

August 2011 IASE 1

All Statistics are Socially Constructed

MILO SCHIELD,
Augsburg College

*US Rep: International Statistical Literacy Project
VP National Numeracy Network
Webmaster [www. StatLit.org](http://www.StatLit.org)*

18 August 2011
IASE Malahide, Ireland
www.StatLit.org/pdf/2011-Schield-IASE-Slides.pdf

August 2011 IASE 2

Statistical Literacy courses: The Next Big Thing?

2009: Survey of all US four-year colleges found that 19% of respondents offered a statistical literacy course.

2010: Wired Magazine said, "*Statistical literacy: A course you should have taken in college.*"

2011: US Supreme Court finds statistical significance is not necessary for causation.

2011: Eight college faculty completed the Augsburg College statistical-literacy teacher-training course online. These teachers analyzed 14 news stories.

August 2011 IASE 3

#1 All Statistics are Socially Constructed

"All statistics, from the best to the worst, are socially constructed.

All statistics are products of choices and compromises that inevitably shape, limit, and distort the outcome."

Joel Best (2002)

August 2011 IASE 4

#2 Four Kinds of Influence on Statistics Take CARE

Context: Influence of factors that were – and were not – taken into account (multivariate analysis)

Assembly: Influence of choices in defining, measuring, comparing and presenting statistics

Randomness: Influence of chance based on choices in type and size of sample

Error (Bias): Influence of factors generating systematic differences between observed and reality.

August 2011 IASE 5

Statistical-Inference Courses Cover Randomness and Error

Statistical educators generally teach that statistics can be influenced by human choices:

Randomness: size and type of sample (random vs. convenience), sampling distribution, etc.

Error/bias: Choices that produce *subject-response bias, measurement bias and sampling bias*. E.g., the target population, the sampled population, the sampling method, the handling of non-responses

August 2011 IASE 6

#3. Statistics Are Influenced by "Context"

Statistical educators agree that statistics are numbers in context, but they don't generally teach that statistics can be influenced by human choices involving "context": choices in

- what to take into account using tables, series, ratios, comparisons and comparisons of ratios
- how to model data
- what factors not taken into account (potential confounders)

August 2011 IASE 7

Hyatt: Close to the US Capital



August 2011 IASE 8

Loudest Animal on Earth



Pond insect 'loudest animal on Earth'

A tiny 'water boatman' insect is the world's loudest animal relative to its body size, according to a new study.

August 2011 IASE 9

Two Per Cent Milk



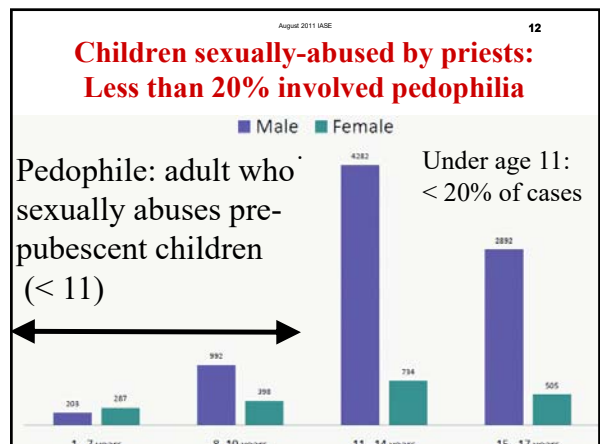
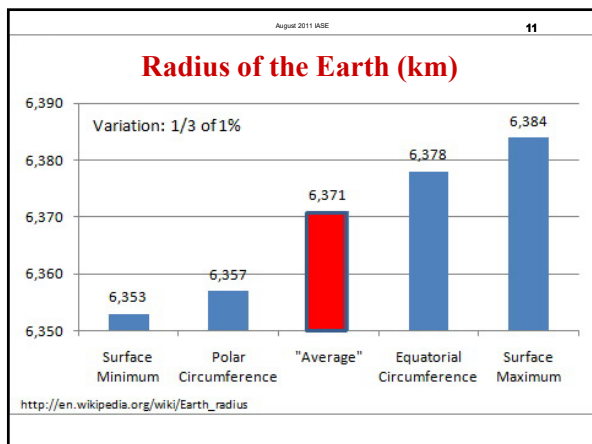
August 2011 IASE 10

#4 Statistics are Influenced by "Assembly"

Statistical educators don't generally teach that statistics can be influenced by human choices in "Assembly": choices in

- definitions of groups/measures*
- comparisons (# more vs. % more)
- presentation (pie chart vs. bar graph)

* aside from mean vs. median...



August 2011 IASE 13

Pure Assembly!!



ON TEENAGERS, ADULT:
Statistics show that
 teen pregnancy
 drops off significantly
 after age 25.
*Mary Anne Tebeo, Republican state senator from Colorado Springs
 (contributed by Harry F. Punsce)*
MONDAY DECEMBER 1999

August 2011 IASE 14

How to Teach Context and Assembly?

Must involve problems that can be assessed.

Augsburg's Statistical Literacy course has created:

- over 300 involving *Context*
- over 100 exercises involving *Assembly*

Consider ways to teach assembly using definitions

August 2011 IASE 15

**Definitions of statistics
 Can be Classified on a Spectrum**

Two Extremes for Comparisons:

Formal: Need no knowledge of ideas

- Fewer tall smokers than smokers...

Material: Need detailed knowledge of ideas

- More autistic boys than autistic girls...

In between: Need some knowledge of ideas

- Men are taller than women on average.
- Women live longer than men on average.

August 2011 IASE 16

**Statistical educators should teach
 definitions as part of study design**

- **Should teach formal comparisons of operational definitions.** Little knowledge of ideas is needed. Like logic/math.
- **Might delay teaching material comparisons of operational definitions.** Requires a detailed knowledge of specialized ideas. Leave this to SMEs.
- **Could selectively teach in-between comparisons of operational definitions.** At least show sensitivity of some statistics to small changes in definition.

August 2011 IASE 17

**Formal Compare: Operational Definitions
 Counts: Criteria-based**

Examples: Which count/total is larger?

- US population; US male population?
- Not employed; Not employed *and* looking for work?
- Use of physical force ; Use *or* threat of physical force?

Principles involving counts or totals:

- *And* phrases and modifiers restrict options – smaller total
- *Or* phrases increase options – larger total

August 2011 IASE 18

**Formal Compare: Operational Definitions
 Counts: Frequency-based**

Examples: Which count is bigger?

- Smoker: smokes *more than* five cigarettes a day
- Smoker: smokes *more than* two cigarettes a day
- Sober: consume *less than* 3 drinks in past week
- Sober: consume *less than* 1 drink in past week

Principles involving ranges:

- *More than X.* Larger X gives smaller count.
- *Less than X.* Larger X gives bigger count.

August 2011 IASE 19

Formal Compare: Operational Definitions
Counts: Duration-based

Examples: Which count is bigger?

- 1a. Smoker: smoked in the past 2 weeks
- 1b. Smoker: smoked in the past 4 weeks
- 2a. Sober: No alcohol in past 2 weeks
- 2b. Sober: No alcohol in past 4 weeks

Principles involving time periods:

- Event in past X periods: Bigger X, bigger number.
- Event-free in past X periods: Bigger X, smaller #.

August 2011 IASE 20

Formal Compare: Operational Definitions
Ratios: Compare Numerators

Which ratio is larger in each group?

1. Percentage of US citizens who are adults; who are adult males; who are adults or are males?
2. The US death rate due to suicide; the US death rate; the US death or emigration rate

Principles: Changes in the Numerator:

And phrases and modifiers restrict – smaller ratios

Or phrases increase options – larger ratios

August 2011 IASE 21

Between Compare: Operational Definitions
Part-Whole Ratios

General: Minimal knowledge of ideas is needed

Examples: Which ratio is bigger?

1. US birth rate: per 1,000 adults or per 1,000 women?
- 2a. Percentage of US males who are in the military
- 2b. Percentage of the US military who are male
- 3a. Percentage of US oil **imports** that are from OPEC
- 3b. Percentage of US oil **usage** that is from OPEC

August 2011 IASE 22

Study Between Operational Definitions:
Part-Whole Ratios

General: Minimal knowledge of ideas is needed

Examples: Which ratio is bigger?

- 1a. Percentage of auto accidents that involve a death
- 1b. Percentage of auto death that involve an accident
- 2a. Accidental death rate per 1,000 US males
- 2b. Male-death rate per 1,000 US accidents
- 3a. Percentage of US adult prisoners who have low IQ
- 3b. Percentage of US low-IQ adults who are in prison

August 2011 IASE 23

Recommendation

Statistical educators must go beyond chance and bias in analyzing the influences on a statistic.

We should encourage *critical thinking* about the choices that could have been made in creating statistics and how sensitive the statistics are to those choices.

We should start with the simple question: “*Where do statistics come from?*” Once people realize that all statistics are constructs, socially-constructed tools, then we can teach statistical literacy as a liberal art (an inductive activity) rather than as a mathematical skill.

August 2011 IASE 24

References

Joel Best Articles Page: www.StatLit.org/Best.htm

Schild, Milo (2007). Teaching the Social Construction of Statistics, 2007 Midwest Sociological Society, Chicago. See www.StatLit.org/pdf/2007SchildMSS.pdf

Schild, Milo (2010). Assessing Statistical Literacy: TAKE CARE in *Assessment Methods in Statistical Education: An International Perspective*. Edited by P. Bidgood, N. Hunt and F. Joliffe. Wiley Publishers, Ch. 11, p. 133-152.

All Statistics are Socially Constructed

MILO SCHIELD,
Augsburg College

US Rep: International Statistical Literacy Project

VP National Numeracy Network

Webmaster [www. StatLit.org](http://www.StatLit.org)

18 August 2011

IASE Malahide, Ireland

www.StatLit.org/pdf/2011-Schield-IASE-Slides.pdf

Statistical Literacy courses: The Next Big Thing?

2009: Survey of all US four-year colleges found that 19% of respondents offered a statistical literacy course.

2010: Wired Magazine said, "*Statistical literacy: A course you should have taken in college.*"

2011: US Supreme Court finds statistical significance is not necessary for causation.

2011: Eight college faculty completed the Augsburg College statistical-literacy teacher-training course online. These teachers analyzed 14 news stories.

#1 All Statistics are Socially Constructed

“All statistics, from the best to the worst, are socially constructed.

All statistics are products of choices and compromises that inevitably shape, limit, and distort the outcome.”

Joel Best (2002)

#2 Four Kinds of Influence on Statistics

Take CARE

Context: Influence of factors that were – and were not – taken into account (multivariate analysis)

Assembly: Influence of choices in defining, measuring, comparing and presenting statistics

Randomness: Influence of chance based on choices in type and size of sample

Error (Bias): Influence of factors generating systematic differences between observed and reality.

Statistical-Inference Courses Cover Randomness and Error

Statistical educators generally teach that statistics can be influenced by human choices:

Randomness: size and type of sample (random vs. convenience), sampling distribution, etc.

Error/bias: Choices that produce *subject-response bias, measurement bias and sampling bias*. E.g., the target population, the sampled population, the sampling method, the handling of non-responses

#3. Statistics Are Influenced by “Context”

Statistical educators agree that statistics are numbers in context, but they don't generally teach that statistics can be influenced by human choices involving “context”: choices in

- what to take into account using tables, series, ratios, comparisons and comparisons of ratios
- how to model data
- what factors not taken into account (potential confounders)

Hyatt: Close to the US Capital



COURTESY OYSTER.COM

Loudest Animal on Earth



Pond insect 'loudest animal on Earth'

A tiny 'water boatman' insect is the world's loudest animal relative to its body size, according to a new study.



Two Per Cent Milk

Fat by Weight

$$2\% = \frac{5}{244}$$

Nutrition Facts	
Serving Size 244 g	
Amount Per Serving	
Calories 122	Calories from Fat 43
% Daily Value*	
Total Fat 5g	7%
Saturated Fat 3g	15%
Trans Fat	
Cholesterol 20mg	7%
Sodium 100mg	4%
Total Carbohydrate 12g	4%
Dietary Fiber 0g	0%
Sugars 12g	
Protein 8g	
Vitamin A 9%	Vitamin C 1%

Fat by Calories

$$\frac{43}{122} = 35\%$$

Fat by Daily Value

$$7\% = \frac{43}{65 \times 9}$$

Divide 43 calories from fat by daily fat calories allowed = 65 grams of fat times 9 calories per gram of fat

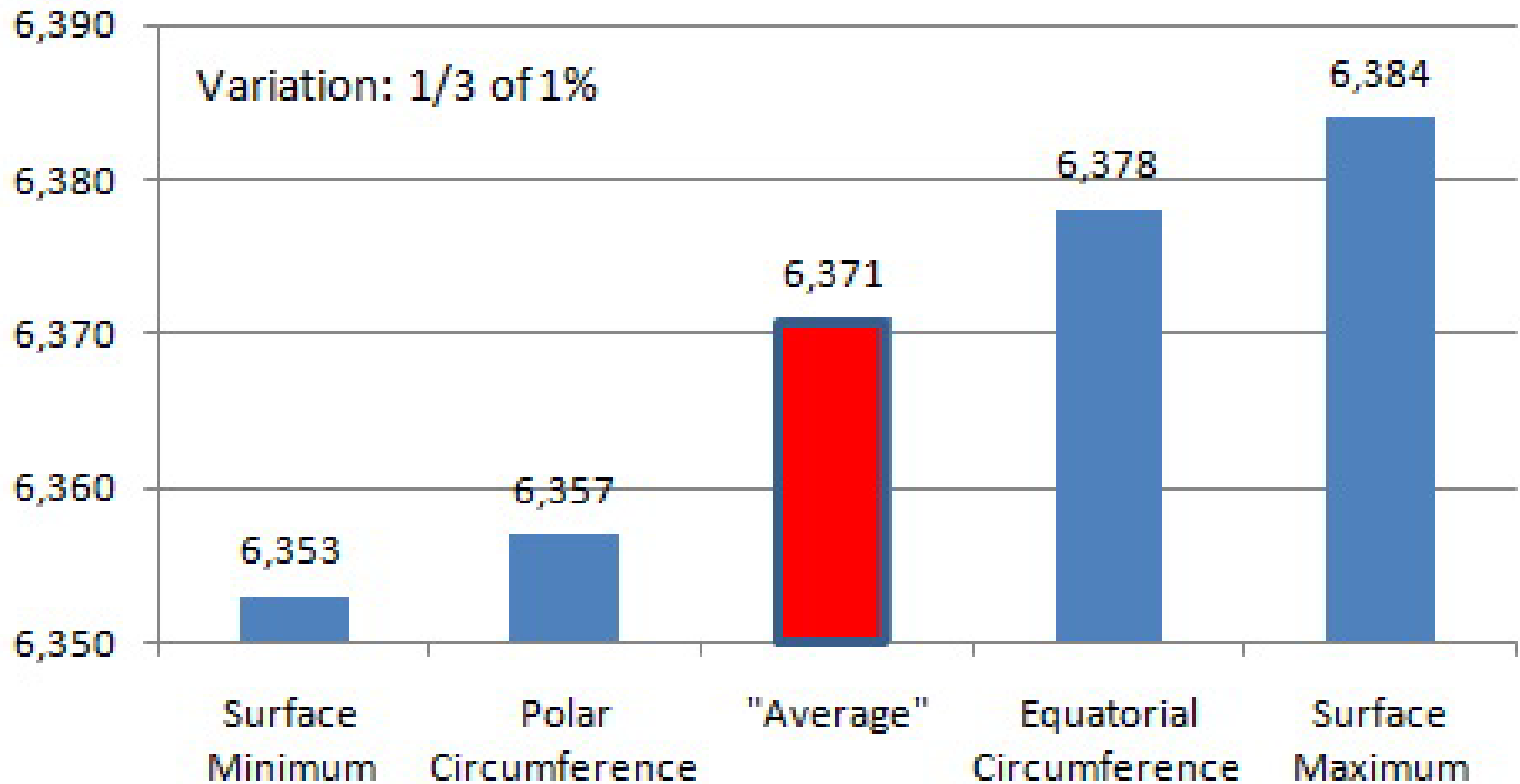
#4 Statistics are Influenced by “Assembly”

Statistical educators don't generally teach that statistics can be influenced by human choices in “Assembly”: choices in

- definitions of groups/measures*
- comparisons (# more vs. % more)
- presentation (pie chart vs. bar graph)

* aside from mean vs. median...

Radius of the Earth (km)



http://en.wikipedia.org/wiki/Earth_radius

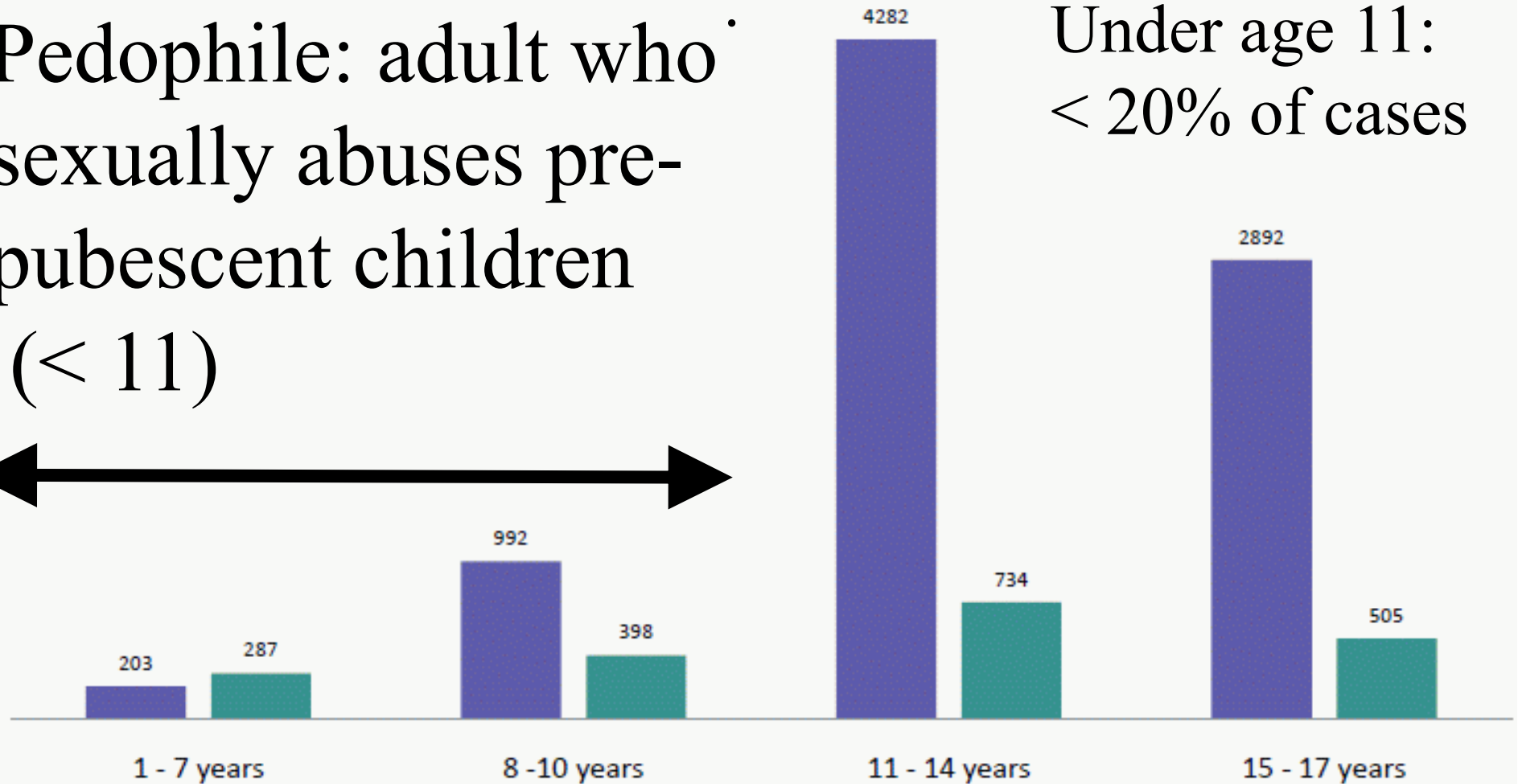
Children sexually-abused by priests: Less than 20% involved pedophilia

Pedophile: adult who sexually abuses pre-pubescent children (< 11)



Under age 11:
< 20% of cases

■ Male ■ Female



Pure Assembly!!

ON TEENAGERS, ADULT:

Statistics show that
teen pregnancy
drops off significantly
after age 25.

*Mary Anne Toledo, Republican state senator from Colorado Springs
(contributed by Harry F. Pomeroy)*

MONDAY

DECEMBER 1999

How to Teach Context and Assembly?

Must involve problems that can be assessed.

Augsburg's Statistical Literacy course has created:

- over 300 involving *Context*
- over 100 exercises involving *Assembly*

Consider ways to teach assembly using definitions

Definitions of statistics Can be Classified on a Spectrum

Two Extremes for Comparisons:

Formal: Need no knowledge of ideas

- Fewer tall smokers than smokers...

Material: Need detailed knowledge of ideas

- More autistic boys than autistic girls...

In between: Need some knowledge of ideas

- Men are taller than women on average.
- Women live longer than men on average.

Statistical educators should teach definitions as part of study design

- **Should teach formal comparisons of operational definitions.** Little knowledge of ideas is needed. Like logic/math.
- **Might delay teaching material comparisons of operational definitions.** Requires a detailed knowledge of specialized ideas. Leave this to SMEs.
- **Could selectively teach in-between comparisons of operational definitions.** At least show sensitivity of some statistics to small changes in definition.

Formal Compare: Operational Definitions

Counts: Criteria-based

Examples: Which count/total is larger?

- US population; US male population?
- Not employed; Not employed *and* looking for work?
- Use of physical force ; Use *or* threat of physical force?

Principles involving counts or totals:

- *And* phrases and modifiers restrict options – smaller total
- *Or* phrases increase options – larger total

Formal Compare: Operational Definitions

Counts: Frequency-based

Examples: Which count is bigger?

1a. Smoker: smokes *more than* five cigarettes a day

1b. Smoker: smokes *more than* two cigarettes a day

2a. Sober: consume *less than* 3 drinks in past week

2b. Sober: consume *less than* 1 drink in past week

Principles involving ranges:

- *More than X.* Larger X gives smaller count.
- *Less than X.* Larger X gives bigger count.

Formal Compare: Operational Definitions

Counts: Duration-based

Examples: Which count is bigger?

1a. Smoker: smoked in the past 2 weeks

1b. Smoker: smoked in the past 4 weeks

2a. Sober: No alcohol in past 2 weeks

2b. Sober: No alcohol in past 4 weeks

Principles involving time periods:

- Event in past X periods: Bigger X , bigger number.
- Event-free in past X periods: Bigger X , smaller #.

Formal Compare: Operational Definitions

Ratios: Compare Numerators

Which ratio is larger in each group?

1. **Percentage of US citizens** who are adults;
who are adult males; who are adults or are males?
2. The US death rate due to suicide; the US death rate;
the US death or emigration rate

Principles: Changes in the Numerator:

And phrases and modifiers restrict – smaller ratios

Or phrases increase options – larger ratios

Between Compare: Operational Definitions Part-Whole Ratios

General: Minimal knowledge of ideas is needed

Examples: Which ratio is bigger?

1. US birth rate: per 1,000 adults or per 1,000 women?
- 2a. Percentage of US males who are in the military
- 2b. Percentage of the US military who are male
- 3a. Percentage of US oil **imports** that are from OPEC
- 3b. Percentage of US oil **usage** that is from OPEC

Study Between Operational Definitions: Part-Whole Ratios

General: Minimal knowledge of ideas is needed

Examples: Which ratio is bigger?

1a. Percentage of auto accidents that involve a death

1b. Percentage of auto death that involve an accident

2a. Accidental death rate per 1,000 US males

2b. Male-death rate per 1,000 US accidents

3a. Percentage of US adult prisoners who have low IQ

3b. Percentage of US low-IQ adults who are in prison

Recommendation

Statistical educators must go beyond chance and bias in analyzing the influences on a statistic.

We should encourage *critical thinking* about the choices that could have been made in creating statistics and how sensitive the statistics are to those choices.

We should start with the simple question: “*Where do statistics come from?*” Once people realize that all statistics are constructs, socially-constructed tools, then we can teach statistical literacy as a liberal art (an inductive activity) rather than as a mathematical skill.

References

Joel Best Articles Page: www.StatLit.org/Best.htm

Schild, Milo (2007). Teaching the Social Construction of Statistics, 2007 Midwest Sociological Society, Chicago. See www.StatLit.org/pdf/2007SchildMSS.pdf

Schild, Milo (2010). Assessing Statistical Literacy: TAKE CARE in *Assessment Methods in Statistical Education: An International Perspective*. Edited by P. Bidgood, N. Hunt and F. Joliffe. Wiley Publishers, Ch. 11, p. 133-152.