Supplemental Material

CBE—Life Sciences Education

Wilson and Rigakos

Supplemental Information 1: *Demographics for both focus group used for prompt language validation and students evaluated by SPFA*. (**A**) Focus group demographics including gender, major, and undergraduate levels. Data using this group of students is displayed in Figure 2. (**B**) Demographics of students evaluated by SPFA including gender, major, ethnicity, and class rank that includes both multidisciplinary undergraduate sample and sample from pre flowcharts for middle and high school summer program. Data using this group of students is displayed in Figures 3 and 5-10.

A.

		Count	% of Focus Group
Gender	Male	27	63%
	Female	16	37%
Major	Bussiness (Accounting, Mangement, Finance, Economics, Marketing)	14	33%
	Liberal Arts (Pastorial music, History, Catholic Studies, Commnication, Lanaguages, Political Science, Phycology)	7	16%
	Sports (Marketing, Performance)	6	14%
	Education (Elementary, Special)	5	12%
	Math and Sciences (Biology, Enviromental)	4	9%
	Nursing	3	7%
	Undecided	3	7%
	N/A (High school senior)	1	2%
Year	Year 1 (Freshman)	26	60%
	Year 2 (Sophmore)	13	30%
	Year 3 (Junior)	3	7%
	Year 4 (Senior)	0	0%
	Year 5+	1	2%

B.

		Count	% of Interdisplinary
Gender	Male	21	40%
	Female	31	59%
Major	Bussiness (Accounting, Mangement, Finance, Economics, Marketing)	4	8%
	Liberal Arts (Pastorial music, History, Catholic Studies, Commnication, Lanaguages, Political Science, Phycology)	14	26%
	Sports (Marketing, Performance)	2	4%
	Education (Elementary, Special)	2	4%
	Math and Sciences (Biology, Enviromental)	7	13%
	Nursing	8	15%
	Undecided	2	4%
Year	Middle School	6	11%
	High School	7	13%
	Year 1 (Freshman)	5	9%
	Year 2 (Sophmore)	23	43%
	Year 3 (Junior)	8	15%
	Year 4 (Senior)	3	6%
Ethnicity	Asian	8	15%
	Black	5	9%
	White	34	64%
	Hispanic	3	6%
	Other	1	2%
	Multiple	1	2%

Supplemental Information 2: *Example curriculum comparing scientific process and scientific method using Understanding Science teaching model.*

Scientific Method vs Scientific Process

Part 1:

Compare and contrast two different depiction of the scientific process in groups of 3. The first depiction in **Figure 1** is linear and is what is generally taught as the scientific method. The second depiction (**Figure 2**) is a circular chart that can be found at undsci.berkeley.edu/article/scienceflowchart. This circular chart is interactive so you can move your curser over the different circles to see what ideas are contained to help you in comparing and contrasting the these representations. As you are comparing and contrasting these two figures answer the questions below. Everybody should keep notes on the discussion and the answers to the questions in their lab notebook.

Questions:

- What ideas are the same between Figure 1 and 2?

- What ideas or elements are in Figure 2 that are not in Figure 1?

- Is there things about the formatting or appearance between Figure 1 and 2 that you think are different? What do these differences mean in terms of the scientific process?

- Which do you think is a more realistic representation of what happens while doing research? Why?

- What parts of Figure 2 are you confused about in their place in the scientific process?

- Why do you think that people outside of science would not think of the scientific process as depicted in figure 2?

Figure 1: “Scientific Method”

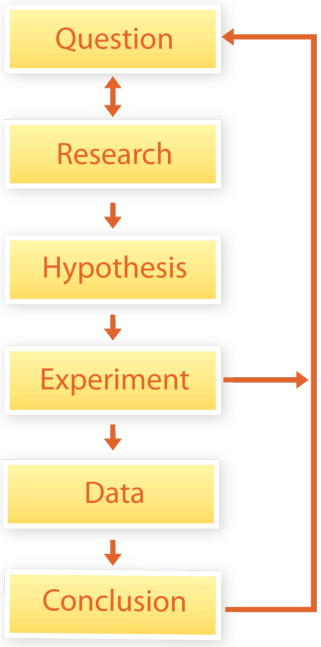
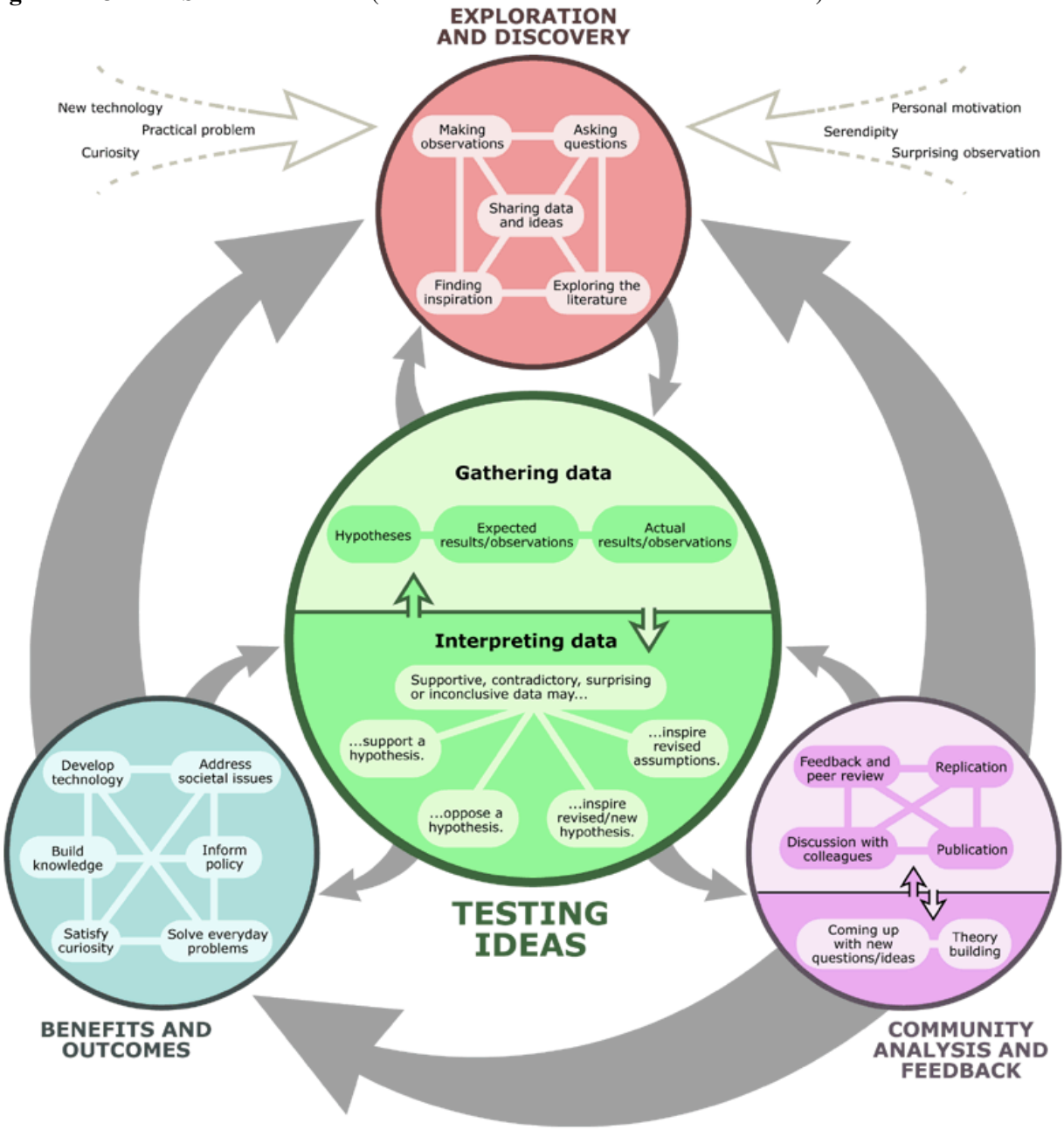


Figure 2: Overall Scientific Process (interactive version on website listed above)



Supplemental Information 3: *List of targeted scientific process activities used for summer research program.* Summer research students varied in the amount of time on campus from 2 weeks to 6 weeks. The students present for 6 weeks got all the following actives whereas the students present only for 2 weeks got a smaller subset.

- Controls
    - Worksheet to identify control group, independent variable, dependent variable
    - Activity to determine possible standardized variables for a purposed research study on plant growth
    - Case study on controls and standardized variables in clinical trails
- Sampling
    - Activity of pulling test results to determine how many samples are needed to reach a conclusion based on blood test
- Bias
    - Worksheet to determine type of bias present in scenarios used to assess amount of teenage smoking (sampling vs measurement vs interpretation bias)
- Experimental Design
    - Case study
        - Autism treatment options: Detecting and dissecting pseudoscience
- Qualitative vs Quantitative Data
    - Activity:
        - Brainstorm types and kinds of qualitative and quantitative data that could be collected on example plants provided
    - Mini case studies
        - Interview qualitative data: Recovering from cardiac survey
        - Turning qualitative data to quantitative data: images to numbers
        - Qualitative Data: Anatomy images
- Scientific model systems
    - Model organisms examples with listing of type of studies performed in the organism for single cell organisms, invertebrates, vertebrates, and plants
    - Results in model organism may or may not match what is being modeled
        - Thalidomine to treat lung cancer: Mice vs Human
        - Testing cholera in C. elegans to reduce vaccine symptoms in humans
- Society and science
    - NIH (National Institutes of Health) and NSF (National Science Foundation) Funding and Scientific Research: Sequester Case Study
    - Judicial Case Study: U.S. top court bars patents on human genes unless synthetic

Supplemental Information 4: *Instructions for rubric use.*

Rubric dimension number 1 assesses the connections. The rater tallies the lines and the arrows that connect ideas in the flowchart counting the lines, the one-direction arrows, and the two-direction arrows independently. For the purpose of this count, two items connected by two arrows going in opposite directions, should be counted as a single two-direction arrow instead of multiple one-direction arrows. The determination of ratings for this dimension of the rubric is partially subjective, as the suggested differentiation between the different ratings does not account for all possible permutations of types and number of connections. A numbered or bulleted list, for example, would have no connections between ideas. Likewise, a combination of lines, one direction, and two direction arrows connecting ideas is not accounted for in the rubric ratings. The delineation between ratings is presented in relative terms, as the goal of this dimension is to see a change in the types of connections used and not necessarily the overall number.

Rubric dimension number 2 is a measure of the numeracy of experimental design ideas including general terms like question, hypothesis, experiment, results, and conclusion. It also includes ideas that are important to appropriate and interpretable experimental design (data analysis, variables, controls, ect). Each set of items for this section need to be counted and placed in its corresponding row. For this section, and any following, each idea (either a word or phrase) can only be counted once. If it is counted as part of the experimental design section, for example, it cannot also be counted in the nature of science section. Each set has a suggestive, but not an exhaustive list of items that would count towards each section. After determining the number of items in each row the rater then would sum them together to determine the overall number of items in the experimental design section. This is placed on the first row in the section. Using this sum of items you would then determine the rating by finding the range of that applies to your calculated value.

Rubric dimension number 3 documents the numeracy of reasons for doing science. The terms listed here are just examples, so other terms may be judged acceptable. For this dimension, the reasons given need to be counted and the rating awarded according to the total tally.

Rubric dimension number 4 measures the student's understanding of the "nature of science." The first sub-category addresses the idea that a problem or a question can not be solved by performing one experiment, and will instead likely need multiple lines of evidence (different types of experiments). The science and society part of this ranking is aimed at the recognition that science is influenced by society and that society influences science. The last criterion for nature of science is a representation that a problem is not solved or evaluated by a single lab, but that there is a role for the whole scientific community. For all of these criteria, the terms listed here are just examples. For this dimension the items pertaining to the listed ideas need to be counted and the rating given according to the total number of items that represent "nature of science" ideas.

Rubric dimension number 5 measures the overall layout of items. This interconnectivity rating documents the flow of the items together. This dimension is only documented by a rating and does not include any items. The rating of 1 is suggested to be a purely linear arrangement of the items but also includes flowcharts that lack arrows. The rating of 2 is an intermediate rating that is generally given if the flowchart isn't completely linear and can include an arrow at the end that returns to the beginning. This occurs, for example, in many depictions of the scientific method where, after the hypothesis is rejected, an arrow points back up to the beginning. The rating of 3 shows a possible organization of the items where they are connected in more of a circular arrangement but is unidirectional and doesn't have non-adjacent items connected. The rating of 4 is intermediate between 3 and 5. The rating of 5 shows a completely connected flowchart with flow represented in multiple directions, no specific start, and items are connected even when they are not next to each other in the flowchart.

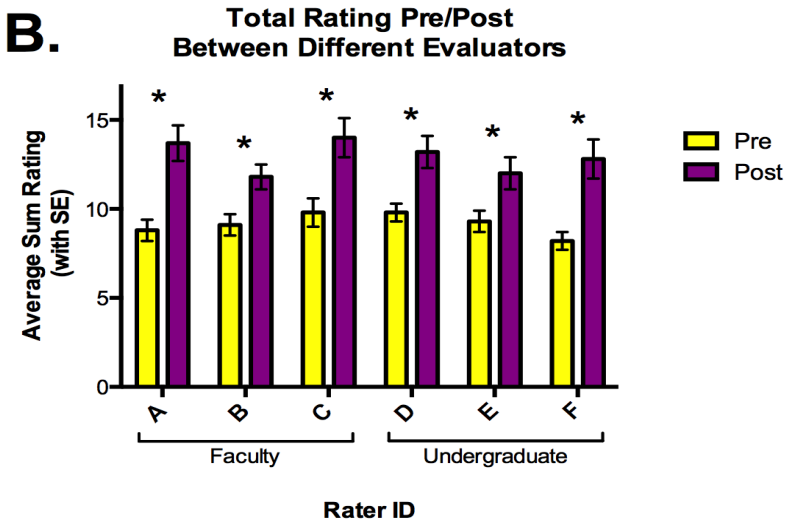
Rubric dimension 6 isn't a stand-alone dimension. Instead, it is an overarching parameter covering the overall item count and rating totals. The first part of this dimension is to sum together all the items from dimensions 1-4; put this total in the first column. The next part is to sum all the ratings for dimensions 1-5; this total goes in the second merged column under the ratings.

Supplemental Information 5: *Rubric has high inter-rater reliability*. (**A**) Calculated variance for rating and item number scores between primary faculty evaluator (evaluator ID A) and other faculty and undergraduate evaluators. A positive variance indicates higher sum rating or sum item number and negative variance indicates lower rating or item number scores. The variance for each flowchart was then averaged to achieve the average variance for each evaluator. The average variance for all evaluators is also displayed. (**B**) Average sum rating for each of 6 evaluators. (**C**) Average sum item number for each of 6 evaluators.

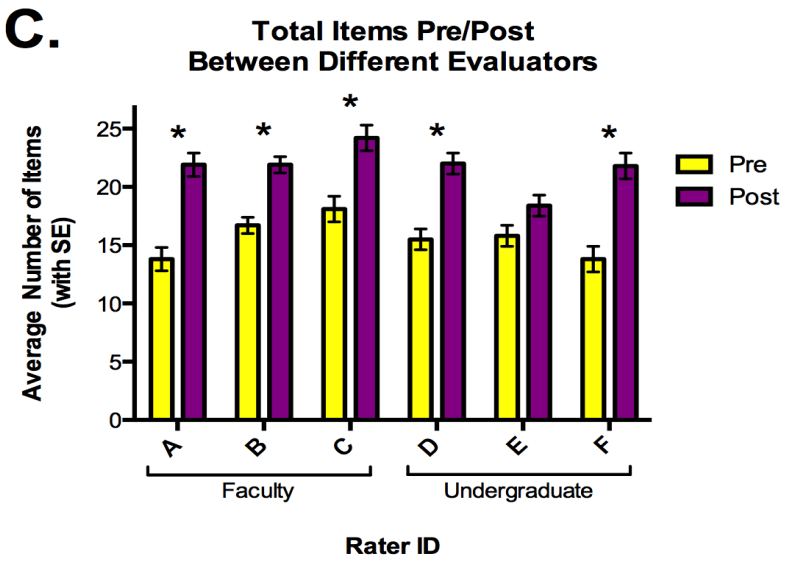
A.

	Evaluator ID					Average Variance
	Faculty		Undergraduate			
	B	C	D	E	F	
Sum Rating	-0.8	0.7	0.3	-0.6	-0.7	-0.2
Sum Item Number	1.2	3.3	.08	-0.8	-0.1	0.9

B.



C.



Supplemental Information 6: *Example flowchart with multiple structures*. These examples are from the multidisciplinary sample. Each structure present is boxed in blue. The distinction that makes a flowchart have multiple structures is having terms or phrases not connected.