

V0c 2022 Schield ICOTS 1

## Association vs. Causation; Disparity vs. Discrimination

---

**Milo Schield**  
 University of New Mexico  
 Statistical Literacy Coordinator  
 Fellow, American Statistical Association  
 Past-President, National Numeracy Network  
 Elected Member, International Statistical Institute  
 US Representative: International Statistical Literacy Project  
 Papers by topic: [www.StatLit.org/Schild-Pubs.htm](http://www.StatLit.org/Schild-Pubs.htm)

August 15, 2022  
 International Conference on Teaching Statistics  
 Slides: [www.StatLit.org/pdf/2022-Schild-ICOTS-slides.pdf](http://www.StatLit.org/pdf/2022-Schild-ICOTS-slides.pdf)  
 Paper: [www.StatLit.org/pdf/2022-Schild-ICOTS.pdf](http://www.StatLit.org/pdf/2022-Schild-ICOTS.pdf)

V0c 2022 Schield ICOTS 2

## Today's student need to study Statistics

---

**Disparities in**

- Education, suspensions and graduation
- Policing, crime, sentencing and prison
- Wages, income, assets, loans and wealth
- Health, health care, homicides and deaths

**Disparities by**  
 gender, race, ethnicity, religion, politics, age, etc.

**All of these rely on statistics: social statistics.**

V0c 2022 Schield ICOTS 3

## Confounding is the elephant in observational statistics

---

Teachers know it.

Not in intro. statistics or research methods.

It should be taught in an introductory course.

There isn't time in traditional statistics.



V0c 2022 Schield ICOTS 4

## Our Mission

---

Students need to be able to read and interpret social statistics in order to evaluate today's arguments.



Students need to understand "take into account"

Statistical educators need to offer a confounder-based statistical literacy course!

V0c 2022 Schield ICOTS 5

## University of New Mexico is offering a new course!

---

Taught 7 sections (200 students) in 2021-22



**Statistical Literacy**

MATH 1300 (3)  
 Participants will study the social statistics encountered by consumers. Investigate the story behind the statistics. Study the influences on social statistics. Study the techniques used to control these influences. Strong focus on confounding.

Meets New Mexico General Education Curriculum Area 2: Mathematics and Statistics.



V0c 2022 Schield ICOTS 6

## National Numeracy Network Meeting @ UNM: Oct 21-23.

---



Go to [www.NNN-US.org](http://www.NNN-US.org)  
 Hybrid: In-person and virtual

Friday Keynote: Dr. Larry Lesser.  
*Aligning Goals to Audience:  
 Context for Statistical Literacy*

*Talks by teachers  
 Talks by students*



V0c 2022 Schield ICOTS 7

### Correlation Does Not Imply Causation

---

This admonition is unhelpful in two ways:

Correlation measures two-factor co-variation. Two-group comparisons are more common.

‘Imply’ in math means ‘sufficient’.  
 ‘Imply’ in everyday usage means ‘supports’.

V0c 2022 Schield ICOTS 8

### Association is Not Causation

---

This mantra has its own problems.

What does ‘not’ mean?

- Never? If not ‘never’ what does it mean?

So how do journalists distinguish these?

Schild and Raymond (2009) examined the titles for 2,000 news stories.

V0c 2022 Schield ICOTS 9

### Association vs. Causation: Everyday Media

---

A-B-C GRAMMAR

Between: ???

Association: See

Causation: Do

Journalists live on ‘the edge’.

- Always implying; never asserting (alleged).
- Always prevaricating; never lying (ideally).

V0c 2022 Schield ICOTS 10

### News Story Headlines: A: 2% C: 2%

---

A: Association	C: Causation
Asserts an association;	Asserts causation;
Says "what"	Asserts "how" *
associated/association	cause, create, produce
correlation	effect, result, consequence
Two-group comparisons:	Sufficient: prevent, stop
"Women live longer than men"	"If X, then Y will happen"
"Men more likely to drink beer"	Contra-factual

Two-group comparisons: more common than two-factor covariation.  
 \* These are common usage, but not “etched in stone.”

V0c 2022 Schield ICOTS 11

### News Story Headlines: B: 66% C: 2%

---

B: Between	C: Causation
Asserts an association but suggest causation	Asserts causation; Asserts "how" *
increases, raises, ups; cut	cause, create, produce
"As x ↑, y ↓"; "more x, less y"	effect, result, consequence
before/after; linked, factor	Sufficient: prevent, stop
leads to; causal factor	"If X, then Y will happen"
due to, because of	Contra-factual

V0c 2022 Schield ICOTS 12

### Moral Causation: Difference vs. Discrimination

---

A: Association	C: Causation (moral)
Math Differences:	Immoral Differences:
Count/Rate/Amount	Evaluative or Judgemental
different, unequal	inequity/inequitable
Rank: first, second, last	unfair/unjust/undeserved
Superlatives: highest/lowest	discriminate: with prejudice
Comparatives: more, higher, times as much, percent more	discrimination*
	racism/sexism

V0c 2022 Schield ICOTS 13

### Moral Causation Disparity vs. Discrimination

B: Between (moral)	C: Causation (moral)
Descriptive Differences with a Moral Connotation	Immoral Differences: Evaluative or Judgemental
unequal/inequality	inequity/inequitable
disproportionate	unfair/unjust/undeserved
discriminate: discern difference	discriminate: with prejudice
disparity / disparate impact	discrimination*
over/under represented	racism/sexism

\* Includes unintended/structural/systemic  
 'Disparate impact' is very close to 'systemic discrimination'

V0c 2022 Schield ICOTS 14

### Crude Association

Black-white income gap: \$37,000 – a huge disparity  
 Crude association: 45% more for whites than for blacks.

Mean Income 2020	Family Structure			Standardized	
	All	Married	Unmarried	% Married	All
White	\$118,388	\$133,585	\$66,800	77.2%	\$115,628
Black	\$81,537	\$114,860	\$52,564	46.5%	\$98,110
Gap	\$36,851	\$18,725	\$14,236		Combined \$17,518

Notice the difference in the "mix". 77% vs. 47%  
 Solution? *Standardize*. Give both groups the same mix.

Result after *taking into account* 'family structure'.  
 >Black-white income gap: \$18,000  
 >Mean family income: 18% more for whites than for blacks.

V0c 2022 Schield ICOTS 15

### Crude Associations: Standardization Results

US Mean Family Income (2020)		
	Before	After
White	\$118,388	\$115,628
Black	\$81,537	\$98,110
Difference	\$36,851	\$17,518

Taking into account family structure  
 eliminated 52% of black-white income gap

V0c 2022 Schield ICOTS 16

### Black-White Income Gap: Discussion

Does this prove that much, if not most, of the original black-white income gap is *not* due to discrimination? No!

First, statisticians have no expertise in saying whether a disparity is caused by discrimination. Second, statisticians recognize that discrimination in some other area could create the observed disparity in family structure.

Suppose that the criminal justice system discriminates against black men and women. People in prison are less likely to get married or to stay married.

V0c 2022 Schield ICOTS 17

### Some Disparities do not involve Discrimination

99% of married families are heterosexual.  
 93% of prison inmates are male.  
 75% of arrests are male  
 But 50% of population is male.  
 Does this disparity *prove* sexual discrimination? No!

25% of those arrested are blacks.  
 But 13% of the population is black.  
 Does this disparity *prove* racial discrimination? No!

V0c 2022 Schield ICOTS 18

### Conclusion #1

Statisticians have no expertise in whether an association is causation: a disparity is discrimination.

Statisticians do have expertise in evaluating the resilience of an association to being influenced.

Nothing in this paper is designed to show that discrimination does not exist.

Statisticians are not judges in such matters. The statistician's role is to question and evaluate the statistics involved in such arguments.

V0c 2022 Schield ICOTS 19

## Conclusion #2

---

Statisticians may not want to address “hot topics”.

But our students need statistical literacy in order to read, interpret and evaluate the statistics – the social statistics – they encounter in everyday life.

Statistical educators are fully prepared to teach what it means to ‘take into account’ a related factor.

That’s our job!

V0c 2022 Schield ICOTS 20

## Our Mission

---

Students need to be able to read and interpret social statistics in order to evaluate today’s arguments.



Students need to understand “take into account”

Statistical educators need to offer a confounder-based statistical literacy course!

V0c 2022 Schield ICOTS 21

## National Numeracy Network Meeting @ UNM: Oct 21-23.

---



Go to [www.NNN-US.org](http://www.NNN-US.org)  
 Hybrid: In-person and virtual

Friday Keynote: Dr. Larry Lesser.  
*Aligning Goals to Audience:  
 Context for Statistical Literacy*

*Talks by teachers*  
*Talks by students*



V0c 2022 Schield ICOTS 22

## References

---

Schield, M. and R. Raymond (2009). Distinguishing Association from Causation in Media Headlines. [www.statlit.org/pdf/2009SchieldRaymondASA.pdf](http://www.statlit.org/pdf/2009SchieldRaymondASA.pdf)

Schield, M. (2017). GAISE 2016 Promotes Statistical Literacy. [www.StatLit.org/pdf/2017-Schield-SERJ.pdf](http://www.StatLit.org/pdf/2017-Schield-SERJ.pdf)

Schield, M. (2021). Statistical Literacy: Teaching Confounding. [www.statlit.org/pdf/2021-Schield-USCOTS.pdf](http://www.statlit.org/pdf/2021-Schield-USCOTS.pdf)

Schield, M. (2022). Statistical Literacy: Thinking Critically about Confounding. [www.statlit.org/pdf/2022-Schield-CCT.pdf](http://www.statlit.org/pdf/2022-Schield-CCT.pdf)

Schield, M. (2022). Statistical Literacy: Seven Simple Questions for Policy Makers. [www.statlit.org/pdf/2022-Schield-SJIAOS.pdf](http://www.statlit.org/pdf/2022-Schield-SJIAOS.pdf)

# **Association vs. Causation; Disparity vs. Discrimination**

---

**Milo Schield**

**University of New Mexico**

**Statistical Literacy Coordinator**

**Fellow, American Statistical Association**

**Past-President: National Numeracy Network**

**Elected Member, International Statistical Institute**

**US Representative: International Statistical Literacy Project**

**Papers by topic: [www.StatLit.org/Schild-Pubs.htm](http://www.StatLit.org/Schild-Pubs.htm)**

**August 15, 2022**

**International Conference on Teaching Statistics**

**Slides: [www.StatLit.org/pdf/2022-Schild-ICOTS-slides.pdf](http://www.StatLit.org/pdf/2022-Schild-ICOTS-slides.pdf)**

**Paper: [www.StatLit.org/pdf/2022-Schild-ICOTS.pdf](http://www.StatLit.org/pdf/2022-Schild-ICOTS.pdf)**

# **Today's student need to study Statistics**

---

## **Disparities in**

- Education, suspensions and graduation
- Policing, crime, sentencing and prison
- Wages, income, assets, loans and wealth
- Health, health care, homicides and deaths

## **Disparities by**

gender, race, ethnicity, religion, politics, age, etc.

**All of these rely on statistics: social statistics.**

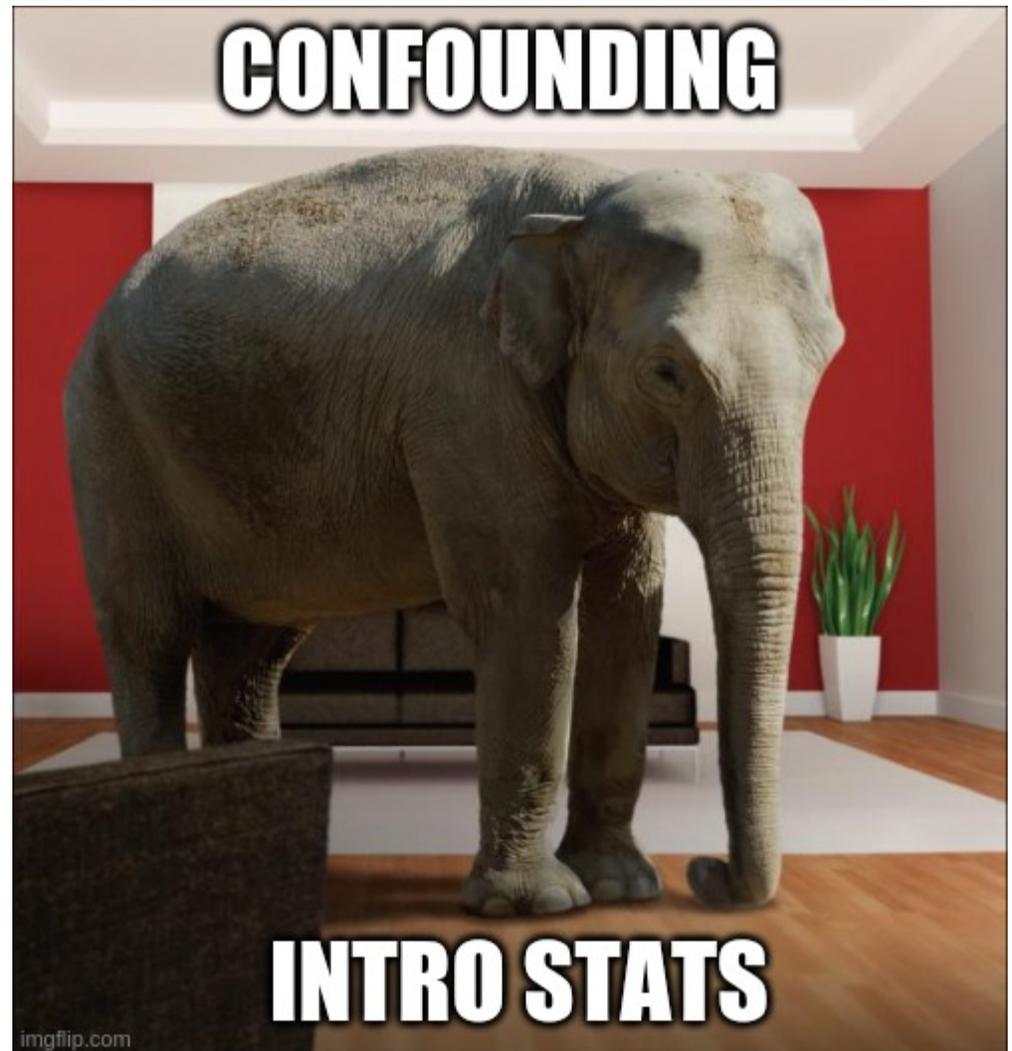
# **Confounding is the elephant in observational statistics**

Teachers know it.

Not in intro. statistics  
or research methods.

It should be taught in  
an introductory course.

There isn't time in  
traditional statistics.



# Our Mission

---

Students need to be able to read and interpret social statistics in order to evaluate today's arguments.



Students need to understand “take into account”

Statistical educators need to offer a  
confounder-based statistical literacy course!

# University of New Mexico is offering a new course!

---

Taught 7 sections (200 students) in 2021-22



## Statistical Literacy



### MATH 1300 (3)

Participants will study the social statistics encountered by consumers. Investigate the story behind the statistics. Study the influences on social statistics. Study the techniques used to control these influences. Strong focus on confounding.

Meets New Mexico General Education Curriculum Area 2: Mathematics and Statistics.

# **National Numeracy Network Meeting @ UNM: Oct 21-23.**

---



Go to [www.NNN-US.org](http://www.NNN-US.org)  
Hybrid: In-person and virtual



Friday Keynote: Dr. Larry Lesser.

*Aligning Goals to Audience:  
Context for Statistical Literacy*

*Talks by teachers*

*Talks by students*

# **Correlation Does Not Imply Causation**

---

This admonition is unhelpful in two ways:

Correlation measures two-factor co-variation.  
Two-group comparisons are more common.

‘Imply’ in math means ‘sufficient’.

‘Imply’ in everyday usage means ‘supports’.

# **Association is Not Causation**

---

This mantra has its own problems.

What does 'not' mean?

- Never? If not 'never' what does it mean?

So how do journalists distinguish these?

Schild and Raymond (2009) examined the titles for 2,000 news stories.

# Association vs. Causation: Everyday Media

---

A-B-C GRAMMAR

Between: ???

Association: See

Causation: Do

Journalists live on ‘the edge’.

- Always implying; never asserting (alleged).
- Always prevaricating; never lying (ideally).

# News Story Headlines:

## A: 2%

## C: 2%

A: Association
Asserts an association;
Says "what"
associated/association
correlation
Two-group comparisons:
"Women live longer than men"
"Men more likely to drink beer"

C: Causation
Asserts causation;
Asserts "how" *
cause, create, produce
effect, result, consequence
Sufficient: prevent, stop
"If X, then Y will happen"
Contra-factual

Two-group comparisons: more common than two-factor covariation.

\* These are common usage, but not "etched in stone."

# News Story Headlines:

**B: 66%** **C: 2%**

## B: Between

Asserts an association  
but suggest causation

increases, raises, ups; cut  
"As  $x \uparrow$ ,  $y \downarrow$ "; "more x, less y"  
before/after; linked, factor  
leads to; causal factor  
due to, because of

## C: Causation

Asserts causation;  
Asserts "how" \*

cause, create, produce  
effect, result, consequence  
Sufficient: prevent, stop  
"If X, then Y will happen"  
Contra-factual

# Moral Causation: Difference vs. Discrimination

A: Association	C: Causation (moral)
Math Differences:	Immoral Differences:
Count/Rate/Amount	Evaluative or Judgemental
different, unequal	inequity/inequitable
Rank: first, second, last	unfair/unjust/undeserved
Superlatives: highest/lowest	discriminate: with prejudice
Comparatives: more, higher,	discrimination*
times as much, percent more	racism/sexism

# Moral Causation

## Disparity vs. Discrimination

---

<b>B: Between (moral)</b>	<b>C: Causation (moral)</b>
Descriptive Differences with a Moral Connotation	Immoral Differences: Evaluative or Judgemental
unequal/inequality	inequity/inequitable
disproportionate	unfair/unjust/undeserved
discriminate: discern difference	discriminate: with prejudice
disparity / disparate impact	discrimination*
over/under represented	racism/sexism

\* Includes unintended/structural/systemic

‘Disparate impact’ is very close to ‘systemic discrimination’

# Crude Association

Black-white income gap: \$37,000 – a *huge* disparity  
*Crude association*: 45% more for whites than for blacks.

Mean Income 2020		----- Family Structure -----			--- Standardized ---	
Families	All	Married	Unmarried	%Married	% Married	All
White	\$118,388	\$133,585	\$66,800	77.2%	73.1%	\$115,628
Black	\$81,537	\$114,860	\$52,564	46.5%	73.1%	\$98,110
Gap	\$36,851	\$18,725	\$14,236		Combined	\$17,518

Notice the difference in the “mix”. 77% vs. 47%  
 Solution? *Standardize*. Give both groups the same mix.

Result after *taking into account* ‘family structure’.

>Black-white income gap: \$18,000

>Mean family income: 18% more for whites than for blacks.

# Crude Associations: Standardization Results

---

US Mean Family Income (2020)		
	Before	After
White	\$118,388	\$115,628
Black	\$81,537	\$98,110
Difference	\$36,851	\$17,518
Taking into account family structure eliminated 52% of black-white income gap		

# **Black-White Income Gap: Discussion**

---

Does this prove that much, if not most, of the original black–white income gap is *not* due to discrimination? No!

First, statisticians have no expertise in saying whether a disparity is caused by discrimination. Second, statisticians recognize that discrimination in some other area could create the observed disparity in family structure.

Suppose that the criminal justice system discriminates against black men and women. People in prison are less likely to get married or to stay married.

# **Some Disparities do not involve Discrimination**

---

99% of married families are heterosexual.

93% of prison inmates are male.

75% of arrests are male

But 50% of population is male.

Does this disparity *prove* sexual discrimination? No!

25% of those arrested are blacks.

But 13% of the population is black.

Does this disparity *prove* racial discrimination? No!

# Conclusion #1

---

Statisticians have no expertise in whether an association is causation: a disparity is discrimination.

Statisticians do have expertise in evaluating the resilience of an association to being influenced.

Nothing in this paper is designed to show that discrimination does not exist.

Statisticians are not judges in such matters. The statistician's role is to question and evaluate the statistics involved in such arguments.

## **Conclusion #2**

---

Statisticians may not want to address “hot topics”.

But our students need statistical literacy in order to read, interpret and evaluate the statistics – the social statistics – they encounter in everyday life.

Statistical educators are fully prepared to teach what it means to ‘take into account’ a related factor.

That’s our job!

# Our Mission

---

Students need to be able to read and interpret social statistics in order to evaluate today's arguments.



Students need to understand “take into account”

Statistical educators need to offer a  
confounder-based statistical literacy course!

# **National Numeracy Network Meeting @ UNM: Oct 21-23.**

---



Go to [www.NNN-US.org](http://www.NNN-US.org)  
Hybrid: In-person and virtual



Friday Keynote: Dr. Larry Lesser.

*Aligning Goals to Audience:  
Context for Statistical Literacy*

*Talks by teachers*

*Talks by students*

# References

---

Schild, M. and R. Raymond (2009). Distinguishing Association from Causation in Media Headlines.

[www.statlit.org/pdf/2009SchildRaymondASA.pdf](http://www.statlit.org/pdf/2009SchildRaymondASA.pdf)

Schild, M. (2017). GAISE 2016 Promotes Statistical Literacy.

[www.StatLit.org/pdf/2017-Schild-SERJ.pdf](http://www.StatLit.org/pdf/2017-Schild-SERJ.pdf)

Schild, M. (2021). Statistical Literacy: Teaching Confounding.

[www.statlit.org/pdf/2021-Schild-USCOTS.pdf](http://www.statlit.org/pdf/2021-Schild-USCOTS.pdf)

Schild, M. (2022). Statistical Literacy: Thinking Critically about Confounding. [www.statlit.org/pdf/2022-Schild-CCT.pdf](http://www.statlit.org/pdf/2022-Schild-CCT.pdf)

Schild, M. (2022). Statistical Literacy: Seven Simple Questions for Policy Makers. [www.statlit.org/pdf/2022-Schild-SJIAOS.pdf](http://www.statlit.org/pdf/2022-Schild-SJIAOS.pdf)