AUGSBURG STUDENT EVALUATIONS OF STAT 102: SOCIAL STATISTICS FOR DECISION MAKERS

Milo Schield, Augsburg College. Minneapolis, MN.

Abstract: In 2016, the ASA endorsed the update of the GAISE guidelines. In a paper delivered at the 2016 IASE Roundtable in Berlin, Schield argued that these new GAISE guidelines required a new course which he labeled as Stat 102: Social Statistics for Decision Makers. Since 2013, Augsburg College has offered a catalog course that is arguably GAISE 2016 compliant and henceforth referred to as STAT 102. It is focused on multivariable thinking about social statistics obtained from observational studies where confounding and assembly are important influences. At the end of the course, students were asked to list the ideas and concepts they found most helpful. This list was summarized into 20 statements and the same students were asked to rate them in their importance. This paper presents their results. Students identified 'confounding', 'hypothetical thinking', 'statistics are numbers in context' and the 'Take CARE' framework as the four most important ideas in the course. Students see value in this course. In an anonymous survey with 100% participation, 61% of the 105 students taking this course in the last two years agreed that "this course should be required by all college students for graduation."

Keywords: Statistical literacy, GAISE 2016, statistical education

1. The Course and the Students

The ASA (2016) has endorsed the 2016 update to the GAISE guidelines. These updated guidelines call for educators to "give students experience with multivariable thinking." This entails more focus on observational studies, social statistics and confounding

In a paper presented at the IASE Roundtable in Berlin, Schield (2016) reviewed the 2016 update to the GAISE guidelines and proposed that statistical educators should support offering a new introductory course: STAT 102: Social Statistics for Decision Makers. This new course would complement STAT 101 (the traditional inference-based research-methods course) and STAT 100 (Statistical Literacy). Of the 26 attendees who evaluated this claim, 64% agreed or strongly agreed on offering STAT 102.

Since 2013, Augsburg College has offered a course (MIS264: Statistical Literacy for Managers) that is arguably GAISE 2016 compliant and a prototype of a STAT 102 course. This course is taken by business students majoring in Management, Marketing, MIS or International Business.

Schield (2004 and 2010) has described this course in detail. This course is different. It has different goals and a different approach from traditional inference-based statistics. The goal is to understand all the various kinds of influence on a statistic – not just randomness and bias. All these influences are classified under four categories: Confounding, Assembly or Assumptions, Randomness and Error/Bias. Students are reminded that the first letter of these four categories forms "CARE". Students are encouraged to "Take CARE" when dealing with statistics.

Appendix A gives a graphical overview of each of these sources of influence on a statistic and a brief description of some of the related topics.

This course is front-loaded. The big ideas are presented up front. The rest of the course involves unpacking these big ideas to expose more details. But every source of influence on a statistic fits under one of these headings. This gives students structure in a course with dozens of different influences. Without this, students can become overwhelmed. But this is all conjecture.

The empirical question is "Do the students find this framework valuable?" The purpose of this paper is to answer that question.

2. Student Evaluations of Topics

At the end of the spring 2016 course, students were required to complete two surveys using the anonymous Questionnaire in Moodle. They were given 10 points for completion. All 22 students completed both surveys.

The first survey asked students to identify the most important ideas, topics or principles in the course. Students had to write in their answers. There were no prompts or cues.

From the student responses to the first survey, their answers were grouped into 20 categories. In the second survey, students were given a choice of the 20 items and asked to pick the top six (Table 1), top three (Table 2) and the top one (Table 3). These results are presented in Appendix B.

3. Analysis of Student Evaluations of Topics

The topic students choose as the single-most important in the course was "Take CARE" where CARE is an acronym standing for the four kinds of influence on a statistic: Confounding, Assembly or Assumptions, Randomness and Error/bias.

Although the student choose "Take CARE" as the most important idea, that doesn't mean they were voting for "Take CARE" per se. These students have never been exposed to any other framework. Their vote could be in support of any over-arching classification of all the various influences on a statistic. It might involve three topics; it might involve five topics. It might even involve four topics, where at least two are different.

Their vote does strongly support the use of a classification with just a few categories. All too often, those supporting quantitative literacy or numeracy have lists of 10, 15, 20 or 25 items that one should consider in evaluating numbers or statistics. Schield (2008).

This four-category approach was used by Tintle et al (2016). Their four categories were Significance (How strong is the evidence?), Generalization (How broadly do the results apply?), Estimation (How large is the effect?), and Causation (Can we say what caused the effect?). Their four-categories were introduced in the first chapter and provided structure for the entire course.

Two topics were ranked in second-place as being the single most-important. They are confounding and hypothetical thinking. Hypothetical thinking was described as thinking about "plausible confounders" and "plausible definitions". Hypothetical thinking and confounding are closely linked. A confounder is a third variable that wasn't included in the original data or else it was included, but it wasn't controlled for or taken into account.

Decision makers are unlikely to read the original research report. They may see a press release or a brief summary in a journal. Neither of these is likely to mention plausible confounders. Decision makers need to think! What third factor could have a bigger influence on the outcome than the predictor in the association? And then, could this third factor be linked to the predictor in the association? Once one has a plausible confounder, then the burden of proof shifts onto those making the argument. If they don't mention it, then the decision maker has good reason to say, "I'm not convinced."

The next topic (fourth place for the single-most important topic was "Statistics are more than numbers. They include the context." This is arguably the most important topic, since everything else rests on this fact. But students aren't prone to thinking in essentials.

It is most interesting to note that 'Assembly' was ranked #2 in the Top-Three list, but did not appear when choosing the single most important topic.

4. Student Evaluations of the Course

All of the students who have taken this STAT 102 course have been required to answer an anonymous survey. Here are four questions and their answers for the 105 students that have taken this course in the last two academic years.

Q7. How valuable is this course in helping you read and interpret everyday statistics?
No value 2%; Some value 11%; Fair value 40%; Highly valuable 47%.

Q8. How helpful was this course in improving your critical thinking skills?
Not helpful 3%; Somewhat helpful 18%; Very helpful 50%; Extremely helpful 29%.

Q9. Would you recommend this course to a friend?
Definitely not 2%; Probably not 4%; Can't say 9%; Probably 50%; Definitely 36%

Q10. This course should be required by all college students for graduation:
Strongly disagree 3%; Disagree 10%; Neutral 26%; Agree 41%; Strongly Agree 20%

5. Analysis of Student Evaluations of the Course

The student answers to the four questions on the value of this course are most informative. Consider Q7 and Q8: 87% said the course helped them read and interpret everyday statistics; 79% said this course improved their critical thinking skills. But these two question have no cost. Students know what the teacher is looking for. Now consider Q9: 86% said they would recommend this course to a friend. This is a better question because it involves an action, but it may still be influenced by subject bias: students know what the teacher wants to hear.

Now consider Q10: 61% agreed that this course should be required by all college students for graduation (13% disagreed). These answers may be influenced by subject bias: students know what the instructor wants. But now the stakes are much higher. This 61% is arguably a more accurate measure of the overall value these students see in this course.

What percentage of the students taking traditional inference-based statistics would agree that it should be required by all college students for graduation?

6. Future Work

Since the study of most important topics only involved 22 students, it should be extended. While one might argue that additional topics (e.g., randomized assignment, placebo effect, stratification, standardization, multivariate regression and Simpson's paradox) should be included, this would undermine this approach which was to rely solely on what students choose to report as being important and memorable for them.

Note that study design was chosen by only three of the students in the top six list so it ranked #15. Given the importance of study design (especially randomized controlled trials) and effect size, it seems that there should be more emphasis on these topics in the future.

7. Conclusion

For Augsburg students, the most important idea in their STAT 102 course was this: Having an over-arching framework that includes all types of influence on a statistic classified into four groups. In their particular case, it was "Take CARE" where C stands for Confounding, A for Assembly, R for Randomness and E for Error (bias). .

For Augsburg students the second-more important idea was a tie between confounding and hypothetical thinking. These are closely linked. Confounders are not typically measured, so one must think hypothetically to deal with them.

Since confounding is all but absent from a traditional research-inference course, the students ranking supports the need for a separate course: STAT 102: Social Statistics for Decision Makers.

Sixty one percent of the 105 students taking this course in the last two years agree or strongly agree that this kind of course should be "required for graduation by all college students." This is an incredible statistic! It strongly supports the idea of offering a separate introductory statistics course: STAT 102 Social Statistics for Decision Makers.

At the IASE Roundtable in Berlin, Schield (2016) argued that statistical education should offer three introductory statistics courses: Stat100 Statistical Literacy, Stat 101 Traditional inference-based statistics and Stat 102 Social Statistics for Decision Makers. Of the attendees who evaluated this claim, 64% agreed or strongly agreed on offering STAT 102. This is also an incredibly strong vote for change in a discipline that is not noted for supporting disciplinary change in content!

Now we have data showing that students and statistical educators see value in offering STAT 102: Social Statistics for Decision Makers. The next step is to develop new textbooks: STAT 102 textbooks that are GAISE 2016 compliant. Ultimately statistical education may finally measure up to the MacNaughton (2004) goal for an introductory statistics course: “to give students a lasting appreciation for the vital role of statistics.”

ACKNOWLEDGEMENTS

To Marc Isaacson and Cynthia Schield for comments and suggestions. To the hundreds of Augsburg students who have given feedback on this course: a work in progress. To Mercedes V. Talley, Program Director at the W. M. Keck Foundation, for seeing and supporting statistical literacy as having the potential for changing statistics education.

REFERENCES

ASA GAISE Update (2016). Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report 2016. Copy at www.amstat.org/asa/files/pdfs/GAISE/GaiseCollege_Full.pdf

MacNaughton, Donald (2004). The Introductory Statistics Course: The Entity-Property Approach. Posted at www.MatStat.com/teach.

McKenzie, John, Jr. (2004). Conveying the Core Concepts. *ASA Proceedings of the Section on Statistical Education*. P. 2755-2757. Copy at www.statlit.org/pdf/2004McKenzieASA.pdf

Schield, Milo (2004). Statistical Literacy Curriculum Design. *IASE Curricular Development in Statistics*. Statistics Education Roundtable, Sweden. Copy at www.statlit.org/pdf/2004SchieldIASE.pdf

Schield, Milo (2008). Analyzing Numbers in the News: A Structured Critical-Thinking Approach. National Numeracy Network. www.statlit.org/pdf/2008SchieldNNN.pdf

Schield, Milo (2010). Assessing Statistical Literacy: TAKE CARE in *Assessment Methods in Statistical Education: An International Perspective*. Hunt & Joliffe, Eds. Excerpts: www.statlit.org/pdf/2010SchieldExcerptsAssessingStatisticalLiteracy.pdf

Schield, Milo (2016). Offering STAT 102: Social Statistics for Decision Makers. Invited presentation: IASE Roundtable, Berlin. www.statlit.org/pdf/2016-Schield-IASE.pdf

Sharpe, Norean, Richard De Veaux and Paul Velleman (2014) *Business Statistics: A First Course*. Third edition. Pearson Publishing.

Tintle, Nathan, Beth Chance, George Cobb, Allan Rossman, Soma Roy, Todd Swanson and Jill VanderStoep (2016). *Introduction to Statistical Investigations*. Wiley.

Utts, Jessica (2014). *Seeing Through Statistics*. Fourth edition. Wadsworth Publishing.

Appendix A: An Overview of the Four Types of Influence

Students are taught to "Take CARE" when dealing with statistics. Each of the four letters in CARE is related to a specific category or kind of influences. Here is a brief overview of these four categories. Figure 1 indicates some of the influences under Confounding.

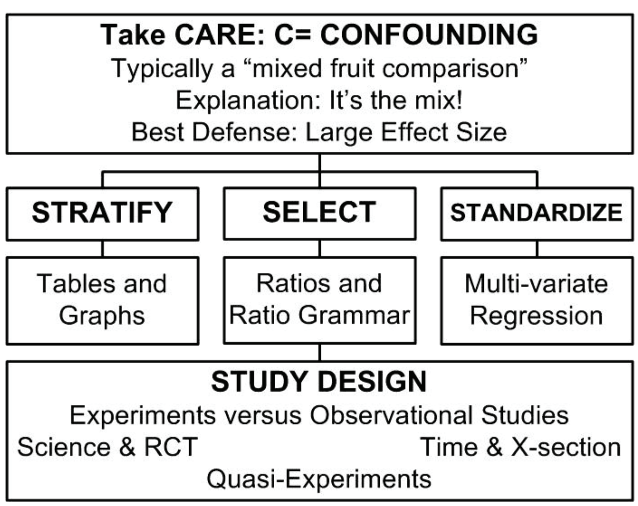

Figure 1: Take CARE. Confounding

The 2016 GAISE update mentions selection, stratification and standardization as ways to control for confounding. Effect size and study design also control for confounding.

- Effect size: The larger the effect size, the more difficult (the less likely) for a confounder to nullify or reverse an observed association.
- Selection: Selecting a give value of a confounder so there is no difference in mixture. This selection is often indicated in comparison grammars as a common whole. Selection of the study design is shown separately.
- Stratify: Creating separate columns and rows in tables (creating separate series in charts and graphs) thereby eliminating any difference in mixture within a series.
- Standardize: Modifying a mixed-fruit comparison (an apples and oranges comparison) so that both groups have the same mixture of the confounder. Here are three ways: (1) Using ratios to transform a comparison of total expenditures to a comparison of expenditures per student, per patient or per prisoner. (2) Standardizing weighted averages, rates or percentages so they have the same confounder mix. (3) Using multivariate regression to control for related factors.
- Study Design: Study design is an effective method of controlling for various kinds of confounders. Scientific experiments (repeatable) and Randomized Controlled Trials experiments (on heterogeneous subjects that are not readily repeatable): these are presented as the best kinds of studies since they control for many – if not most – confounders. Observational studies (longitudinal and cross-sectional) are presented as the worst kinds of studies since the researcher has no control. Quasi experiments are presented as intermediate since the researcher has some control on where and when an intervention occurs.

Figure 2 indicates some of the influences under Assembly.

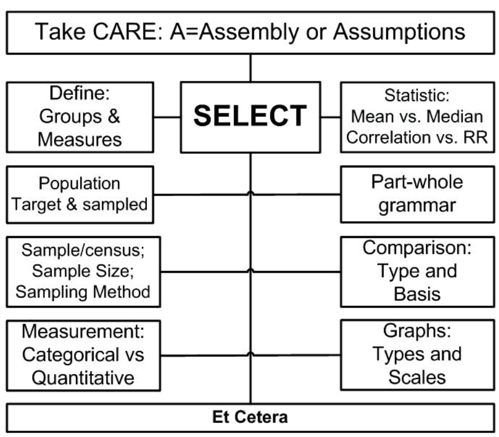

Figure 2: Take CARE. Assembly

There are host of choices involved in conducting a survey, study or experiment. All too often the decision maker has no idea of what choices were made. Decision makers are required to think hypothetically about what choices were available and how a particular choice could have influenced the size of a statistic or the size and direction of a comparison of statistics. The "Et Cetera" at the bottom is a constant reminder that this list is open-ended.

Figure 3 indicates some of the influences under Randomness. Strictly speaking, randomness is not a causal factor. But as an epistemic construct it plays the same role as a causal factor in reality in that explains and predicts.

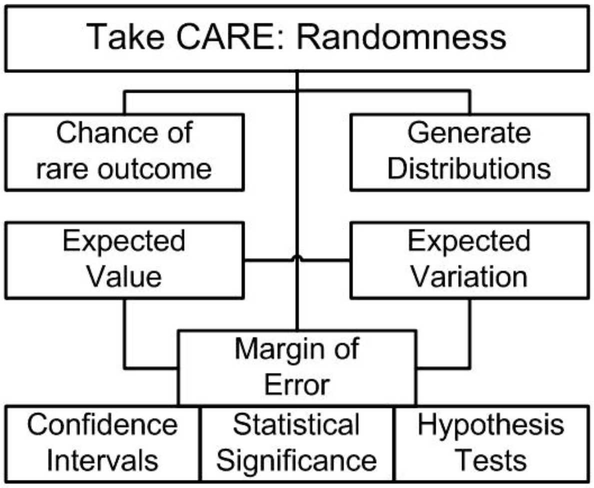

Figure 3: Take CARE. Randomness

Randomness can be used to generate various distributions such as the Normal and the Log-Normal. These distributions have an empirical basis in reality. Randomness is the basis for the Law of Very Large Numbers: the unlikely is almost certain given enough tries. Randomness can be used to generate various distributions such as the Normal and the Log-Normal.

By combining the expected value from a random sample, process or assignment with the expected variation at a given level of confidence or probability the expected margin of error can be derived. From this one can derive confidence intervals, analyze the results of hypothesis tests and determine the region to be described as being 'statistically significant.' Sharpe et al (2014) and Utts (2014) have authored textbooks in which students are not shown the details of deriving the sampling distribution from the binomial.

Figure 4 indicates some of the influences under Error (Bias).

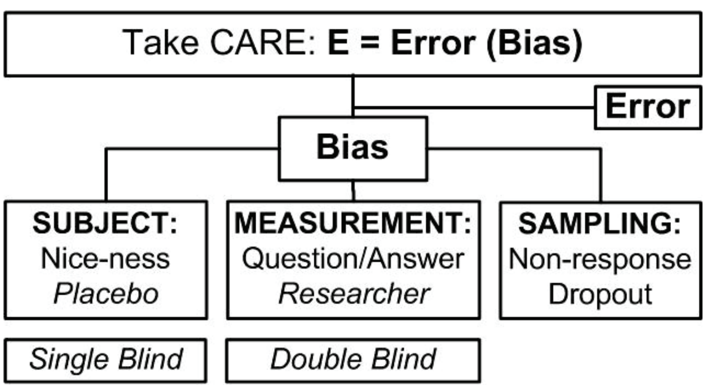

Figure 4: Take CARE. Error/Bias)

This is a typical handling of bias with one exception: confounding is treated separately from sampling bias. Confounding can occur in comparing variables in entire populations. Sampling bias only occurs in samples.

Appendix B: Students Replies

In spring 2016, Augsburg students in MIS 264 were asked to make a list of those items they considered the most important in this course. From this open-ended, free-form list, 20 topics were identified. These 20 were presented to the same students and they were asked to select the top six, the top three and the one most important of the 20 topics. .

In choosing the top-six, the 22 students had 6 choices each for a total of 132 choices. The items were presented in alphabetical order. The results are shown in Table 1.

Table 1: The Six Most-Important Topics in Statistical Literacy for Managers

	The six most important topics in Statistical Literacy for Managers	
Rank		Count
1	Take CARE: Confounding, Assembly, Randomness and Error/bias	13
2	Hypothetical thinking: plausible confounders, plausible definitions.	12
3	Named Ratios and Ratio grammar; Percent, Percentages, Rates	11
3	Read tables and graphs	11
3	Statistics are more than numbers. They include the context	11
6	Confounding	9
7	Assembly: Choose how things defined, measured, compared, presented	8
7	Bias: Subject, Measurement (Researcher) and Sampling Bias	8
7	Percentiles, percentiles & percentage points	8
10	Compare ratios: Likely grammar	7
10	Part-whole	7
12	Association-causation (Luck-skill) including the grammar	6
12	Bias: Placebo, Single blind; double blind	6
14	Random: Confidence Intervals, Statistical Significance	4
15	Random (Margin of error)	3
15	Study design/types: Experiments vs. Obs Study, Random Assign, etc.	3
17	Attributed to: Speculative statistics, New form of association	2
17	Mean-Median-Mode	2
19	Random-Coincidence. Law of very large numbers	1
20	Z-scores: Control for two kinds of variation	0
	TOTAL	132

Using the same list, students were asked to vote on the three most important ideas in Statistical Literacy. The 22 students had 3 choices each for a total of 66 choices. The results are shown in Table 2.

Some topics moved down in moving from the top six to the top three. "Read Tables and Graphs" was #4 in the top six, but dropped to #10 in the top three. Bias was #8 in the top six, but dropped to #12 in the top three.

Some topics moved up in shifting from the top six to the top three. "Confounding was #6 in the top six list, but was #4 in the top three list.

Table 2: The Three Most-Important Topics in Statistical Literacy for Managers

	Top 3 most important topics in Statistical Literacy for Managers	
Rank		Count
1	Take CARE: Confounding, Assembly, Randomness and Error/bias	12
2	Assembly: Choose how things defined, measured, compared, presented	10
3	Statistics are more than numbers. They include the context	7
4	Confounding	6
4	Hypothetical thinking: plausible confounders, plausible definitions.	6
6	Attributed to: Speculative statistics, New form of association	5
6	Named Ratios and Ratio grammar; Percent, Percentages, Rates	5
8	Association-causation (Luck-skill) including the grammar	3
8	Part-whole	3
8	Read tables and graphs	3
11	Mean-Median-Mode	2
12	Bias: Placebo, Single blind; double blind	1
12	Compare ratios: Likely grammar	1
12	Percentiles, percentiles & percentage points	1
12	Random: Confidence Intervals, Statistical Significance	1
	TOTAL	66

Note that there were no votes for five of the categories in this top-three list.

Students were then asked to select the one topic they considered the most important.

Table 3: The Most-Important Topic in Statistical Literacy for Managers

	The most important topics in Statistical Literacy for Managers	
Rank		Count
1	Take CARE: Confounding, Assembly, Randomness and Error/bias	10
2	Confounding	3
2	Hypothetical thinking: plausible confounders, plausible definitions.	3
4	Statistics are more than numbers. They include the context	2
5	Association-causation (Luck-skill) including the grammar	1
5	Bias: Placebo, Single blind; double blind	1
5	Named Ratios and Ratio grammar; Percent, Percentages, Rates	1
5	Read tables and graphs	1
	TOTAL	22

Note that there were no votes for 12 of the 20 topics in this most important topic list.

Finally students were asked to identify the three topics they considered most difficult from the original list of 20 topics.

Table 4: The Three Most-Difficult Topics in Statistical Literacy for Managers

	The top three most difficult topics in Statistical Literacy	
Rank		Count
1	Z-scores: Control for two kinds of variation	9
2	Random (Margin of error)	7
2	Random: Confidence Intervals, Statistical Significance	7
4	Named Ratios and Ratio grammar; Percent, Percentages, Rates	5
5	Hypothetical thinking: plausible confounders, plausible definitions.	4
5	Percentiles, percentiles & percentage points	4
7	Attributed to: Speculative statistics, New form of association	3
7	Bias: Placebo, Single blind; double blind	3
7	Bias: Subject, Measurement (Researcher) and Sampling Bias	3
7	Compare ratios: Likely grammar	3
7	Part-whole	3
12	Statistics are more than numbers. They include the context	3
12	Association-causation (Luck-skill) including the grammar	2
12	Confounding	2
12	Random-Coincidence. Law of very large numbers	2
12	Read tables and graphs	2
12	Study design/types: Experiments vs. Obs Study: Random Assign, etc.	2
18	Assembly: Choose how things defined, measured, compared, presented.	1
18	Mean-Median-Mode	1
	TOTAL	66

Students thought z-scores, margin of error and confidence intervals & hypothesis tests were the three most difficult topics. I was surprised to see that they voted z-scores as #1. I can understand that margin of error, confidence intervals and hypothesis tests are demanding topics. They were among the most difficult topics in the McKenzie (2004) survey.

One could have asked the students to rank their choices, but that would introduce an additional layer of analysis in determining the final ranking. This count-based approach has the virtue of transparency and simplicity.