Utilizing distributed practice assignments to develop students' statistical literacy in an online class

V.N. Vimal Rao [1,2]

Introduction

The course

HSCI 2117 - Introduction to Statistics for the Health Sciences at George Washington University's School of Medicine and Health Sciences Department of Clinical Research and Leadership

- First offered in the Spring 2016 term
- Originally conceived to closely follow the *consensus curriculum* (Cobb, 2007)

Course details:
- 8-week online course
- Offered six times per year
- Sections limited to 20 students
- Mainly military students and pre-nursing students

Course syllabus, from Spring 2017 to Spring 2019

Module 1:
- Descriptive statistics
- Introduction to probability
- Probability distributions

Module 2:
- Sampling distributions and estimation
- Univariate hypothesis tests
- Bivariate hypothesis tests

Required coursework:
- 6 homework problem sets
- 4 discussion boards
- 2 exams
- Capstone activity: read and discuss a research paper that utilizes statistical tools learned in the course

Recent changes:
- alignment with modern guidelines and research on tertiary level statistics education (GAISE, 2016)
- adoption of a new curriculum
- focus on statistical literacy

Students' capstone struggles

- too much focus on study design
- difficulty interpreting summary tables
- difficulty linking evidence to claims

Motivation

The goal

Develop statistical literacy by emphasizing (Gal, 2000):
- study design and measurement validity
- statistical analyses, data visualizations and summaries
- justifying claims with statistical evidence

A potential problem

Too heavy *extraneous cognitive load* (Paas, Renkl, & Sweller, 2003; Sweller, 1988):
- novice learners struggle to integrate new information with prior knowledge (Kirschner, Sweller, & Clark, 2006)
- repeatedly introducing new contexts interferes with students' learning of the new statistical concepts and tools

A potential solution

Distributed practice of skills:
- learning is spaced or divided across multiple sessions or sets of activities (Rohrer & Taylor, 2006)
- benefits from *reminding* effect (Benjamin & Tullis, 2010)

Requires a context that will:
- facilitate student *buy-in* (Newfeld, 2016) and increase student engagement (Neumann, Hood, & Neumann, 2013)
- allow for deep and meaningful interactions based on students' experiences and interests (Weiland, 2017)
- allow students to make connections between their own knowledge and the analyses (Wild & Pfannkuch, 1999)

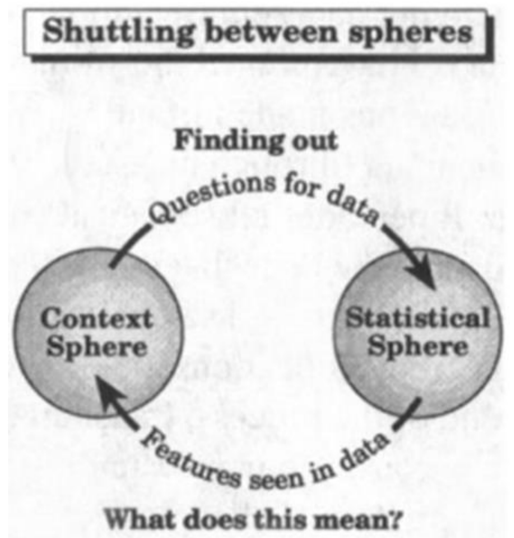
Distributed practice assignments

Combine the theories of distributed practice, cognitive load, and statistics education:
- Use the same context in multiple problem sets
- Invoke the *reminding* effect by building on previous work
- Reduce *extraneous cognitive load*
- Students can focus on learning new statistical tools
- Engagement with the context will progressive deepen

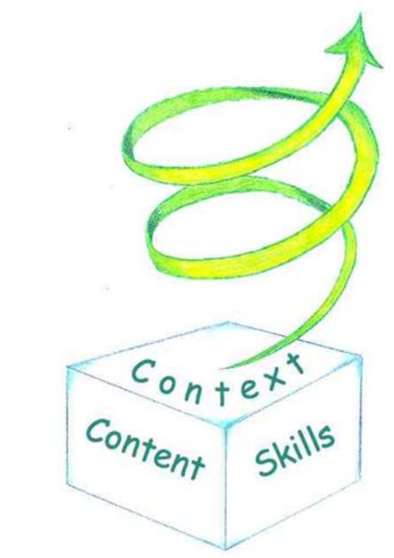
The strategy

All problem sets will utilize the Season Effect Dataset (Ngendahimana, 2016) and require students to practice:
- manipulating datasets for the required analysis
- calculating and analysing summary statistics
- interpreting and communicating results of analyses

The capstone activity will require students to read and discuss the dataset's published study (Turan et al., 2015).



+



Creation

Season Effect Dataset variable use by assignment

	Assignment					
	HW1	HW2	HW3	HW4	HW5	HW6
BMI				✓		✓
Diabetes		✓				✓
DurationSurgery	✓		✓	✓	✓	✓
Season	✓	✓			✓	
SSI		✓	✓		✓	✓

Problem set questions referring to DurationSurgery

Problem set 1 Calculate the following summary statistics for DurationSurgery from the Seasonal Effect data: mean, variance, standard deviation, minimum, median, maximum, range. Show all work, and any formulas used. (https://youtu.be/2rEhWFhSqnI) (20%)

Problem set 3 The average duration of surgery (in hours) in all patients in the Seasonal Effect data set is approximately 3.581 with a standard deviation of approximately 1.946. The duration of surgery values seems to follow a normal distribution. Estimate the percentage of surgeries that took longer than 6 hours using the normal probability distribution, P(Duration>6). Show all work. (20%) (Note: the actual percentage of surgeries that took longer than 6 hours is 10.24%.)

Problem set 4 In the Seasonal Effect data set, the average surgery duration is approximately 3.581 hours with a sample standard deviation of approximately 1.946. A surgeon has 20 surgeries similar to those in the study scheduled for the next week.
Use the Central Limit Theorem to calculate the probability that the surgeon will have an AVERAGE surgery time less than 2 hours. Hint: Determine what is the average of average duration for surgeries and the standard error for average duration for 20 surgeries. (15%)

Problem set 5 The following table provides summary statistics for the DurationSurgery based on whether or not patients contracted an SSI from the Seasonal Effect data set. One of the researchers is curious whether there is evidence to suggest that surgery duration was longer in patients who contracted SSIs. Use the following information to conduct the following hypothesis test:
- A one-tail T-test for a two-sample difference in means at the 99% confidence level
- with Null Hypothesis that the average surgery duration in patients that did contract SSIs is equal to the average surgery duration in patients that did not contract SSIs
- and with Alternate Hypothesis that the average surgery duration in patients that did contract SSIs is greater than the average surgery duration in patients that did not contract SSIs

Problem set 6 Determine the coefficients of the least squares regression line for the following model between BMI and duration of surgery: $Duration_i = a + b * BMI_i + \varepsilon_i, where\ \varepsilon_i\ is\ a\ random\ error\ term$ (20%)

Reaction

Research Questions

Does this strategy have positive effects on:
- increasing student engagement with statistical tasks?
- increasing student *buy-in* to the value of statistics?
- reducing *extraneous cognitive load* and interference with instruction of statistical tools and concepts?

Does this strategy lead to:
- progressively higher levels of context interaction?
- improvement in problem set grades focused on the assessment and use of statistical tools?
- improvement in exam grades focused on the assessment of statistical reasoning and thinking?
- increased depth of conversation in discussion boards focused on the practice of statistical literacy?

Future analyses

- Do students experience context fatigue?
- Does using only one context limit students' learning?
- Would distributing reading the research study throughout the course improve students' literacy?
- Can interleaving multiple contexts improve literacy?
- Does a formal evaluation of student interactions through a statistical literacy framework (Budgett & Pfannkuch, 2010) show significant student growth?

References

Benjamin, A. S., & Tullis, J. (2010). What makes distributed practice effective?. *Cognitive psychology*, *61*(3), 228-247.
Budgett, S., & Pfannkuch, M. (2010). Assessing Students' Statistical Literacy. In P. Bidgood, N. Hunt, & F. Jolliffe (Eds.), *Assessment Methods in Statistical Education: An International Perspective*, (pp. 103-121). Chichester, U.K.: John Wiley & Sons.
Cobb, G. W. (2007). The introductory statistics course: A Ptolemaic curriculum?. *Technology innovations in statistics education*, *1*(1).
Crissinger, B.R. (2015). The effect of distributed practice in undergraduate statistics homework sets: A randomized trial. *Journal of Statistics Education*, *23*(3).
Ebbinghaus, H. (1885). *Über das gedächtnis: untersuchungen zur experimentellen psychologie*. Duncker & Humblot.
GAISE College Report ASA Revision Committee, "Guidelines for Assessment and Instruction in Statistics Education College Report 2016," http://www.amstat.org/education/gaise
Gal, I. (2000). *Adult Numeracy Development: Theory, Research, Practice*. Cresskill, NJ: Hampton Press.
Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational psychologist*, *41*(2), 75-86.
Neumann, D.L., Hood, M., Neumann, M.M. (2013). Using real-life data when teaching statistics: Student perceptions of this strategy in an introductory statistics course. *Statistics Education Research Journal*, *12*(2), 59-70.
Newfeld, D. (2016). A first assignment to create student buy-in in an introductory business statistics course. *Teaching Statistics*, *38*(3), 87-90.
Ngendahimana, D., "Season Effect Dataset", TSHS Resources Portal (2016). Available at https://www.causeweb.org/tshs/season-effect/
Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational psychologist*, *38*(1), 1-4.
Rohrer, D., & Taylor, K. (2006). The effects of overlearning and distributed practice on the retention of mathematics knowledge. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *20*(9), 1209-1224.
Rumsey, D. J. (2002). Statistical literacy as a goal for introductory statistics courses. *Journal of Statistics Education*, *10*(3).
Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science*, *12*(2), 257-285.
Turan, O. A., Babazade, R., Eshraghi, Y., You, J., Turan, A., & Remzi, F. (2015). Season and vitamin D status do not affect probability for surgical site infection after colorectal surgery. *European Surgery*, *47*(6), 341-345.
Weiland, T. (2017). The importance of context in task selection. *Teaching Statistics*, *39*(1), 20-25.
Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International statistical review*, *67*(3), 223-248.

COLLEGE OF EDUCATION + HUMAN DEVELOPMENT

[1] Department of Educational Psychology, College of Education and Human Development
University of Minnesota, Minneapolis, MN 55455, USA

[2] Department of Clinical Research and Leadership, School of Medicine and Health Sciences
George Washington University, Washington, DC 20052, USA

UNIVERSITY OF MINNESOTA
Driven to Discover℠