V1 Schield: 2019 NNN Annual Meeting 1

What is Numeracy

by
Milo Schield, NNN VP

NNN Annual Meeting
Austin CC Oct 13, 2019

www.StatLit.org/pdf/
2019-Schield-NNN-Slides.pdf

V1 Schield: 2019 NNN Annual Meeting 2

Steen (1997) *Why Numbers Count*

Numeracy is the new literacy of our age.

Whatever this phrase [Q/L] may mean—and as the essays in this volume testify, it means very different things to different people.

It [Q/L] requires a working synthesis of literacy and numeracy; it evolves with technology; and it both shapes and is shaped by society.

V1 Schield: 2019 NNN Annual Meeting 3

Steen (1997) *Why Numbers Count*

Regardless of name—numeracy, mathematics, quantitative literacy, or the derisive "'rithmetic," —this kind of literacy is widely recognized as of fundamental importance.

Yet beyond "the basics," there is little agreement about specific goals appropriate for tomorrow's world. No wonder, then, that we have made so little progress in achieving numeracy.

Mere Literacy is not Enough. George Cobb

V1 Schield: 2019 NNN Annual Meeting 4

Steen (1997) *Why Numbers Count*

In short an *innumerate* citizen today is as vulnerable as the illiterate peasant of Gutenberg's time.

Although the widespread availability of data should enrich public discourse, inevitable over-simplifications and misinterpretations may ultimately cheapen it. ... Instead of enhancing Jeffersonian democracy, *limited numeracy* can easily shift the balance to a technocracy.

Innumeracy thus becomes another means of disenfranchisement: by reinforcing the idea that truth is relative and unknowable, people with the least defenses against charlatans will be most vulnerable.

V1 Schield: 2019 NNN Annual Meeting 5

Steen (2001) *Mathematics and Democracy*

Numeracy is not so much about understanding abstract concepts as about applying elementary tools in sophisticated settings.

Numeracy is not the same as mathematics, nor is it an alternative to mathematics.

Rather, it is an equal and supporting partner in helping students learn to cope with the quantitative demands of modern society

V1 Schield: 2019 NNN Annual Meeting 6

Perspectives on *Numeracy*

NUMERACY

Quantitative Reasoning	Quantitative Literacy	Statistical Literacy	Statistical Reasoning

NUMERACY

What parts of math?
Topics Skills Models Patterns Thinking

Perspectives on *Numeracy*

NUMERACY

Which quantitative topics?
Arithmetic Algebra Geometry

NUMERACY

Audience: Which Majors???

Math/ STEM	Social Science	Professions Biz/Econ/etc	Humanities English, Art

Perspectives on *Numeracy*

NUMERACY

Citizens
High School College Grads

NUMERACY

Studies the quantitative skills needed by educated citizens in a modern democracy.

What is the Most Important Thing to Know about Statistics?

"All statistics are socially constructed."
Joel Best, author *Lies, Damned Lies & Statistics*

Doesn't mean that

- there is no reality, all statistics are imaginary

Means that people create statistics like diamonds

What is the Most Important Thing to Know about Statistics?

Statistics are socially constructed: the products of social activities.

There's a tendency in our culture to believe that statistics—that numbers—are little nuggets of truth.

That we can come upon them and pick them up very much the way a rock collector picks up stones.

What is the Most Important Thing to Know about Statistics?

A better metaphor would be to suggest that statistics are like jewels; that is,

they have to be selected, they have to be cut, they have to be polished, and they have to be placed in settings so that they can be viewed from particular angles.

Joel Best, Sociologist

What follows from being Socially Constructed?

Numbers can't be influenced. 1 + 1 = 2

Statistics can be influenced. 1 + 1 may equal 2

One gallon of antifreeze and one gallon of water do not yield two gallons. The combination of large and small molecules takes up less space.

V1 Schield: 2019 NNN Annual Meeting 13

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 typically evidence of causation.

V1 Schield: 2019 NNN Annual Meeting 14

A-B-C Words: A = Association

Statistical association: an observable connection.

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

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

V1 Schield: 2019 NNN Annual Meeting 15

A-B-C Words: C = Causation

Causation: Lightning *caused* (*resulted in*) the fire.
Insomnia is a side *effect*.

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*.

Schield and Raymond (2009). www.StatLit.org/pdf/2009SchieldRaymondASA.pdf

V1 Schield: 2019 NNN Annual Meeting 16

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 increased *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.

V1 Schield: 2019 NNN Annual Meeting 17

A-B-C Words: Distribution in Headlines

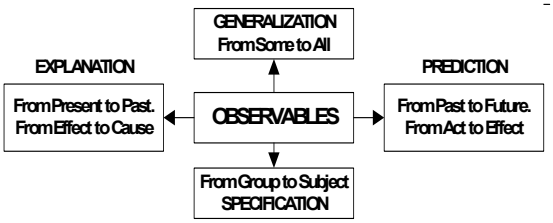
Of the 2,000 news headlines analyzed[6],
71% involved A, B or C.

Of those headlines involving A, B or C,
- **86% were "between" claims**,
- 11% sufficiency, 3% causation, 3% association.

6. Schield and Raymond (2009).

V1 Schield: 2019 NNN Annual Meeting 18

Critical Thinking Inferences



Association is Not Causation.

V1 Schield: 2019 NNN Annual Meeting 19

Prediction: Its hopeless! This bear is faster than we are

2nd hunter: No, it's not hopeless.

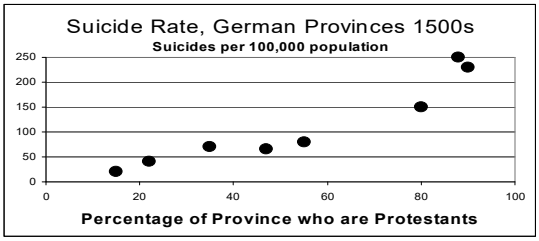
I don't have to outrun the bear.

I just have to outrun you!

V1 Schield: 2019 NNN Annual Meeting 20

Specification: Cross-Level Inference

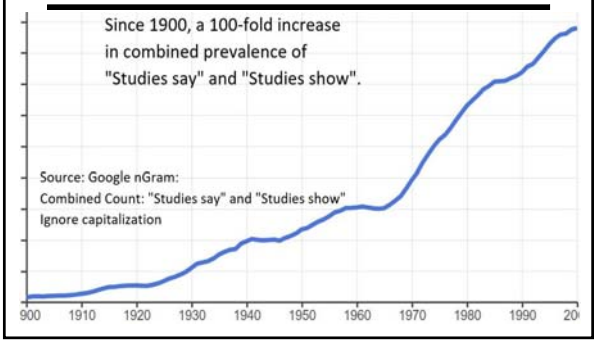
Protestants more likely to suicide than Catholics?



No. Reverse was true here. Catholics more likely.

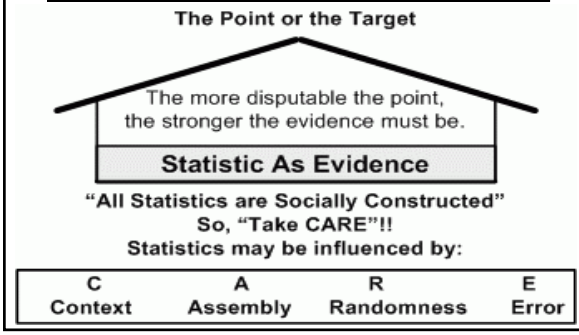
V1 Schield: 2019 NNN Annual Meeting 21

Statistical Studies as Evidence in Arguments



V1 Schield: 2019 NNN Annual Meeting 22

Statistical Literacy Studies Statistics in Arguments



V1 Schield: 2019 NNN Annual Meeting 23

Take C.A.R.E Four Influences on Statistics

Statistics are influenced by

C = Confounding: By related factors.

A = Assembly: By definitions and presentation.

R = Randomness: By uncertainty or chance

E = Error: By mistakes or bias.

Assembly is the etcetera category.

V1 Schield: 2019 NNN Annual Meeting 24

Confounding Without Statistics

A father and his children were on a subway.

The children were out of control: jumping on seats, yelling, and throwing things.

The father did nothing.

He slumped forward looking down at the floor, his head between his hands.

Finally an unhappy onlooker called on the father to take control of his kids.

V1 Schield: 2019 NNN Annual Meeting 25

Confounding Without Statistics

The father looked up sadly and said:

We just left the hospital where their mom died.

Immediately the negative judgments were transformed into pity for this family.

The onlookers were confused – confounded – by a confounder: the death of the kids' mom.

V1 Schield: 2019 NNN Annual Meeting 26

Statistical Influences: Confounding #1

Adults who shave their faces tend to be taller than those who shave their legs.

Does face-shaving cause tallness? ___
Is this association confusing? ___
Is it confused by some outside factor? ____.
What is it? ____

Gender. Men tend to be taller than women.
Men are more likely to shave their face.

V1 Schield: 2019 NNN Annual Meeting 27

Statistical Influences: Confounding #2

People that read home and fashion magazines are more likely to get pregnant than people that read car and sport magazines.

Are the magazines causing pregnancy? ___
Is the association confused by an outside factor?
What is this outside factor? _____

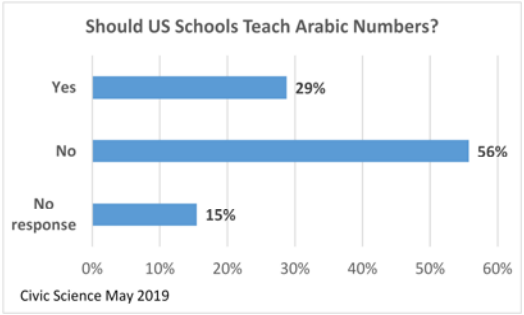
Gender! Women can get pregnant (men can't)
Gals more likely to read home/fashion magazines

V1 Schield: 2019 NNN Annual Meeting 28

Assembly Fueled Brexit? Gross vs. Net (50%):

V1 Schield: 2019 NNN Annual Meeting 29

Assembly: Sharia math, then Sharia law!!!



V1 Schield: 2019 NNN Annual Meeting 30

Statistical Influences : Assembly #1

Claims about college students:

Administration: 80% of are 'satisfied'.
Students: 70% are not 'satisfied'.

Same data:
Happy (30%), OK (50%), Unhappy (20%)

Q. Who is correct?
A. Both are. Different definitions of 'satisfied'.

V1 Schield: 2019 NNN Annual Meeting 31

Statistical Influences : Assembly #2

Living with AIDs

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

Two claims about groups living with AIDS:

1. More blacks than whites.
2. More whites than blacks.

V1 Schield: 2019 NNN Annual Meeting 32

Assembly: Presentation

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 Schield: 2019 NNN Annual Meeting 33

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 Schield: 2019 NNN Annual Meeting

Stat Influences : Randomness

Sport Illustrated Curse

Those people featured as *best* by Sports Illustrated seem cursed!

They don't do as well the next year. Why not?

Being the best one year is often coincidence.

V1 Schield: 2019 NNN Annual Meeting 35

Statistical Literacy deals with Ignorance

How was data collected?
C: What factors not taken into account?
A: How were things counted, measured, grouped?
R: How small is the group?
E: Are subjects telling the truth?

Often we don't know! We are ignorant!
We are not omniscient.

The solution? Think hypothetically! Plausible?

V1 Schield: 2019 NNN Annual Meeting 36

Controlling Confounding: 'Control Of' vs. 'Control for'

1. Study design indicates "control of"
2. Comparisons and ratios indicate "control for"

Control for:

1. Comparisons control for a relevant basis
2. Ratios control for size of group
3. Comparisons of ratios control for both

V1 Schield: 2019 NNN Annual Meeting 37

Associations: Two Kinds

Two-group comparisons:
- Women live longer than men

Two-factor covariation: ordinal & quantitative
- The more height, the more weight
- As height increases, weight increases

- As height increases by *x*, weight increases by *y*
- For every additional *x* in X, Y increases by *y*.

V1 Schield: 2019 NNN Annual Meeting 38

Associations: Snapshot vs. Movie

Two-factor covariation:
- As weight increases, height increases
- For every additional pound, height increases by a fifth of an inch.

Two interpretations:
- Snapshot: Change in focus (other people)
- Movie: Internal change (eat more pizza)

V1 Schield: 2019 NNN Annual Meeting 39

Two-Group Comparisons: Math vs. Ordinary English

Arithmetic (Assembly):
- Six is three times two. [Math speak]
- Six is three times *as much as* two. [English]
- Six is *200% more* than two.
- Six is *two times more* than two.

Two exceptions:
- 6% is 4% more than 2%?
- 6% is 200% more than 2%. [(6% - 2%)/2%]
- 6% is 4 percentage points more than 2%.

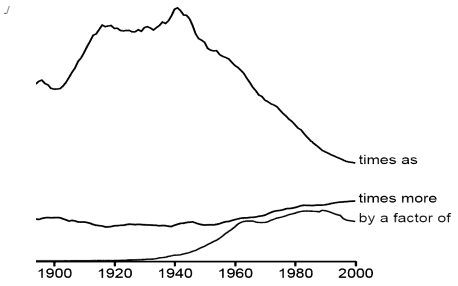
V1 Schield: 2019 NNN Annual Meeting 40

Two-Group Comparisons: Math vs. Ordinary English

Two is four times less than eight. [Sometimes]
- Ok when the subject cannot go negative: Revenues, incomes, sizes, weights, prices of houses/groceries,
- Ambiguous when it easily goes negative: Profits, temperatures, bank balances

V1 Schield: 2019 NNN Annual Meeting 41

Prevalence of Comparisons Google Ngrams



V1 Schield: 2019 NNN Annual Meeting 42

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 Schield: 2019 NNN Annual Meeting 43

#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 Schield: 2019 NNN Annual Meeting 44

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

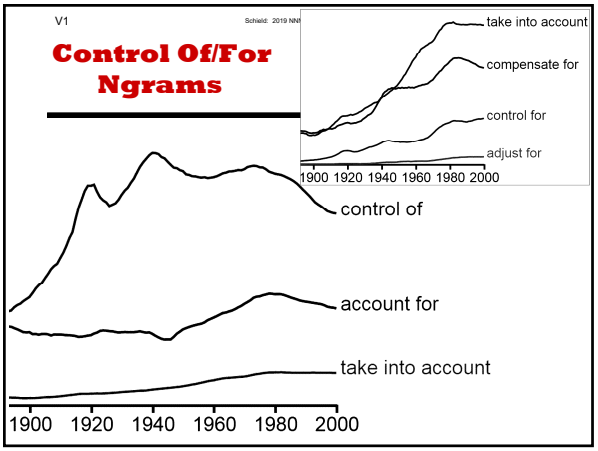
V1 Schield: 2019 NNN Annual Meeting 45

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

V1 Schield: 2019 NNN

Control Of/For Ngrams



V1 Schield: 2019 NNN Annual Meeting 47

Crude Associations

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

More likely to get pregnant: Younger adults

- that are shorter
- that don't shave daily
- That have longer hair

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

V1 Schield: 2019 NNN Annual Meeting 48

Statistical Literacy and Grammar

Grammar of comparisons

Grammar of named ratio families:
percent, percentage, rate and chance

Grammar for comparing named ratios: Likely

V1 Schield: 2019 NNN Annual Meeting 49

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*

V1 Schield: 2019 NNN Annual Meeting 50

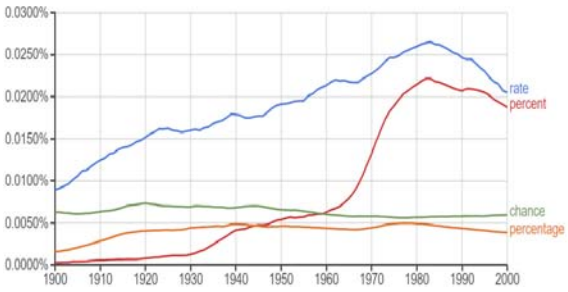
RATIOS (Using Prepositions)

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*
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)

V1 Schield: 2019 NNN Annual Meeting 51

Prevalence of Named Ratios: nGrams



V1 Schield: 2019 NNN Annual Meeting 52

Two Kinds of Percents

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

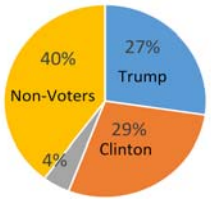
1. One child's share of the candy.
2. Lifespan 100% longer: US than Swaziland
3. Advertisement: "40% off"
4. Mafia interest rate: 10% per month

V1 Schield: 2019 NNN Annual Meeting 53

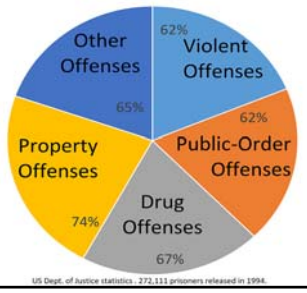
Part-Whole Using Pie Charts Valid vs. Invalid

Of all adults.

2016 US Presidential Election



Recidivism Rate: US Prisoners



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

V1 Schield: 2019 NNN Annual Meeting 54

Four Different Grammars: Percent, Percentage, Rate, Chance

1. 40% of adults did not vote [for president].
   Among adults, 40% did not vote [for President].
2. The *percentage* of adults who didn't vote was 40%
   The percentage of non-voters among adults was 40%
3. The non-voter *rate* of|for adults was 40%.
   The rate of non-voters among adults was 40%.
4. There is a 40% *chance* that an adult was a non-voter.
   Adults had a 40% chance of not voting.

Converting: From Percent to Percentage

1. 60% of adults voted.
   The percentage of adults who voted is 60%.
2. 60% of male adults voted.
   Percentage of male adults who voted is 60%.
3. 60% of adults who are men voted.
   Percentage of adults **who are men who voted** is 60%.
   [Convert #3 to #2 first.]

Confusion of the Inverse: Exchanging Part with Whole

1. Simple reversal:
   "The percentage of men who are in the military" versus
   "the percentage of the military who are men".
2. Tricky grammar reversal:
   "The percentage of smokers who are women" versus
   "The percentage of smokers among women".
3. Plausible claim, but the inverse is what is needed:
   "Most CEOs had a pet as a child" versus
   "Must children who had a pet became CEOs" or
   "Children who had a pet are more likely to become CEOs".

Conclusion

Students need a better understanding of the words and ideas involving statistics in arguments.

Statistical Literacy should be taught across the curriculum.

Learning this takes time – lots of time

Teaching this is not easy, but it is important!

Literacy is at least as important as the math!

Statistical Literacy is Quantitative Rhetoric

Deals with statistics in arguments.

Much of today's 'fake news' involves the use or misuse of statistics in arguments.

Students need statistical literacy in order to understand and evaluate the claims being made.

Students need statistical literacy to become critical thinkers in a complex modern democracy.

What is Numeracy

by
Milo Schield, NNN VP

NNN Annual Meeting
Austin CC Oct 13, 2019

www.StatLit.org/pdf/
2019-Schield-NNN-Slides.pdf

Steen (1997)
Why Numbers Count

Numeracy is the new literacy of our age.

Whatever this phrase [Q/L] may mean—and as the essays in this volume testify, it means very different things to different people.

It [Q/L] requires a working synthesis of literacy and numeracy; it evolves with technology; and it both shapes and is shaped by society.

Steen (1997)
Why Numbers Count

Regardless of name—numeracy, mathematics, quantitative literacy, or the derisive "'rithmetic," —this kind of literacy is widely recognized as of fundamental importance.

Yet beyond "the basics," there is little agreement about specific goals appropriate for tomorrow's world. No wonder, then, that we have made so little progress in achieving numeracy.

Mere Literacy is not Enough. George Cobb

Steen (1997)
Why Numbers Count

In short an *innumerate* citizen today is as vulnerable as the illiterate peasant of Gutenberg's time.

Although the widespread availability of data should enrich public discourse, inevitable over-simplifications and misinterpretations may ultimately cheapen it. ... Instead of enhancing Jeffersonian democracy, *limited numeracy* can easily shift the balance to a technocracy.

Innumeracy thus becomes another means of disenfranchisement: by reinforcing the idea that truth is relative and unknowable, people with the least defenses against charlatans will be most vulnerable.

Steen (2001)
Mathematics and Democracy

Numeracy is not so much about understanding abstract concepts as about applying elementary tools in sophisticated settings.

Numeracy is not the same as mathematics, nor is it an alternative to mathematics.

Rather, it is an equal and supporting partner in helping students learn to cope with the quantitative demands of modern society

Perspectives on *Numeracy*

NUMERACY

Quantitative Reasoning	Quantitative Literacy	Statistical Literacy	Statistical Reasoning

NUMERACY

What parts of math?

Topics Skills Models Patterns Thinking

Perspectives on *Numeracy*

NUMERACY

Which quantitative topics?

Arithmetic Algebra Geometry

NUMERACY

Audience: Which Majors???

Math/ STEM	Social Science	Professions Biz/Econ/etc	Humanities English, Art

Perspectives on *Numeracy*

NUMERACY

Citizens

High School College Grads

NUMERACY

Studies the quantitative skills needed by educated citizens in a modern democracy.

What is the Most Important Thing to Know about Statistics?

"All statistics are socially constructed."
Joel Best, author *Lies, Damned Lies & Statistics*

Doesn't mean that

- there is no reality, all statistics are imaginary

Means that people create statistics like diamonds

What is the Most Important Thing to Know about Statistics?

Statistics are socially constructed: the products of social activities.

There's a tendency in our culture to believe that statistics—that numbers—are little nuggets of truth.

That we can come upon them and pick them up very much the way a rock collector picks up stones.

What is the Most Important Thing to Know about Statistics?

A better metaphor would be to suggest that statistics are like jewels; that is,

they have to be selected, they have to be cut, they have to be polished, and they have to be placed in settings so that they can be viewed from particular angles.

Joel Best, Sociologist

What follows from being Socially Constructed?

Numbers can't be influenced. 1 + 1 = 2

Statistics can be influenced. 1 + 1 may equal 2

One gallon of antifreeze and one gallon of water do not yield two gallons. The combination of large and small molecules takes up less space.

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 typically evidence of causation.

A-B-C Words: A = Association

Statistical association: an observable connection.

Association:

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

Prediction:

- Graduating from high school *predicts* success in life.

A-B-C Words: C = Causation

Causation: Lightning *caused* (*resulted in*) the fire.
Insomnia is a side *effect*.

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*.

Schield and Raymond (2009). www.StatLit.org/pdf/2009SchieldRaymondASA.pdf

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 increased *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

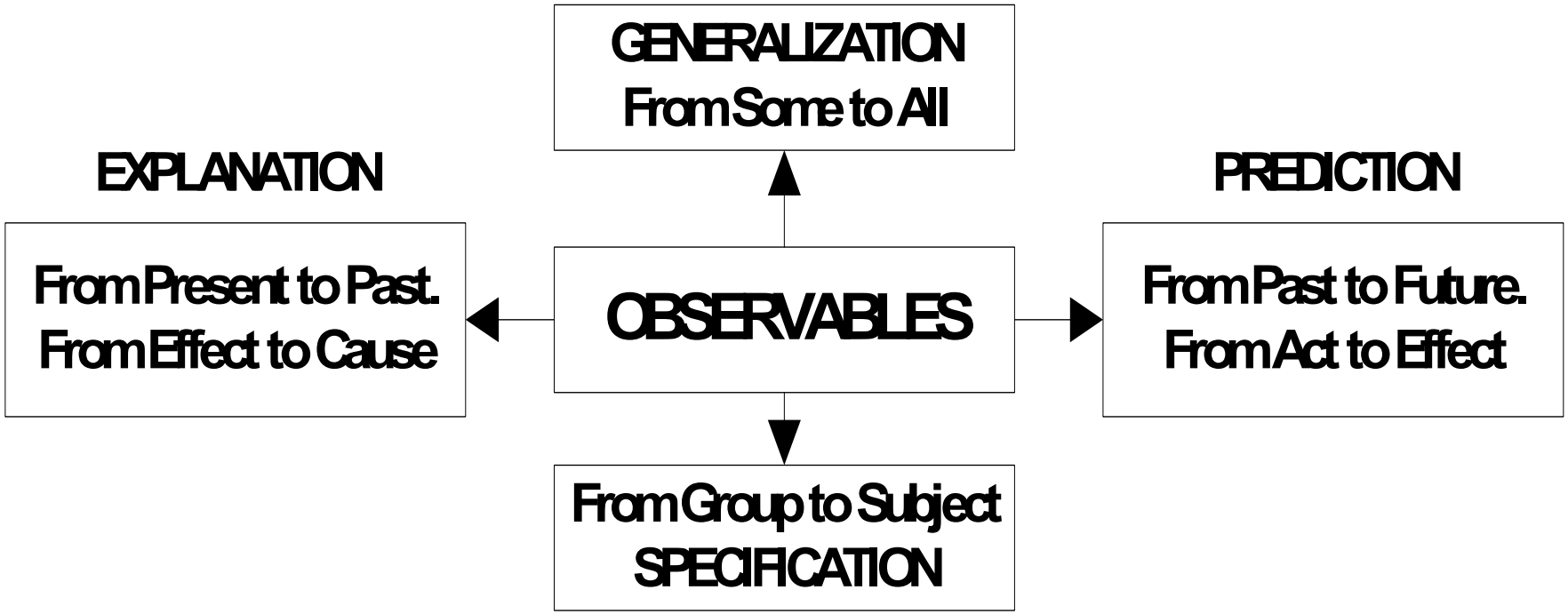
Of the 2,000 news headlines analyzed[6],
71% involved A, B or C.

Of those headlines involving A, B or C,

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

6. Schield and Raymond (2009).

Critical Thinking Inferences



Association is Not Causation.

Prediction: Its hopeless! This bear is faster than we are

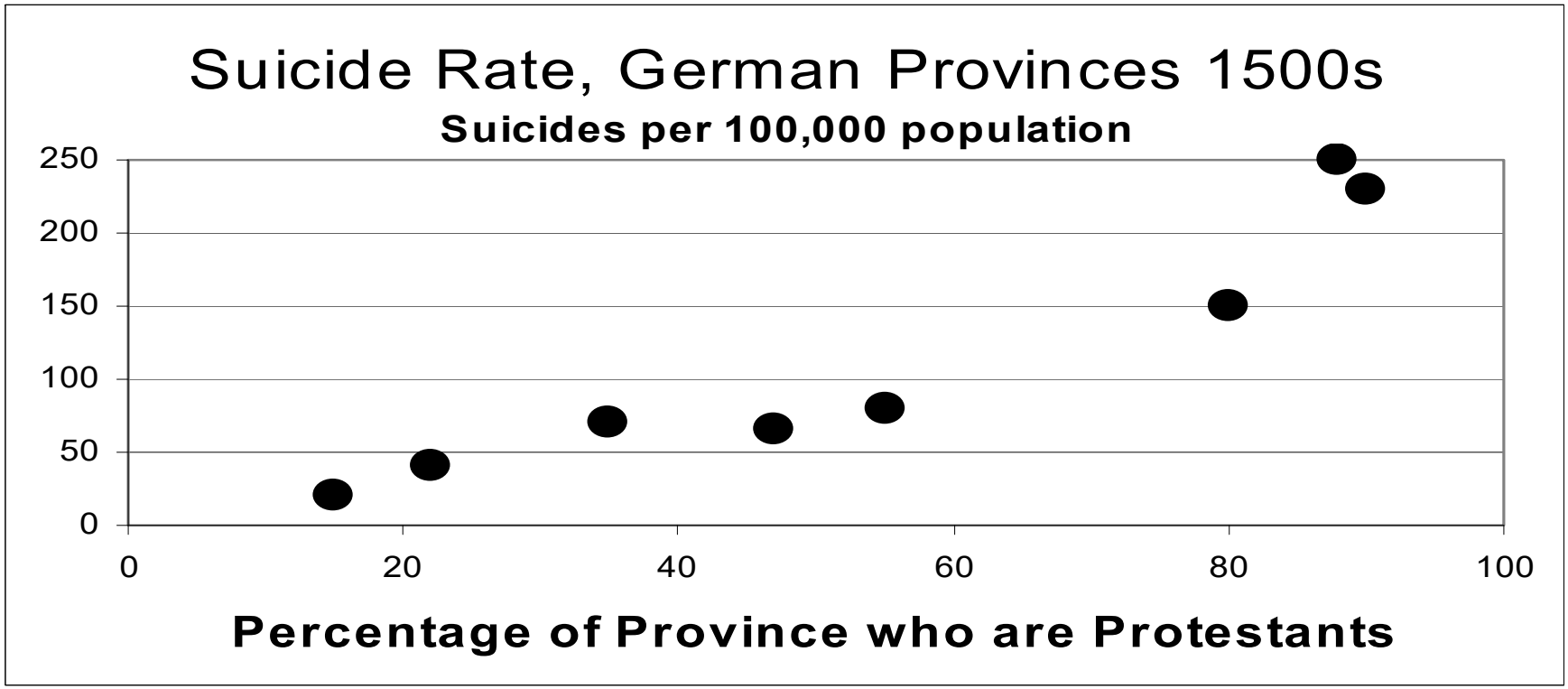
2nd hunter: No, it's not hopeless.

I don't have to outrun the bear.

I just have to outrun you!

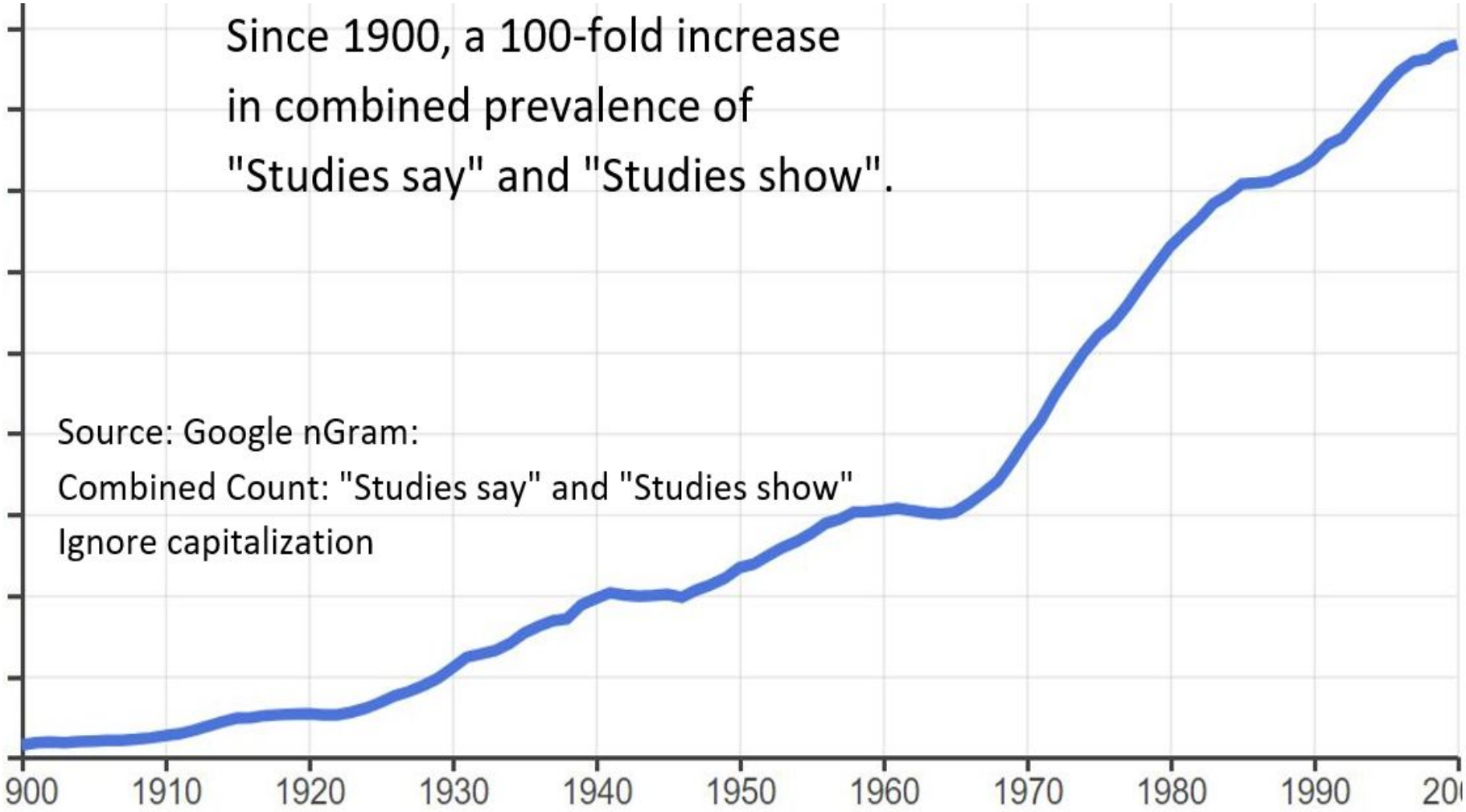
Specification: Cross-Level Inference

Protestants more likely to suicide than Catholics?

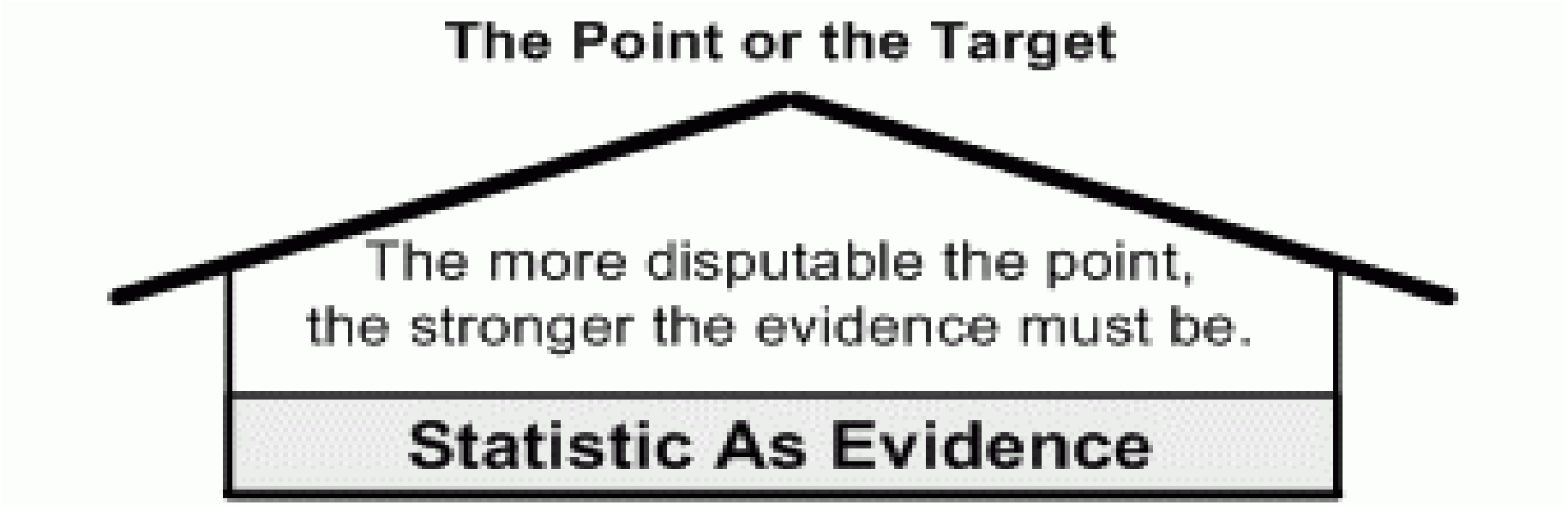


No. Reverse was true here. Catholics more likely.

Statistical Studies as Evidence in Arguments



Statistical Literacy
Studies Statistics in Arguments



"All Statistics are Socially Constructed"
So, "Take CARE"!!
Statistics may be influenced by:

C	A	R	E
Context	Assembly	Randomness	Error

Take C.A.R.E
Four Influences on Statistics

Statistics are influenced by

C = *Confounding*: By related factors.

A = *Assembly:* By definitions and presentation.

R = *Randomness:* By uncertainty or chance

E = *Error:* By mistakes or bias.

Assembly is the etcetera category.

Confounding Without Statistics

A father and his children were on a subway.

The children were out of control: jumping on seats, yelling, and throwing things.

The father did nothing.

He slumped forward looking down at the floor, his head between his hands.

Finally an unhappy onlooker called on the father to take control of his kids.

Confounding Without Statistics

The father looked up sadly and said:

We just left the hospital where their mom died.

Immediately the negative judgments were transformed into pity for this family.

The onlookers were confused – confounded – by a confounder: the death of the kids' mom.

Statistical Influences: Confounding #1

Adults who shave their faces tend to be taller than those who shave their legs.

Does face-shaving cause tallness? ___
Is this association confusing? ___
Is it confused by some outside factor? ____.
What is it? ____

Gender. Men tend to be taller than women.
Men are more likely to shave their face.

Statistical Influences: Confounding #2

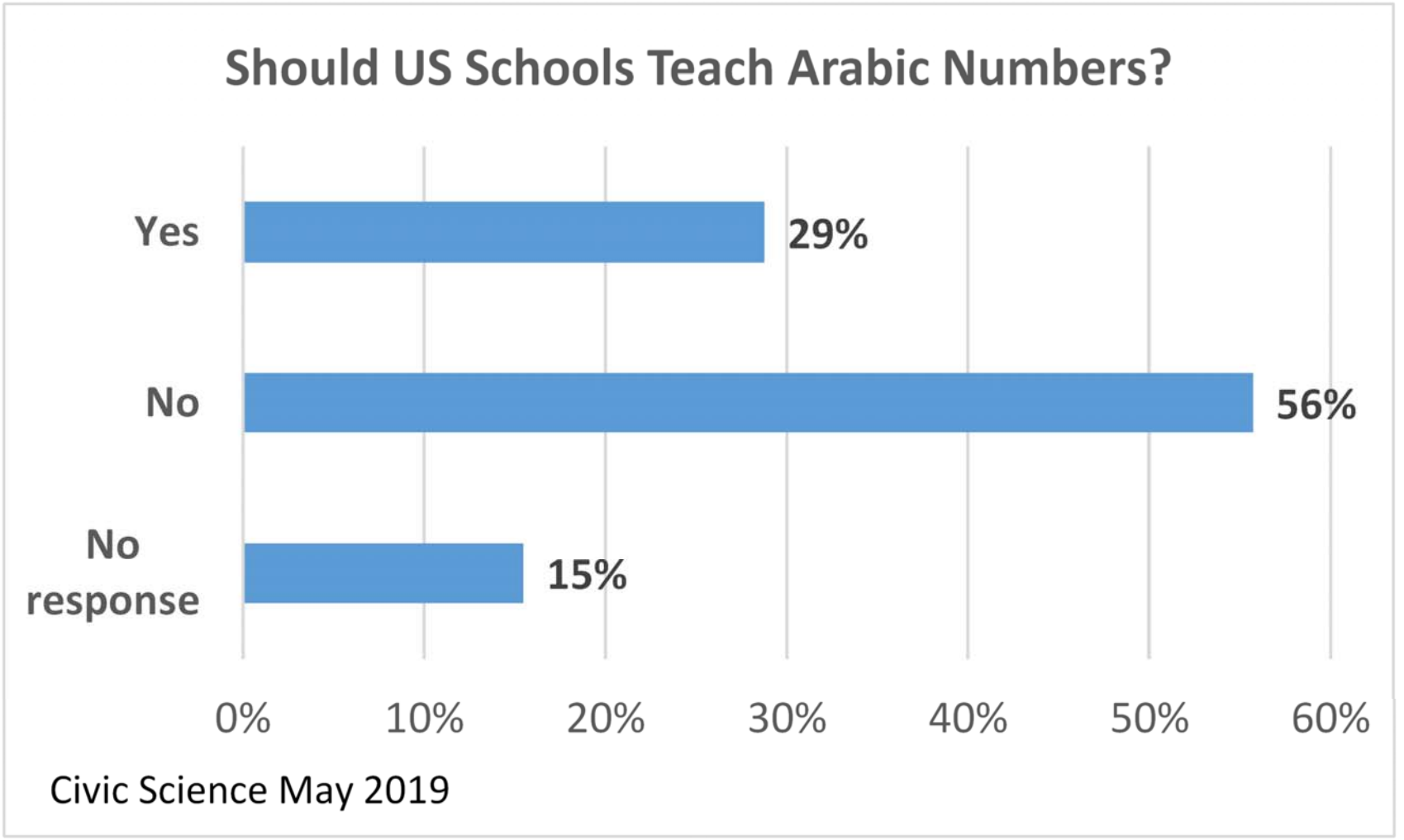
People that read home and fashion magazines are more likely to get pregnant than people that read car and sport magazines.

Are the magazines causing pregnancy? ___
Is the association confused by an outside factor?
What is this outside factor? _____

Gender! Women can get pregnant (men can't)
Gals more likely to read home/fashion magazines

Assembly Fueled Brexit?
Gross vs. Net (50%):

Assembly:
Sharia math, then Sharia law!!!



Statistical Influences : Assembly #1

Claims about college students:

Administration: 80% of are 'satisfied'.
Students: 70% are not 'satisfied'.

Same data:
Happy (30%), OK (50%), Unhappy (20%)

Q. Who is correct?
A. Both are. Different definitions of 'satisfied'.

Statistical Influences : Assembly #2

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

Two claims about groups living with AIDS:

1. More blacks than whites.
2. More whites than blacks.

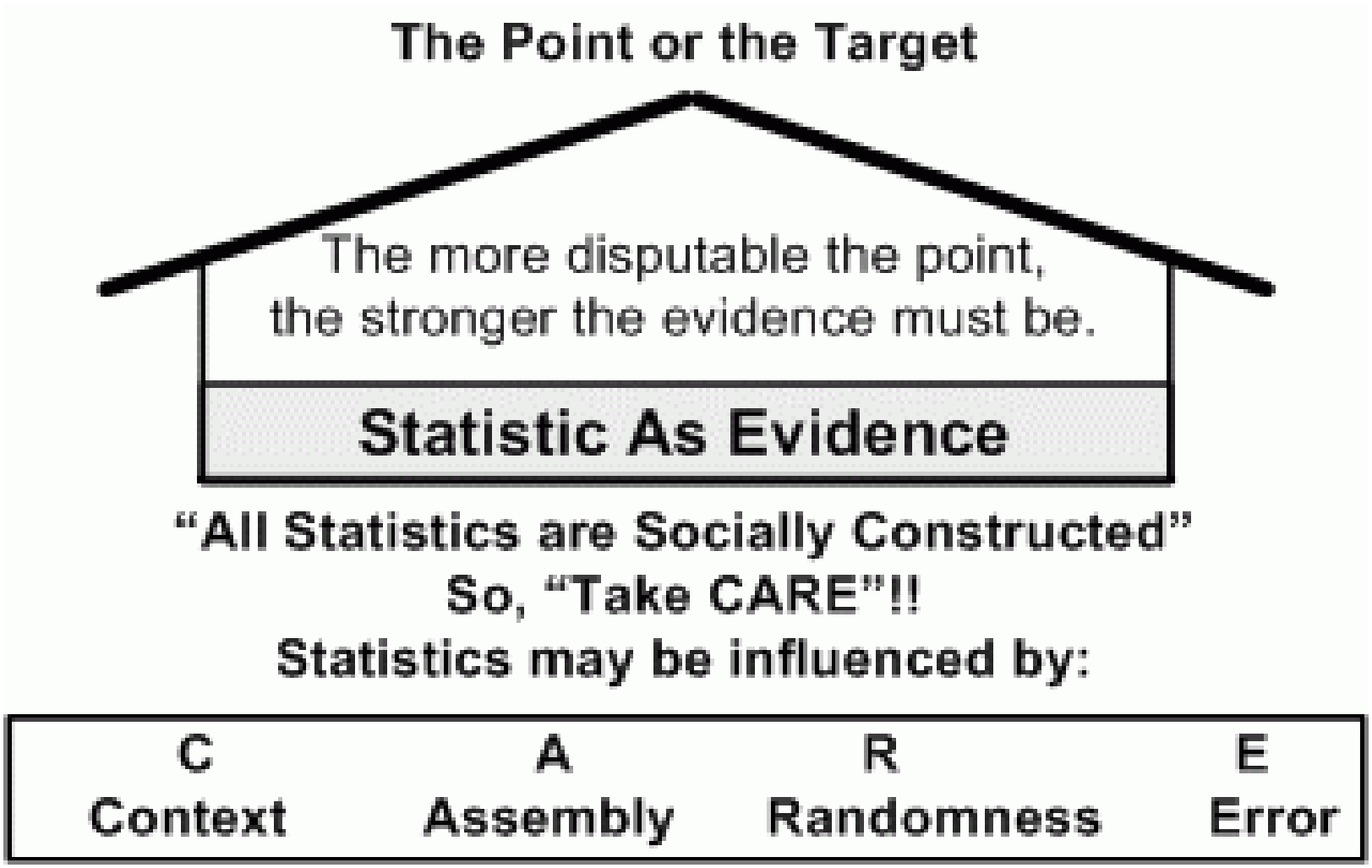
Assembly: Presentation

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.

Stat Literacy studies Stats as Evidence in Arguments

./



Stat Influences : Randomness

Sport Illustrated Curse

Those people featured as *best* by Sports Illustrated seem cursed!

They don't do as well the next year. Why not?

Being the best one year is often coincidence.

Statistical Literacy deals with Ignorance

How was data collected?
C: What factors not taken into account?
A: How were things counted, measured, grouped?
R: How small is the group?
E: Are subjects telling the truth?

Often we don't know! We are ignorant!
We are not omniscient.

The solution? Think hypothetically! Plausible?

Controlling Confounding: 'Control Of' vs. 'Control for'

1. Study design indicates "control of"
2. Comparisons and ratios indicate "control for"

Control for:

1. Comparisons control for a relevant basis
2. Ratios control for size of group
3. Comparisons of ratios control for both

Associations: Two Kinds

Two-group comparisons:

- Women live longer than men

Two-factor covariation: ordinal & quantitative

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

- As height increases by *x*, weight increases by *y*
- For every additional *x* in X, Y increases by *y*.

Associations: Snapshot vs. Movie

Two-factor covariation:

- As weight increases, height increases
- For every additional pound, height increases by a fifth of an inch.

Two interpretations:

- Snapshot: Change in focus (other people)
- Movie: Internal change (eat more pizza)

Two-Group Comparisons: Math vs. Ordinary English

Arithmetic (Assembly):

- Six is three times two. [Math speak]
- Six is three times *as much as* two. [English]
- Six is *200% more* than two.
- Six is *two times more* than two.

Two exceptions:

- 6% is 4% more than 2%?
- 6% is 200% more than 2%. [(6% - 2%)/2%]
- 6% is 4 percentage points more than 2%.

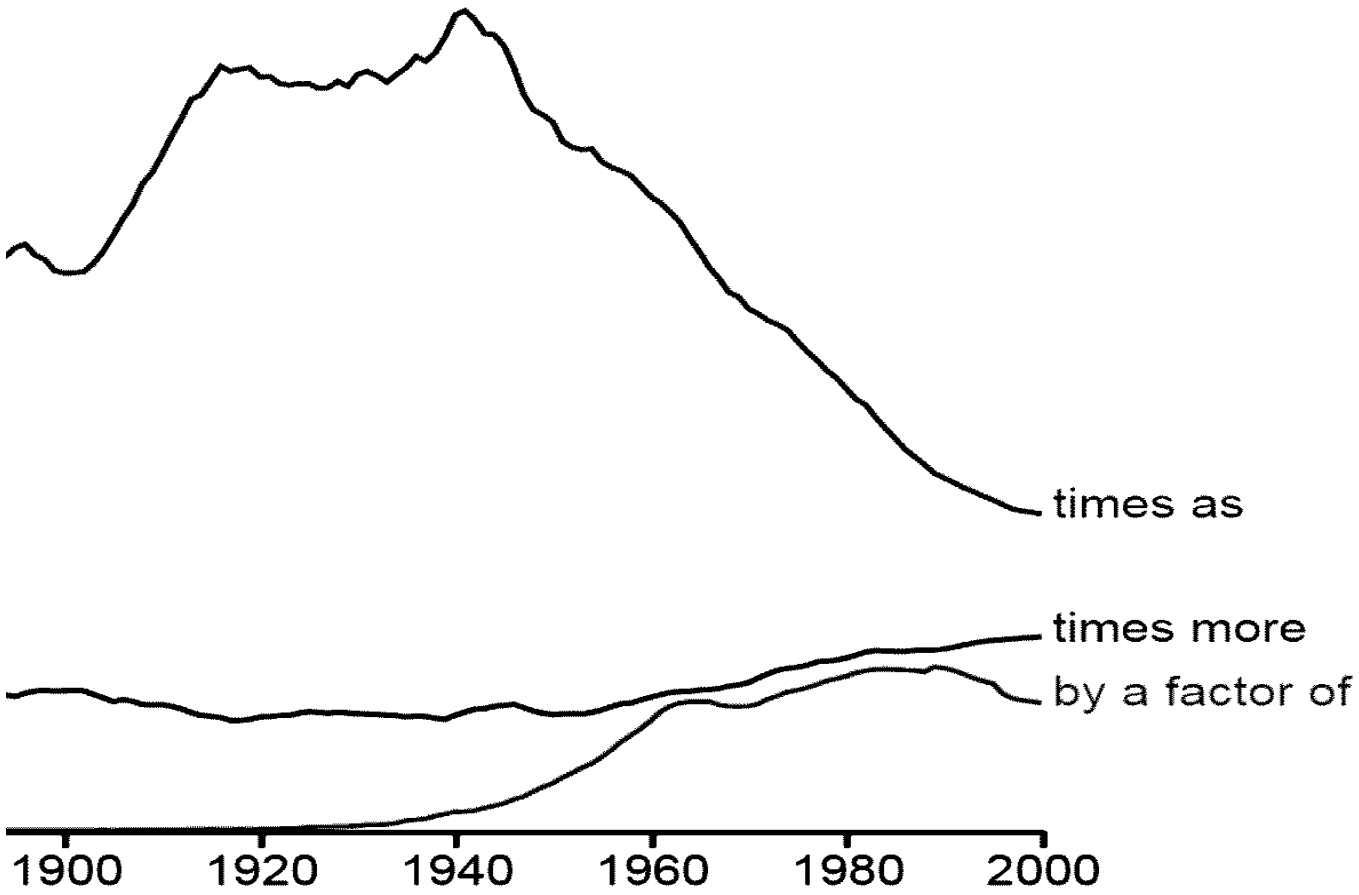
Two-Group Comparisons: Math vs. Ordinary English

Two is four times less than eight. [Sometimes]

- Ok when the subject cannot go negative: Revenues, incomes, sizes, weights, prices of houses/groceries,
- Ambiguous when it easily goes negative: Profits, temperatures, bank balances

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).

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

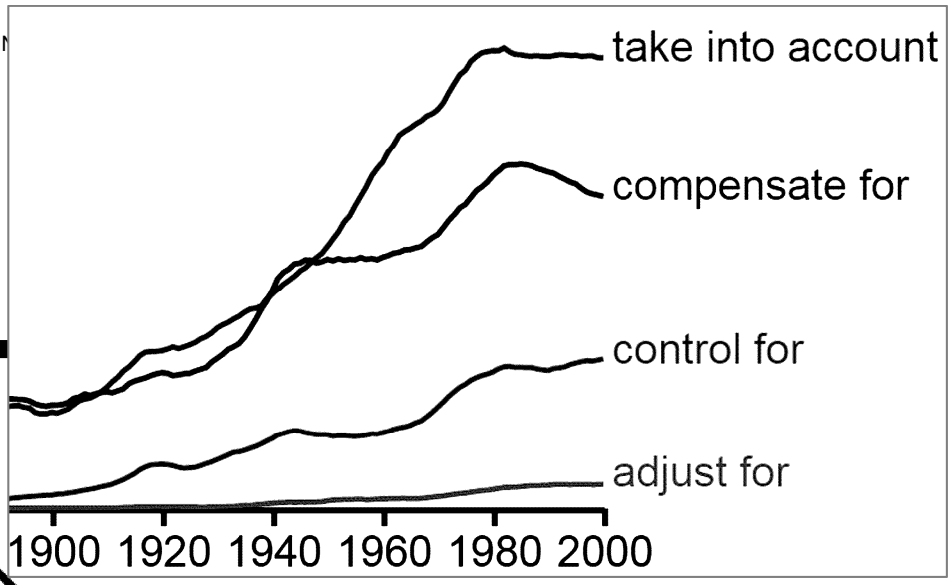
Controlling Confounding: Control For

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

V1

Schield: 2019 NNM

Control Of/For Ngrams



control of

account for

take into account

1900 1920 1940 1960 1980 2000

Crude Associations

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

More likely to get pregnant: Younger adults

- that are shorter
- that don't shave daily
- That have longer hair

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

Statistical Literacy and Grammar

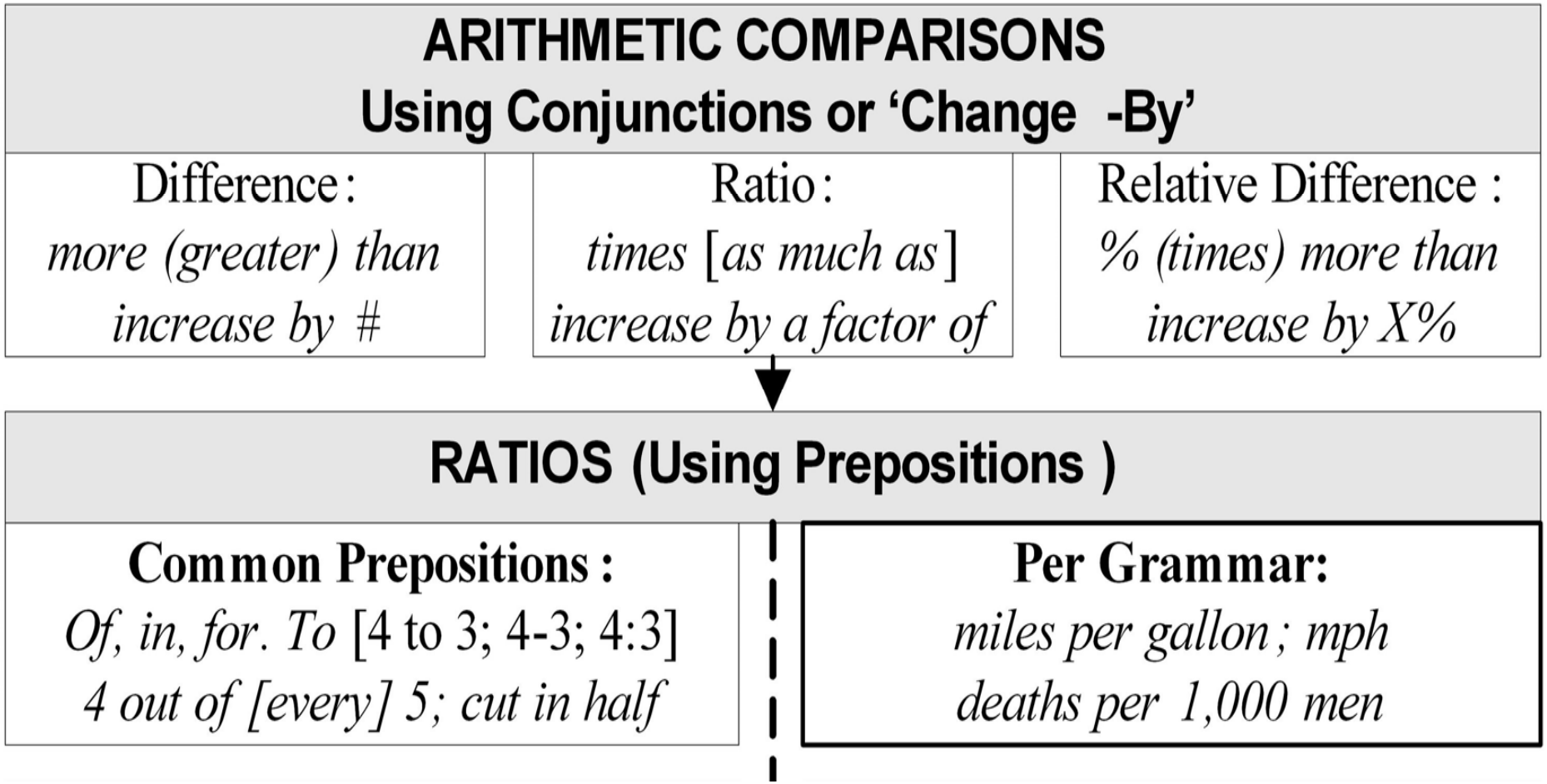
Grammar of comparisons

Grammar of named ratio families:

percent, percentage, rate and chance

Grammar for comparing named ratios: Likely

From Comparisons to Ratios: Using Prepositions



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

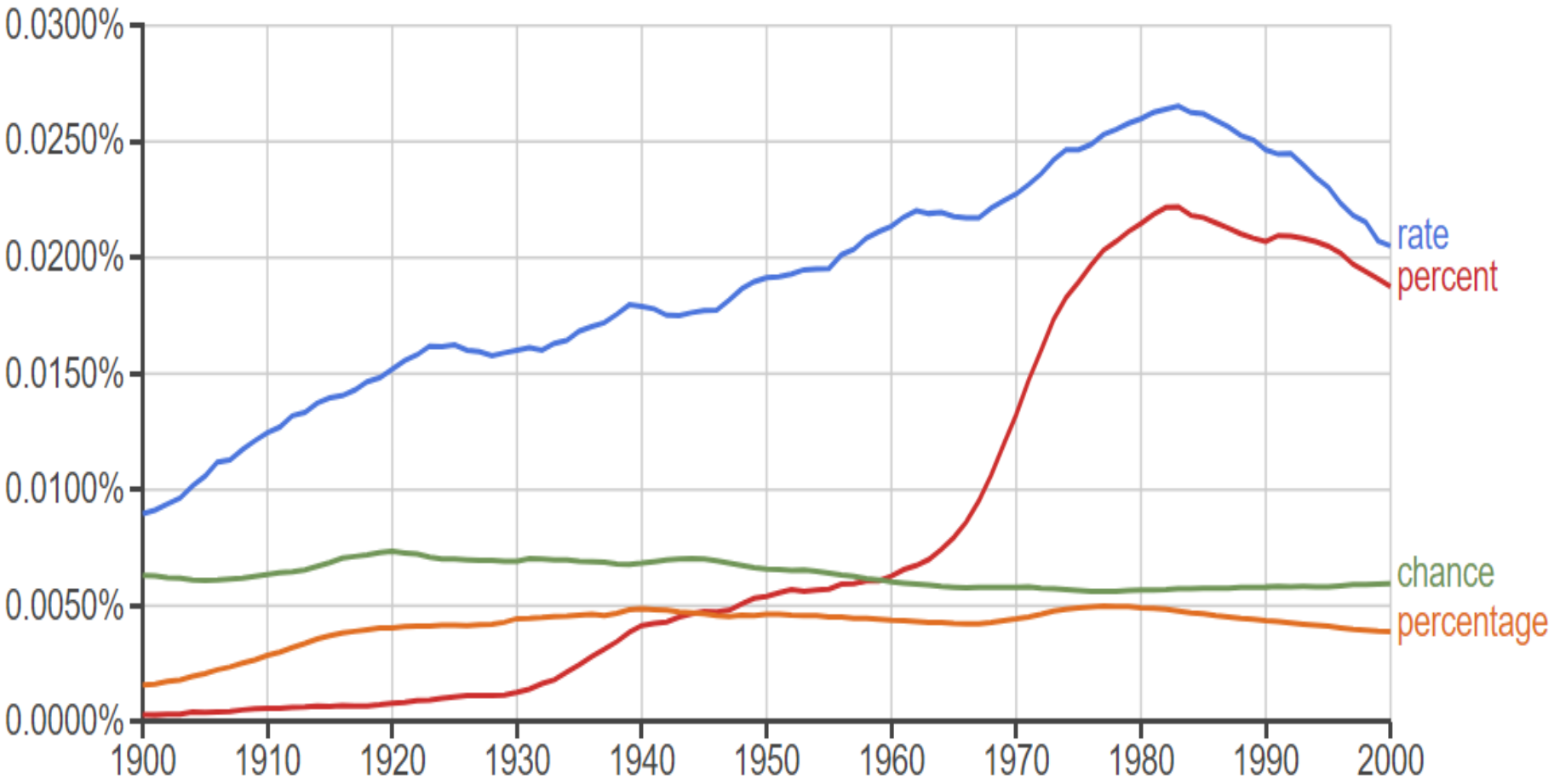
Percentage Grammar: fraction/share
percentage of men who bet

Chance Grammar: odds/risk/probability
chance of [our] winning;
chance that we will win
chance to win ; chance for a win

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: nGrams



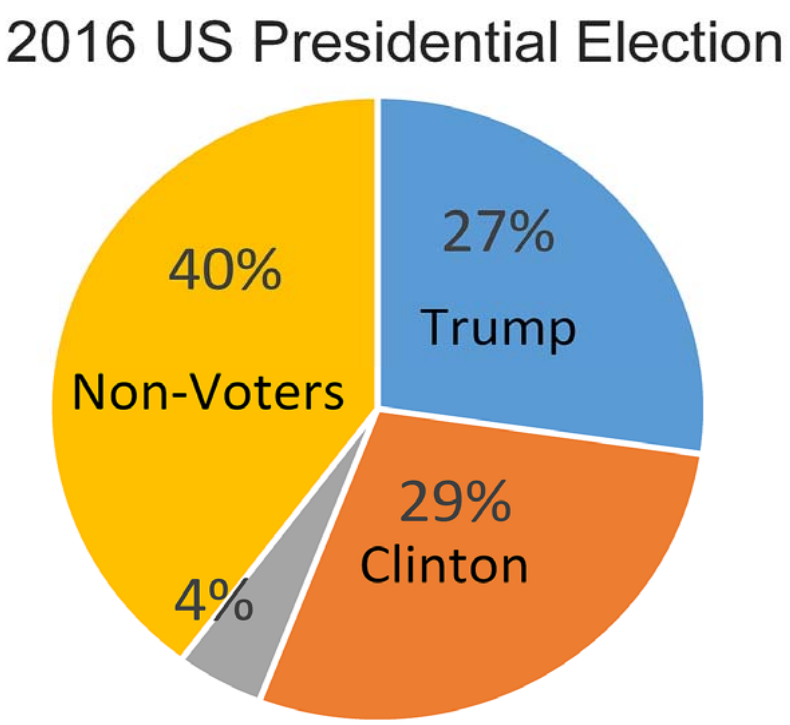
Two Kinds of Percents

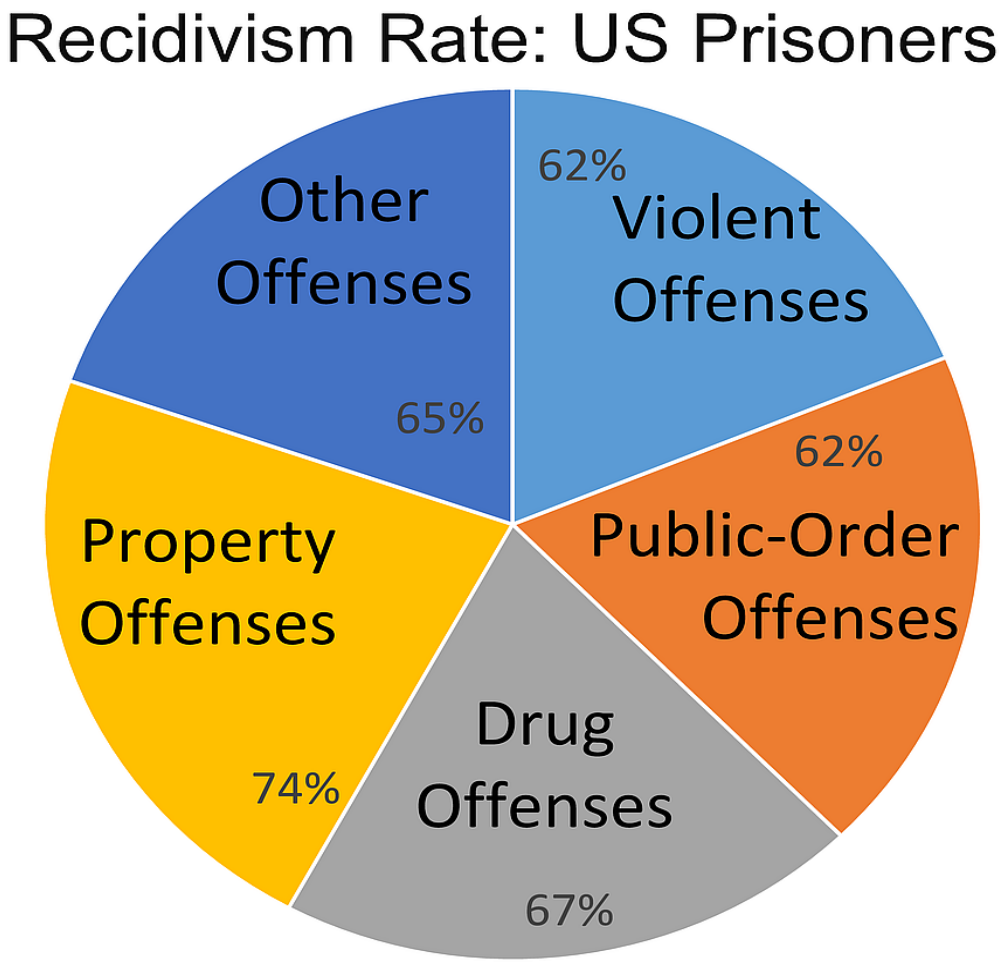
Which kind of percents are these:
part-whole or percent compare?

1. One child's share of the candy.
2. Lifespan 100% longer: US than Swaziland
3. Advertisement: "40% off“
4. Mafia interest rate: 10% per month

Part-Whole Using Pie Charts
Valid vs. Invalid

Of all adults.





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

Four Different Grammars: Percent, Percentage, Rate, Chance

1. 40% of adults did not vote [for president].
   Among adults, 40% did not vote [for President].

2. The *percentage* of adults who didn't vote was 40%
   The percentage of non-voters among adults was 40%

3. The non-voter *rate* of|for adults was 40%.
   The rate of non-voters among adults was 40%.

4. There is a 40% *chance* that an adult was a non-voter.
   Adults had a 40% chance of not voting.

Converting:
From Percent to Percentage

1. 60% of adults voted.
   The percentage of adults who voted is 60%.

2. 60% of male adults voted.
   Percentage of male adults who voted is 60%.

3. 60% of adults who are men voted.
   Percentage of adults **who are men who voted** is 60%.
   [Convert #3 to #2 first.]

Confusion of the Inverse: Exchanging Part with Whole

1. Simple reversal:
   "The percentage of men who are in the military" versus
   "the percentage of the military who are men".

2. Tricky grammar reversal:
   "The percentage of smokers who are women" versus
   "The percentage of smokers among women".

3. Plausible claim, but the inverse is what is needed:
   "Most CEOs had a pet as a child" versus
   "Must children who had a pet became CEOs" or
   "Children who had a pet are more likely to become CEOs".

Conclusion

Students need a better understanding of the words and ideas involving statistics in arguments.

Statistical Literacy should be taught across the curriculum.

Learning this takes time – lots of time

Teaching this is not easy, but it is important!

Literacy is at least as important as the math!

Statistical Literacy is Quantitative Rhetoric

Deals with statistics in arguments.

Much of today's 'fake news' involves the use or misuse of statistics in arguments.

Students need statistical literacy in order to understand and evaluate the claims being made.

Students need statistical literacy to become critical thinkers in a complex modern democracy.