Guidelines for Designing a Simulation Exercise

- Simulations are a form of **scientific discovery learning** which has the following parts or phases (Eckhardt, et. al., 2013):
  - **Prediction** – stating a possible simulation outcome and/or forming hypotheses
  - **Conducting** – performing the simulated experiment and "collecting data"
  - **Reasoning** – drawing conclusions from the simulation and/or comparing predictions to the outcome of the experiment
- **Providing Appropriate Instructional Support (with example from the sampling distribution and CLT simulations):**
  - **Interpretive Support (Prediction phase)** – Ask the student a directed question that activates prior knowledge and leads them to form appropriate hypotheses. This incites **germane cognitive load** (uses *expectancy driven methods*) and decreases the chances of passive learning common in automated activities (Liu, 2010). Done ***prior*** to the start of the simulation.
    - **Example**: "Given that the data comes from a distribution with a mean of 10 and a standard deviation of 2, what is the mean of ALL sample means from samples of size 20? What is the standard deviation of the sample mean from ALL the samples of size 20? If we were to collect only 100 samples of size 20, would the average of the means from these samples be exactly the same as the mean and standard deviation stated above? Explain briefly."
  - **Experimental Support (Conducting phase):**
    - Keep the simulation as simple as possible so that students are not overwhelmed by the process of constructing the simulated data sets as this may "induce a high working memory load that can be detrimental to learning" (Kirschner et. al., 2006).
      - **Example**: For the simulation exercise, provide the data sheet set up with 20 rows and 100 columns that are already "linked." The students are only responsible for generating the formula to construct a random sample ONCE and select "Standardize Attributes" in the column information dialog box to copy this same formula to all 100 columns.
    - Reduce visual scanning when the results of multiple tasks are being referenced to reduce **extraneous cognitive load** (Liu, 2010).
      - **Example**: In the CLT simulation, students are told to paste the corresponding histograms and normal quantile plots in specific places in the Lab document that will be submitted. Questions about interpreting these plots specifically reference exactly where the plots should have been inserted (*ie*: "Compare the distribution of the sample means generated from samples of size 5 (found in #10), samples of size 20 (found in #15), and samples of size 75 (found in #20).")

Veronica P. Hupper, Ph.D
Mathematics and Statistics Department, University of New Hampshire

- In the assignment, ask questions that have students gradually work towards the understanding of important simulation parameters (Eckhardt et. al., 2013).
  - **Example**: "The resulting data sheet contains how many simulated samples? Each sample contains how many simulated observations? Briefly explain how you know."
  - **Example**: Walk the students through the steps to use the software to find the mean of each of the 100 simulated samples and save this to a separate data sheet. Use methods that the students are already familiar with in the software to then find the mean and standard deviation of the 100 simulated sample means.
- Students at the introductory level easily understand basic graphical representations such as histograms (Blume & Royall, 2003), and so these should be used to present the information in the simulation whenever possible.
  - **Example**: When demonstrating the CLT, have the students construct a histogram and/or normal quantile plot of the means resulting from 100 simulated samples for each of 3 sample sizes simulated and compare the shapes of the distribution.

➢ **Reflective Support (Reasoning phase)** – Prompt students after the simulation and data analyses to reflect on how the simulation relates to the hypotheses made prior to generating the simulated data.

- **Example**: Specifically ask students if the results from the simulation (the average of the 100 sample means or the histograms of the sample means for the different sample sizes) agree with what they predicted.

References:

Blume, J.D. & Royall, R.M. (2003). Illustrating the Law of Large Numbers (and Confidence Intervals). *The American Statistician* **57,** 51 – 57.

Eckhardt, M., Urhahne, D., Conrad, O., & Harms, U. (2013). How effective is instructional support for learning with computer simulations? *Instructional Science.* **41**, 105 – 124.

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 Psychology.* **41**, 75 – 86.

Liu, T. (2010). Developing Simulation-based Computer Assisted Learning to Correct Students' Statistical Misconceptions based on Cognitive Conflict Theory, using "Correlation" as an Example. *Journal of Educational Technology.* **13**, 180 – 192.

Veronica P. Hupper, Ph.D
Mathematics and Statistics Department, University of New Hampshire