Project-SET Sampling Variability Learning Progression Released 10/5/12 [1,2]

		GAISE Framework[3]			
		Formulate Question	**Collect Data**	**Analyze Data**	**Interpret Results**
Loop 1[4]		a. Set up a statistical question to explore a population parameter of interest by either (1) a teacher bringing in a claim to explore, or (2) a student making a claim or posing a question to explore	a. Describe methods to obtain repeated samples of the same size b. Take repeated samples of the same size "by hand" (e.g., collect packs of M&Ms and find the proportion of green in each pack, roll a die 20 times and find the number of times an even number is rolled in each set of 20 rolls)	a. Notice that different samples will give different summary statistics b. Notate the "pattern" with which the summary statistic varies by making a dot plot	a. Informally relate the sample statistic to the population parameter where (1) the population parameter is known, and (2) the population parameter is not known
Loop 2[5]		a. Set up a statistical question to explore a population parameter of interest (the question could be the same as in loop 1 but be explored in more depth in loop 2)	a. Take repeated samples of the same size (technology may be used)	a. Clarify that the summary statistic varies in a predictable way, where predictable does not mean that we can predict the next summary statistic, but that we can predict the distribution of the summary statistic b. Define and construct an approximation to the sampling distribution	a. Draw inferential conclusions based on the sample statistic about the population parameter (e.g., the sampling distribution of the statistic illustrates that C is a plausible value for the population parameter of interest)
Loop 3	**Repeat loops 1-2 with different size random samples to compare what happens**		a. Take three different sets of repeated samples with: (1) small sample size, (2) medium sample size, and (3) large sample size	a. For each set of repeated samples (small, medium, large), construct an approximation to the sampling distribution	a. Explicitly describe what happens to the shape of the sampling distribution as the size of the sample varies

Loop 4		a. Set up a statistical question to explore a population mean or proportion	a. Take repeated samples of the same size	a. Construct an approximation to the sampling distribution and relate it to the Central Limit Theorem b. Provide quantitative descriptions of the variability of the sampling distribution c. Compute the standard error	a. Use the Central Limit Theorem to draw conclusions about the population and answer the posed statistical question b. Describe what happens to the sampling distribution as the sample size increases
Loop 5[6]		a. Set up a statistical question to explore a population parameter b. Set up a simulation using software by (1) outlining and explicitly describing how the simulation model will be set up, (2) describing why real data is not used to answer the question	a. Run the simulation	b. Use software to illustrate the approximation to the sampling distribution	a. Draw inferential conclusions about the posed questions by using the sampling distribution b. Review why it was necessary to use a simulation in order to answer the posed statistical question
Loop 6[7]		a. Set up a statistical question to explore a population (each student poses their own question)	a. Take one single sample (do not combine results as a class)	a. Find the value of the sample statistic and use its sampling distribution to make inferences about the population parameter b. Discuss the strength of the evidence to answer the posed question[8]	a. Understand that knowing the sampling distribution enables us to make inferences about the population parameter from a single random sample and multiple samples are not required b. Discuss the notion of certainty and uncertainty in the context of inference

[1]The prerequisite knowledge needed to work through the LP consists of three items. A student must be able to: (1) Define and understand the sample/population relationship (e.g., a sample is a portion of the population and a sample can be a representative portion of the population), (2) Define population parameters and summary statistics, (3) Define sampling methods (e.g., define how to obtain a representative sample)

[2] The Project-SET LP is organized around four dimensions: (1) the progression of the topic within a loop, (2) the sophistication on each loop (3) the alignment with GAISE, and (4) the alignment with the CCSS. The loops are meant to illustrate the different depths of the concept of sampling variability. We hypothesize that a student would be guided through loop 1. As the student is ready, then he/she would be guided and progress through loop 2, 3, 4, 5, and 6.

[3] The Project-SET LP directly aligns with the GAISE Framework. To illustrate the alignment, the LP is organized around the four GAISE components depicted as the columns of the LP.

[4] The Project-SET LP directly addresses the high school Common Core State Standard "Making Inferences & Justifying Conclusions" S-IC. Notes are made throughout the LP in order to illustrate the specific points in which the LP aligns with the high school CCSS. It is important to note that because the CCSS includes a lot of statistical concepts in middle school, this LP also aligns with several standards for those grades. For example, simulations are covered in the middle school CCSS. Loop 5 illustrates how to progress students through a simulation.

[5] Loop 2 of the Project-SET LP aligns with CCSS S-IC.1 " Understand statistics as a process for making inferences about population parameters based on a random sample from that population."

[6] Loop 5 of the Project-SET LP aligns with CCSS S-IC.2 "Decide if a specified model is consistent with results from a given data-generating process using simulation."

[7]The entry for Interpret Results of Loop 6 directly aligns with CCSS S-CI.1 " Understand statistics as a process for making inferences about population parameters based on a random sample from that population."

[8]This point in the LP is a natural transition to discussing p-values and hypothesis testing in an following lesson. As such, this LP serves as a lead in to the introduction of more formal discussions of inference.