Statistical Literacy for General Education

Milo Schield, New College of Florida

Abstract:
College students are concerned about social disparities and their use as evidence of systemic discrimination. These arguments usually involve simple statistical comparisons prefaced by "studies say". These statistics are numbers in context where the context matters. Most college students don't know how to evaluate these statistical comparisons. Students need an introductory general-education course that helps them think critically about everyday statistics.

Statistical Literacy focuses on how everyday statistics are constructed, influenced, and manipulated. Statistical Literacy is different: less than a 30% overlap with the traditional statistics course. Traditional statistics focuses on random variation; statistical literacy focuses more on systematic differences. Students learn that comparisons of rates and percentages can change in size and direction after taking into account the influence of related factors: confounders. For example, in 2022, vaccinated cases were more likely to die of COVID than unvaccinated cases. This comparison reversed after taking into account (controlling for) age, After taking into account the influence of a related factor, a statistically-significant difference can become insignificant (and vice versa).

Most college students (and faculty) have never seen these changes in any math or statistics course. Students work multivariate problem such as these using simple weighted averages. No computers or algebra required. The course goal is to help students to read, interpret and evaluate the everyday statistics they will encounter throughout life as citizens.

This courses emphasizes the linguistic difference between association and causation (disparity and discrimination). Students deal with controversial social topics involving gender, race and class in ways that social justice warriors, equal-justice advocates and non-combatants find helpful.

Students use ordinary English to describe and compare rates and percentages presented in tables, graphs and texts. Augsburg University received a grant from the W. M. Keck Foundation "to develop Statistical Literacy as an interdisciplinary curriculum in the liberal arts." The resulting textbook and course were taught by seven Augsburg faculty to over a thousand students in their General Education curriculum.

The course was adopted by the University of New Mexico (UNM) and has been taught as satisfying a mathematics requirement in their General Education curriculum by six faculty to over 600 students. Students value this course. This is the first time they've been taught what it means to "take something into account quantitatively". At least a quarter agreed or strongly agreed that this math course should be required of all college students for graduation. Half at UNM said it did more to help their critical thinking than any course that had taken.

The AAC&U should develop a Value rubric for Statistical Literacy and advocate for statistical literacy.

INTRODUCTION:
Today's arguments are different.

- Studies show that every inch of height is worth an extra $1,000 in income per year.
- A new paper casts doubt on masks as a surefire COVID precaution…
- Research shows that a "higher consumption of coffee is linked to a longer life."
- Studies show that accurate statistics aren't any better than those you make up.
- New study reveals that stress from tying to follow every new study shortens life span.

Today's students don't know:

- the difference between *association* and *causation* (*disparity* and *discrimination*),
- the difference between *hard* science and *soft* science,
- the difference between a *crude* comparison and an *adjusted* comparison,
- how to interpret a table of percentages,
- how to describe and compare percentages and rates
    - using percent, percentage, rate, and chance grammars
    - as presented in statements, tables, and graphs,
- what it means to "*take something into account*" quantitatively or
- how to evaluate statistics as evidence for causal claims.

This inability to deal with today's arguments is the biggest weakness in General Education!

Statistical Literacy (Stat 100) is a new and different course.

- Designed for General Education: education for all college students,
- Designed for students in non-quantitative majors: English, history, music, etc.,
- Satisfies a math requirement in General Education at Univ. of New Mexico.
- Critical thinking course requiring hypothetical thinking about plausible confounders.
- Uses ordinary English (*small changes in syntax, can create big changes in semantics*),
- Focus on everyday statistics presented in the media as evidence in arguments.

Statistical Literacy (Stat 100) has some good credentials:

- Funded by a W. M. Keck Foundation grant: "to develop statistical literacy as an interdisciplinary curriculum in the liberal arts".
- Taught at Augsburg University (GST200) by eight faculty to over a thousand students.
- Taught at the U. of New Mexico (Math 1300) by six faculty to more than 200 students/year since 2021. Satisfies the mathematics requirement in the New Mexico General Education curriculum.
- Taught at New College of Florida (STAN 2720) since 2023.
- Textbook (*Statistical Literacy: Critical Thinking about Everyday Statistics*) published by Kendall-Hunt in 2023.
- Designed by Milo Schield, PhD Rice, Fellow with the American Statistical Association and US Representative of the International Statistical Literacy Project. Author of the textbook. Author of 100+research papers which have received over a thousand citations.
- Taken by over a thousand students and taught by over a dozen teachers.

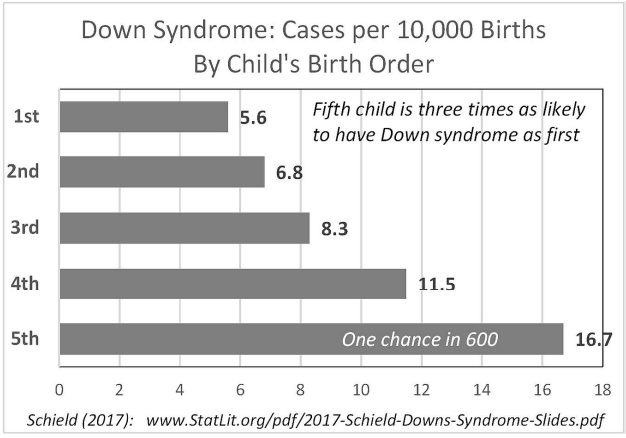
Statistical Literacy (Stat 100) has some unique credentials:

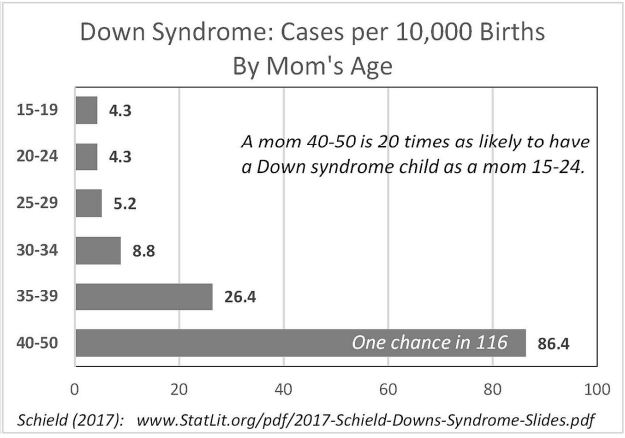
- Highly valued by first year-students. See Schield (2022c).
- Doesn't use calculus, algebra or computers.
- Strong focus on everyday statistics: typically social, economic or political statistics.
- Strong focus on multivariate thinking (per the ASA GAISE 2016 Update).
- Strong focus on confounding.

What is confounding? Confounding means confusing (found with.)

- Example: *People who shave their face tend to be taller than those who shave their legs.* We all recognize that sex is correlated with whether we shave our face or our legs, and we know that sex is correlated with height.

Consider a real case: the chance of a child being born with Down syndrome.





- Down syndrome is three times as likely for the 5th child as for the first child.
- Down syndrome is 20 times as likely for moms over 40 as for moms under 25.

Birth order and mom's age are correlated. Later children are more likely to have an older mom.

Confounding is the "elephant" in introductory statistics.

- Statistical educators know it.
- Confounding is covered in advanced textbooks and courses.
- Not found in most introductory statistics or research methods textbooks or courses.
- It should be taught to today's students.
- There isn't time. The introductory course is already packed full. There's no extra room.
- Teaching confounding complicates things. If students knew how sensitive our results are to confounding, they would have less respect for our discipline.

Solution to Confounding: 'Take into Account' using 'Per'

- Comparisons of amounts can be confounded by the sizes of the groups
  E.g., California has more unemployed people than North Dakota.
- Solution: form a ratio or rate to take into account the size of the group.
  E.g., North Dakota may have a higher unemployment rate than California.

Compare 1996 prison expenses in California (CA) with those in New York (NY).

PRISON EXPENSE			
State	Total$	# Inmates	$ per Inmate
CA	$2.9B	136K	$21,385
NY	$1.9B	69K	$28,426
50% more			25% less

Figure 1: Prison Expense: Amount and Per Inmate

- California spends 50% more on prison expense than New York.
- California has more prison inmates than New York.
- California spends 25% less on prison expense *per inmate* than New York.

Conclusion: Seeing this reversal (in going from a comparison of counts to a comparison of ratios) is not very significant. We are all generally aware that it can happen.

Q. Could a comparison of ratios be influenced by taking something else into account?

A. Most of us don't know. We've never seen it happen. My answer is "Yes"!!

Solution to Confounding: Take Into Account using Standardization
Consider the 2019 family incomes of Blacks and Whites shown in Figure 2. A family is a household involving two or more people related by blood or law.

The vertical axis is average family income. The left side shows the average income for White (67k) and Black (53k) families headed by a single parent. The right side shows the average incomes for White (134k) and Black (115k) families headed by a married couple.

The diagonal lines connecting the end points are the weighted average lines. The value of the weighted average depends upon the mixture of single-parent and married couple families.

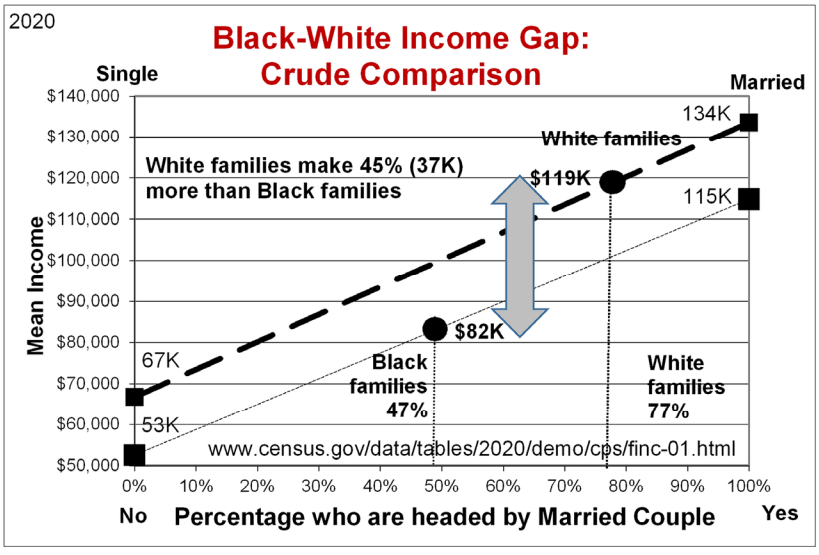


Figure 2: Black-White-Income Gap: Crude Comparison

The horizontal axis at the bottom shows the percentage of families that are headed by a married couple. On the far right, 100% of families are headed by a married couple. On the far left, zero% (none) of the families are headed by a married couple. All of the families are headed by a single parent.

Next, we need the data on these percentages for White and Black families. The Census Bureau data gives those two percentages as well.

- 77% of White families are headed by a married couple.
- 47% of Black families are headed by a married couple.

Given those two percentages, we can draw lines vertically to intersect the two weighted-average lines: the bottom line for Black families; the top line for White families.

Result: The average family income for Whites is $119K ($82K for Black families).

To review, in 2019, White families made 45% (37K) more than Black families: a big disparity.

But this comparison is a crude comparison. This doesn't mean it is fake news. It means it is a 'mixed fruit' comparison; an apples and oranges comparison. The two groups have a different mix: a different percentage of families that are headed by a married couple.

This crude comparison does not 'control for' this difference in family structure.

Figure 3 illustrates standardizing: giving both groups the same mix of the confounder.

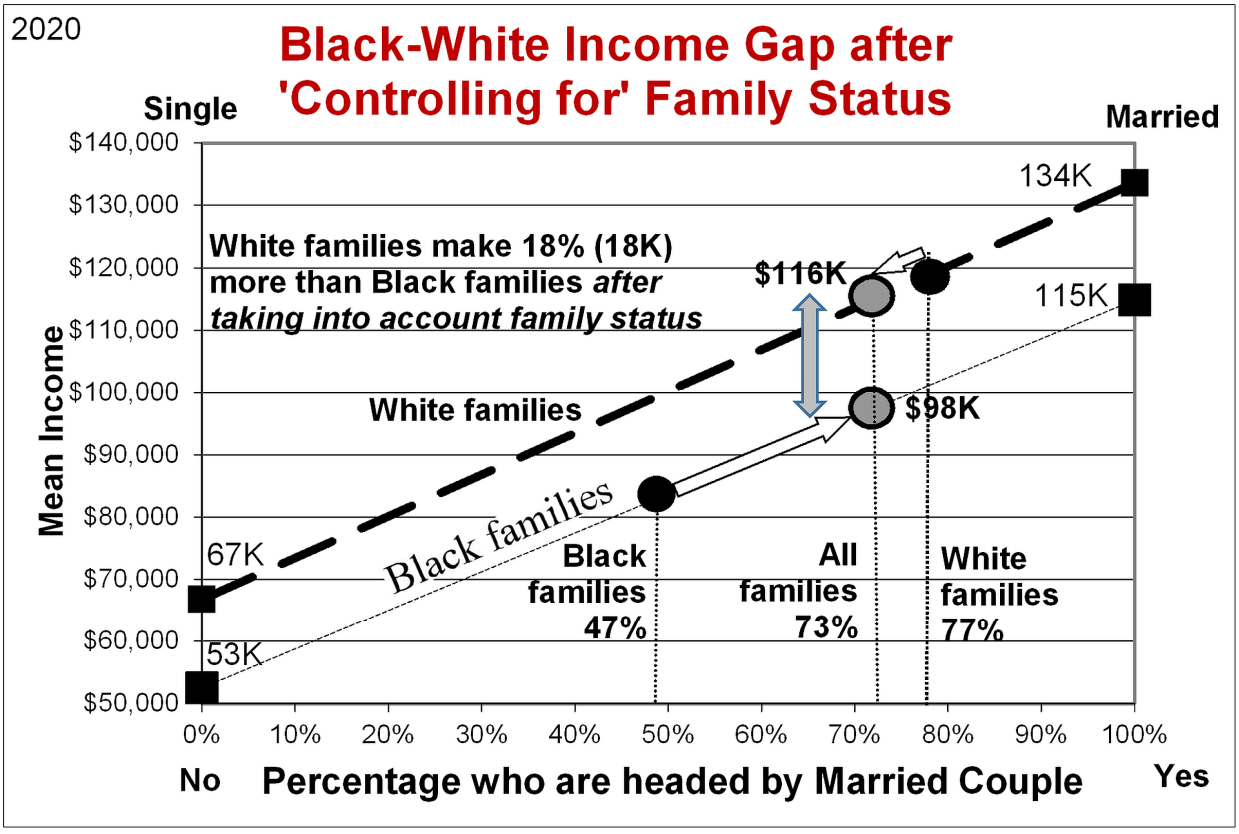


Figure 3: Black-White-Income Gap: After Standardization

In order to convert this mixed fruit comparison into a same-fruit comparison, we need to give both groups the same mix: the same percentage of families headed by a married couple.

Here are three ways to do this.

1. Give Black families the same mixture as that for White families.
2. Give White families the same mixture as that for Black families.
3. Give both groups the same mixture as found in their combination.

The first two are simpler, since just one of the two weighted averages needs to be re-calculated. However, the third is what is done in multivariate analysis, so that is illustrated here.

In this case, 77% of all families are headed by a married couple. Applying this to Black families increases their average income to 98K. Applying this to White families, decreases their average income to $134K.

In 2019, *after taking into account* family structure, White families make 18% ($18K) more than Black families.

Summarizing the before-after change.

In 2019, the Black-White income gap was 37K. But after *taking into account* family structure, that gap was 18K. So, [controlling for] family structure *explains* 56% of Black-White income gap: (37K-18K)/37K.

Q. Does this prove that most of the black-white income gap is NOT due to discrimination?

A. NO! First, statistical educators have no special expertise in saying whether a difference – a disparity – is caused by discrimination. But they do have expertise in identifying the influence a measured confounder can have on a ratio or on a comparison of ratios.

Q. Could the 56% eliminated by the difference in family structure still be due to discrimination – systemic discrimination?

A, YES!!

Explain: Suppose that Blacks were more likely to be found guilty and sentenced to prison than Whites. Suppose that Blacks were more likely to commit that kind of crime. While in prison, inmates are unlikely to get married. If they were married, they are more likely to get divorced. So, the lower percentage of Black families headed by a married couple could be due to discrimination.

Having shown how a comparison of ratios can be influenced by taking into account a measured confounder, consider this very different course in more detail.

Compare statistical literacy (Stat100) with traditional statistics (Stat 101).

The left side of Figure 4 shows the three components of statistical literacy. It is much more than just another mathematics course. But it not a communications course with a statistical focus. The emphasis on communication is the only way to see if students can read and interpret statistics: conditional probabilities. All too often, this is much more difficult than one might think.

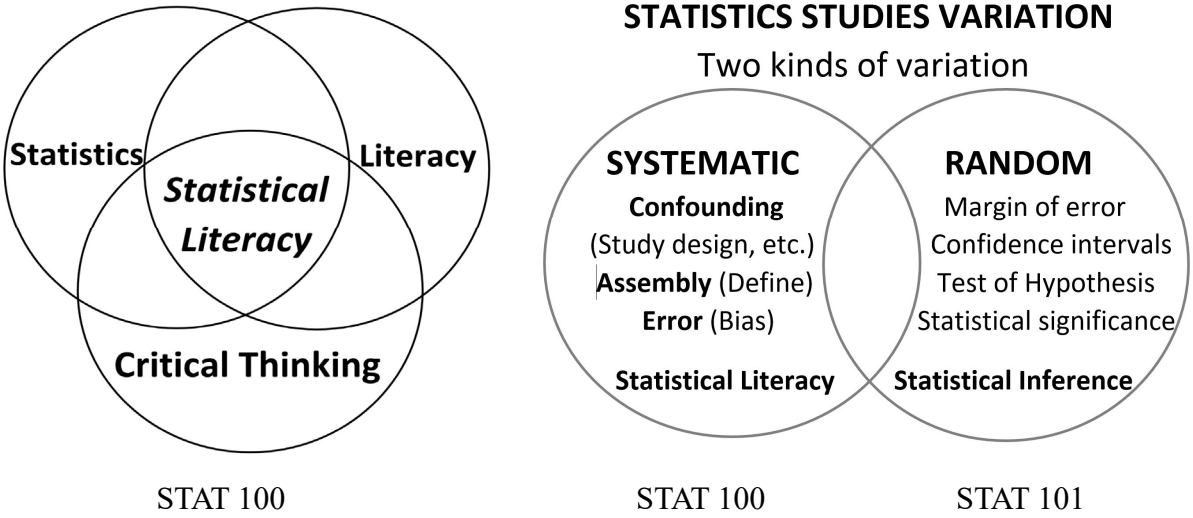


Figure 4: Statistical Literacy Components (left); Two Kinds of Variation (right)

The right of Figure 4 compares statistical literacy (Stat 100) with traditional statistics (Stat 101). They both study variation. Stat 101 focuses primarily on random variation; Stat 100 focuses primarily on systematic variation. These two courses complement each other. Undergraduate statistics majors at the University of New Mexico are required to take both courses.

Statistical Literacy is different from Information Literacy and Data Literacy. Schield (2004).

What are Statistics?

Statistics are different from numbers! Statistics are numbers in context. In math, one plus one is two. But in statistics, one bunny plus one bunny can "result in" more than two bunnies. One ice cube plus one ice cube can "result in" less than two ice cubes. The reality matters.

Do we say, “Lies, damned lies and numbers”? No! . We say “Lies, damned lies and statistics.

Numbers are like rocks; statistics are like diamonds: shaped and polished to look good.

Statistics: Four Important Things	
1	Statistics are numbers in context
2	Statistics can be influenced
3	Statistics are socially constructed
4	"Take Care" with statistics

For more on the social construction of statistics, read Best (2001). To summarize, there's no story behind the numbers, but there is almost always a story behind the everyday statistics.

Statistics in arguments: Take CARE

Arguments are mental. I visualize an argument as being like a house. The basement supports the walls; the walls support the roof, and the structure supports the point of the roof. The statistics are typically the evidence in the basement.

Recall that statistics can be influenced, so the best advice is to "Take care" when dealing with statistics. There lots of things that can influence a statistic; there are lots of ways a statistic can be influenced. In this course, all of the influence have been grouped into four categories. The first letter of each category matches the four letters in CARE: C, A, R, and E.

C = **Confounding**: Statistics are influenced by related factors.
A = **Assembly**/Assumptions: Statistics are influenced by other choices.
R = **Randomness**: Statistics are randomly influenced by chance.
E = **Error**: Statistics are systematically influenced by mistakes or bias.

Here are some examples of each kind of influence.

C = CONFOUNDING:

We already talked about the problem of confounding. How the Black-White income gap can be influenced that whether one takes into account family structure. There are other ways that ward off the influence of confounding.

Study design can ward off the influence of confounding. When you hear the phrase, "studies say", or "the data says", or "statistics show", they all refer to a certain kind of study design. In the next figure, the six kinds of study design are classified into two groups: experiments and observational studies.

In experiments, there is a "doing" by the researcher (or by nature) so the subjects are the objects of the doing. In observational studies, the researcher is passive and merely observes. All six can involve surveys as a technique for gathering data, but typically surveys are observational studies.

SCIENCE SAYS		
Hard Science		**Soft Science**
Experiments		**Observational studies**
Ideal experments		Longitudinal studies
Randomized trials		Cross-sectional studies
Quasi-experiments		Anecdotal reports

Statistics obtained from experiments are hard science: these statistics are highly resistant to being influenced. Statistics obtained from observational studies are soft science; they are less resistant to being influenced.

When Dr. Fauici said we should "follow the science", that science was largely soft science. That doesn't mean it was fake news. It was presumably the best available at that time. But as soft science, it was more likely to be updated as new information became available. Whereas the effectiveness of the COVID drugs was based on hard science – a form of randomized trials.

A=ASSEMBLY.

Assembly involves all the choices that influence a statistic and are not covered elsewhere under CARE. Assembling statistics is like 'wordsmithing' or 'lawyering.' It can be done professionally, opportunistically or malevolently.

- Zillow (model): This house is worth $250,000. This statistic is a model-based prediction.
- Definitions: ‘adult’, ‘pedophile’, ‘endangered species.’ The Northern spotted owl is an endangered species, but the Southern spotted owl is not an endangered species. The two groups can interbreed and have fertile offspring, but they are geographically separated. The endangered species acts mandated that geographic subspecies could be considered as “endangered species”.
- Wording: Happy *at* college; happy *with* college. The numbers associated with these groups can be quite different.
- Time period: Discrimination in last week, month or year. If I want a small number, I will use a shorter time frame. If I want a bigger number, I will use a longer time frame.
- Income: $68K per household or $86K per family. A family is a household involving two or more people related by blood or law. If I want a lower number, I will quote incomes per household. If I want a higher number, I will quote income per family.

Rather than review randomness and Error/bias in detail, here are the highlights.

R = RANDOMNESS.

Randomness affects a statistic in three different contexts: Extreme data, big data and small data.

Extreme data involves statistics for the best or the worst: the extremes. Extremes are often due to chance – coincidence. As such we don't expect them to repeat. See the Sports Illustrated jinx.

Big data creates a new avenue for randomness as explained by the Law of Very Large Numbers. Qualitatively, the unlikely is all but certain given enough tries. The frequency and severity of unexplainable coincidences will increase as the data size increases. Quantitatively, if something has one chance in a million, the outcome is expected (more likely than not) given a million tries.

Small data: The smaller the data set, the more easily the average can be influenced by random selection or random assignment. This is the core of the traditional introductory statistics course.

E = Error / bias

Errors include the confusion of the inverse; confusing association with causation, disparity with discrimination. Bias is typically classified as subject, measurement and sampling bias. Bias is well-covered in the traditional statistics course.

STUDENT EVALUATIONS AND COMMENTS

The ultimate goal is to have the students taking statistics value their education. Here is some data that indicates this course is achieving that goal.

1st year	4th year	Category
95%	87%	Valuable in helping you read the news
65%	79%	Helpful in developing critical thinking
32%	61%	Should be required for graduation

Source 1st year students: Schield (2022B). Source 4th year students: Schield (2016).

UNM students were asked which course did more to develop their critical thinking: statistical literacy or the best course they had ever taken in developing their critical thinking. Half of the respondents said statistical literacy did more to develop their critical thinking than any other course they had taken. (Schield, unpublished)

Student comments:

- "*the first time I feel like I'd actually use a math class outside of the classroom regularly*"
- "*I enjoyed critical thinking and the news stories*."

Given this strong vote of appreciation by students, doesn't that mean this course will be taught? No. Here are reasons why mathematics and statistics faculty may not be interested.

Mathematics-Statistics faculty may not actively support teaching Statistical Literacy.

Mathematics faculty may not be interested. Here are some reasons:

1. *The course lacks rigor.* Yes! Low mathematical rigor, but high conceptual rigor.
2. *There is no textbook.* In 2023, Kendall Hunt published Schield's *Statistical Literacy* textbook (484 pages)). It includes lots of real social statistics (194 tables, 66 graphs).
3. *Mth-stat faculty don't know how to teach critical thinking; they don't feel comfortable teaching critical thinking.* The instructor controls the emphasis on critical thinking. The textbook package includes over a thousand on-line, multiple-choice exercises. The one-line writing exercises have rules that allow for ease in manual grading.
4. *Math-stat faculty would rather teach quantitative reasoning or quantitative literacy with its focus on analytic distributions.* In QR/QL, teaching math (the calculation) is the goal. In teaching statistical literacy, the mathematics (calculation) is the means – not the end.

Statistics faculty may be reluctant to teach statistical literacy

1. *Statistics faculty would rather continue teaching the traditional statistical inference course.* Doing this maintains the strong link between statistics and mathematics. It maximizes the number of faculty who can teach this course if needed.
2. *Statistics faculty are reluctant to teach confounding because it is closely connected to causation*. Early statisticians got burned by using statistical association as evidence of causal connections in upholding eugenics. After that, they agree that saying "correlation is not causation" or "association is not causation" was their line in the sand.

3. *Statistics faculty are reluctant to teach confounding because it might bring their discipline into disrepute.* As one statistical educator said, "If people knew how easily statistics are influenced, they would have less respect for out discipline."
4. *Statistics faculty would rather on passive predictions than on causal explanations.* Passive predictions are rigorous; causal explanations lack a strong mathematical foundation.
5. *Statistics faculty have no competitive pressure to change what they teach.* More college students take a statistics course than any other course offered in college – aside from possibly first year English. Historically, the pressure for change has come from within. Today their primary call for change is to include parts of data science.
6. *Young faculty seeking tenure have little incentive to teach or publish statistical literacy.* Statistical education is not as respected among statistical educators as mathematically based research.

Given their allegiance to mathematical calculations and the associated inertia and reluctance by mathematical and statistical educators to take on critical thinking, it is unlikely that they will actively support teaching statistical literacy. If the General Education curriculum is to be updated for the 21st century, mathematical and statistical educators may need to be persuaded externally.

RECOMMENDATIONS:

The AAC&U should ….

- Invite a Statistical Literacy workshop at 2025 meeting
- Host a conference on Statistical Literacy
- Dedicate an issue of *Liberal Education* to Statistical Literacy
- Create a VALUE rubric for Statistical Literacy
- Encourage all colleges to offer statistical literacy in their general education curriculum.

Why?

- Demonstrate leadership in updating General Education
- Increase awareness among colleges and universities
- Facilitate dialogue between math-stat educators and general education needs
- Improve the connection between the social sciences and the humanities
- Overcome math-stat faculty reluctance/resistance to offering Statistical Literacy
- Help all students (and adults) deal with 21st century arguments involving disparities and discrimination (inequalities and inequities).

Bibliography

Best. J. (2001). *Damned Lies and Statistics*. University of California Press.

Schield, M. (2004). Information Literacy, Statistical Literacy and Data Literacy. *Iassist Quarterly (IQ)*. International Association of Social Science Statistics and Information Technology (IASSIST) Pg 6-11. Copy at www.statlit.org/pdf/2004-Schield-IASSIST.pdf

Schield, M (2016). Augsburg Student Evaluations of Stat 102: Social Statistics for Decision Makers. *Proceedings of the Section on Statistical Education.* Pg 3741-3750. Copy at www.statlit.org/pdf/2016-Schield-ASA.pdf

Schield, M. (2021). Statistical Literacy: Teaching Confounding. *US Conference on Teaching Statistics* (USCOTS). Copy at www.statlit.org/pdf/2021-Schield-USCOTS.pdf

Schield, M. (2022a). Statistical Literacy: Critical Thinking about Confounding. *Conference on Critical Thinking*. Copy at www.statlit.org/pdf/2022-Schield-CCT.pdf

Schield, M. (2022b). Statistical Literacy UNM Math 1300: First Year Results. *ASA Proceedings of the Section on Statistical and Data Science Education*. P. 1235-1265. Copy at www.statlit.org/pdf/2022-Schield-ASA.pdf

Schield, M. (2022c). Statistical literacy: Seven simple questions for policymakers. *Statistical Journal of the IAOS* 38 (2022) 471–475 471. DOI 10.3233/SJI-2209. Copy at www.statlit.org/pdf/2022-Schield-SJIAOS.pdf

Schield, M. (2023). *Statistical Literacy: Critical Thinking about Everyday Statistics.* Textbook published by Kendall-Hunt Higher Education. https://he.kendallhunt.com/product/statistical-literacy-2023-critical-thinking-about-everyday-statistics

Schield papers organized by topic through 2022: See www.StatLit.org/Schield-pubs.htm

Words: 3,645