UNIVERSITY OF TORONTO FACULTY OF ARTS & SCIENCE | Department of Statistical Sciences | Human Biology UNIVERSITY OF TORONTO | Immunology UNIVERSITY OF TORONTO

Preparing tomorrow's life sciences researchers

Dr. Bethany White
Associate Professor, Teaching Stream
Department of Statistical Sciences, University of Toronto

Dr. Jasty Singh
Assistant Professor, Teaching Stream
Department of Immunology, University of Toronto

United States Conference on Teaching Statistics

Course Beginnings

PROBLEM: Define the research question.
- Goals of the study?
- Who/what/when is the question about?

PLAN: Decide how to carry out the study.
- Selection & how to study
- Measurements?
- How will decisions made impact Data, Analysis & Conclusion stages?

DATA: Collect and record info & create summaries.
- Data storage/organization?
- Issues with data quality?
- What graphs, numbers and tables?

ANALYSIS: Extract meaning from the data.
- Emerging patterns from graphs, numbers, tables?
- Which statistical procedures?
- What are the statistical findings?

CONCLUSION: Interpret & communicate statistical findings in context of research Q.
- Are there limitations to the results?
- Do the results generate follow-up research questions?

Adapted from: Wild & Pfannkuch (1999) & MacKay & Oldford (2000)

Life Scientist	Statistician
Relevance and authenticity of research & data	Identifies/explains sources of error in research design & data analysis
Encourages good life sciences research practices	Promotes good statistical practice

Course Design

Learning Outcome

1	2	3	4	5	6	7	8
See the relevance of statistical issues in all stages of the life sciences research process.	Select appropriate statistical methods to address basic life sciences research problems.	Use statistical software to explore data and create numerical/graphical summaries that address research problems.	Use statistical software to conduct appropriate statistical inference procedures to address a research problem.	Draw scientific conclusions from graphical/numerical summaries of data and inferential procedures.	Identify strengths and weaknesses in study designs and analyses in published life sciences research.	Design studies to address basic life sciences research problems.	Recognize when standard statistical procedures are not appropriate and know to seek statistical expertise.

Course Activities

- Lectures
- R Video Modules
- iClicker Questions
- Class Discussions

Assessments

- Laboratories
- Group Project

- Final Exam

STA288 Statistics & Scientific Inquiry in the Life Sciences

➤ **Non-traditional** intro statistics course

⬆ concepts ⬇ calculations
⬆ simulation ⬇ probability

Aligns with GAISE (ASA, 2016)

We had to confront statistical errors in research repeatedly during the course!

Prevalent Statistical errors in research

- Widespread misuse and misinterpretation of statistics, especially in the life sciences. (*Weissgerber et al. 2016*)
- Error rates of 38%+ have been reported by many authors in recent decades. (*Allen 2015*)

Why?

- "inappropriate reasoning about statistical ideas is widespread and persistent, similar at all age levels (even among some experienced researchers), and quite difficult to change" *(Garfield & Ben-Zvi. 2007)*
- Most "misuses of statistics are inadvertent and are from a lack of knowledge or planning" but some are deliberate to "achieve a desired statistical result." (*Thiese et al. 2015*)

Need for improved/more training

e.g *Gardenier & Resnik 2002, Weissgerber et al. 2016, Baker 2016*

Summer 2018

- Literature review
- Formulated questions & study planning

Fall 2018

- Survey development (lit, draft, review, consultations)
- Ethics protocol

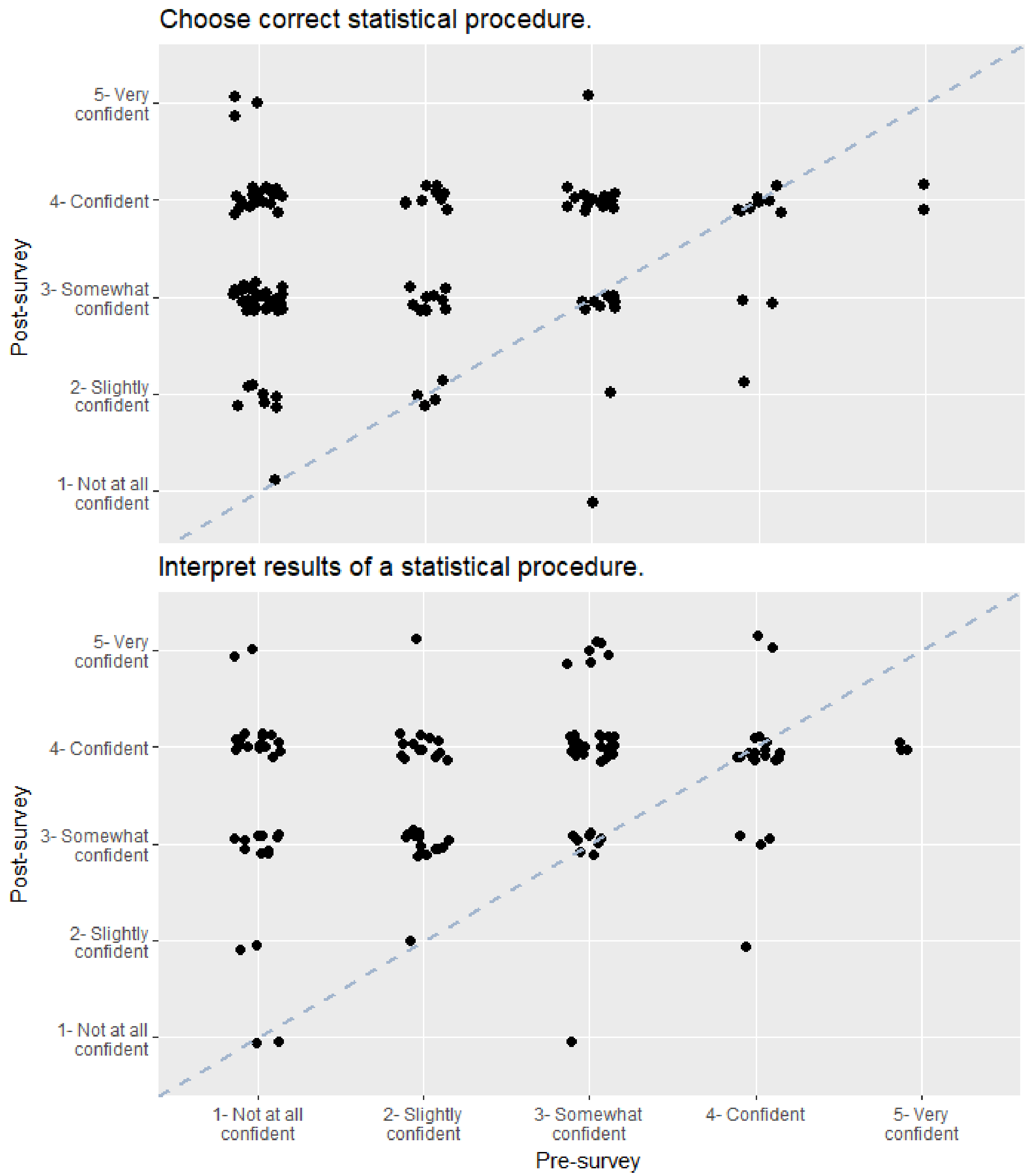
Winter 2019

- Pre-course survey & consent (Jan 7-23)
- Post-course survey (Mar 29-Apr 5)

UNIVERSITY OF TORONTO FACULTY OF ARTS & SCIENCE | Department of Statistical Sciences | Human Biology UNIVERSITY OF TORONTO | Immunology UNIVERSITY OF TORONTO

Self-efficacy

- **77%** of students more confident to *choose* correct statistical procedure at post
- **74%** of students more confident to *interpret* results at post



Engaging with statistics in research

A neuroscience student volunteering in a lab classified rod terminals in the retina as either bipolar (+) or not bipolar (-). Using a total of six mice (three for each genotype, either "wild-type (+/+)" or "Pikachurin knock-out (-/-)"), this student examined whether the proportions of the two rod terminals differ between wild-type (+/+) and Pikachurin (i.e., a protein involved in photoreceptor formation) knock-out (-/-) mice.

What can we conclude from the student's Chi-Square (χ^2) test (Fig.1)?

Fig. 1

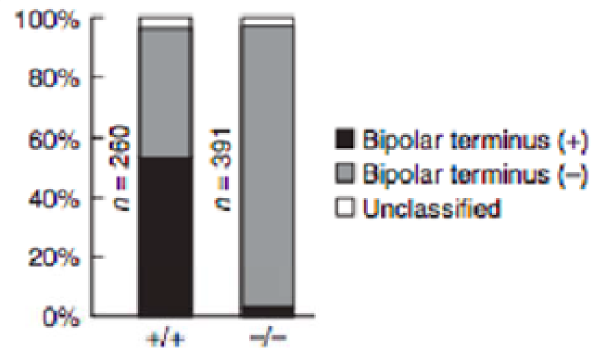
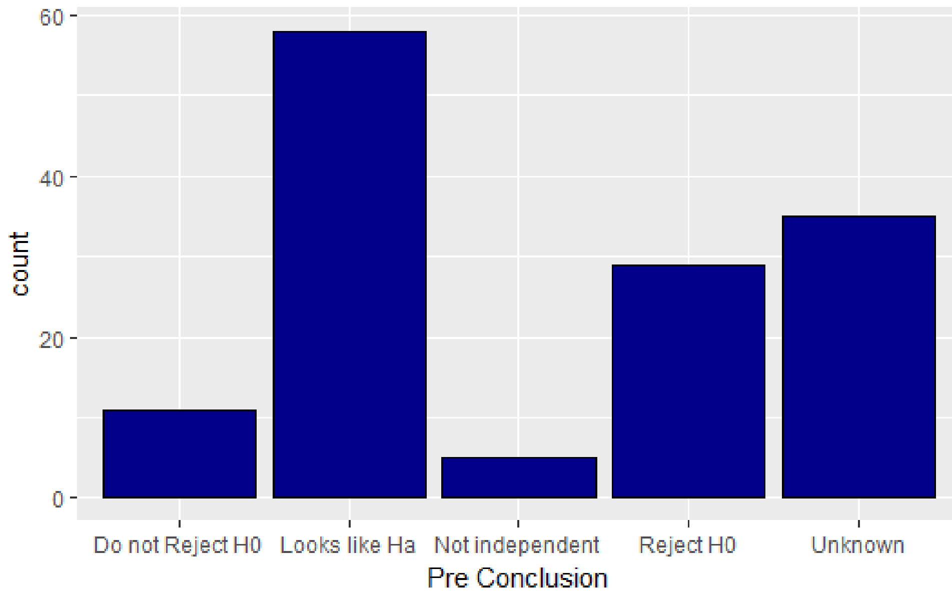


Fig 1: Quantatative analysis of bipolar dendrites in the wild-type(+/+) and *Pikachurin*(-/-) mouse retina. 260 and 391 measurements were taken from the 3 mice in the wild-type and knock-out groups, respectively. $\chi 2$ Test; P-value<0.001

a) Mice with the Pikachurin knock-out (-/-) tend to have a smaller proportion of bipolar terminus (+) than wild-type mice, so this proportion seems to depend on genotype.
b) There is not a statistically significant difference in the proportions of bipolar terminus (+) for wild-type (+/+) and Pikachurin knock-out (-/-) mice, so the proportion does not seem to vary based on genotype.
c) There is evidence against equality of the proportions of bipolar terminus (+) in wild-type (+/+) and Pikachurin knock-out (-/-) mice, suggesting this proportion differs based on genotype.
d) We cannot conclude anything from this statistical test because the measurements are not independent.
e) I do not know

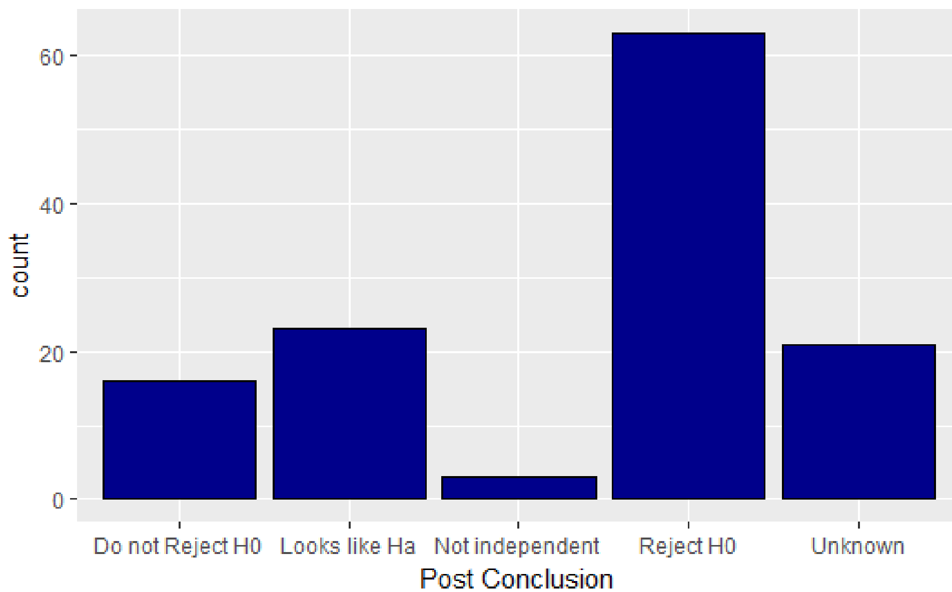
Engaging with statistics in research

PRE

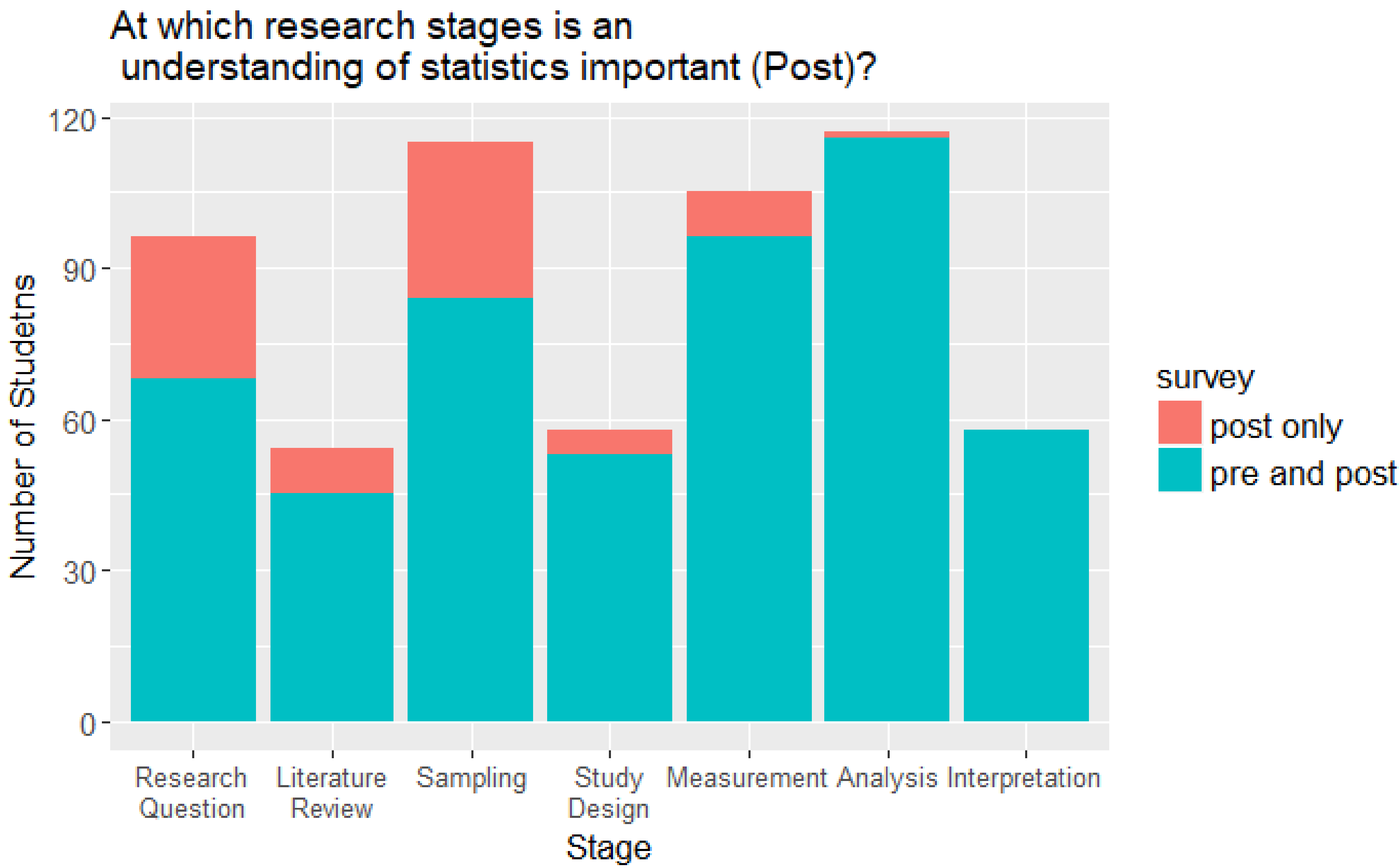


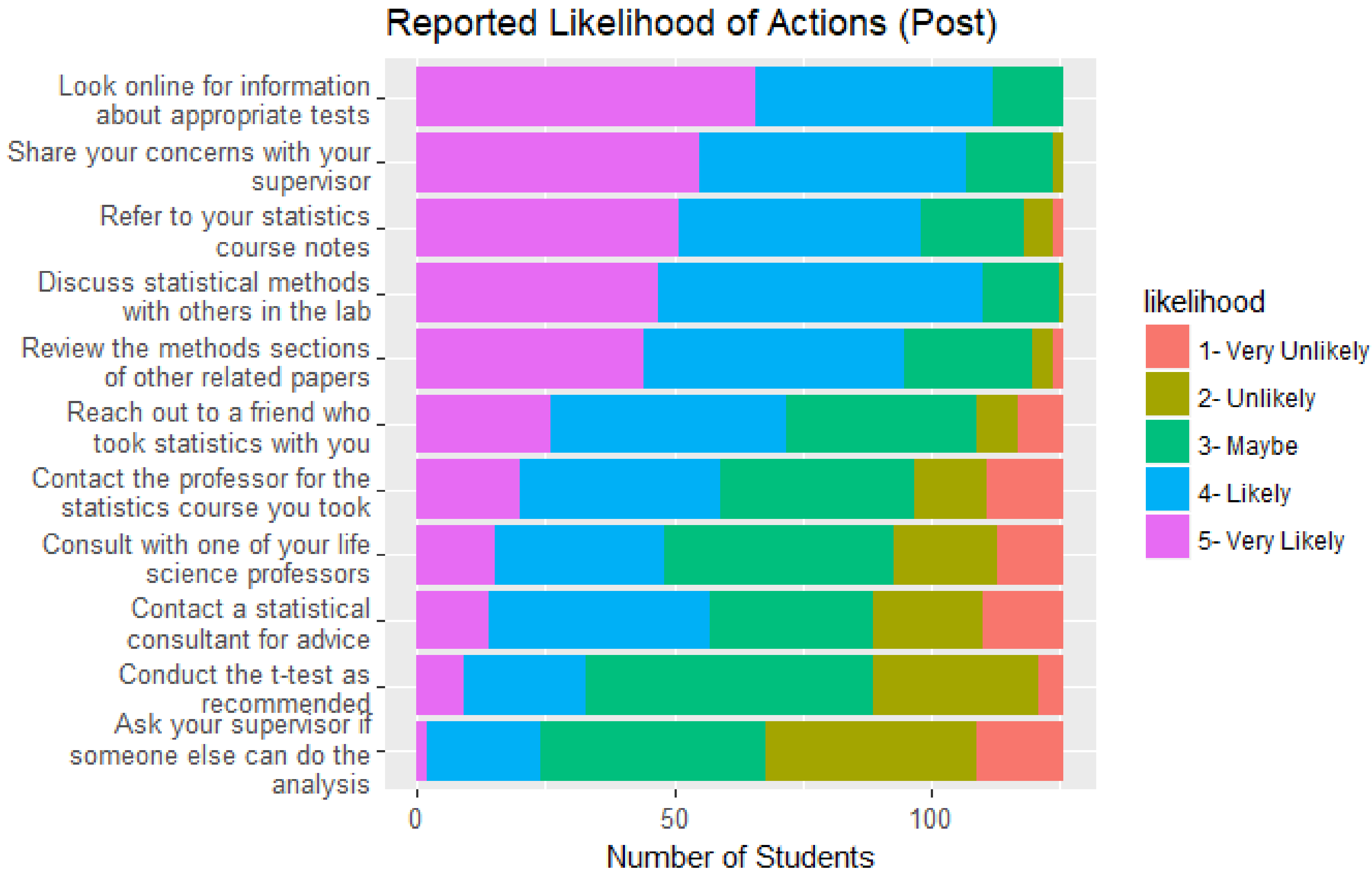
n=126

POST



Course Learning Outcome: *See the relevance of statistical issues in all stages of the life sciences research process.*




Reported Likelihood of Actions (Post)
Look online for information about appropriate tests
Share your concerns with your supervisor
Refer to your statistics course notes
Discuss statistical methods with others in the lab
Review the methods sections of other related papers
Reach out to a friend who took statistics with you
Contact the professor for the statistics course you took
Consult with one of your life science professors
Contact a statistical consultant for advice
Conduct the t-test as recommended
Ask your supervisor if someone else can do the analysis
0
50
100
Number of Students
likelihood
1- Very Unlikely
2- Unlikely
3- Maybe
4- Likely
5- Very Likely

Insights from our teaching experience & research…

- **Collaborative (multidisciplinary) teaching** makes for a richer, more authentic quantitative learning experience for students.
- **The prevalence of statistical errors** in life sciences research is alarmingly **high**.
- From study:
  - Improvement in self-efficacy to choose correct statistical procedure and interpret results.
  - Many students still not able to recognize when standard methods not appropriate at end of course and do not all see the relevance of statistics to all stages of scientific inquiry process.
- **One statistics course is not nearly enough!** If we only have one to work with, the most important course-level learning outcome is that students "*Recognize when standard statistical procedures are not appropriate and know to seek statistical expertise early in the research process.*"

References

Allen, B (2015) "Healthy And Unhealthy Statistics: Examining The Impact Of Erroneous Statistical Analyses In Health-Related Research". *Electronic Thesis and Dissertation Repository, The University of Western Ontario*. 3119. Available online at https://ir.lib.uwo.ca/etd/3119/.

Baker, M (2016) "1,500 scientists lift the lid on reproducibility" *Nature*. 533, 452-454.

Bulmer, M and Haladyn, JK (2011) "Life on an Island: a Simulated Population to Support Student Projects in Statistics". *Technology Innovations in Statistics Education*, 5(1).

GAISE College Report ASA Revision Committee (2016) "Guidelines for Assessment and Instruction in Statistics Education College Report 2016". Available online at https://www.amstat.org/asa/files/pdfs/GAISE/GaiseCollege_Full.pdf.

Gardenier, J and Resnik, D (2002) "The Misuse of Statistics: Concepts, Tools, and a Research Agenda". *Accountability in Research: Policies and Quality Assurance*, 9(2), 65-74.

Garfield, J & Ben-Zvi, D (2007) "How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics" *International Statistical Review*. 75, 3, 372-396.

MacKay, R.J. & Oldford, W (2000). "Scientific Method, Statistical Method and the Speed of Light." *Statistical Science*. 15(3).

Thiese, MS, Zachary, CA, Walker, SD (2015) "The misuse and abuse of statistics in biomedical research" *Biochemia Medica*. 25(1), 5-11.

Weissgerber TL, Garovic VD, Milin-Lazovic JS, Winham SJ, Obradovic Z, Trzeciakowski JP, Milic, NM. (2016) "Reinventing Biostatistics Education for Basic Scientists". *PLoS Biol* 14(4): e1002430. doi:10.1371/journal.pbio.1002430.

Wild, C & Pfannkuch, M (1999) "Statistical Thinking in Empirical Enquiry" *International Statistical Review*, 67,3, 223-265.