Using a “Study of Studies” to Help Statistics Students Assess Research Findings

Azka Javaid, Xiaofei Wang, and Nicholas J. Horton

The American Statistical Association (ASA) Undergraduate Guidelines Workgroup in the Curriculum Guidelines for Undergraduate Programs in Statistical Science emphasizes the importance of study design as an essential skill for undergraduate programs in statistics (ASA, 2014). The guidelines also stress communication skills, along with teamwork and collaboration, as vital elements for statistical practice. In addition, the Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report notes the importance of concepts like bias and causal inference in study design (ASA, 2016). To achieve these goals, the GAISE College Report includes several recommendations, one of which reiterates the importance of fostering active learning through discussions.

This article describes an activity that is appropriate for students in introductory and intermediate statistics courses to practice interpreting research results and scrutinizing the design and analysis of studies. The activity includes a component of group work to improve students’ communication and collaboration skills.

Traditional textbooks used in the introductory statistics curriculum stress the importance of survey design. As an example, the fourth edition of Intro Stats (IS) reinforces concepts relevant to experimental design, including bias, randomization, and sample size in Chapter 12 (“Sample Surveys”) (De Veaux, Velleman, and Bock, 2013).

Additional sample design concepts, such as observational studies, control groups, statistical significance, and confounding variables, are presented in Chapter 13 (“Experiments and Observational Studies”) of Intro Stats. Other textbooks (e.g., OpenIntro Statistics, 2015) follow a similar approach.

How can textbook readings be reinforced in a class? This activity can help students explore aspects of design, assess research findings in published papers, and critique representations and interpretations of original research.

Implementation

The students received copies of the one-page “Diner Beware” column, which was read aloud by the class. Next, they were split into groups of two to four students and each received a copy of one of the 12 research articles cited in “Diner Beware.” The research articles ranged in length from 4 to 28 pages, with an average of 10 pages.

The students were asked to skim the research article and, as a group, summarize the original research study design (i.e., describe the study’s sample design, determine whether the study was randomized or...
| Title                                                                 | Author                  | Publication Date       |
|----------------------------------------------------------------------|-------------------------|------------------------|
| You Can Be Too Beautiful                                             | James Hamblin           | March 2013              |
| The Queen Bee’s Guide to Parenting                                   | Lindsey Abrams          | April 2013              |
| Various Ways You Might Accidentally Get Drunk                         | James Hamblin           | May 2013                |
| The Unexpected Ways a Fetus is Shaped by a Mother’s Environment      | Lindsey Abrams          | June 2013               |
| The Worst Time to Have Surgery                                       | James Hamblin           | July/August 2013        |
| Is There Really Such a Thing as a “Workaholic”?                      | Jordan Weissmann        | September 2013          |
| Violence Is Contagious                                               | Rebecca J. Rosen        | October 2013            |
| Why You Look Like Your Dog                                           | Sarah Yager             | November 2013           |
| How Women Change Men                                                 | Sarah Yager             | December 2013           |
| Who Cheats—and Why                                                  | Julie Beck              | January/February 2014   |
| Why You Can’t Keep a Secret                                         | Sarah Yager             | March 2014              |
| The Optimal Office                                                   | Julie Beck              | April 2014              |
| Our Gullible Brains                                                 | Sarah Yager             | May 2014                |
| Funny or Die                                                         | Julie Beck              | June 2014               |
| What is Art?                                                         | Matthew Hutson          | July/August 2014        |
| How to Look Smart                                                    | Julie Beck              | September 2014          |
| Status Anxiety                                                       | Matthew Hutson          | October 2014            |
| Keeping the Faith                                                    | Emma Green              | November 2014           |
| Faking It                                                            | Julie Beck              | December 2014           |
| You Are Just Like Me!                                                | Matthew Hutson          | January/February 2015   |
| The Secret of Superstition                                           | Matthew Hutson          | March 2015              |
| Diner Beware                                                        | Bourree Lam             | April 2015              |
| When Emotional Intelligence Goes Wrong                               | Andrew Giambonne        | May 2015                |
| The Hypocrisy of Professional Ethicists                             | Emma Green              | June 2015               |
| Palm Reading Is Real?                                                | Eleanor Smith           | July/August 2015        |
| A Scientific Look at Bad Science                                     | Bourree Lam             | September 2015          |
| Why We Compete                                                       | Matthew Hutson          | October 2015            |
| The Strange Origins of Urban Legends                                 | Matthew Hutson          | November 2015           |
| Why You Bought That Ugly Sweater                                     | Eleanor Smith           | December 2015           |
| A Strategic Guide to Swearing                                        | Stephanie Hayes         | January/February 2016   |
| People Are Pretty Bad At Reading Faces                               | Naomi Sharp             | March 2016              |
| CEOs Behaving Badly                                                 | Alyza Sebenius          | April 2016              |
| How to Boost on the Sly                                             | Matthew Hutson          | May 2016                |
| Life Isn’t Fair                                                     | Matthew Hutson          | June 2016               |
| The Science of Beer Goggles                                         | Stephanie Hayes         | July/August 2016        |
| The Charisma Effect                                                  | Matthew Hutson          | September 2016          |
| Do People Need Small Talk to Be Happy?                               | Stephanie Hayes         | October 2016            |
| How Voters Respond to Electoral Defeat                               | Ben Rowen               | November 2016           |
| Why Kids Need Recess                                                 | Alia Wong               | December 2016           |
| Awesomeness Is Everything                                            | Matthew Hutson          | January/February 2017   |
| Unsafe at Any Speed                                                  | Jake Pelini             | March 2017              |
| How to Buy Happiness                                                | Isabella Kwai           | April 2017              |
| Puppy Love                                                           | Katherine Riley         | May 2017                |
| Boredom Is Good for You                                             | Jude Stewart            | June 2017               |
| How Vanity Could Save the Planet                                     | Matthew Hutson          | July/August 2017        |
| How to Cut in Line                                                  | Jude Stewart            | September 2017          |
| When the Mind Wanders                                                | Jake Pelini             | October 2017            |
Students made a brief set of slides summarizing their original article using RMarkdown (Bau-
mer, Çetinkaya Rundel, Bray, Loi, and Horton, 2014). The RMark-
down slides were then shared with the class via RPubs, a platform for web publishing from RStudio (the slides could be submitted in other ways, such as by e-mailing the instructor). Lastly, students were given 5 to 10 minutes to present their findings. The student presenta-
tions were intended to improve communication skills and give students experience with techno-
logical innovations like RPubs.

An example of this process can be presented with the research article titled “Odors and Consumer Behavior in a Restaurant” (Guéguen and Petr, 2006). Guéguen and Petr’s work analyzed the effect of lemon and lavender scents on the duration of time and the amount of money spent by customers in a restaurant. They carried out their study from 8 p.m.–11 p.m. on three Saturdays in May with 88 patrons, and hypoth-
esized that lavender is considered a relaxing odor while lemon is a stimulating odor.

Another example is presented by the research article titled “The Impact of Background Music on the Behavior of Restaurant Patrons” (Milliman, 1986). Milliman explored how background music affects the behavior of restaurant patrons, focusing on factors such as mood and attitude towards the dining experience. The study aimed to understand how different music styles and settings influence customer satisfaction and loyalty in the restaurant environment.

### Table 2—12 papers in Bourree Lam’s “Study of Studies” on Restaurants and Dining, the Atlantic (April 2015)

| Title and Author | Publication |
|------------------|-------------|
| Odors and Consumer Behavior in a Restaurant | *International Journal of Hospitality Management* |
| Plate Size and Color Suggestibility | *Journal of Consumer Research* |
| Assessing the Influence of the Color of the Plate on the Perception of a Complex Food in a Restaurant Setting | *Flavour* |
| Dining in the Dark | *Appetite* |
| The Effect of Musical Style on Restaurant Customers’ Spending | *Environment and Behavior* |
| The Influence of Background Music on the Behavior of Restaurant Patrons | *Journal of Consumer Research* |
| The Impact of Restaurant Table Characteristics on Meal Duration and Spending | *Cornell Hotel and Restaurant Administration Quarterly* |
| How a Crowded Restaurant Affects Consumers’ Attribution Behavior | *International Journal of Hospitality Management* |
| Lower Buffet Prices Lead to Less Taste Satisfaction | *Journal of Sensory Studies* |
| Determinants and Consequences of Female Attractiveness and Sexiness | *Archives of Sexual Behavior* |
| Effect of Server Posture on Restaurant Tipping | *Journal of Applied Social Psychology* |
| Effect on Restaurant Tipping of Male and Female Servers Drawing a Happy, Smiling Face on the Backs of Customers’ Checks | *Journal of Applied Social Psychology* |
Lam summarizes Guéguen and Petr’s research article with only this: “… particular scents also have an effect: diners who got a whiff of lavender stayed longer and spent more than those who smelled lemon, or no scent” (Lam, 2015). Similarly, Lam provides a terse summary for Kimes and Robson’s article: “Diners at banquettes stayed the longest… Diners at bad tables—next to the kitchen door, say—spent nearly as much as others but soon fled.” The students were asked to reconcile these statements with the conclusions in the original research articles. Lam’s summaries are intentionally terse: The format provides an opportunity to show students the pitfalls of taking short news articles at face value.

Results
The activity was conducted with introductory and intermediate statistics students at Amherst College during the fall 2015 and spring 2016 academic semesters. The Amherst College Institutional Review Board (IRB) approved this study. On average, 20–25 students in each class engaged in the activity. Approximately 80 minutes were allotted to the activity.

In summary, students correctly identified basic conceptual elements in the designs of the original studies. These elements include sample size, the research question, conclusion, and classification of the study as observational or randomized. Many students were skeptical of the brief claims about the original studies given in the “Study of Studies.” For example, student work correctly identified Guéguen and Petr’s sample size of patrons from a small pizzeria in Brittany, France. The students also describe how “lavender, but not lemon, increased the length of stay of customers and the amount of purchasing,” which indicates that the students picked up on Guéguen and Petr’s hypothesis and research conclusions.

The students criticized the way the conclusions were portrayed in the “Diner Beware” article, pointing out that the “Diner Beware” summary does not account for the possibility of “cultural bias/geographical bias.” Geographical bias stems from the fact that the study was only conducted in a small town in France, so the conclusions regarding scent and customer spending behavior may not generalize well to people of non-French heritage or individuals from urban areas.

In their article, Guéguen and Petr acknowledge that a small sample size and the use of only one restaurant are limitations of their study; students picked up on these caveats. Students also recognized that “limiting the study to three Saturdays in May between 8 p.m.–11 p.m. further creates sampling bias (targets a specific population).” Daytime and weekday visitors are evidently not represented. Moreover, since there was no replication, it is highly possible that another factor may have confounded the results.

Student analysis of Kimes and Robson’s article also revealed comprehension of the research’s design. In their analysis, the students correctly identified the sample size of 1,413 and the single-blinded nature of the study, since in students’ words, “the participants did not know the true nature of the experiment.” The students expressed skepticism regarding the causal statements made in Lam’s article regarding Kimes and Robson’s study, considering the observational nature of the study and the fact that Kimes and Robson “excluded some information, like the bar and patio seating” and that they “only took data from busy times.”

Kimes and Robson’s limitations stem from the fact that they only used one restaurant to draw conclusions, a shortcoming that relates to the limitation students picked up in regard to the limited focus of the study (i.e., inattention paid to less-busy hours).

Another student group summarized the limitation of Lam’s synthesis as the inability to generalize the original research’s findings, since the “conclusion for this specific restaurant may not apply to all restaurants.”

Students’ propensity for critique allows for challenging conventions, which produces a skepticism- and curiosity-driven outlook. This outlook, though, may have to be calibrated since student skepticism may be excessive.

Discussion
In an activity that linked summaries of research studies with published scientific papers, students generally accurately reported the original research’s study design—in particular, the study’s sample size; whether it was observational or experimental; and the general hypothesis, as well as the overarching conclusions. Students were often critical of the extremely terse representations of the original research in Lam’s “Diner Beware” article in the Atlantic’s “Study of Studies” column. This is not surprising, given that the goal of the “Study of Studies” is to introduce provocative or idiosyncratic research findings, not to review or assess them comprehensively.

If time permits, the instructor might spend some time on debunking misplaced criticism,
ensuring that students have a thorough understanding of the original research, can acknowledge credible published findings, and do not develop “knee-jerk” skepticism.

Overall, this activity was implemented successfully. It raised awareness about study design and secondary representations of original research. The activity can be undertaken with introductory and intermediate statistics students in a single class period, and may help improve communication skills by fostering discussion about experimental design.

We recommend that the study be undertaken after one lecture or more in study design. Conducting the study after few lectures would provide an informal student assessment and in the process, help reinforce previously learned study design concepts.

Numerous other articles published in the “Study of Studies” column could be used in the same way as the “Diner Beware” article (see Table 1 for a comprehensive list of candidate articles).

Acknowledgments

The authors would like to acknowledge support from the Gregory Call Fund of the Amherst College Dean of the Faculty. The authors thank Tasheena Narraidoo for her assistance.

Further Reading

ASA. 2014. Curriculum guidelines for undergraduate programs in statistical science. www.amstat.org/asa/education/Curriculum-Guidelines-for-Undergraduate-Programs-in-Statistical-Science.aspx.

About the Authors

Azka Javaid is a data scientist at IBM Watson Platform for Health Analytics. She received her bachelor’s degree in computer science and statistics from Amherst College in 2017. Her interests include R software development and data engineering. In her free time, she enjoys data wrangling and building Shiny applications. She has published work on assessing race and socioeconomic disparities in critical care settings.

Xiaofei Wang is a lecturer/research scholar in the Department of Statistics and Data Science at Yale University. She obtained her doctorate in statistics from Yale University in 2014. Her current interests span statistical consulting, Bayesian statistics, statistical computing, statistics education, and machine learning.

Nicholas Horton is a professor of statistics at Amherst College, with interests in longitudinal regression, missing data methods, statistical computing, and statistical and data science education. He co-authored a series of books about statistical computing and data science. He is an accredited statistician, a fellow of the American Statistical Association, former member of the ASA board of directors, and past chair of the ASA’s section on Statistical Education.

ASA. 2014. Revised guidelines for assessment and instruction in statistics education (GAISE) college report. www.amstat.org/education/gaise.

Baumer, B., Çetinkaya Rundel, M., Bray, A., Loi, L., and Horton, N.J. 2014. R markdown: Integrating a reproducible analysis tool into introductory statistics. Technology Innovations in Statistics Education 8:1–29.

De Veaux, D., Velleman, P.F., and Bock, D.E. 2013. Intro Stats (4th ed.). Pearson.

Diez, D.M., Barr, C. D., and Çetinkaya Rundel, M. 2015. OpenIntro Statistics (3rd ed.).

Guéguen, N., and Petr, C. 2006. Odors and consumer behavior in a restaurant. International Journal of Hospitality Management 25:335–339.

Kimes, S.E., and Robson, S.K.A. 2004. The impact of restaurant table characteristics on meal duration and spending. Cornell Hotel and Restaurant Administration Quarterly 45:333–346.

Lam, B. 2015. Diner beware: How restaurants trick you into eating less and spending more. The Atlantic 4:29.