Improving Undergraduate Scientific Explanations: Exploring the Role of Data Literacy Skills in Scientific Reasoning

Kathleen Browne, Andrea Drewes, Gabriela Smalley (Rider Univ.), and Sage Lichtenwalner (Rutgers Univ.)

Abstract. Undergraduate students' scientific reasoning skills are challenging to develop. Evaluating data is essential, yet a conceptual task found by many students to be very difficult. This project explores strategies to address gaps in the knowledge base that provides guidance about practical pedagogical approaches to developing students' scientific reasoning. We are testing the effectiveness of an instructional framework designed to help students link scientific reasoning with data literacy skills to improve their proficiency in composing evidence-based scientific explanations: Data Description-Claim-Evidence-Reasoning **(DCER).** The framework will be used:

- to create, test, and revise instructional strategies and assessments;
- in introductory oceanography classes at Rider University, NJ;
- with data visualizations of authentic data from NSF-funded Ocean Observatory Initiative (OOI) and other data sets;
- to test interventions that guide students to routinely
  - ✓ describe data (trends, patterns, ranges, outliers, similarities, differences, etc.) **[Data Descriptions-D]**
  - ✓ draw conclusions about the data and relevant phenomena, **[Claim-C]**
  - ✓ and support those conclusions with scientific reasoning that includes proper evidence tied to the students' understanding of relevant science concepts (**Evidence-E** and **Reasoning-R**).

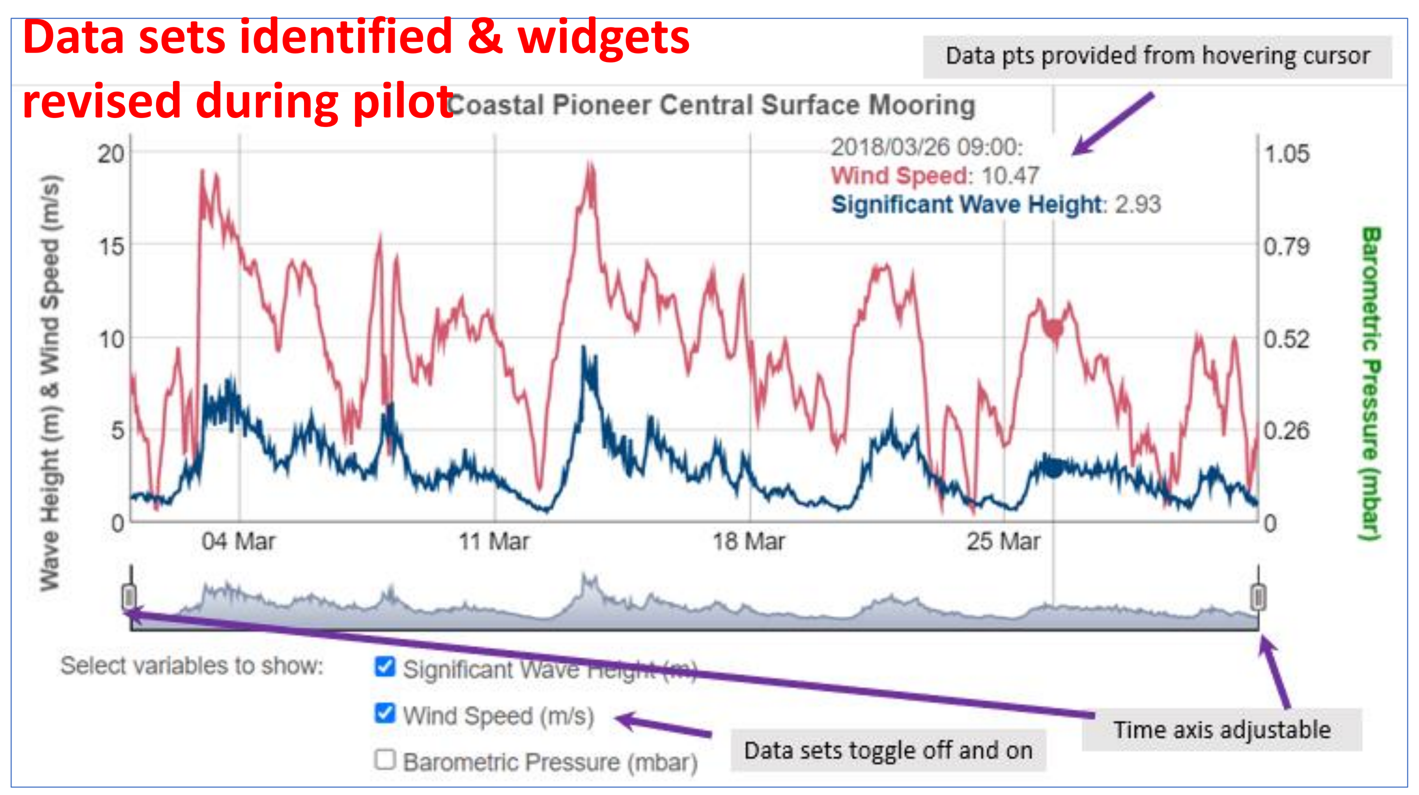
Study Timeline

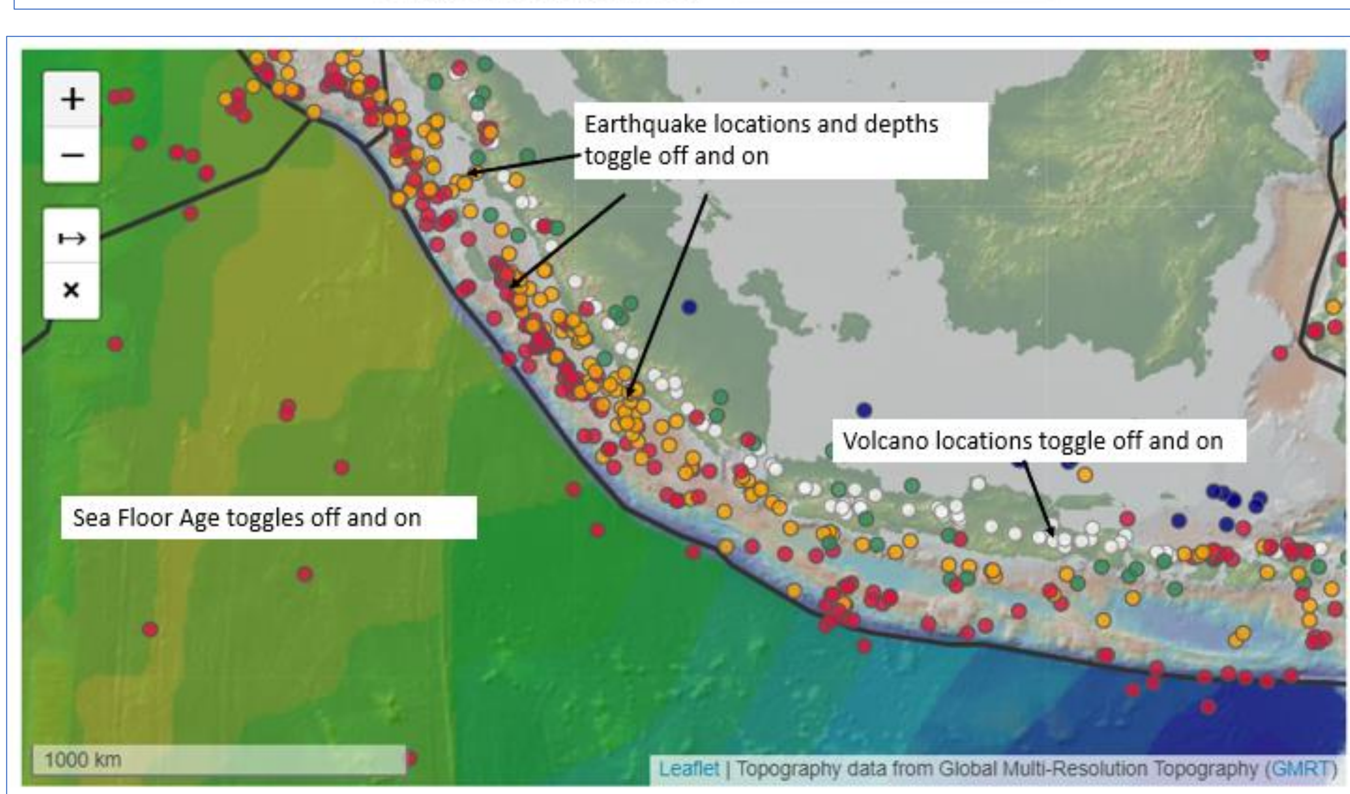
Year 1 (2020-2021; fall, spring and summer): lesson plans, interactive data visualizations, interview protocols, and assessments piloted and revised

Year 2 & 3 (2021-2023): Data collection (fall & spring semesters) & Analyses (summers)

Example Web-based Interactive Data Sets

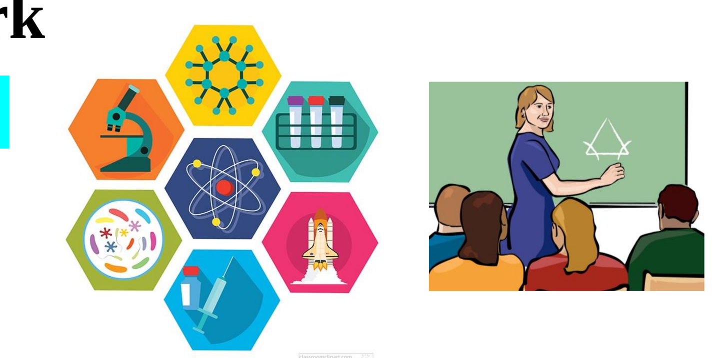
Data sets identified & widgets revised during pilot





DCER Instructional Framework

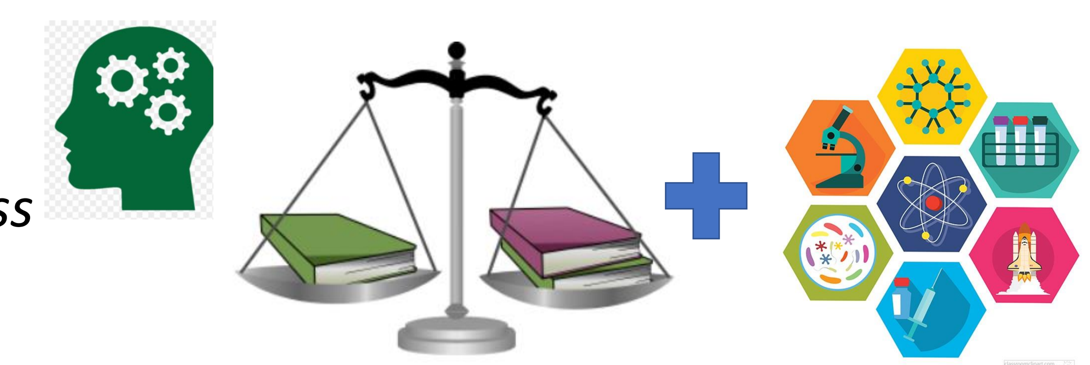
Instruction to guide learning about ocean concepts.



Instruction to guide students to study data and compose descriptions of patterns and variability in authentic data sets.

Students draw conclusions from the data, discuss their ideas and reasonings to tie specific quantitative evidence to their understanding of relevant science concepts with guidance as.

Finally, students compose individual explanations in class "data studies" & exams.



Student S1B03 Results

Exam 1. Between the 4 data sets I realized that there is a lack of pattern. There are many shallow earthquakes within the area but not very many volcanoes at all. The sediment is mostly thin with some thicker areas and the sea floor is about half extremely old and half younger. These data sets tell me that the plates in this boundary have definitely shifted a good bit in order to form the many shallow earthquakes that have occurred within the area circled in the top left map. After considering all of this information and taking a close look at each data set, I believe that the plate boundary that occurs in the Indian Ocean is a transform boundary. Transform boundaries include many shallow earthquakes and very little volcanoes. Since the sea floor in this area is very old, there have been many shifts between the plates over the million years that it has been there and that is why there are shallow earthquakes across this area. The lines that the earthquakes form in the top left map are the areas that the plates slid past each other under the ocean. All of this evidence between earthquakes, volcanoes, sediment thickness, and seafloor age, lead me to believe that in the maps above of the Indian Ocean there is a transform boundary.

	D	C	E	R
Total 4.5	1.5	1	1	1

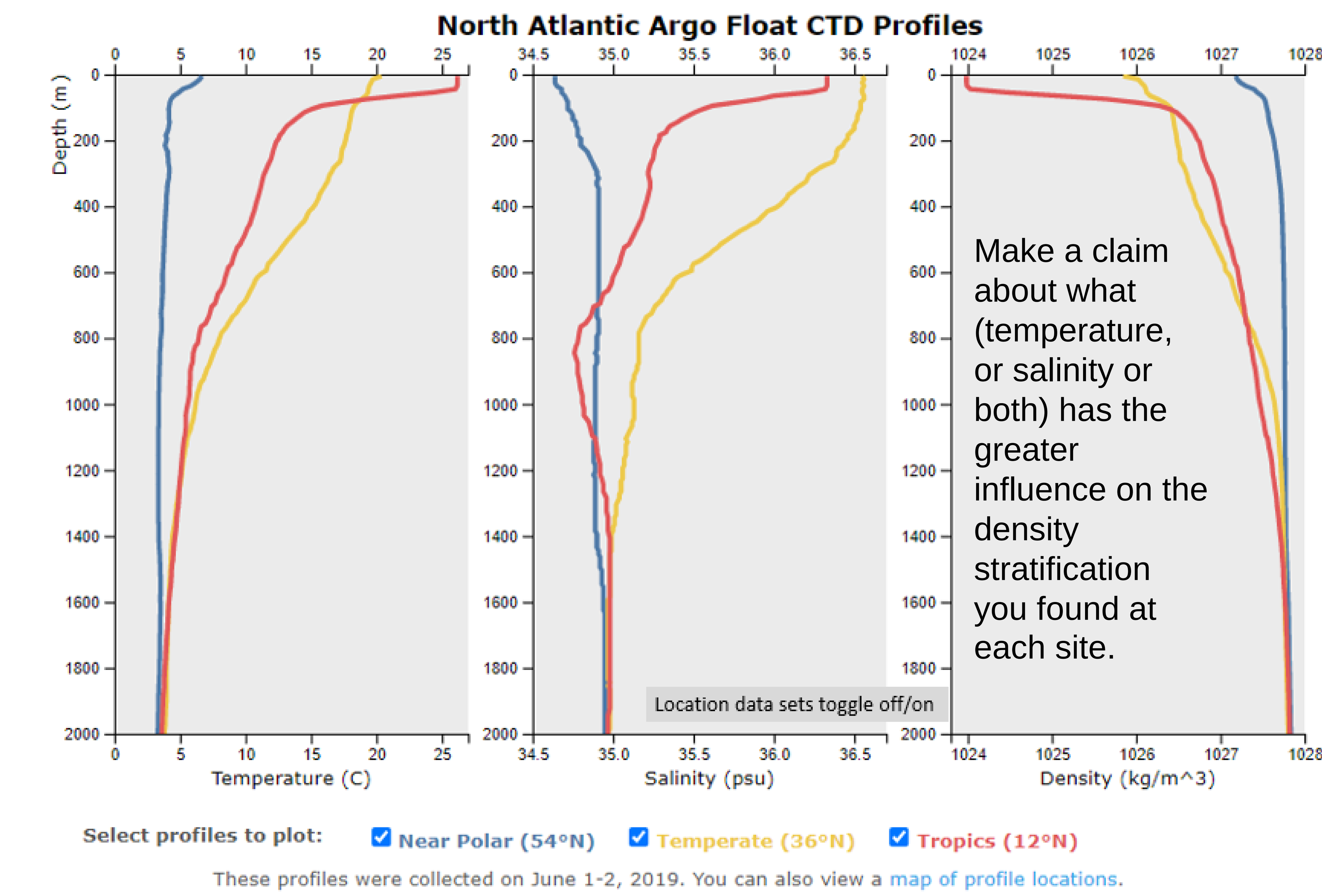
Exam 2. A) The annual trends of temperate primary production are that in the cold-warm months of spring and warm-cold months of fall primary productivity is higher. These trends change annually because the temperate area is not always hot or cold, it changes. The annual trends of tropical primary productivity are that the trend stays the same annually. This is because the tropical region stays warm-hot temperatures.
B) These regions' annual trends are different because of their different temperatures and the amount of sunlight as well as nutrients that are able to work together at the surface in order to allow primary producers to photosynthesize.
C) My conclusion stated in B relates to A because the primary producers in the temperate and tropic regions are not able to thrive in the same ways. For example, the trends in the temperate area show that chlorophyll is at its highest during the cold-warm spring because that is when there is no thermocline layer in the ocean, since the surface water is still colder during those months. Between March and April in the temperate region, chlorophyll is highest because nutrients are able to move to the surface and there is enough sun to allow for photosynthesis. The tropic region is different because the temperature is warmer there annually, which is why there is no change in primary productivity. The warmer temperatures in this region keep the surface water of the ocean warm, which creates the thermocline layer and blocks nutrients from moving to the surface. Although there is lots of sun in this region year-around, the primary producers cannot photosynthesize unless they have both sunlight and nutrients.

	D	C	E	R
Total 8.5	1.5	3	2	2

Rubric to assess explanations (red = revisions made to the rubric from pilot)

Proficiency level	Descriptions	D w/out R? Y/N	Conclusion	Evidence	Reasoning
	First, thoroughly describes all trends, patterns, similarities, and/or differences etc. ("T/P/D/S etc.") in the data .		*A statement that answers the original question/problem. If no question is provided, students make their own conclusion from the data studied.*	*Scientific data that support the conclusion. The data need to be appropriate and sufficient to support the conclusion.*	*A justification that connects the evidence to the conclusion. It shows why the data count as evidence by using appropriate and sufficient scientific principles (addressed in this class); also includes reasoning for any data that are not relevant to the conclusion.*
0	None provided; completely irrelevant		None provided; completely irrelevant	None provided; completely irrelevant	None provided; completely irrelevant
1 (weak)	Provides few "T/P/D/S etc." descriptions or mostly inappropriate descriptions and/or level of detail of descriptions.		Provides an inappropriate conclusion given the data used; or restates a data description.	Provides very little (when more is present) and/or inappropriate evidence (evidence that does not support the conclusion). 1a very little evid; 1b inaccurate evid	Provides very little or inappropriate reasoning. 1a - only incl. evidence; 1b - only incl. sci principles; 1c - both evid and principles but are connected inaccurately
2 (passing)	Provides some appropriate "T/P/D/S etc." descriptions; but not all; or level of detail is inappropriate for some described.		Provides an appropriate, but insufficient conclusion. If only one conclusion is needed, this score could be used where students reference the proper feature etc. but does not name it specifcially (etc.).	Provides appropriate, but insufficient evidence (given all data present). Or provides sufficient, but some inappropriate evidence. 2a appro but insuff; 2b suff but inappro	Provides reasoning that connects the evidence to the conclusion. Some, but insufficient scientific principles or justification for why the evidence supports the conclusion.
3 (strong)	Completely describes all "T/P/D/S etc." at an appropriate level of detail. With quantitative details included when available. If interps or explanations included, they are ignored.		Provides an appropriate and sufficient conclusion.	Provides appropriate and sufficient evidence that includes some relevant specific quanititative information and pattern descriptions.	Provides reasoning that connects the multiple pieces (when available) of evidence to the conclusion. Includes appropriate and sufficient scientific principles to explain why the evidence supports the conclusion.

Example Data Set and In-Class Exercise Instructions



For each of the 2 sites, tropics and polar, follow the questions below to draft thorough written descriptions of data patterns..

In-class activity completed by groups; last question completed by individuals.

Temperature, Salinity, Density

Complete in the table below: What are the minimum and maximum temperature values for the entire profiles? Subtract the bottom from the surface values to calculate the range in temperatures. Do you see a thermocline in the data (Y/N) and if so, at what depths does it occur and what is the range in temperature values in that layer?

	Min & Max	Calculated range	Thermocline, Halocline, or Pycnocline (Y/N)? If so, depths and range in temperature?
Tropics			
Polar			

ii) Describe in detail how the Temp varies from surface to bottom for both profiles including specific quantitative data for both T and depth.

- Tropics:
- Polar:

Condensed instructions to illustrate some of the scaffolding built into work; supplemented with verbal guidance and discussion; instructions piloted and revised in Yr 1

Data comparisons

i) Compare the ranges in salinity, temperature and density for each site; which site has a greater range of each parameter?

ii) For each location (polar and tropics), describe any similarities or differences in the depths of each "-cline".

- Tropics:
- Polar:

iii) Compare the pycnocline for both polar and tropics. How do the depths compare and how do the thicknesses compare?

Complete *individually* in Canvas assignment (compose individually in a word or similar file, then upload it into the T/S/D data study "final" draft answer assignment. Restate your claim. Then, using all relevant information from your work above, compose a scientific explanation that supports your claim. Remember, scientific explanations should tie together all relevant specific data patterns **and your understanding of relevant science concepts introduced for this topic that would explain your claim** in a logical sequence (your "reasoning"!).– *Your answer should take several well integrated sentences to complete this question.*

When you complete your answer, highlight text with the following color coding:

Conclusion/Claim
Yellow- data summaries relevant for your claim ("evidence")
Blue - science background concepts relevant to explain your claim
Green- - text that connects evidence to science concepts