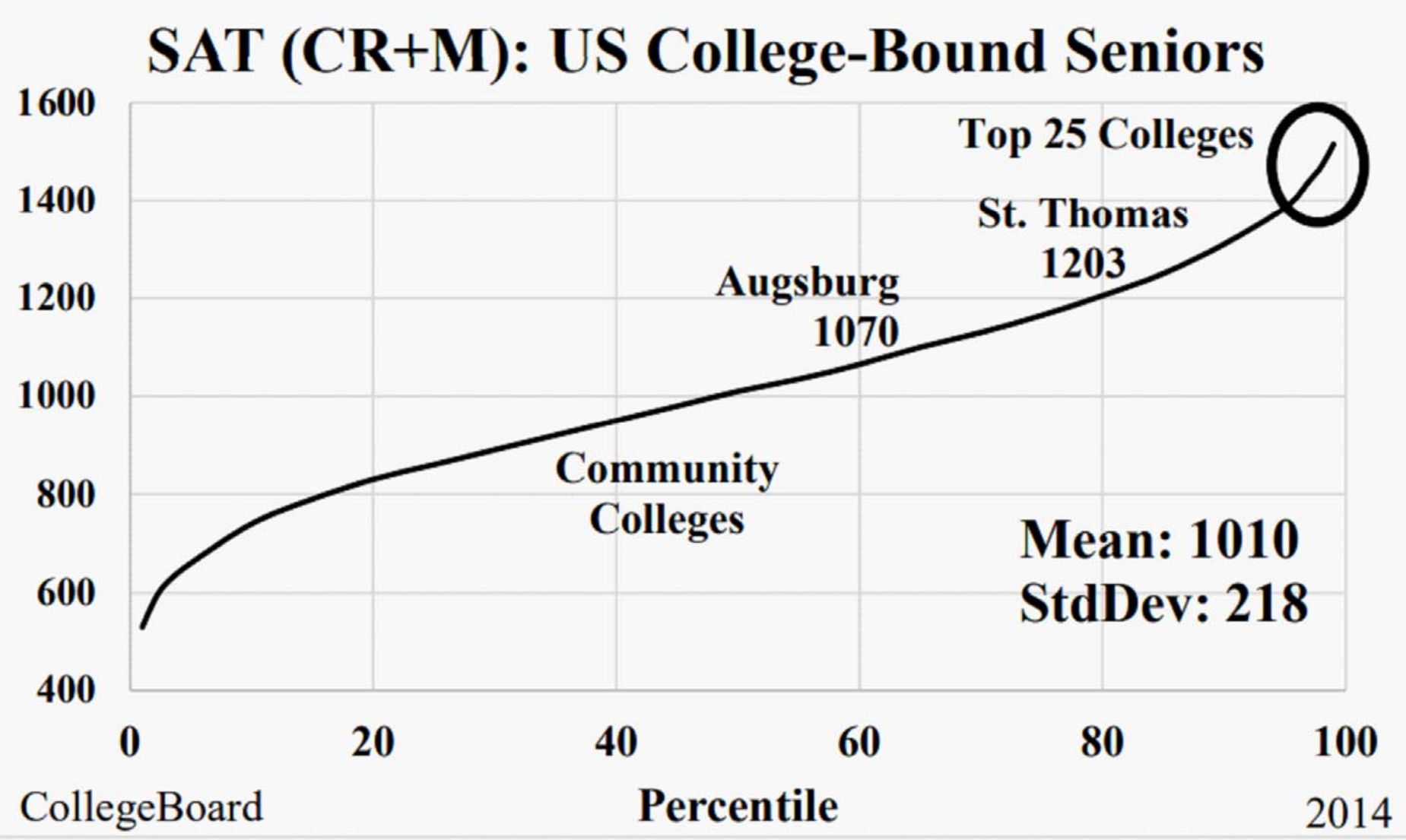
Who takes Intro Statistics at Four-Year Colleges?

Table 1: Distribution of Majors in Stat 101

%	Major
38%	Business or Economics
19%	Social Science or History
13%	Health
10%	Psychology
9%	Engineering
9%	Biological Science
2%	Math or Statistics
100%	All students in these majors

Schield (2016, IASE). Inferred from data in 2012 US Statistical Abstract.

Where are your students?



Schield (2016, IASE)

SAT Math Percentile by Major

SAT MATH	PERCENTILE	MAJOR
613	80%	Math/Stats
585	72%	Physical Sciences
579	70%	Engineering
554	62%	Comp. Science
551	61%	Biological
550	61%	Social Sciences
522	51%	Business
522	51%	English Lang/Lit
506	46%	History
498	43%	Communication
489	40%	Psychology
482	38%	Education

Business Insider (2014), College Board (2015)

SAT Math Scores: Average by Student Major

Percentiles of all those taking the Math SAT

Schield (2016, IASE)

GAISE 2016 Update

The real world is complex and can't be described well by one or two variables.

If students do not have exposure to simple tools for disentangling complex relationships,

they may dismiss statistics as an old-school discipline only suitable for small sample inference of randomized studies.

GAISE 2016 Update

Multivariable thinking is critical to make sense of the observational data around us

- *learn to identify observational studies*
- *learn to consider potential confounding factors*
- *use … stratification … to show confounding*

This report recommends that students be introduced to multivariable thinking, preferably early in the introductory course and not as an afterthought at the end of the course.

Most Important Topics: Student Choices

	The most important topics in Statistical Literacy for Managers
Rank	
1	Take CARE: Confounding, Assembly, Randomness and Error/bias
2	Confounding
2	Hypothetical thinking: plausible confounders, plausible definitions
4	Statistics are more than numbers. They include the context
5	Association-causation (Luck-skill) including the grammar
5	Bias: Placebo, Single blind; double blind
5	Named Ratios and Ratio grammar; Percent, Percentages, Rates
5	Read tables and graphs

Schield (2016, ASA)

A-B-C Words: A = Association

Statistical association is not the same as Basketball Assoc.

Association words assert association explicitly or describe associations involving fixed conditions or unrepeatable events.

Association: Height is *associated* with age in children

Obesity is *correlated* with (related to) diabetes.

Prediction: Graduating from high school *predicts* success in life.

---

*Comparisons: People with degrees earn *more than* those without

Whites have a *higher* risk of suicide *than* blacks.

*Co-variation: *As children get older*, their weight *increases*.

* Manipulation is impossible, or treatment or outcome cannot be repeated.

Schield (2018, SL4DM)

A-B-C Words: C = Causation

Causation words assert causation, sufficiency or contra-factual

Causation: A bomb *caused* the fire. Insomnia is a side *effect*. Lightning *resulted in* a fire. Spark results in a fire.

Sufficient: The more X you do, the more Y *you will get*.

Prevent, *stop*, *end*, *start*, *kill*, *produce*, *cure*, *avoid*, *ban*, *quit*, *block*, *ward off*, *stave off*, *cancel*, *hinder*, or *eliminate*.[6]

Contra-factual: Those who do X *will* get more Y *than if they had not done X*.

A-B-C Words:
B = Between

Between words describe association but imply causation

Verbs: Red wine *cuts* cancer risk. TV *ups* kids' risk of flunking.
Gene X *increases* health risk. Smoking *raises* asthma risk.

Connectors: Nuts *linked to* cancer. Trauma *tied to* heart disease.

Contributor Diet *contributes* to diabetes. Age is *factor* in infertility

Nouns: Spinach is *asthma protector*. Bad water is a *killer*.

Logicals: Anxiety increases *due to* (*because of*) high stake testing

---

*Compare: People who take antidepressants have fewer migraines
Asthma attacks more likely for smokers *than* non-smokers.

*Covariation: *As* teacher pay *increases*, student scores increase.
The more hours worked, the *more likely* a promotion

*Manipulation is possible, and treatment and outcome are repeatable.

A-B-C Words: Distribution in Headlines

Of the 2,000 news headlines analyzed[6],
71% involved A, B or C.

Of those headlines involving A, B or C,

- **86% were "between" claims**,
- 11% sufficiency, 3% causation, 3% association.

6. Schield and Raymond (2009).

Association is not causation

This statement is ambiguous. It can mean:

1 Association is not sufficient to prove causation

2 Association provides no evidence for causation.

Teachers may intend #1; students often hear #2.

A better statement would be:

Association is evidence of causation somewhere.

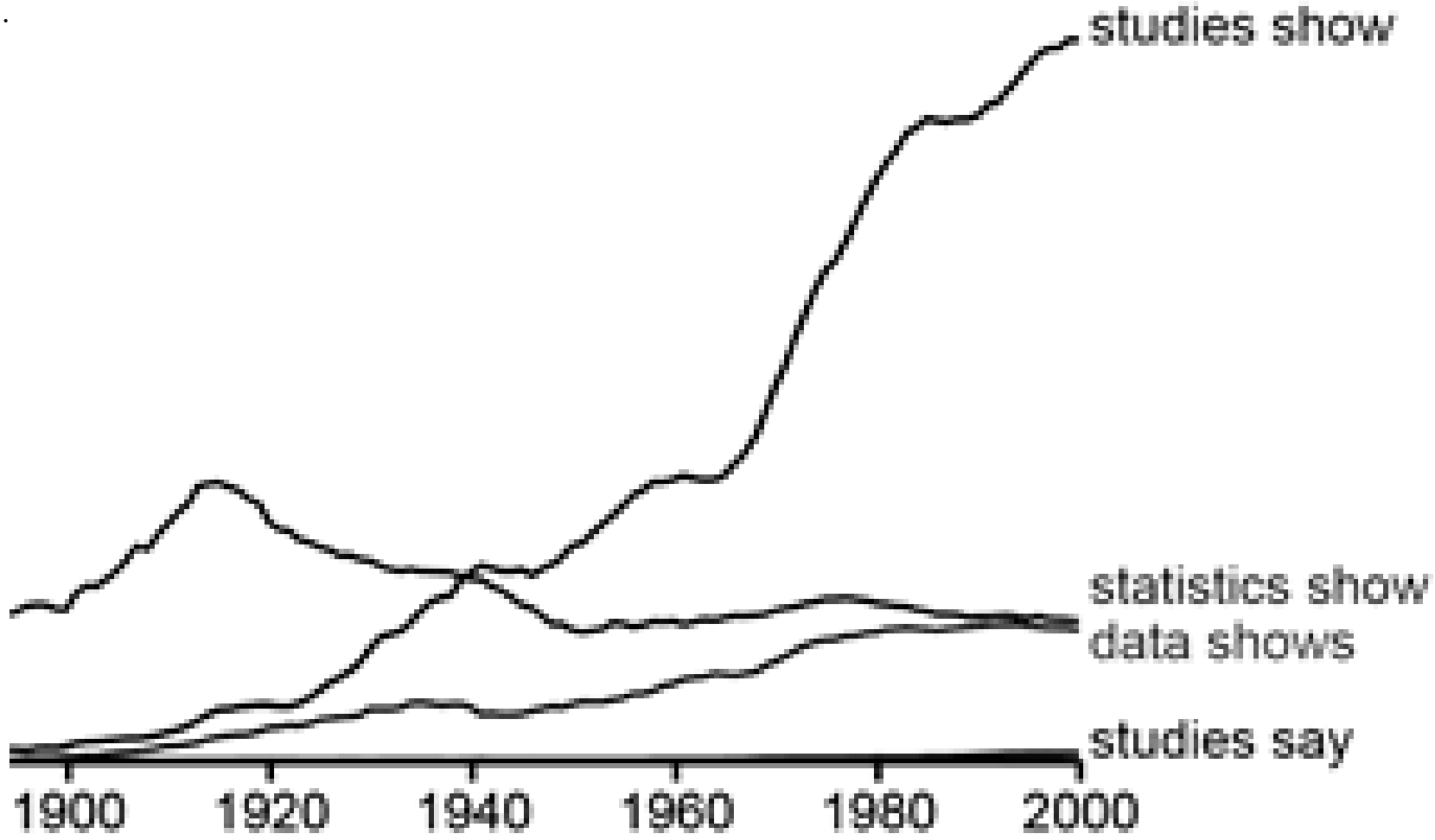
Association is not causation

No idea has stifled the growth of statistical literacy as much as the endless repetition of the words "correlation is not causation".

This phrase seems to be primarily used to suppress intellectual inquiry --
by encouraging the unspoken assumption that correlational knowledge is somehow an inferior form of knowledge.

John Myles White (2010):
www.johnmyleswhite.com/notebook/2010/10/01/three-quarter-truths-correlation-is-not-causation/

Studies are the Primary Unit of Analysis

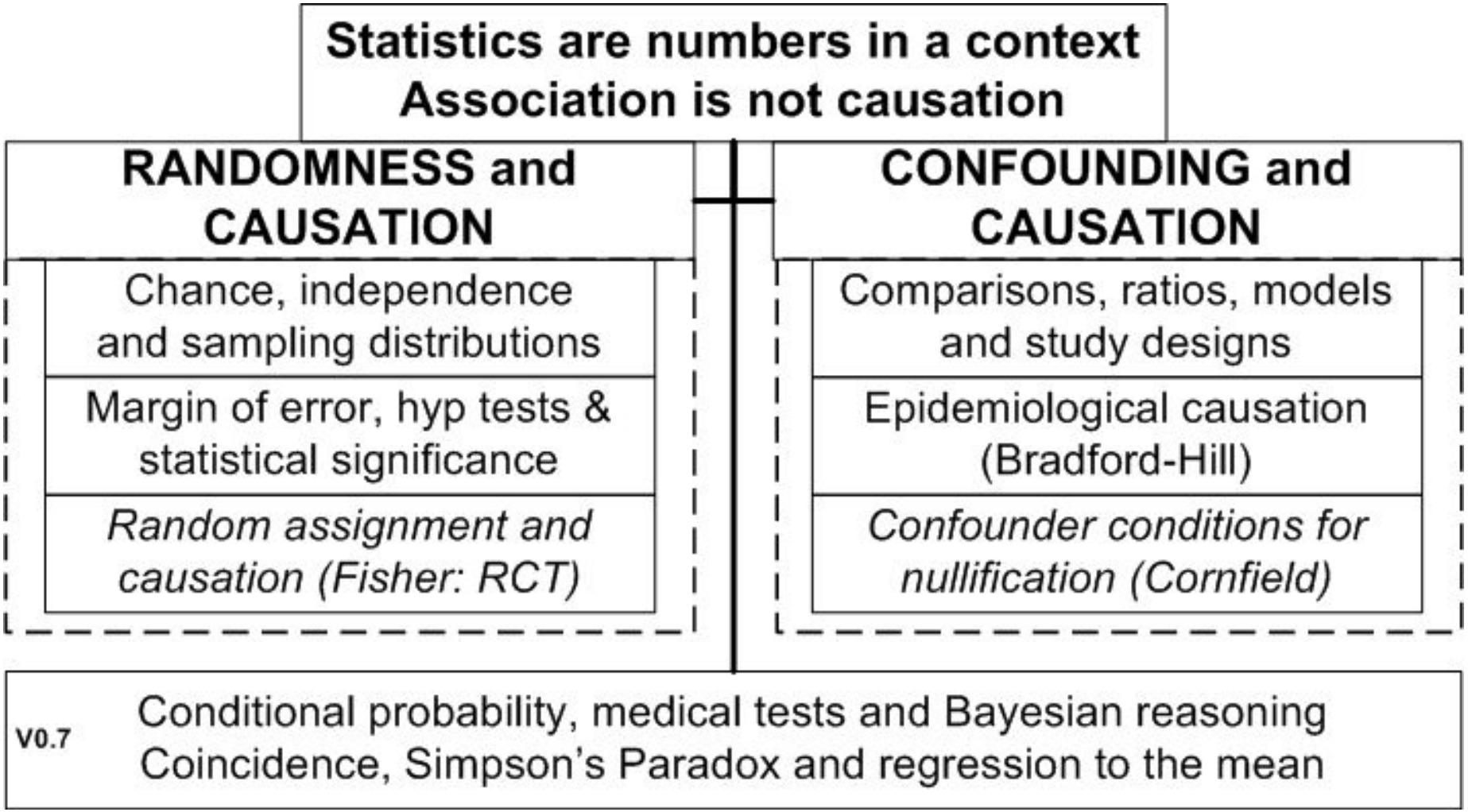


Harvard Case Studies: Title or Abstract

#	INFERENTIAL	
22	"clinical trial"	18
7	"statistical significance"	
4	"statistically significant"	
3	"standard error"	
1	"sampling error"	
1	"margin of error"	
1	"prediction interval"	
1	p-value	
0	"sampling distribution"	
0	"confidence interval"	
0	"null hypothesis"	
0	"reject the null"	
0	"random assignment"	

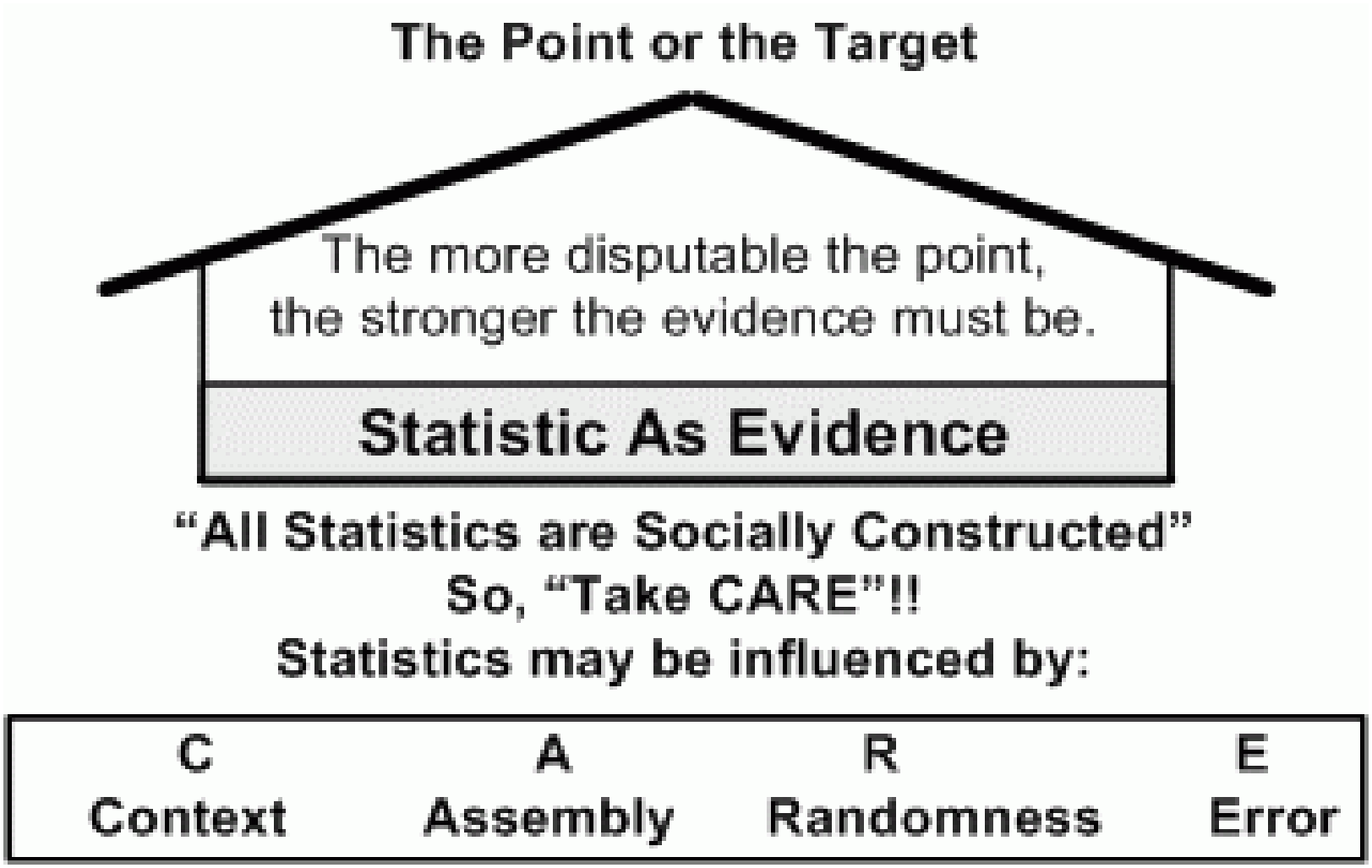
CONTROL/CONFOUND		
2,263	control	
234	"control of"	200
113	"take (ing) into account"	
30	"compensate (ing) for"	
19	"control (ed, ing) for"	
18	confound (er, ing)	
17	"adjust(ed, ing) for"	
3	"sampling bias"	
0	"alternate explanation"	
0	"common cause"	
0	"effect modifier"	
0	"Simpson's paradox"	
0	"lurking variable"	

Statistical Literacy : An Overview



Stat Literacy studies Stats as Evidence in Arguments

./



Statistical Literacy : Assembly

Living with AIDs			
All (1,000)	White (non-Hispanic)	Black (non-Hispanic)	Hispanic
434	150	186	78

Q1. Which group is largest?

Consolidate White (Non-Hispanic) with Hispanic.

Q2. Which group is largest?

Statistical Literacy : Randomness

Five non-quantitative Topics:

1. **Regression to the Mean**
   Sport Illustrated Cover
2. **Statistically significant**
3. **Chance-Related Mistakes:**
   Three Door problem; Birthday problem

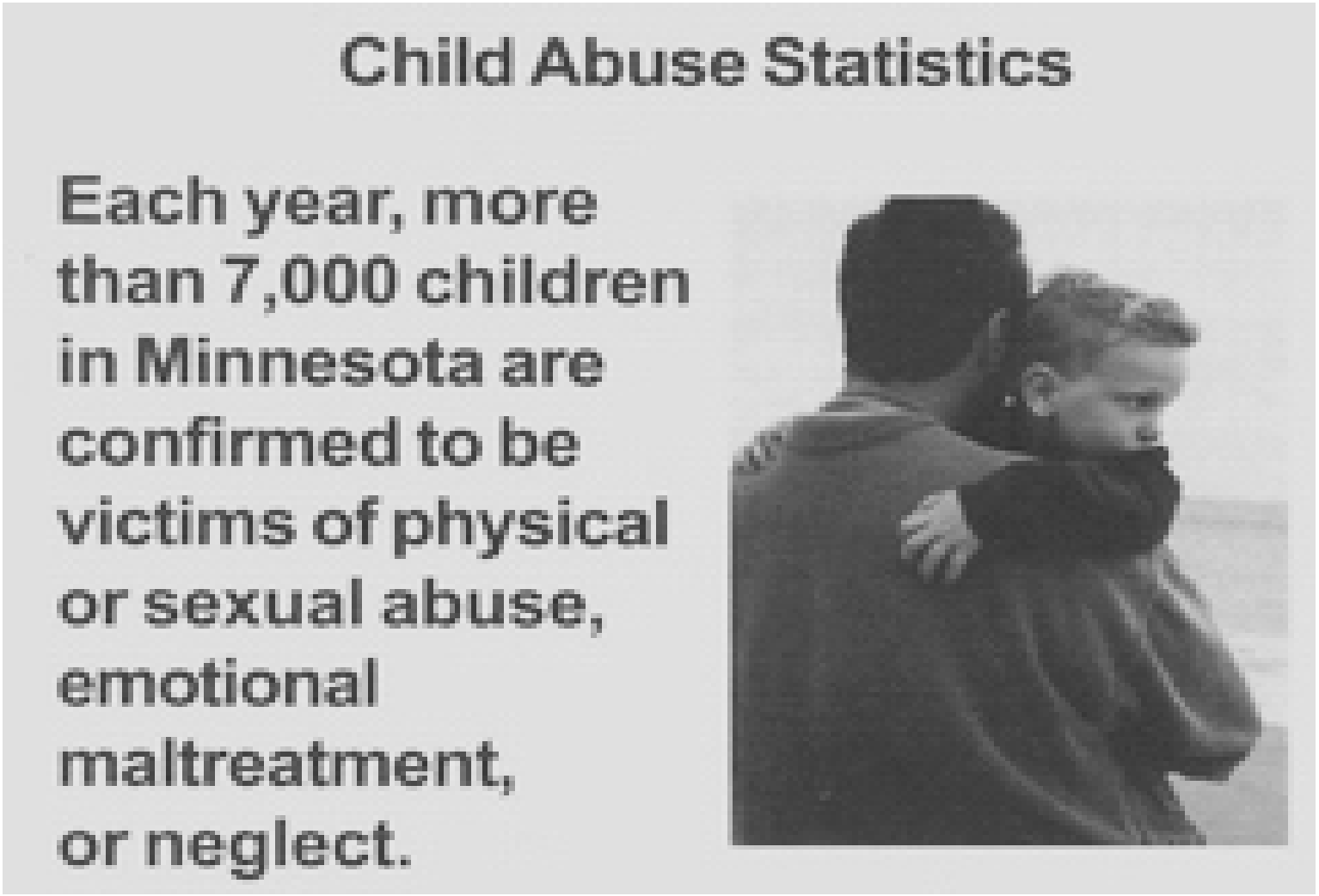
- **Better than chance**
- **Unlikely to be chance**

Statistical Literacy : Error/Bias

Three kinds of error

1. **Subject/respondent error:**
2. **Researcher/measurement error:**
3. **Sampling error:**

Statistical Literacy : Assembly



Statistical Literacy : Recommendation

More college students (over half) take intro statistics than any other course (except English).

One-size fits all is no longer viable. Statistics education must support Stat 101 and 100/102.

Statistics education should (1) support different flavors for different majors, and (2) agree on the contributions of statistics to human knowledge.

/

Willful Ignorance

The past success of statistics has depended on vast, deliberate simplifications amounting to willful ignorance.

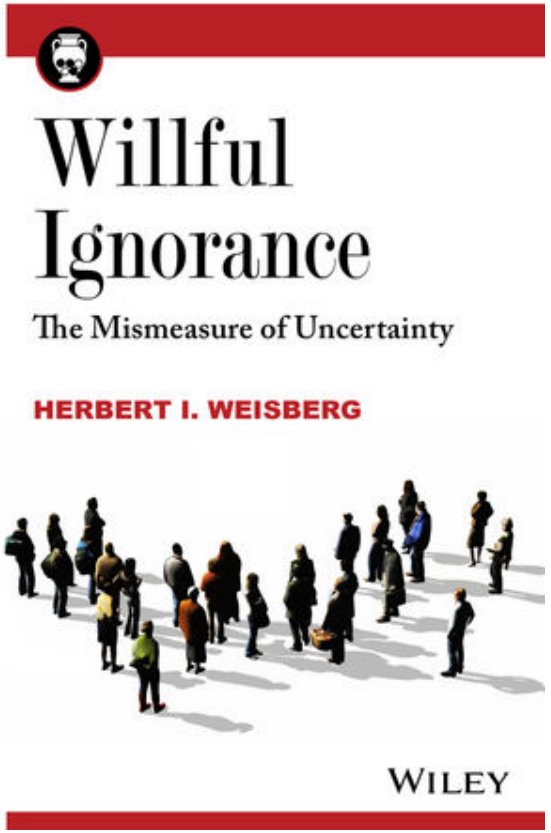
This very success now threatens future advances in medicine, the social sciences, and other fields.

Limitations of existing methods result in frequent reversals of scientific findings/recommendations, to the consternation of scientists and the public.

Herbert I. Weisberg

Willful Ignorance
Herbert Weisberg

The past success of statistics has depended on vast, deliberate simplifications amounting to willful ignorance.



Limitations of existing methods result in frequent reversals of scientific findings and recommendations, to the consternation of scientists and the lay public.

Statistics Literacy For Decision Makers

Statistical Literacy Details Chapter 2

by
Milo Schield

USCOTS Workshop May 16, 2019

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

Take CARE: Details
Chapter 2 Outline

Associations: Comparison and Co-Variation

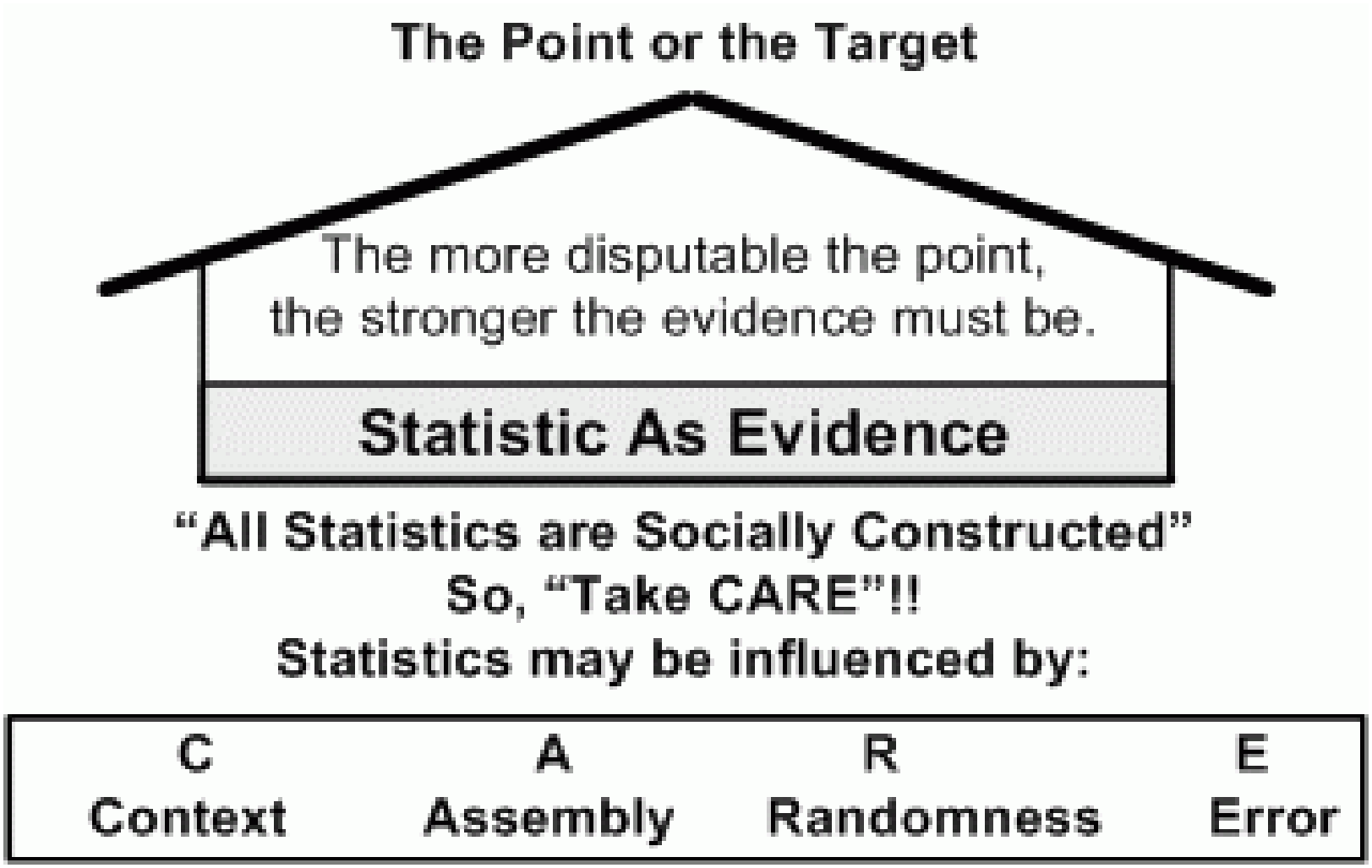
- Comparisons: Ordered and Arithmetic
- Comparisons: Kinds of Arithmetic

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

./



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

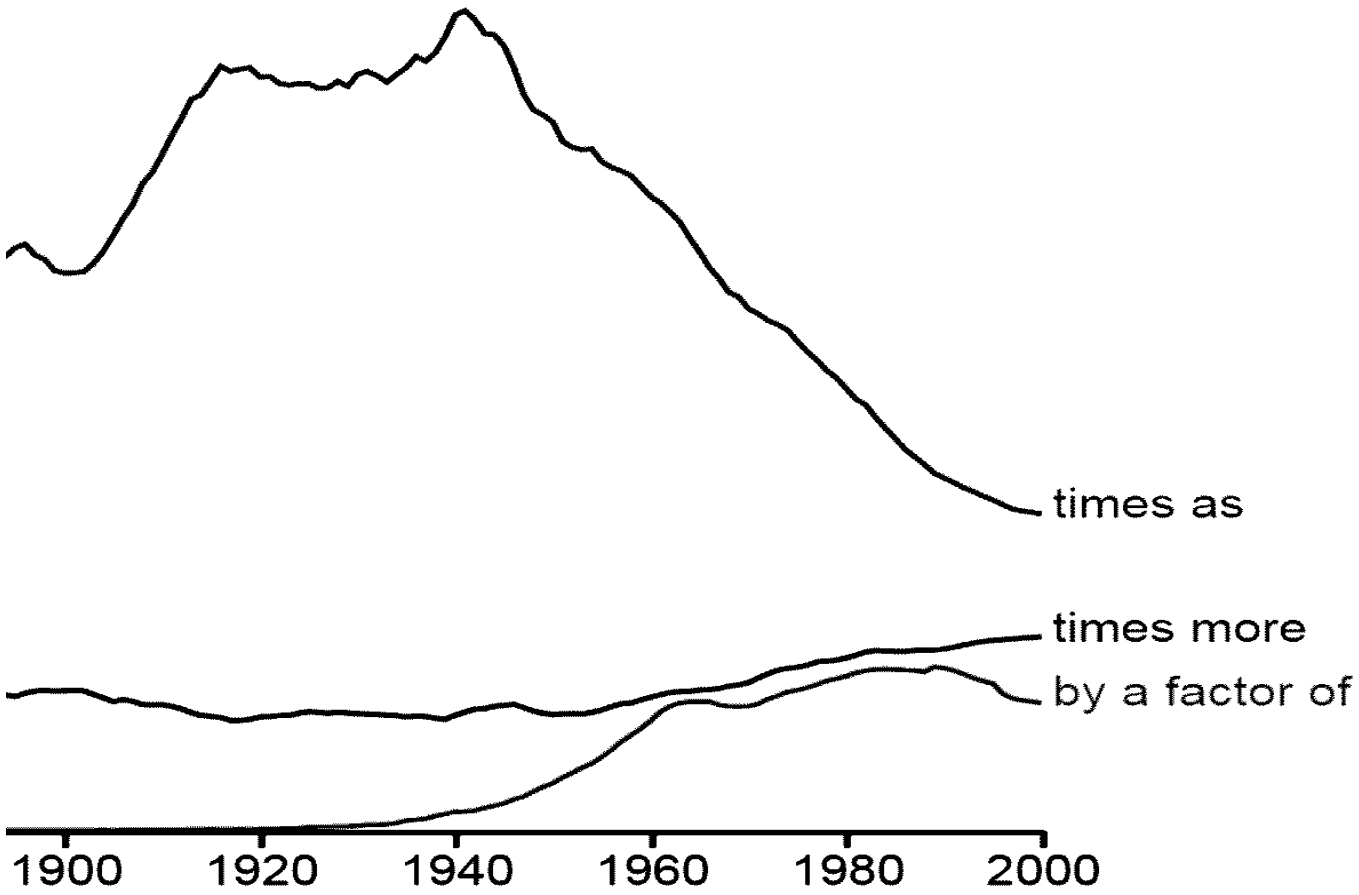
Ordinal (Order): Women live longer than men

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 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

Prevalence of Comparisons Google Ngrams

./



Confounding

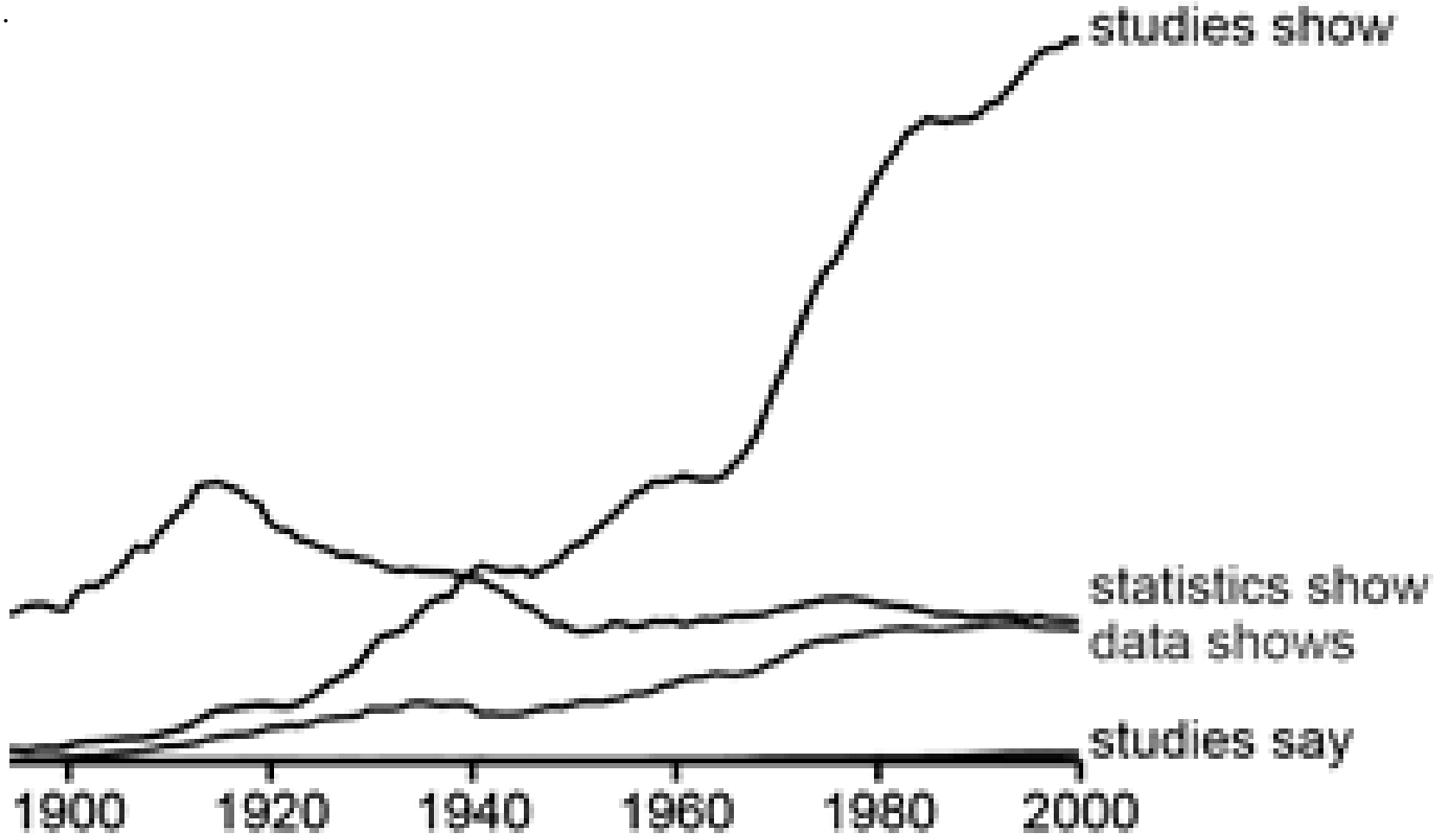
What things block or negate confounders?

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

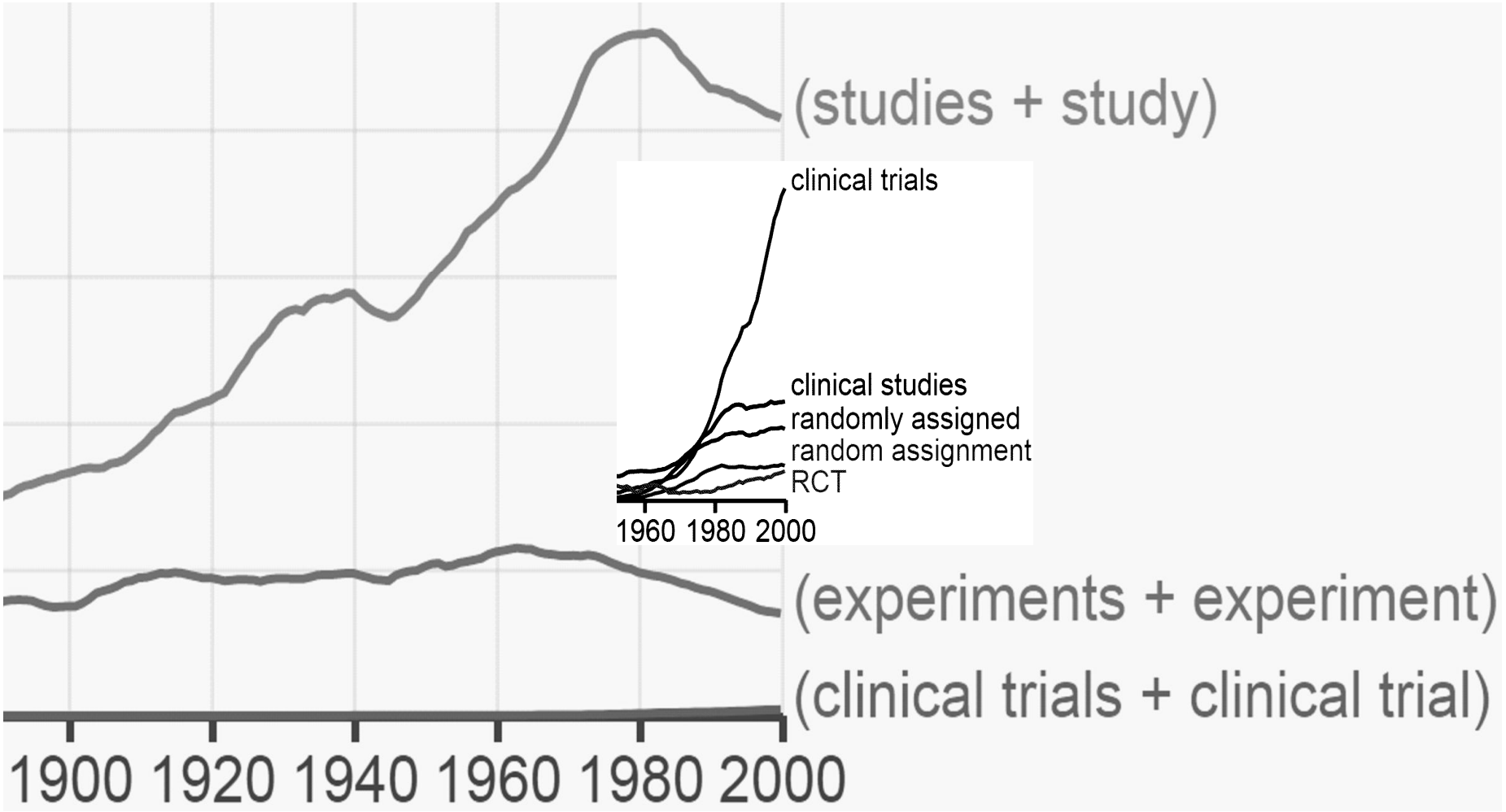


Six Basic Study Designs

Experiment	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.

Study Design Prevalences: Google Ngrams

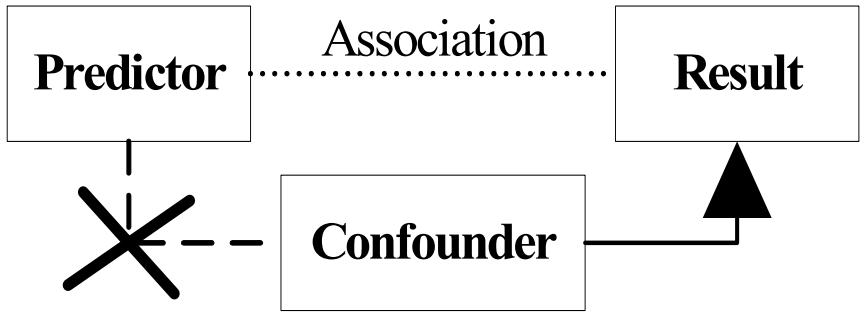


Random Assignment Nullifies Prior Confounding

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."

Experiments Gold std.



RCT Silver std.

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

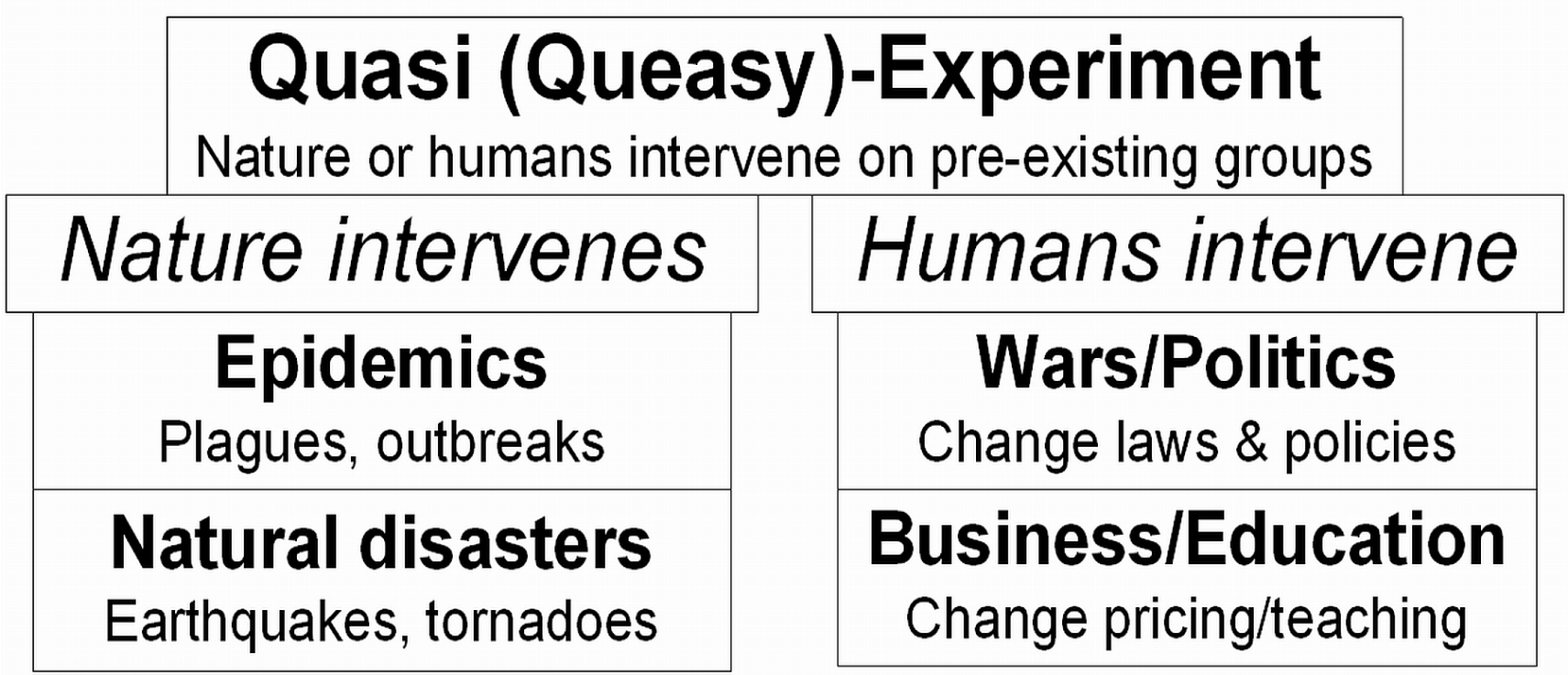
Placebo Effect

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



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 Brouchard 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 Brouchard 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.

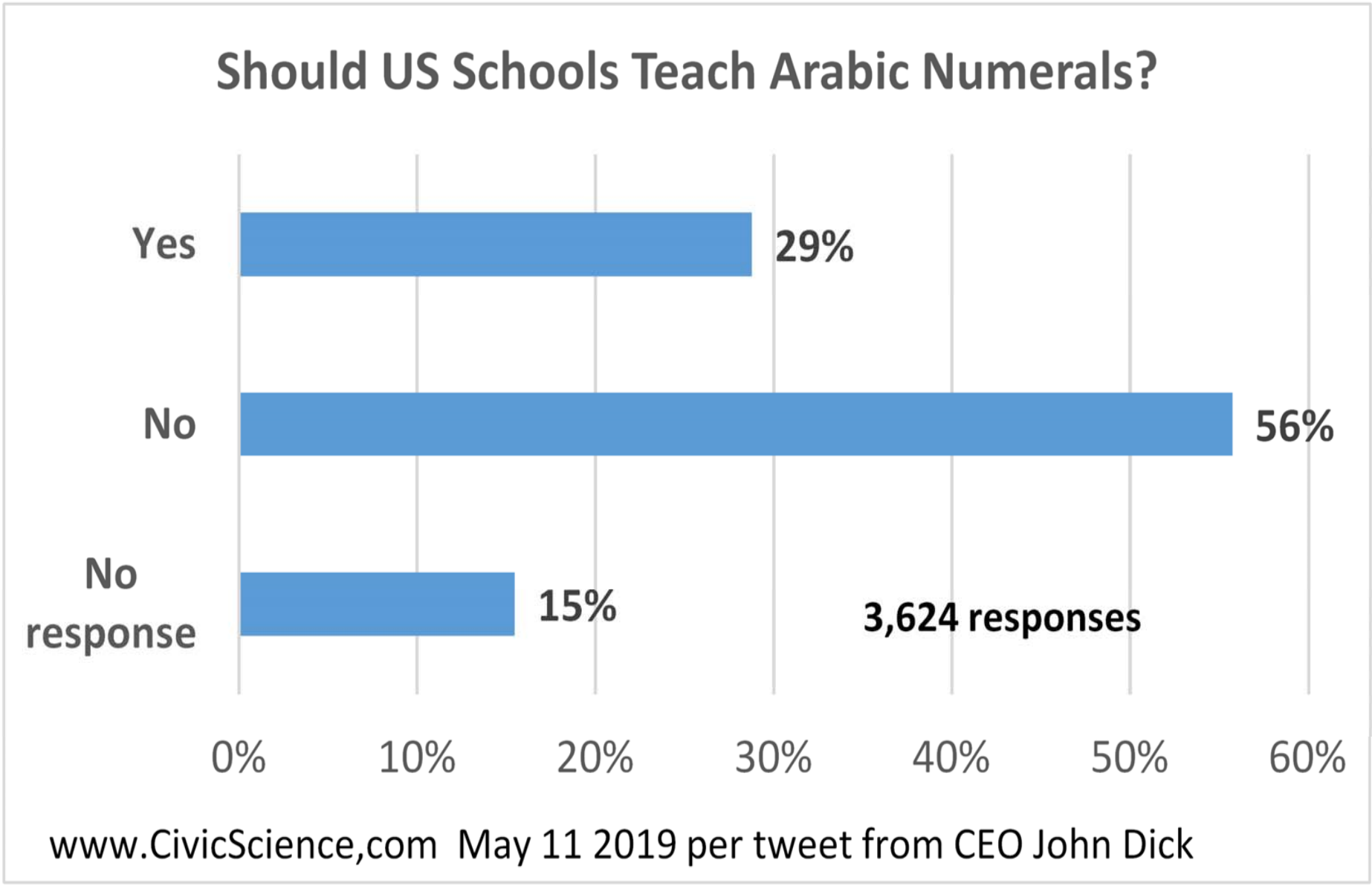
A major decrease in pain was reported by 75% in the treatment group 19% in the control group.

- Natural Health, August, 1998. Page 52.

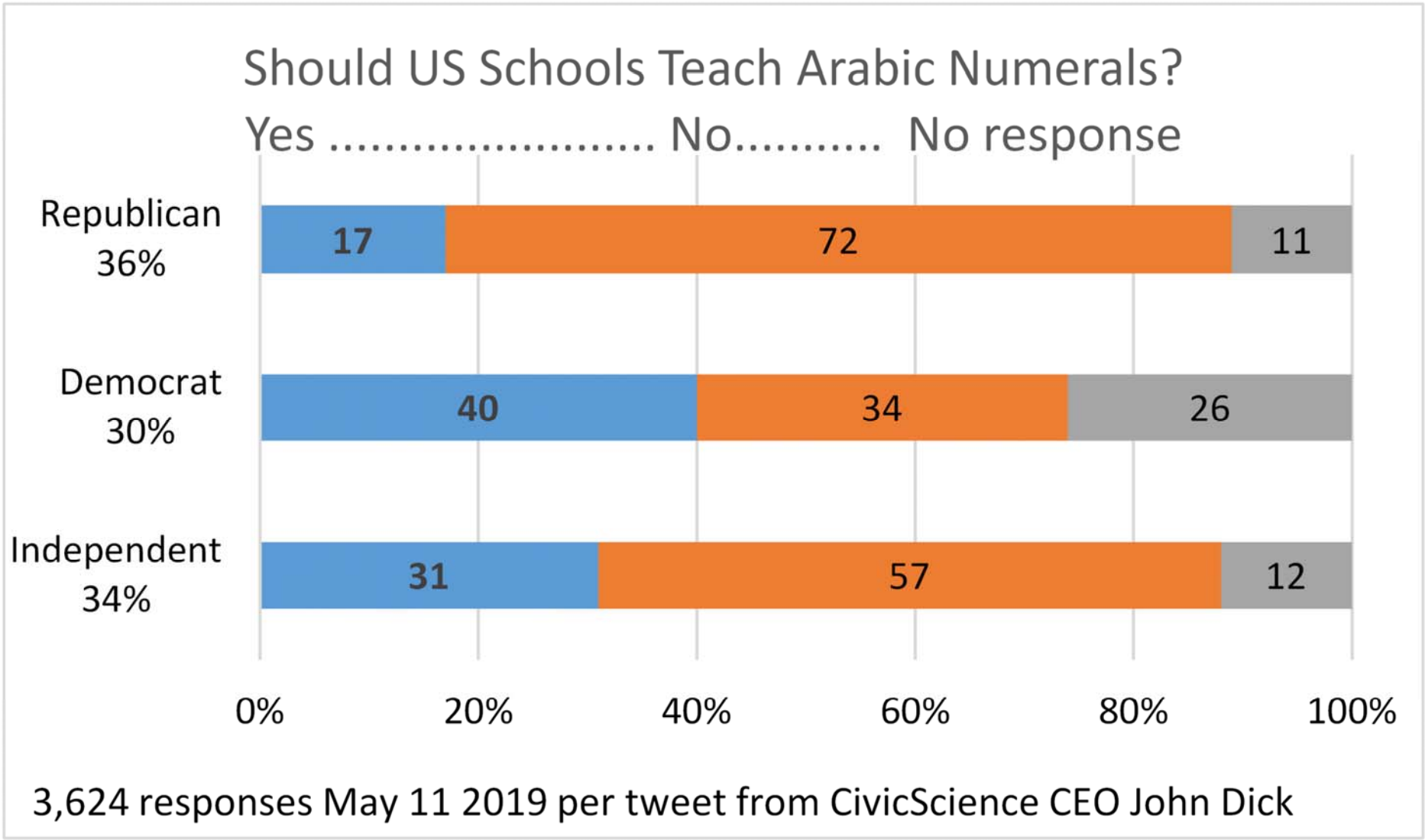
Effect size. Study design.
Hypothetical thinking using Take CARE.

/

Bias or Ignorance?



Bias or Ignorance?



3,624 responses May 11 2019 per tweet from CivicScience CEO John Dick

Statistics Literacy For Decision Makers

Chapter 3: Measurements

by
Milo Schield

Half-Day Workshop
USCOTS May 16, 2019
www.StatLit.org/pdf/2019-Schield-USCOTS-Slides3.pdf

Measurements: Chapter 3 Outline

Distributions

Measures of center

Two-group comparisons of Means & Medians

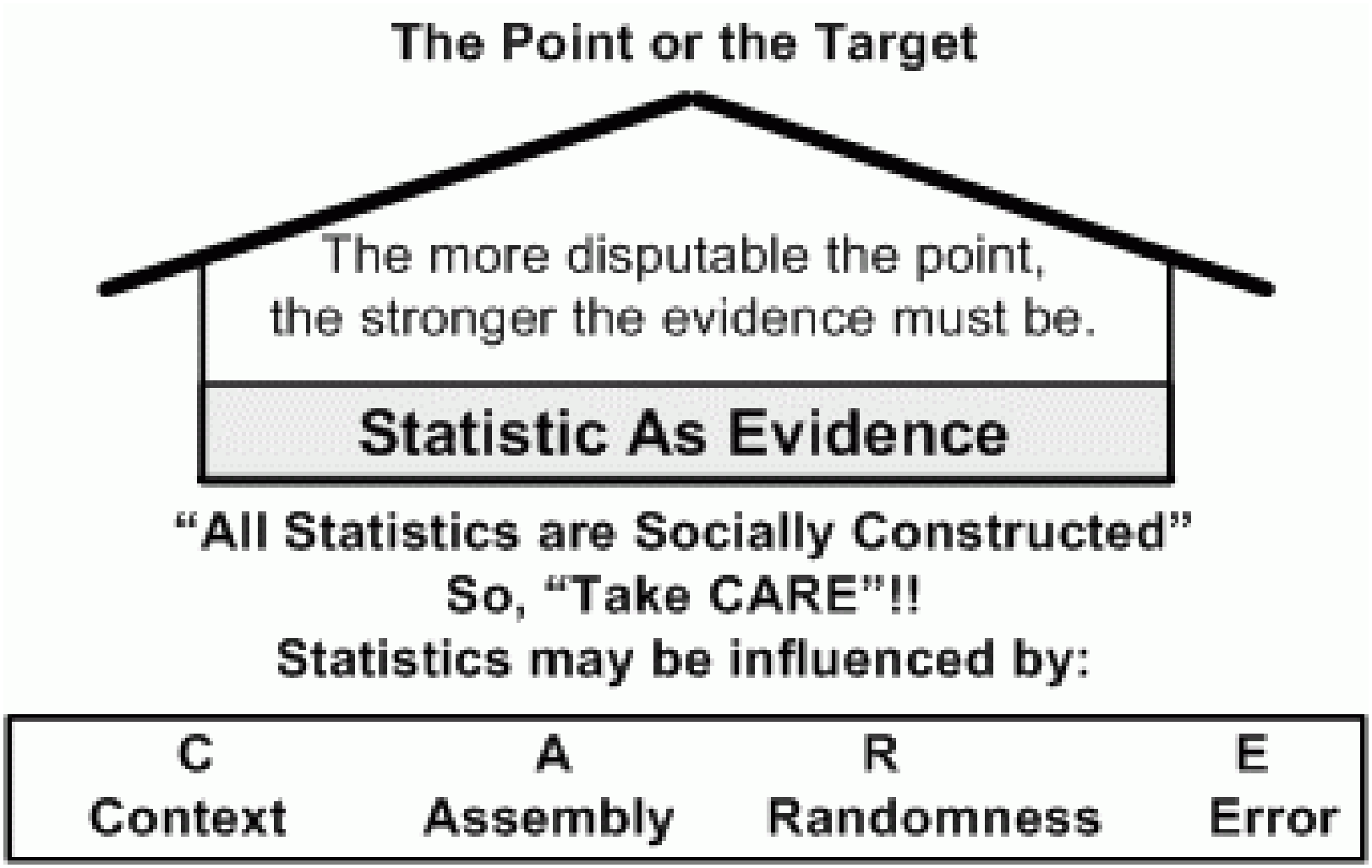
Two-variable co-variation

Spread

Slope and simple regression

Stat Literacy: Study Statistics as Evidence in Arguments

./



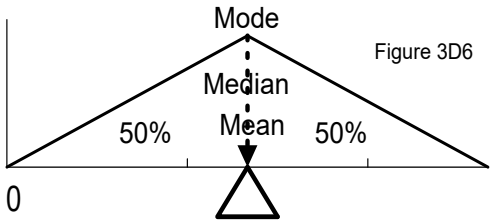


Figure 3D6

Measures of Center

In an asymmetric distribution, mean, median and mode typically align alphabetically with mean most sensitive to extremes. Why?

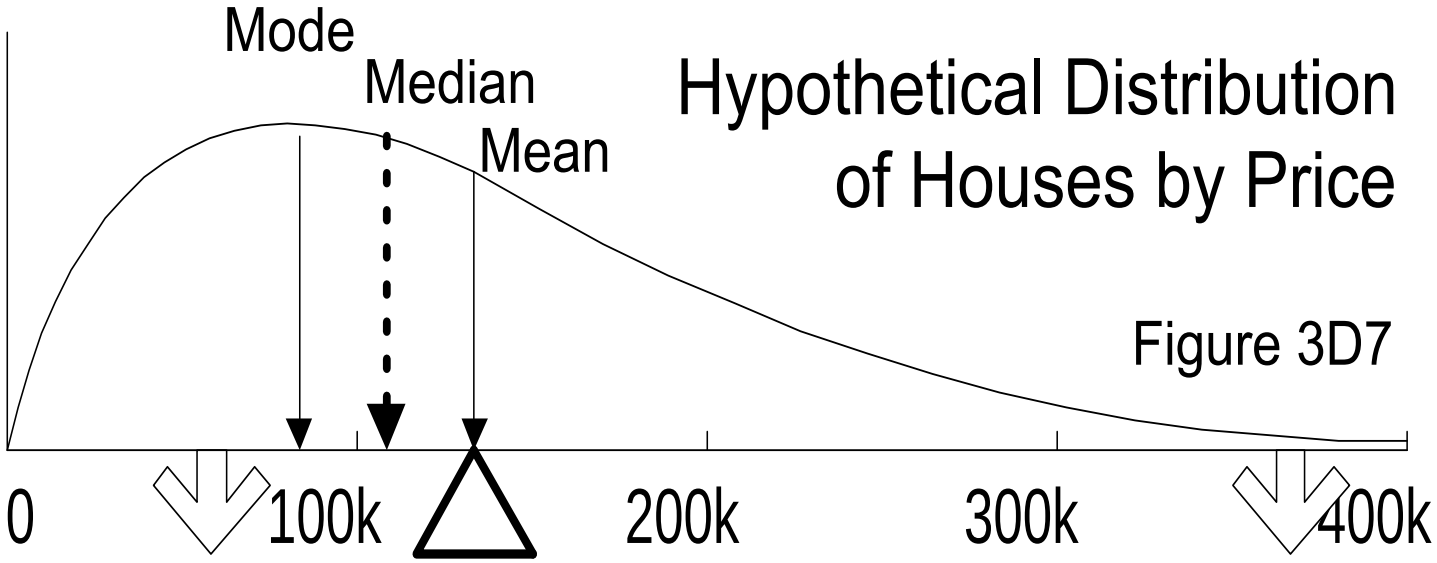


Figure 3D7

Mean, median, mode: Alphabetically. Why?

Suppose that house prices in your town have a positive near-symmetric distribution

Suppose Bill and Melinda Gates move to your town. They built two Mac-Mansions.

How does that change the mode, median and mean of the original distribution?

Mode? Median? Mean?

Most relevant in the short run? In the long-run?

Issues:

1. Mean is more sensitive to outliers.
   Yet statisticians prefer the mean. Why?
2. Omit measure: *City1 income more than City2.*
3. Omit characteristic: *Midtown is a median city.*
4. Assume the mean exists. *1.8 kids per family.*
5. Ambiguity in specifying the group

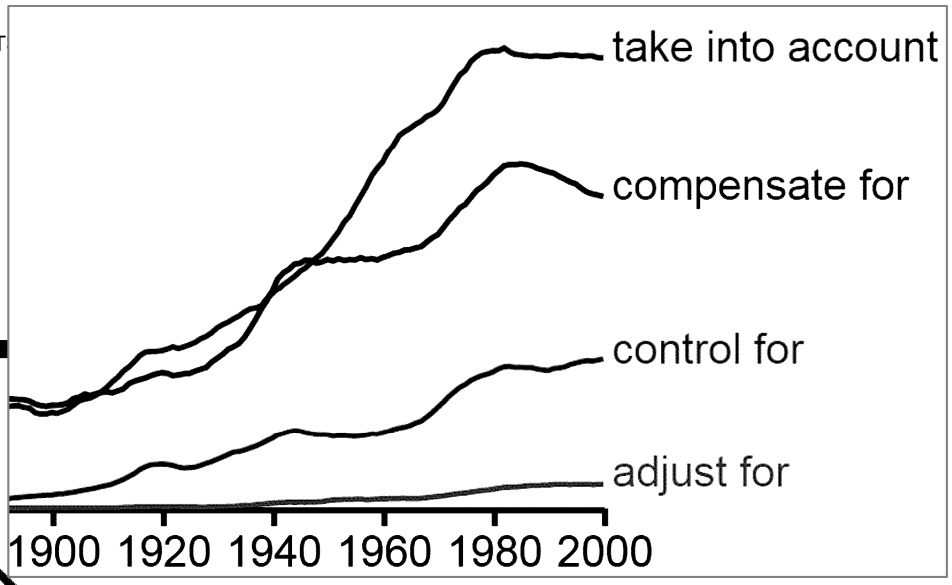
Controlling Confounding: Control Of

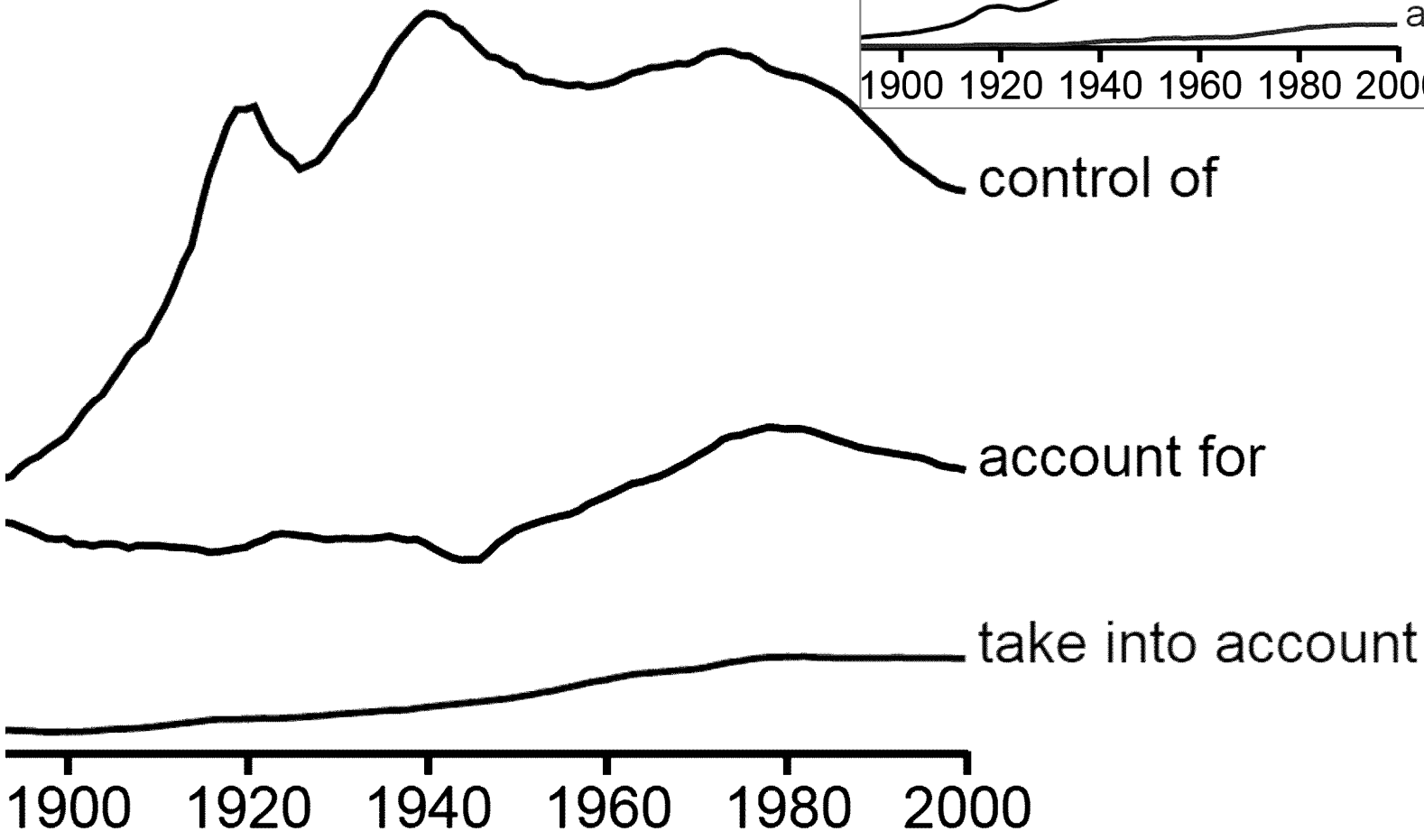
CONTROL OF CONFOUNDERS			
Physical Control (Grade = Quality)			
Experiment		Observational Study	
A+	Scientific	C	Longitudinal
A-	Random Assign	D	Cross-sectional
B	Quasi-Exper	F	Anecdotal story

Controlling Confounding: Control For

CONTROLLING FOR CONFOUNDERS			
Take into account (mental)			
	Can do by hand		*Calculator/Computer*
1	Select/Stratify	4	Linear Regression
2	Form Ratios	5	Logistic Regression
3	Standardize	6	Multivariate Regress

Control Of/For Ngrams





Crude Associations

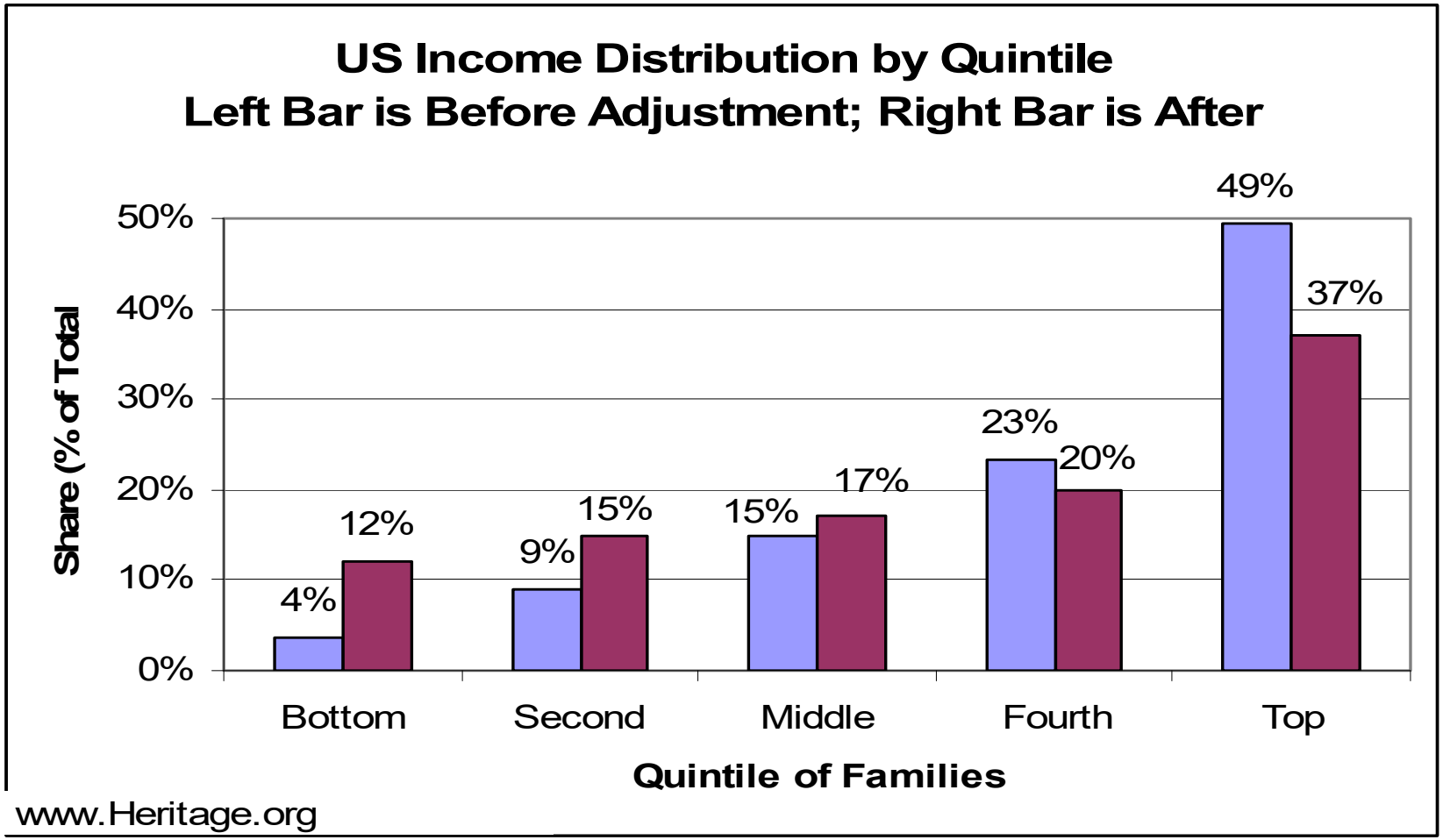
A **crude association** is an association in which nothing else has been taken into account.

Less likely to get pregnant:

- Short young adults than tall.
- Adults that shave daily than those that don't
- Adults with long hair than those with short.

What one takes into account is an assumption. Teachers should say, "Check your assumptions."

Crude Association versus an Adjusted Association



Prison Expense: Crude vs Adjusted Associations

State	Total	# Inmates	Per Inmate
CA	$2.9B	136K	$21,385
NY	$1.9B	69K	$28,426

Total	Per Inmate
50% more	25% less

State	Total	# Inmates	Per Inmate
MD	$481M	21,623	$22,245
KS	$159M	7,148	$22,245

Total	Per Inmate
3 times..	Same

State	Total	# Inmates	Per Inmate
MN	$184M	4,865	$37,825
ME	$48M	1,424	$33,711

Total	Per Inmate
260% more	12% more

Crude Ratio Associations
It's the Mix!!!

Ratio associations can be still be confounded.
Averages are ratios.

NAEP Math 8	Internet Access at Home		
State	All	Yes	No
Virginia (VA)	▲ 275	▼ 282	▼ 258
Texas (TX)	▲ 273	▼ 285	▼ 260

NAEP Math 8	Internet Access at home		
State	All	Yes	No
Virginia (VA)	▲ 275 (100%)	▲ 282 (69%)	▼ 258 (31%)
Texas (TX)	▲ 273 (100%)	▲ 285 (53%)	▼ 260 (47%)

Simpson's Paradox: Time
It's the Mix!!

SAT Verbal flat, but every group improved.

/

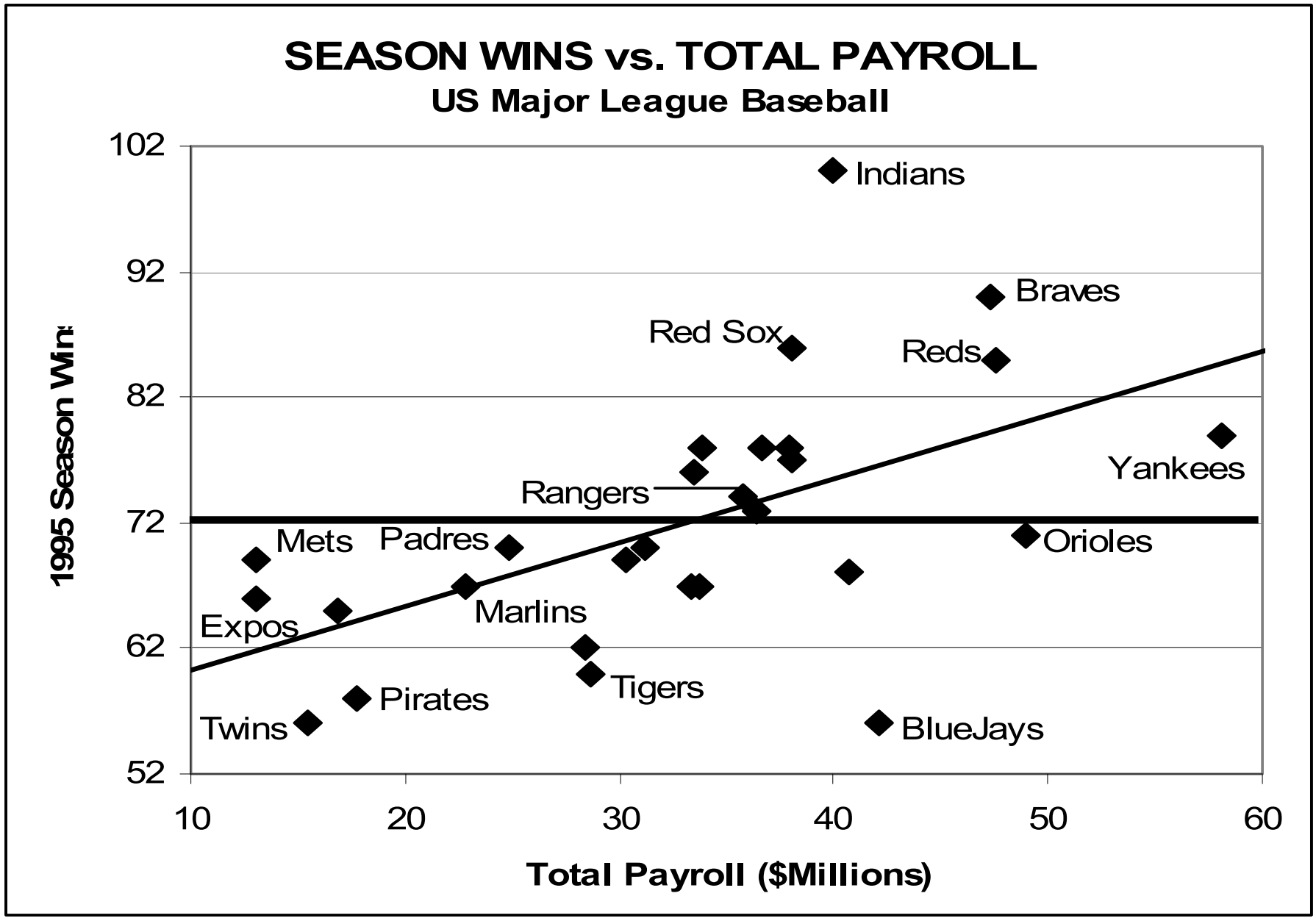
SAT-Verbal	---- Scores ----			---- Distribution ----		
Group	1981	2002*	Chg	1981	2002*	Points
White	519	527	+8	85%	65%	-20
Black	412	431	+19	9%	11%	+2
Asian	474	501	+27	3%	10%	+7
Mexican	438	446	+8	2%	4%	+2
Puerto Rican	437	455	+18	1%	3%	+2
Amer. Indian	471	479	+8	0%	1%	+1
ALL	504	504	0			

Will an Association Reverse? The Cornfield Conditions

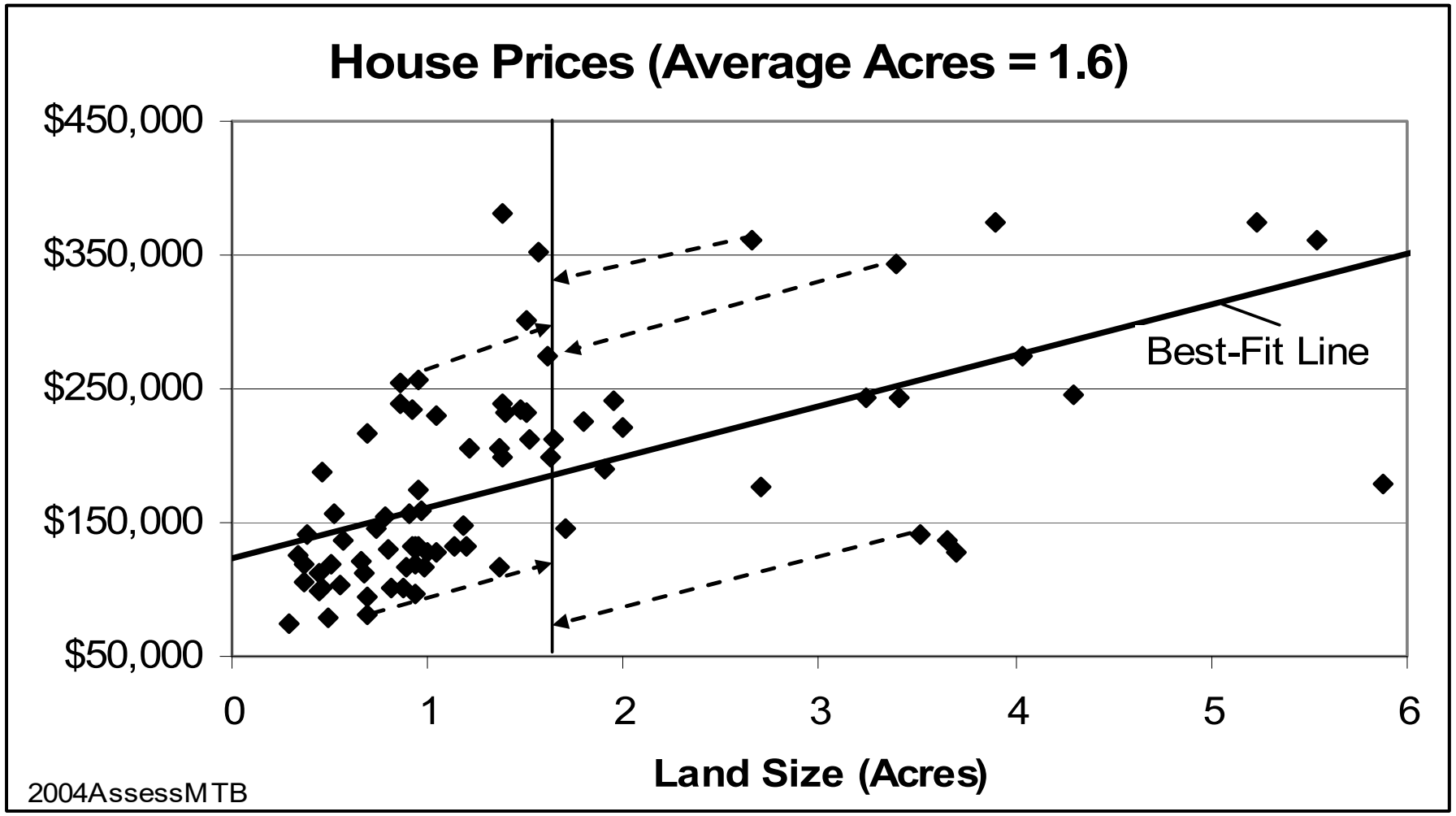
After learning about Simpson's Paradox, one student said, "I'll never trust another statistic." This is cynicism: not a good outcome.

Not all confounders can reverse an association. Jerome Cornfield proved that a confounder association must be "bigger" than the observed.

Cornfield's conditions are one of the three biggest contributions of statistics to human knowledge.


SEASON WINS vs. TOTAL PAYROLL
US Major League Baseball
1995 Season Wins
102
92
82
72
62
52
10
20
30
40
50
60
Total Payroll ($Millions)
Indians
Braves
Red Sox
Reds
Yankees
Rangers
Mets
Padres
Orioles
Expos
Marlins
Tigers
Pirates
Twins
BlueJays

Regression Standardizes



Regression Standardizes An Example:

The data shows that house prices increase by $39,000 per bedroom. This is a crude association.

$16,000 per bedroom if land is *controlled for*,

$9,000 per bedroom after *accounting for* land and house size,

$5,000 after *adjusting for* land, house size, and number of bathrooms.

TV for toddlers interferes with brain growth, says study:

Children under two should not be allowed to watch television because it increases their chances of suffering attention problems later in life, says an American study.

A study of 1,345 children found that each hour spent in front of the set every day increased the risks of attention deficit disorders by 10%.

U.S. journal, *Pediatrics*

Time to Double given Growth Rate

If a child's risk of Attention Deficit Disorder increases by 10% for every extra hour of watching TV, how many hours do they have to watch to double their risk?

Rule of 72*: Time to double = 72 / Rate

72 divided by 10% per hour = 7.2 hours

* Assuming compounding

How to Relate this to Math Colleagues

Don't talk about confounding or effect size.

Talk about assumptions.

- What one controls for is an assumption.
- What one fails to control for is an assumption.

AAU&C Quantitative Literacy VALUE rubric:

Assumptions: Ability to make and evaluate important assumptions in estimation, modeling, and data analysis.

AAC&U Quantitative Literacy VALUE Rubric

Interpretation, Representation, Calculation, Application, **Assumptions**, and Communication

Assumptions: Ability to **make and evaluate** important assumptions in estimation, modeling, and data analysis.

www.statlit.org/pdf/2009QuantitativeLiteracyRubricAACU.pdf
www.aacu.org/peerreivew/2014/summer/RealityCheck

Teaching Statistical Literacy

Chapter 4: Using and Describing Ratios

by
Milo Schield

Half-Day Workshop
USCOTS May 16, 2019
www.StatLit.org/pdf/2019-Schield-USCOTS-Slides4.pdf

Workshop Schedule

Start		Topic
1:00	1	Statistical Literacy Intro
1:30	2	StatLit Details
2:15	3	Measurements
2:45	4	Named Ratio Grammar
3:30	5	Comparing Count Ratios
4:00	6	Untangling Statistics

Ratios: Chapter 4 Outline

Per grammars:

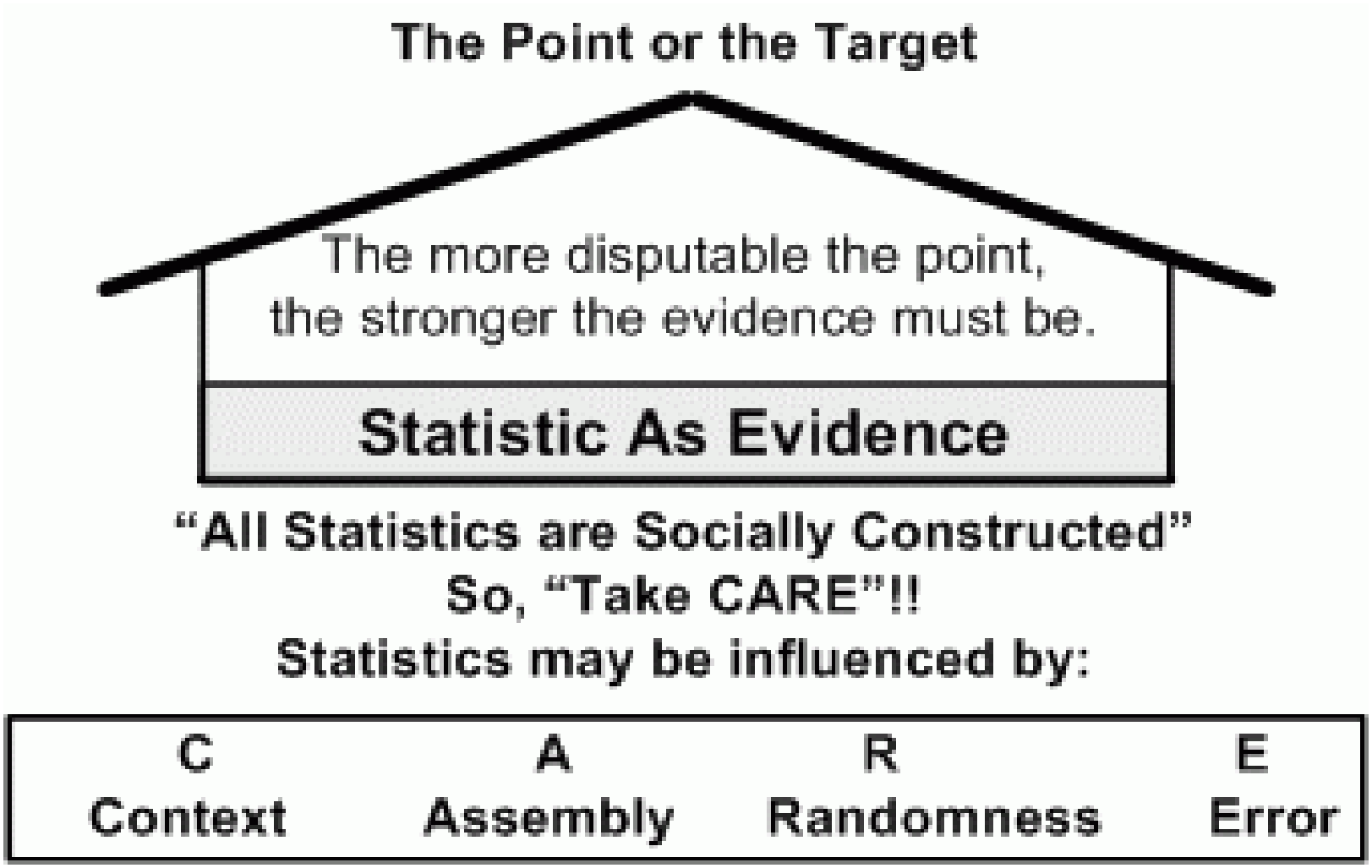
- Percent grammar
- Percentage grammar
- Reading half tables and tables w/o margins
- Rate grammar

Ordinary Preposition grammars:

- Chance grammar
- Ratio grammar

Stat Literacy: Study Statistics as Evidence in Arguments

./



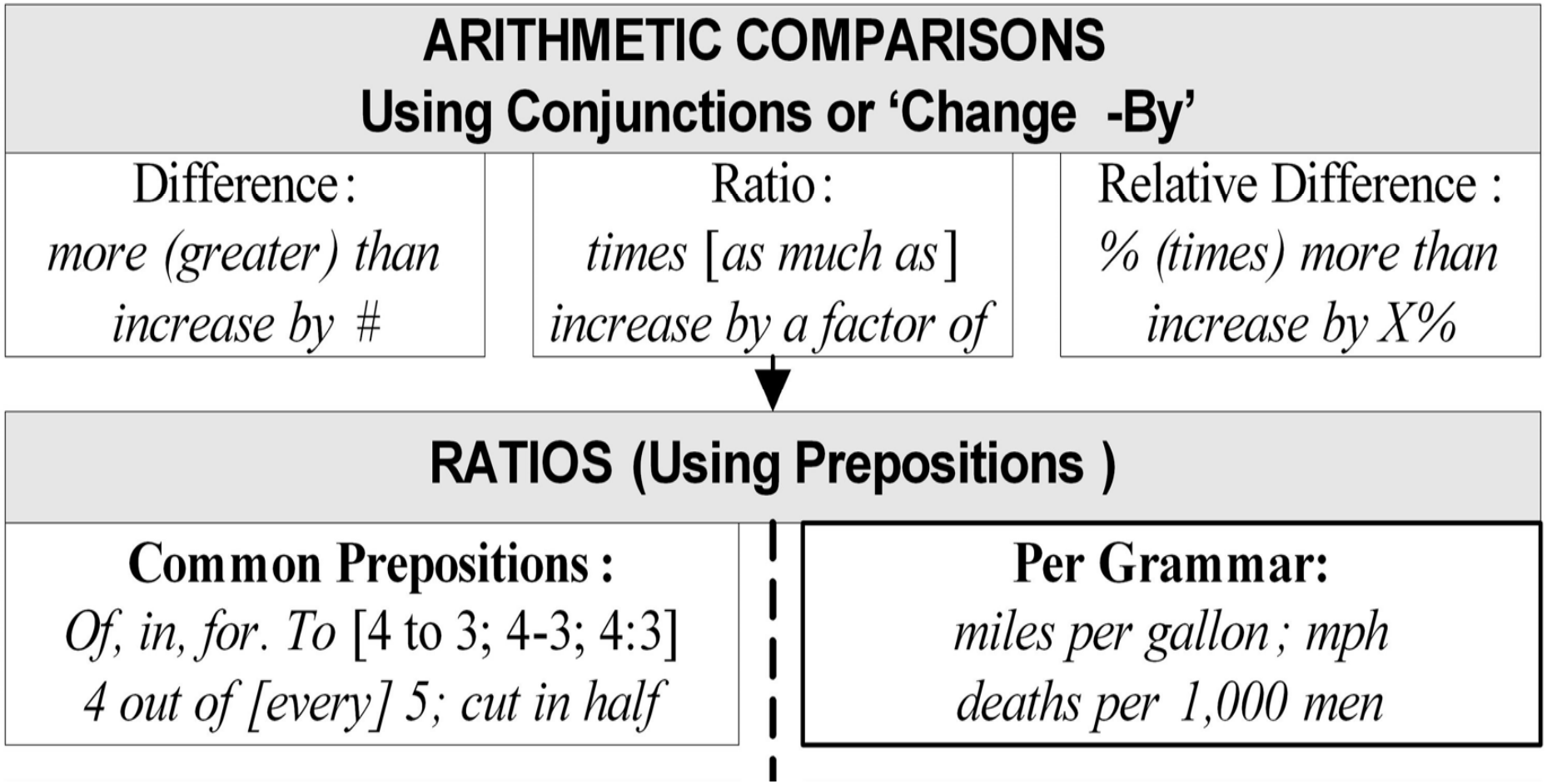
Evaluate these Using Just Assembly/Assumptions

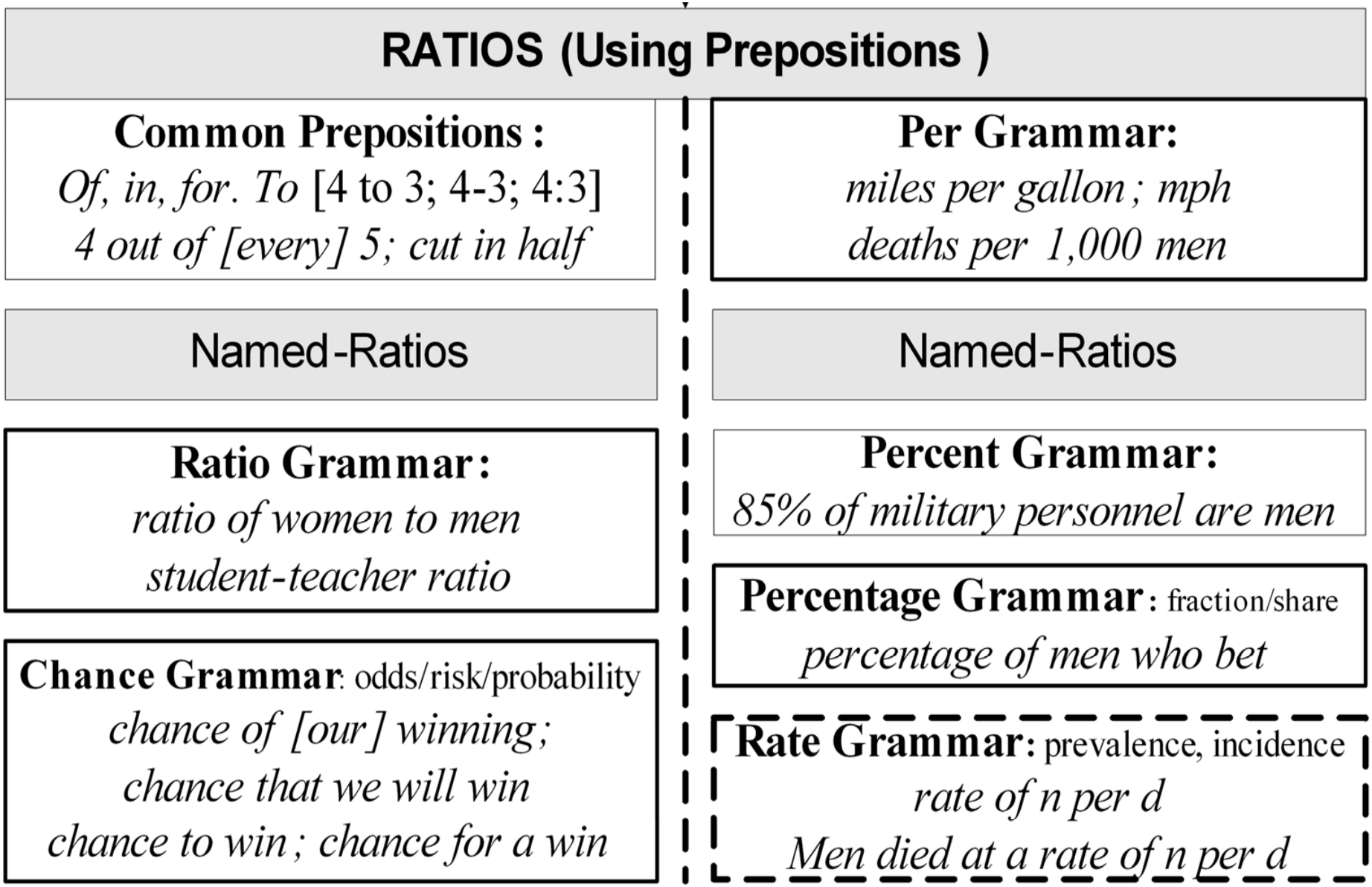
1. One in five children face hunger [2019 billboard in St. Paul]
2. Two absences per month = Likely to fail a grade
3. Ninth-grade attendance better predicts graduation than 8th grade test score
4. Attendance alone explains 31% of the variance in performance
5. Budget cuts lead to deaths in Federal prisons
6. 22 million victims of human trafficking trapped worldwide.
7. The National Rifle Association is a terrorist organization.
8. Ban assault weapons
9. 2016 Memphis. 228 homicides. Down 500 police officers.

Forming Ratios

CONTROLLING FOR CONFOUNDERS			
Take into account (mental)			
	Can do by hand		*Calculator/Computer*
1	Select/Stratify	4	Linear Regression
2	Form Ratios	5	Logistic Regression
3	Standardize	6	Multivariate Regress

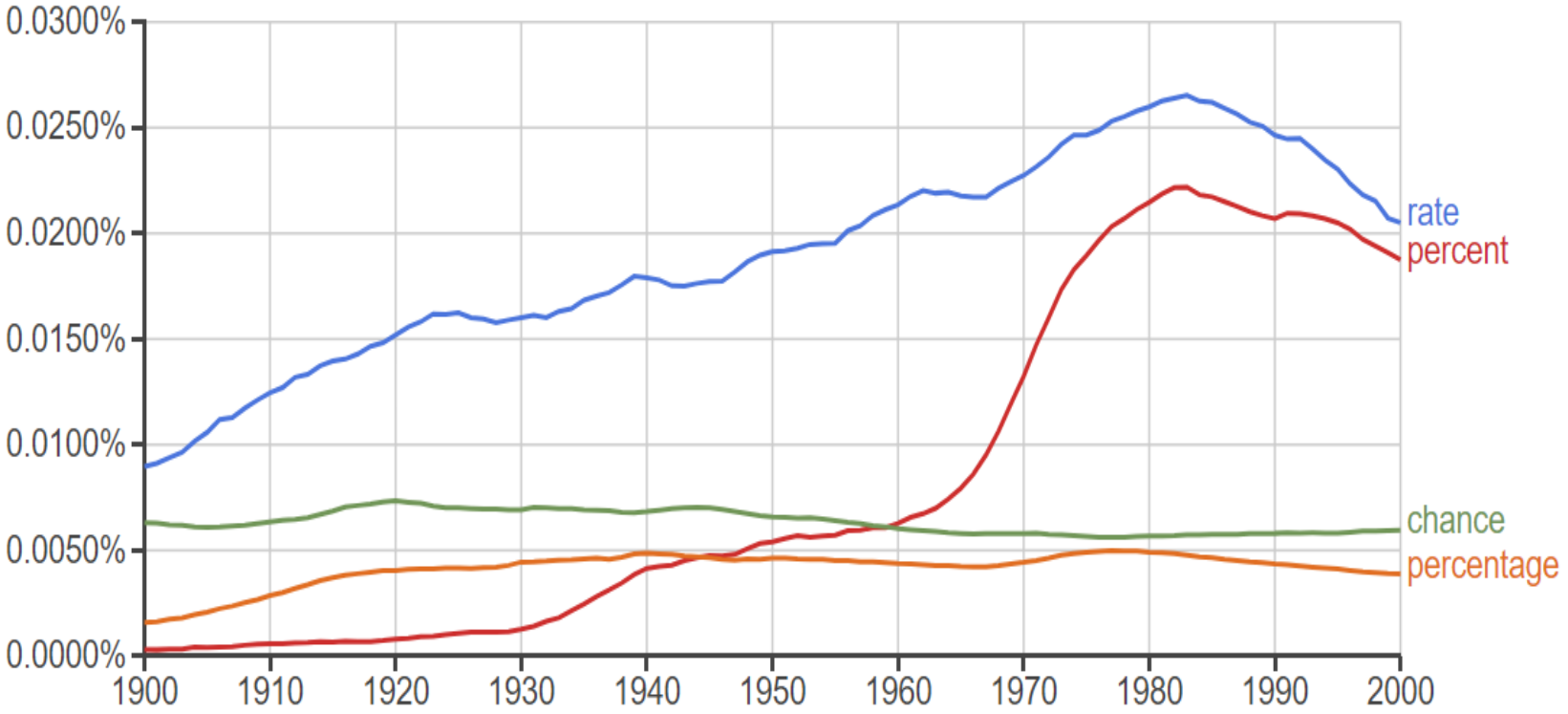
From Comparisons to Ratios: Using Prepositions





Light-edge boxes need clause for part and whole (cannot compare ratios.
Dark-edge boxes have part and whole in phrases (can compare ratios)

Prevalence of Named Ratios



Two Kinds of Percents

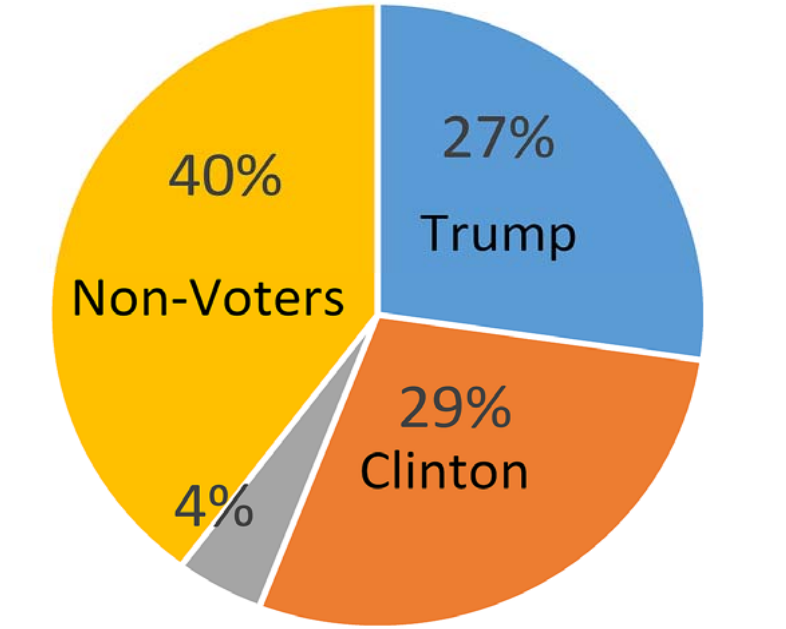
Which kind of percents are these: part-whole or percent compare?

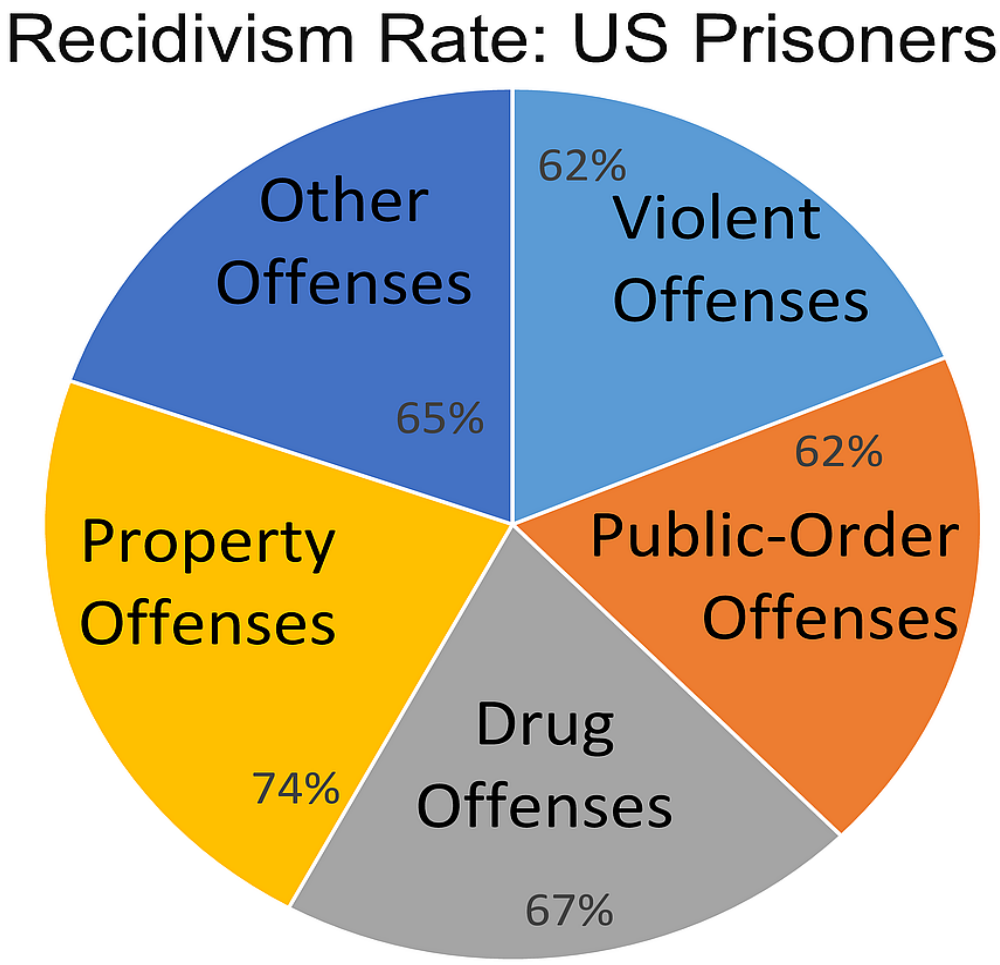
1. The youngest child's share of the candy.
2. Interest charged per year by the Mafia (criminals).
3. People live 100% longer on average in US than in Swaziland.
4. The advertisement said "40% off".

.

Part-Whole Using Pie Charts

Of all adults.





US Dept. of Justice statistics . 272,111 prisoners released in 1994.

Four Different Grammars; Confusion of the Inverse

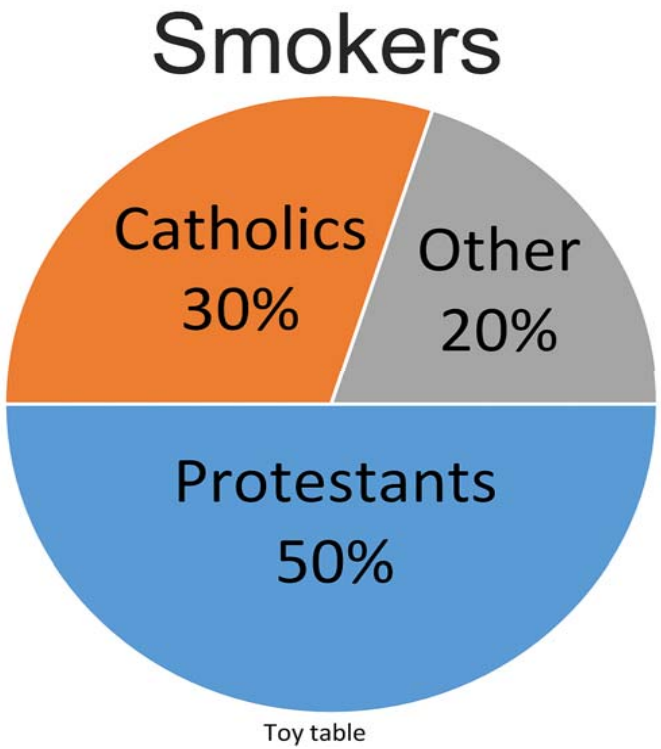
1. 40% of US adults did not vote for president in 2016.
2. The *percentage* of US adults who didn't vote was 40%
3. The non-voter *rate* among US adults in 2016 was 40%.
4. There was a 40% *chance* that an adult was a non-voter.

Confusion of the inverse exchanges part with whole.

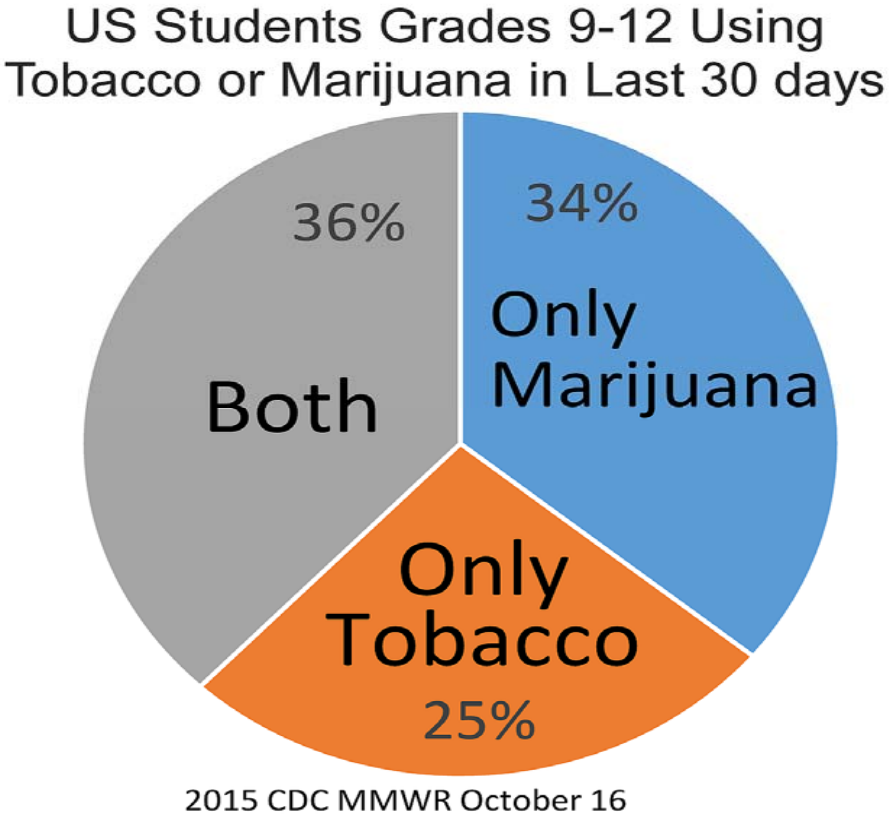
1. "The percentage of men who are in the military"
   .NE. "the percentage of the military who are men".
2. The percentage of smokers among women .NE.
   "the percentage of smokers who are women".

Use Percent Grammar
<X% of Whole are Part>

Describe the 30%



Describe the 36%



Tables: Use Percent Grammar <X% of Whole are Part>

1. What percentage of men are art majors?
2. What percentage of art majors are men?
3. What percentage of students are male art majors?

Students	Men	Women	ALL
Humanities	28	72	100
Arts	4	36	40
Science	48	12	60
ALL	80	120	200

100% Tables: Percent Grammar <X% of Whole are Part>

Describe the 10%

Students	Men	Women	ALL
Humanities	28%	72%	100%
Arts	10%	90%	100%
Science	80%	20%	100%
ALL	40%	60%	100%

Describe the 5%

Students	Men	Women	All
Humanities	35%	60%	50%
Arts	5%	30%	20%
Science	60%	10%	30%
ALL	100%	100%	100%

Use Percent Grammar
<X% of Whole are Part>

Table 33: World Population by Religion and Continent (1996)

(Millions)	Total	Asia	Europe	North Am	Other
Total	5,804	3,513	728	296	1,563
Christian	1,955	303	556	256	1,096
Muslim	1,126	778	32	5	316
Nonreligious	887	753	90	21	44
Hindus	793	787	2	1	4
Buddhists	325	322	2	1	1
Atheists	222	175	41	2	6
All Other	496	395	5	10	96

Table 1333. 1997 U.S. Statistical Abstract.

Percentage Grammar
Four form

1. The percentage of seniors who smoke is 15%.
2. Among seniors, the percentage who smoke is 15%.
3. Among Seniors, the percentage of smokers is 20%.
4. Among men, the percentage of seniors who smoke is 20%

Numbers 3 and 4 are problems.

“Of” introduces whole in percent grammar.

Percentage Grammar
Sports Grammar

Sports grammar is readily understood with a natural whole:

- *percentage of defective cans; percentage of tire failures*

Without a natural whole, sports grammar is ambiguous.

- *percentage of female smokers;*
- *percentage of working males*
- *percentage of infant deaths;*
- *percentage of single mothers*

Half Tables when Parts of 100% Table are Binary

Describe the circled 60%. Use percent grammar.

Class Last Year	Percentage who are Retained	Percentage who are Not Retained	All
Freshman	60%	40%	100%
Sophomore	75%	25%	100%
Junior	90%	10%	100%
Senior	10%	90%	100%
ALL	70%	30%	100%

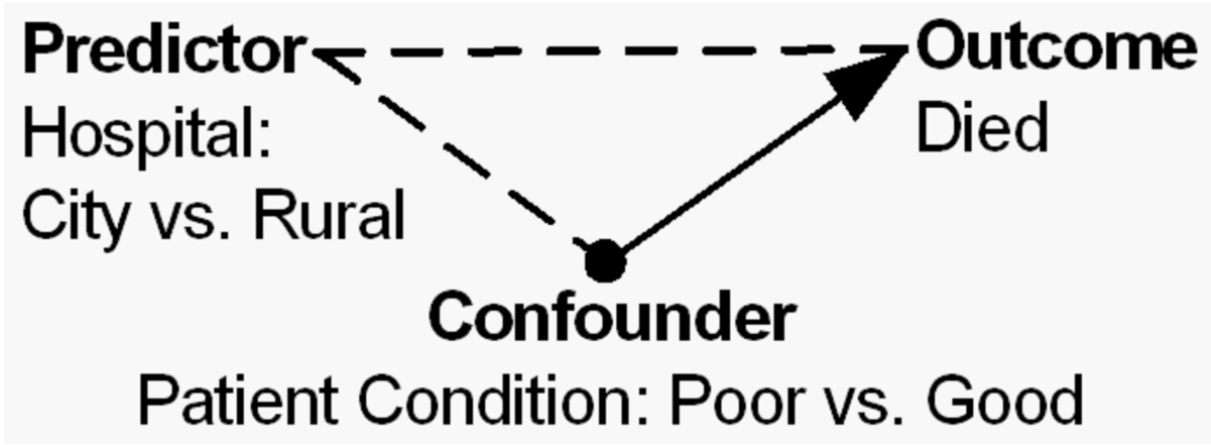
If 60% returned, what percentage did not return?

So, the right two columns are redundant.
Eliminating them will save space!

Confounding

Mortality by Hospital

Hospital	Total	Died	Death Rate
City	1,000	55	5.5%
Rural	1,000	35	3.5%
Both	2,000	90	4.5%



Statistics Literacy For Decision Makers

13: Confounding & Cornfield

by
Milo Schield

Half-Day Workshop
USCOTS May 16, 2019
www.StatLit.org/pdf/2019-Schield-USCOTS-Slides13.pdf

Workshop Schedule

1:00 Ch 1 Statistical Literacy – Introduction
1:30 Ch 2 Statistical Literacy – Details

2:15 Ch 3 Measurements
2:45 Ch 4 Ratios

3:30 Ch 13 Standardizing
4:00 Feedback

.

Confounding: Chapter 13 Outline

Cornfield-Fisher debate

Cornfield conditions

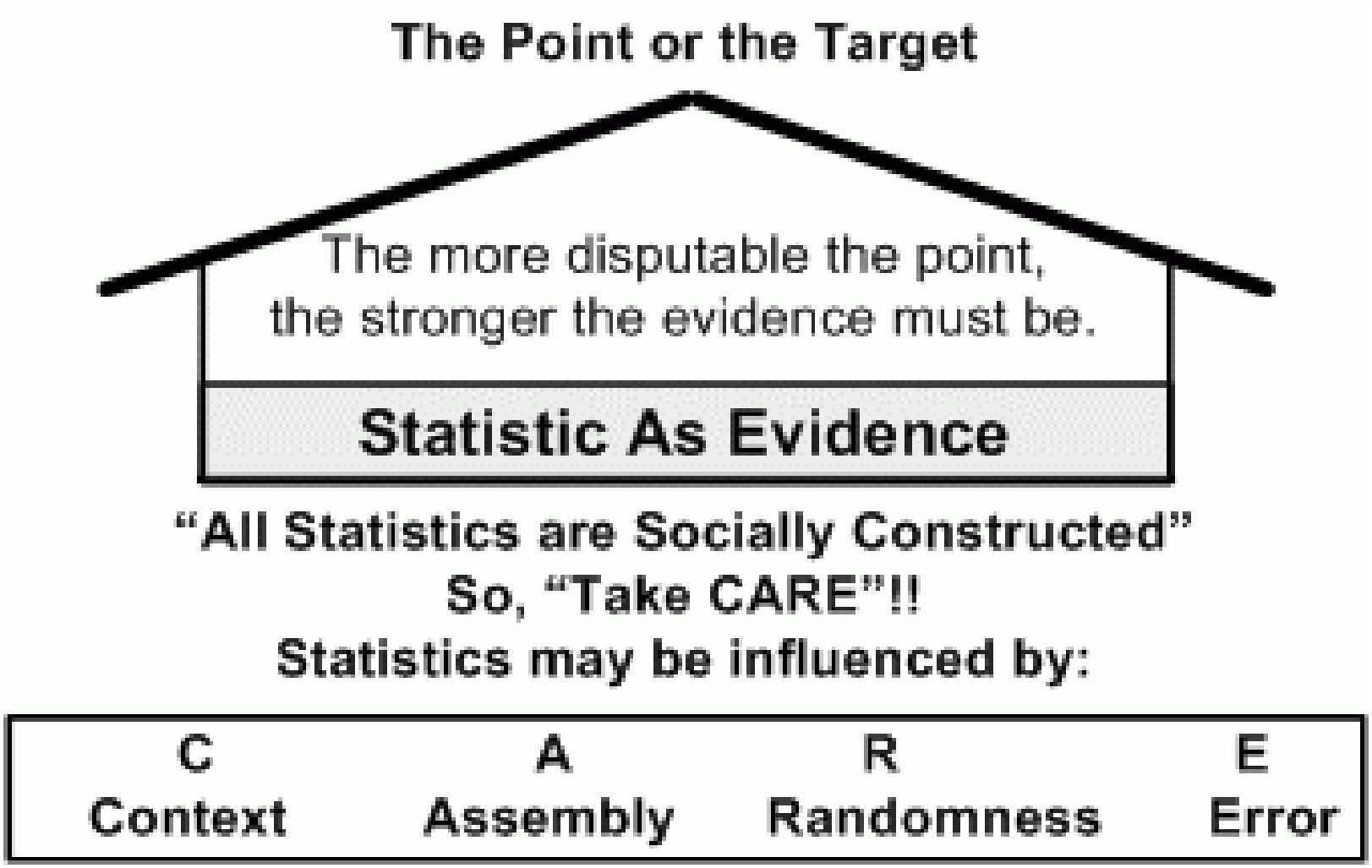
Standardizing percentages, rates and averages

Standardizing percentage & number attributable

Statistical significance and confounding

Stat Literacy: Study Statistics as Evidence in Arguments

./



Cornfield-Fisher Debate

Doctors had noticed the strong association between smoking and lung cancer. Statisticians argued that this evidence strongly supported the claim that smoking was a cause of lung cancer.

Fisher, a smoker, noted that *association is not causation in observational studies*.

Fisher produced data. Identical twins were more likely to share a smoking preference than were fraternal twins. This statistic supported genetics as an alternate explanation for the association.

Cornfield-Fisher Debate

Now when the world's leading statistician says something that every statistician agrees is true, most reasonably-minded statisticians would back off.

And when the world's leading statistician produces data indicating a plausible confounder, it seems incredible that anyone would reply.

Jerome Cornfield did!

Cornfield Conditions

Cornfield **proved** that the relative risk of lung cancer had to be greater for a confounder (e.g., genetics) than for the predictor (e.g., smoking) in order to nullify or reverse the observed association.

Cornfield pointed out that smokers were about 10 times as likely to get lung cancer as non-smokers.

Fisher's data involved a factor of two.

Fisher never replied.

Contributions to Human Knowledge

"Cornfield's minimum effect size is as important to observational studies as is the use of randomized assignment to experimental studies.

No longer could one refute an ostensive causal association by simply asserting that some new factor (such as a genetic factor) might be the true cause.

Now one had to argue that the relative prevalence of this potentially confounding factor was greater than the relative risk for the ostensive cause."

Schield (1999). [This was written 20 years ago!]

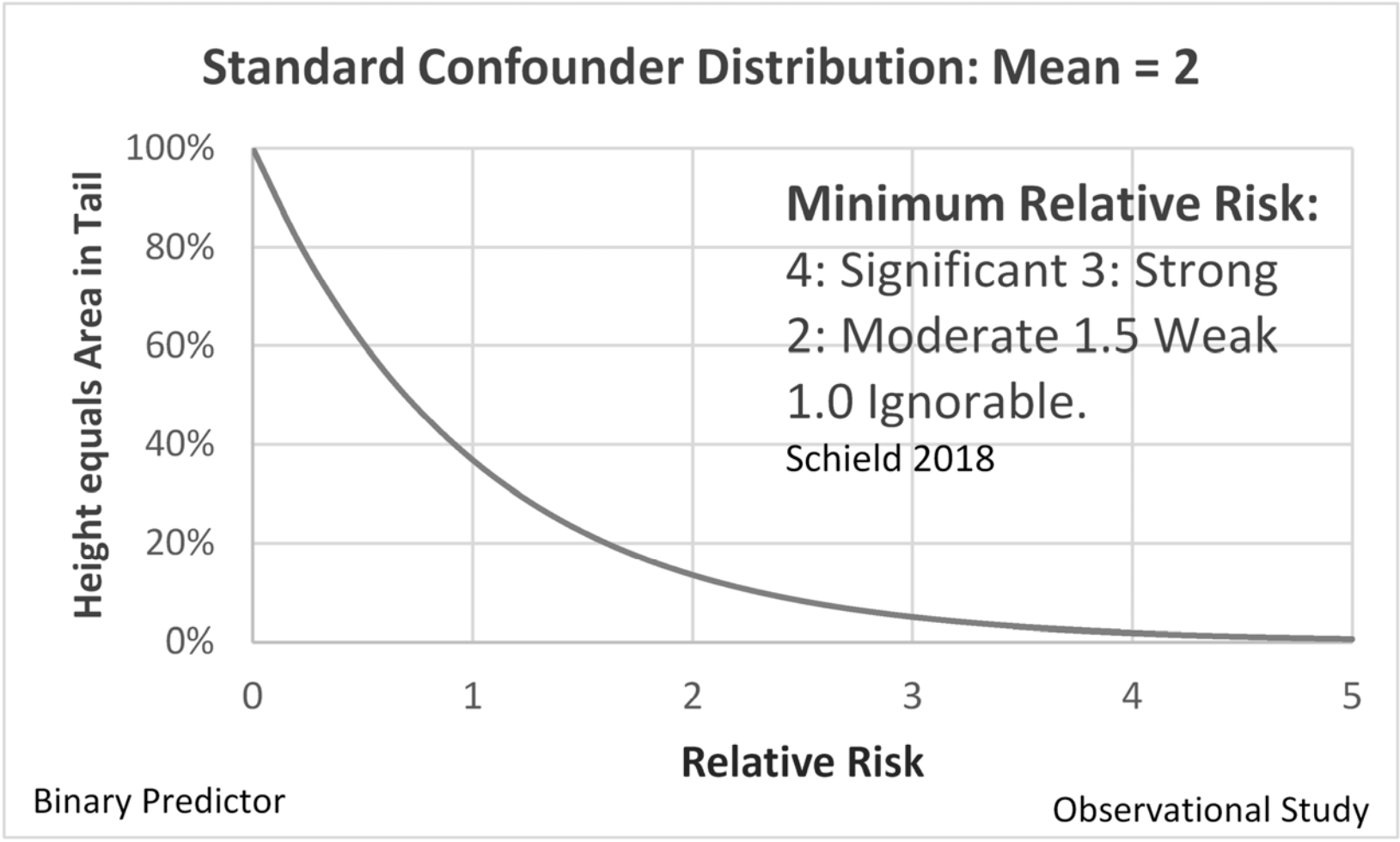
Confounder Distribution

Since confounders may be unknown, there is no way to derive or infer their distribution.

Schield (2018) argued that we needed a standard for confounder: a standard confounder distribution.

He proposed an exponential (one factor determined) with a mean relative risk of 2.
This applied if predictor and confounder are binary.

Confounder Distribution Unknown & Unknowable

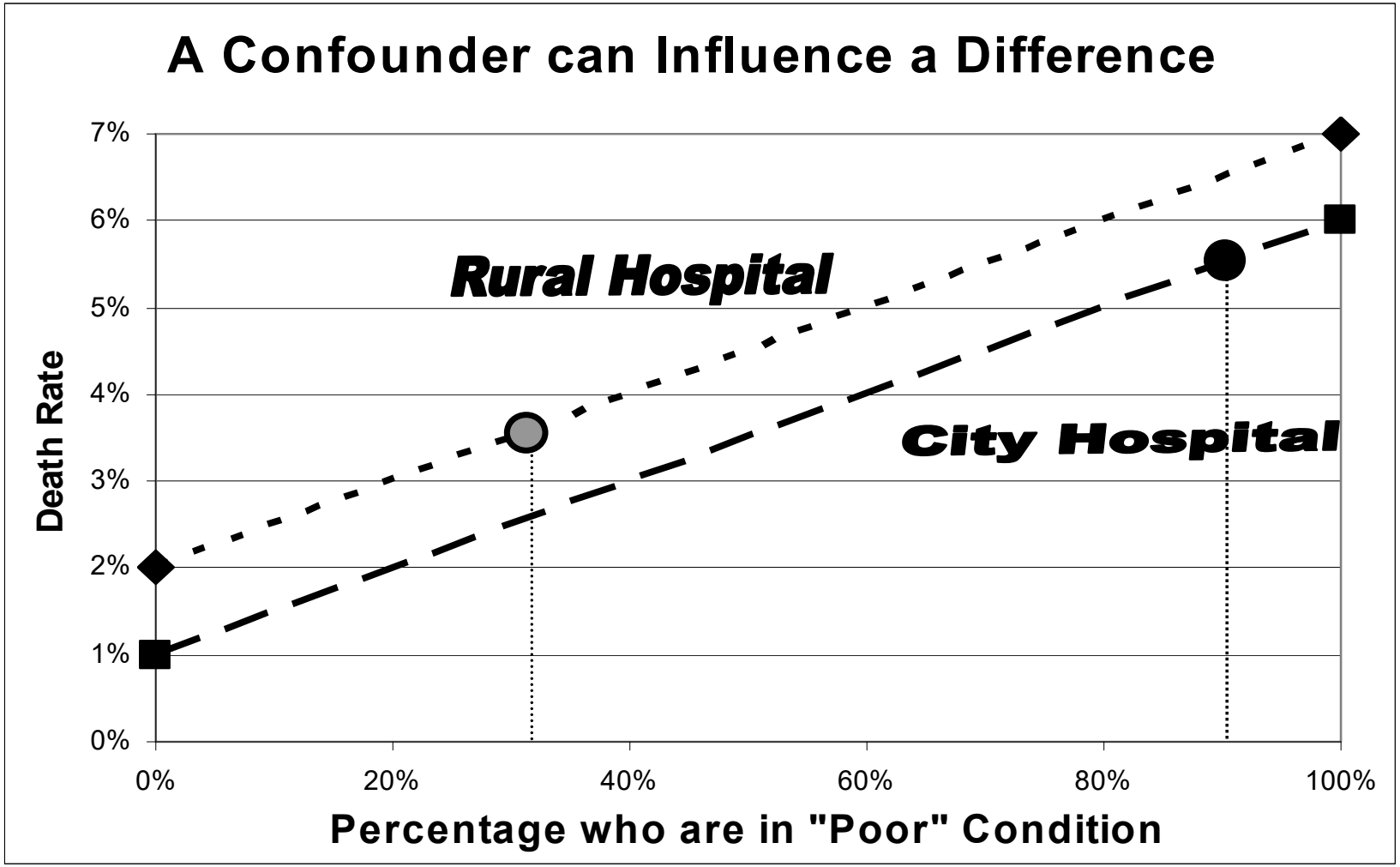


Controlling for a Confounder: Graphical Technique

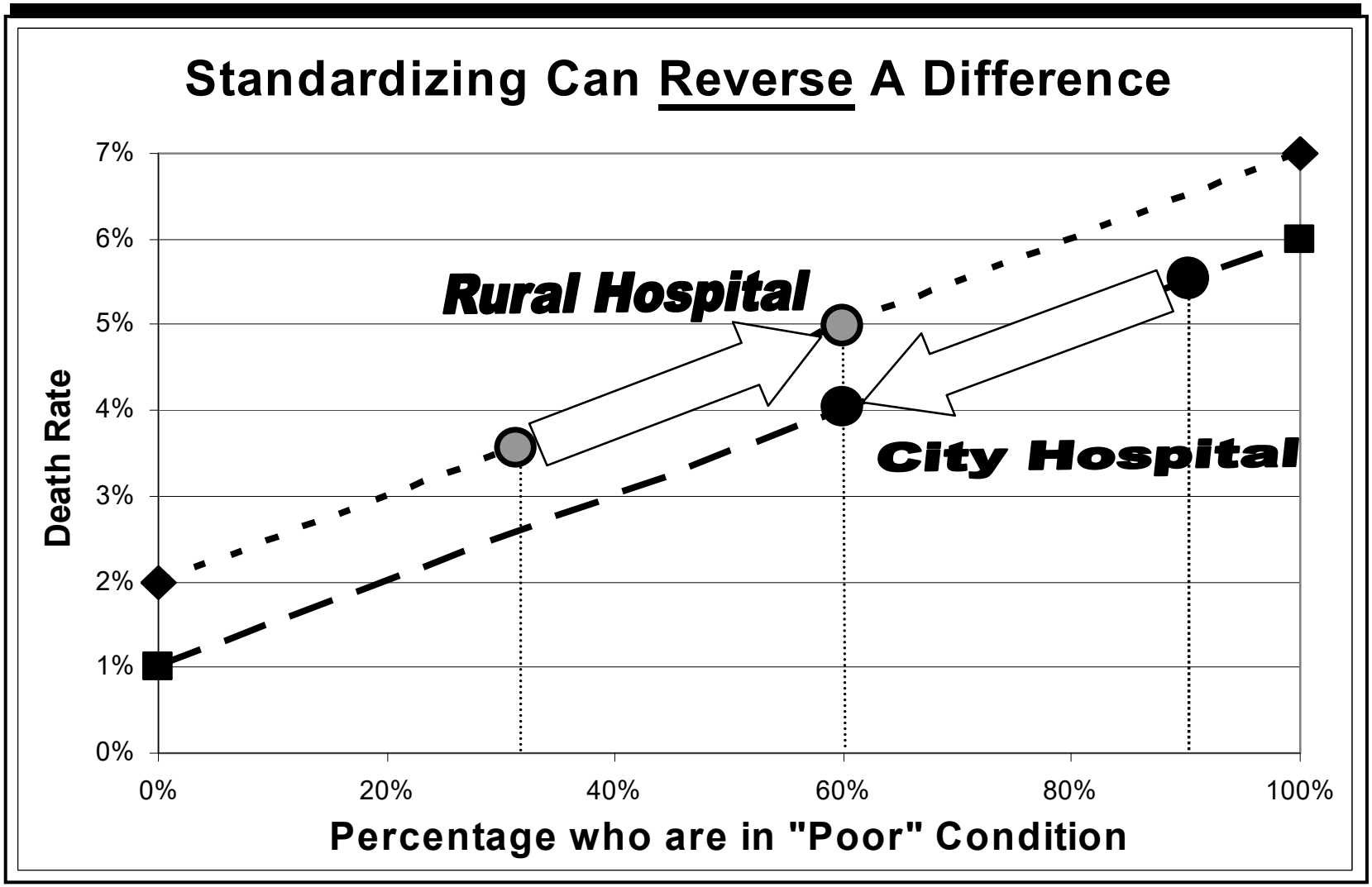
Wainer introduced a simple graphical technique that made the control of a binary confounder a relatively simple matter.

Schield (2006). Presenting Confounding Graphically Using Standardization, *STATS* magazine. www.statlit.org/pdf/2006SchieldSTATS.pdf

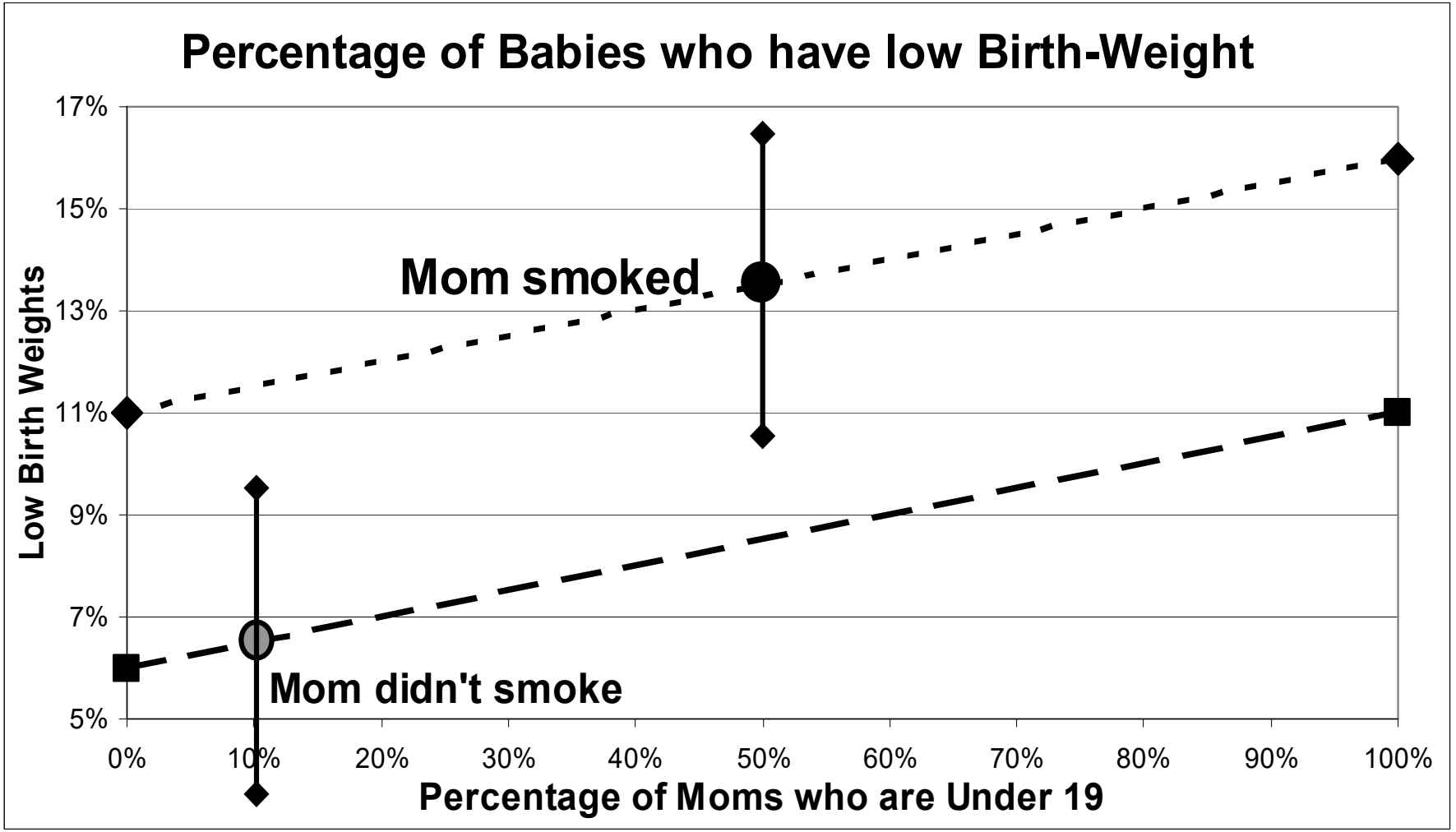
Crude Association: Death Rate: City > Rural



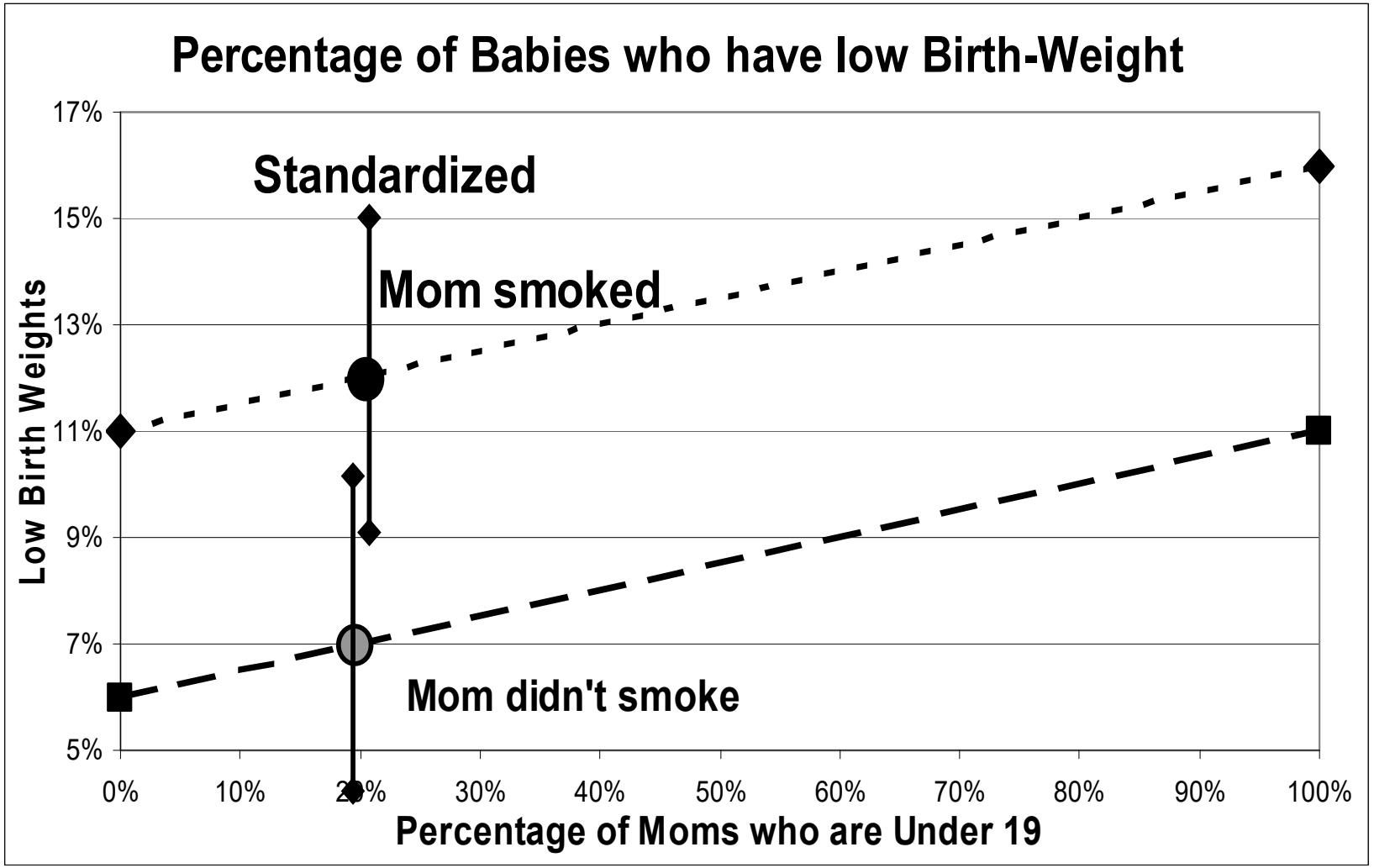
Controlling for a Confounder: Death Rate: City < Rural



Crude Association: Statistically Significant



Standardized Association: Statistically Insignificant



Confounder Effect on Statistical Significance

Controlling for a confounder can transform a statistically-significant association into an association that is statistically insignificant.

Although statistical educators are clearly aware of this, there is nothing in any introductory textbook that alerts students to this possibility.

The failure to show a significance reversal is *statistical negligence*.