Electronic Supplementary Information (ESI) for:
Supporting Submicroscopic Reasoning in Students' Explanations of Absorption Phenomena Using a Simulation-Based Activity

Natalia Spitha†, Yujian Zhang†, Samuel Pazicni*, Carla Morais, Sarah A. Fullington, Amanda R. Buchberger, Pamela S. Doolittle

† *Both authors contributed equally to this work.*
**Corresponding author*

Table of Contents

Section 1: The Beer's Law Simulation activity

S1.1. The BLSim Tool

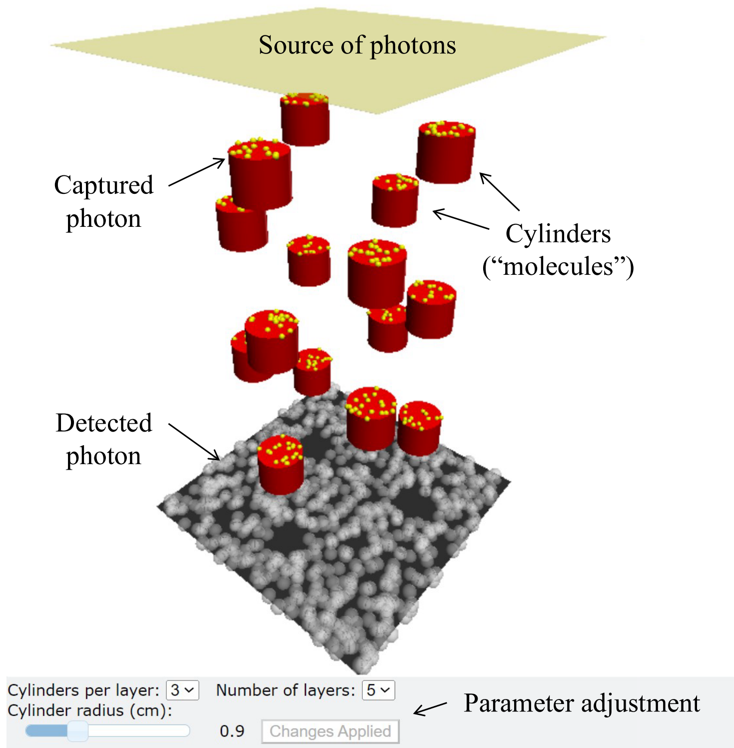


Figure S1. Demonstration of main components in the Beer-Lambert Law simulation tool (annotated screenshot from https://bit.ly/BeerLawSim). The number and cross-sectional area of the cylinders, as well as the number of layers on which cylinders are arranged, can be adjusted in the simulation, which students are instructed to do through various prompts (see Spitha et al., 2021).

S1.2. BLSim Activity Prompts

In a guided-inquiry format, this lab uses a simplified example of what is happening in the cuvette of a spectrophotometer to allow you to provide you a molecular understanding of the (a) fundamental process of absorption and (b) mathematical derivation of Beer's Law.

BACKGROUND

Beer's Law (also known as the Beer-Lambert Law) describes the relationship between the amount of light absorbed by a solution and its concentration. The *absorbance* (A) of a solution of concentration *c* in a cuvette with path length *b* is given by

$$A = -\log\left(I/I_0\right) = \varepsilon bc,$$

where I_0 is the intensity of light entering the solution, *I is the intensity of the light exiting the solution, and* ε is a constant called *molar absorptivity*, which is unique to each solute/solvent system. By now, you have probably held a cuvette and used it to measure the concentration of a solution using the principle of Beer's law. But can you explain exactly what is going on in the cuvette at the molecular level? To help you visualize the process of absorption as well as provide a quantitative basis of what is occurring in the cuvette, you will be interacting with a simulation that was hand-crafted here at [institution redacted].

The applet used in this lab simulates light going through a cuvette (*i.e.,* a sample cell) by dropping small spheres ("photons") through a box filled with cylinders ("solute particles") arranged in a discrete number of layers. When a photon touches the surface of a cylinder, it gets stuck, emulating the absorption of light by a solute. The photons are arranged randomly along a single (x-y) plane and can move uniformly through the sample cell (z-direction). After "crossing" each layer, their positions are always re-randomized along the x-y plane. The photons that are not captured by any cylinders make it through (are transmitted through) the cuvette and are detected by the screen on the bottom (the detector).

On screen, you will see a few buttons (at the top), values you can edit (on the bottom), and a running log about the data obtained from the simulation (very bottom). Every time you reload the simulation/web page, this box will be cleared.

The following controls will be useful:

- **Run/Pause**, which starts or pauses the simulation
- **Reset Photons**, which places the photons back at the "source," allowing you to repeat the experiment. No need to click "Run" again if the simulation is already running.
- **Reset View**, which resets the original view of the box (in case you zoomed in too much and got lost.)
- **Update Cylinders**, which applies any changes you made to the number of layers or number/radius of cylinders and rearranges them in the box.
- **Grid on/off**: Enables/disables a grid visualizing the box. Each square in the grid is 1 centimeter.
- You can also rotate the view of the cylinders by right-clicking and dragging anywhere inside the image.

It is recommended that, before starting the lab, you play around with the controls. See what values are given to you; try to resize the simulation. For good measure, see happens when you don't click **Update Cylinders**. If you understand the simulation controls, it will make it MUCH more efficient for you when collecting data from the simulation. You can do this as part of your pre-laboratory preparation!

Whenever we are collecting data in general, we have always wanted to collect replicates to understand how random error can impact a measurement. Thus, when assessing each set of conditions, make sure you perform at least 5 trials with 1000 photons each (**7–8 trials is recommended**). Also, as you know from learning the scientific process, there are logical steps for you to follow when trying to analyze a problem. Thus, for each of the questions below, using your lab notebook, we recommend the following, already familiar steps:

- Create data tables to collect/log data (*e.g.,* Figure 1)
- Collect data
- Perform calculations
- Critically evaluate the answer (*e.g.,* Does the result make physical sense?)
- Phrase the answer in a sentence

	Radius:		Radius:		Radius:	
Trial #	absorbed		absorbed		absorbed	
1						
2						
3						
4						
5						
6						
7						
8						
Avg.						
St.Dev.						

Example data collection table for experiments run with the simulation.

BEFORE YOU TAKE THE QUIZ

- Watch all associated video content to introduce you to the simulation.
- Understand the basics of absorption spectroscopy, including the instrumental components, such as the light source, the cuvette, the sample, and the detector.
- Recognize Beer's Law and be able to define what each variable means.
- Be able to convert between transmittance, percent transmittance, and absorbance.

PRE-LABORATORY EXERCISES

1. Briefly open the Beer's Law simulation URL (bit.ly/BeerLawSim, case-sensitive!) and take a look at the cylinder illustration and controls. This simulation serves as a model you can use to explain the Beer-Lambert relationship between light intensity (I/I_0), concentration (c), molar absorptivity (ε), and path length (b). Identify and write down in your lab notebook how you think the quantities I, I_0, c, ε, and b are represented in the simulation (that is, which properties of the spheres, cylinders, and box relate to each quantity).

EXPERIMENTAL

Open the Beer's Law Simulation URL. Use the simulation to perform the virtual "experimental trials" listed below, remembering to record your observations, calculations, and predictions in your notebook when prompted (usually, by underlined text). While the post-lab questions are at the end (*RESULTS/CALCULATIONS*), you are welcome to answer and work through post-lab questions as you go along in the lab.

1. For 1 layer with 2 cylinders, record the photons absorbed for *at least three* different values of the cylinder radius with 1000 incident photons for each radius. Make sure you conduct enough trials. See an example data table in Figure 1. In the shaded areas of the table, you may choose another variable that you want to measure or keep track of (*e.g.,* cylinder cross-sectional area, % transmitted, *etc.*).

> *In your lab notebook*: Describe any relationship you observe between the radius of the cylinders and the number of absorbed or transmitted photons.

2. *What do you predict the transmitted number of photons to be for 3 cylinders of radius 1.8 cm on one layer? Explain the reasoning for your prediction.*
   Test your prediction by performing another set of trials.

> If you find it difficult to identify a relationship with just your current sets of trials, or if your prediction was significantly off, feel free to select other radii or numbers of cylinders and run more trials until you feel comfortable you understand how the arrangement and size of cylinders affects the photons transmitted through a layer.

3. For the same number and radius of cylinders per layer (3 cylinders, 1.8 cm), *what do you predict the number of transmitted photons to be when you add a second layer for the photons to pass through? How about when a third layer is added? HINT: Remember that the photons captured by the first layer do not continue to the next layer!* Test your hypothesis by performing another two sets of trials.
4. Run a set of trials by setting the number of layers to 5, choosing a fixed cylinder radius > 1.5 cm, and varying the number of cylinders per layer. Remember to conduct enough trials and to make a table!
5. If you make a plot with the number of cylinders per layer on the x-axis, which dependent variable do you need to plot on the y-axis to obtain a *linear, positive* relationship according to Beer's Law? Obtain 4 data points and test the linearity of your plot. *HINT: the y-variable should NOT just be the number of photons absorbed!*

RESULTS/CALCULATIONS

Fill out the answer sheet for this experiment completely. Create a single PDF of your answer sheet and images of your notebook pages and submit to Canvas. Answer the following post-lab questions as precisely as you can in your lab notebook:

1. From the data collected in your Experimental step 1, determine a mathematical relationship that you can use to calculate the expected fraction of photons absorbed by one layer, given the number and radius of the cylinders as well as the total layer area. Explain why this relationship holds. HINT: What is the probability that a randomly positioned photon will get captured by a cylinder?
2. How does your prediction from Experimental Step 2 compare with the result obtained? Using your result from question 1, explain why your prediction does or does not agree with the results. How do your predictions from Experimental Step 3 compare with the results obtained? Again, using the mathematical relationships you have discovered so far, explain why your prediction does or does not agree with the results. Using the mathematical relationship(s) you have determined, calculate how many transmitted photons one would expect to get in Experimental Step 3 (3 cylinders/layer, 1.8 cm) if (i) a fourth or (ii) fifth layer was added (you don't need to test this with the simulation). Sketch a plot of the expected number transmitted photons against the number of layers. How would you describe the shape of that plot?
3. Actual solutions have a continuous (rather than discrete) arrangement of molecules across their length. Let's now put things in mathematical terms, to see how your above discoveries about cylinders neatly arranged in discrete layers can apply to a continuous situation. Consider the following variables:

- N = number of cylinders *per* cm^3 (not per layer!)
- σ *(sigma)*= cross-sectional area of each cylinder in cm^2
- *A = cross-sectional area of each layer*
- h = height of one layer
- b = total height of box
- *I = number of photons transmitted*
- I_0 *= total number of photons*

Using the above variables:

i. Express the *volume* of each layer.
ii. Express the number of cylinders per layer.
iii. Express the total cross-sectional area occupied by all cylinders in one layer.

4. Based on your observations in Experimental steps 1-3 and Post-lab Question 5, what is the *fractional* change in number of transmitted photons ($\Delta I/I$) passing through a **single layer** of thickness h? (HINT*: It will be negative since photons are 'lost' via absorption*.)

5. Now consider your answer to Post-lab Question 6 for an infinitesimally small change in number photons dI when passing a very thin layer of thickness dh. To find the number of photons transmitted through the entire sample (thickness b, initial number of photons I_0), integrate both sides of that expression with respect to dI and dh. (HINT*: This integration should give you Beer's law!*)

- Math Help #1: $\int_{x=a}^{b} k\, dx = k(b-a)$
- Math Help #2: $\int_{x=a}^{b} \frac{1}{x} dx = \ln(b) - \ln(a)$
- Math Help #3: $\log_a(x) = \frac{\log_b(x)}{\log_b(a)}$
- Thus, Math Help #4: $\log_{10}(x) = \frac{\ln(x)}{\ln(10)}$

6. From Experimental step 5, comment on the linearity of your collected data.
7. How would your plot in Experimental step 5 change if the cylinder radius was increased? What property would this change be analogous to in a solution following Beer's law?

CHALLENGE QUESTIONS FOR YOUR LAB NOTEBOOK

Provide answers to the best of your ability to the following questions. Freely share your ideas and predictions. Any honest attempt to answer the questions will be rewarded full points!

1. After completing this activity, have your answers the Pre-Lab Question 1 changed? If so, revise your answer here. If not, write "See Pre-Lab Question."
2. Discuss to what extent the variables of Beer's Law (intensity, concentration, absorptivity, and path length) are accurately represented in the simulation. Specifically consider the context of an absorption measurement. Can you propose any improvements to how the quantities are represented?
3. Do you have any additional improvements to propose for the simulation?

Section 2: Course differences and their lack of impact on domains of reasoning

The comparison and simulation group in this study each include students from two different analytical courses (“Course 1” having most students majoring or interested in health- or biology-related fields, and “Course 2” having a higher quantitative focus and being primarily targeted at chemistry and engineering majors). In the time leading up to the BLSim activity and the spectroscopy assessments (**Figure S2**), both courses followed a very similar lecture curriculum (led by two different instructors), with students watching pre-recorded lecture videos and attending a synchronous online session to ask questions and/or solve problems with the lecturer. The topics that had been covered in lecture by the time point of the study were measurement and statistics (4 lecture periods for Course 1 and 5 periods for Course 2) and spectrophotometry (4 lecture periods for both courses). In addition to the Beer-Lambert law and its derivation the spectrophotometry lectures discussed concepts such as energy-wavelength relationships, absorption and emission spectra, Jablonski diagrams for representing absorption and emission, and the structural components of spectrophotometers for various applications.

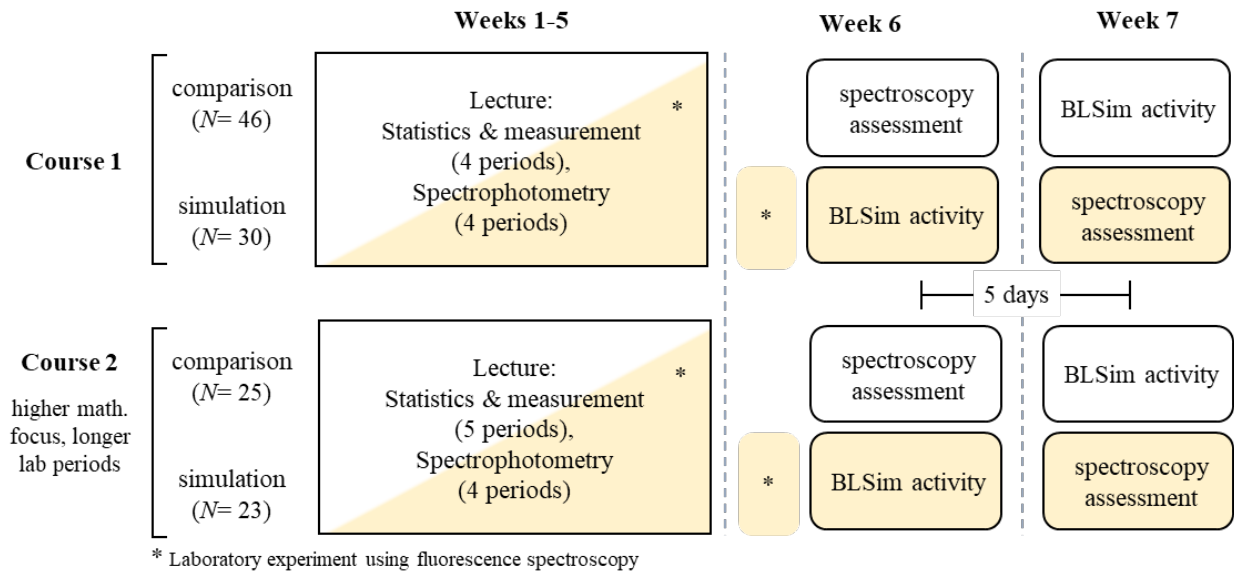


Figure S2. Timeline of spectroscopy-related lecture (rectangles) and lab (rounded rectangles) activities in Courses 1 and 2 leading up to the BLSim activity and the spectroscopy assessment.

Due to the division of students in sections between the two courses, the comparison group had a higher relative representation from students in Course 1 (**Table S1**). To evaluate whether the difference in the “composition” of student interests would have an influence on the differences in reasoning observed between the comparison and simulation groups, we compared the domains of student reasoning (e.g. submicroscopic vs. non-submicroscopic) exhibited by students of either course in the comparison group. If the Course 1 students within the comparison group show a similar distribution of reasoning types as the Course 2 students within the same group, it is reasonable to rule out course differences as a factor for any differences observed between the comparison and simulation groups.

Table S1. Representation of Course 1 and Course 2 in Comparison and Simulation groups

Group	Course 1	Course 2
Comparison	46 (65% of group)	25 (35%)
Simulation	30 (57% of group)	23 (43%)

The types of student reasoning for prompts Q1-Q4 were determined through the coding process described in the main text of the manuscript. **Figure S3** summarizes the code counts for each type of reasoning per course and prompt. A statistical evaluation of this breakdown is not possible due to the low number of counts for specific codes; however, the distribution of codes appears to be qualitatively very similar for students in both courses.

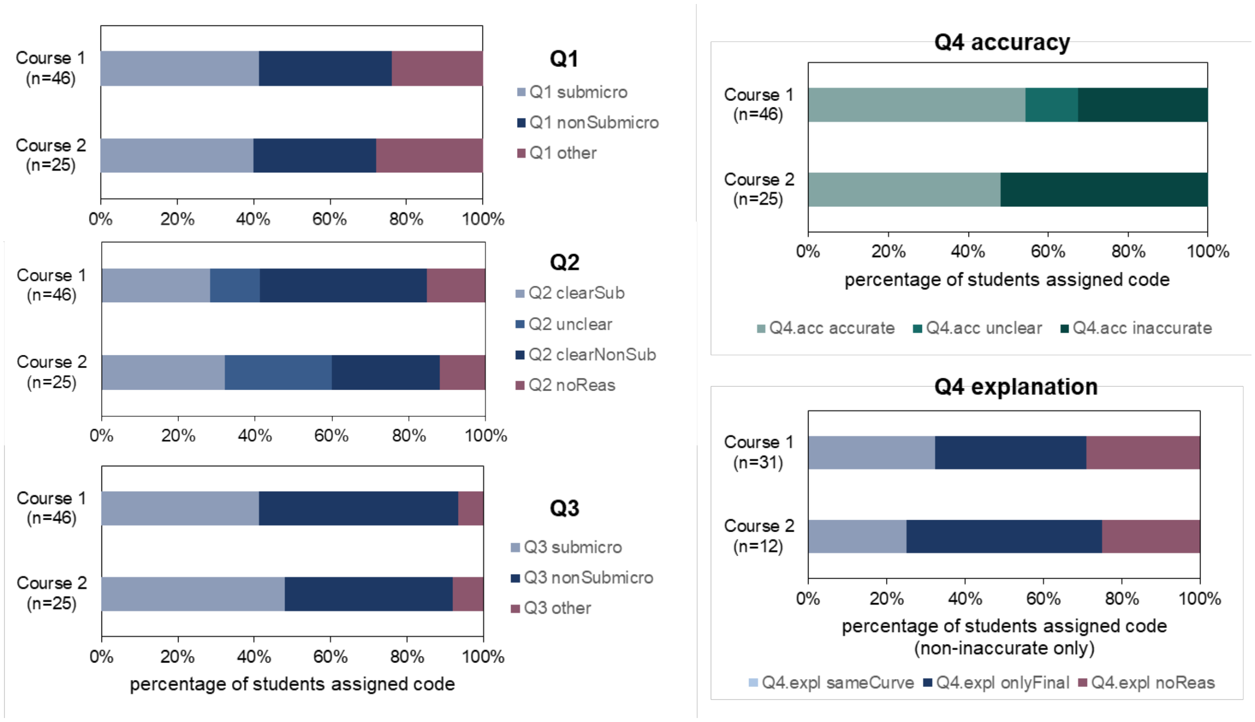


Figure S3. Breakdown of codes observed in the comparison group for prompts Q1-Q4, separated by course.

Section 3: Spectroscopy assessment prompts and their evaluation with the 3D Learning Assessment Protocol (Laverty et al., 2016)

Note the correspondence between the prompt numberings in the main text and the original assessment exercise numbering: Prompt Q1 = Exercise 1b, Prompt Q2 = Exercise 1c, Prompt Q3 = Exercise 2b, Prompt Q4 is Exercises 3a-b.

Design principles and validity of the assessment

The spectroscopy assessment (with three multi-part questions centered on the concentration, absorptivity, and path length variables of the Beer-Lambert Law, was developed intentionally according to the principles of the 3D Learning Assessment Protocol (Laverty et al., 2016). Each question was examined to fit the criteria outlined in the 3D-LAP for having the potential to elicit certain scientific practices. For example, in order to elicit the scientific practice of "Developing and Using Models," Exercise 1 provided students with a phenomenon (light entering and exiting a cuvette), asked them to draw a representation, and asked them to use the representation to explain a claim (i.e. the shape of the intensity profile). To establish response process validity for this assessment, think-aloud interviews (covering all exercises) were conducted with three volunteer analytical chemistry students at the end of the Fall 2020 semester, establishing that the questions were being interpreted as-intended by the researchers.

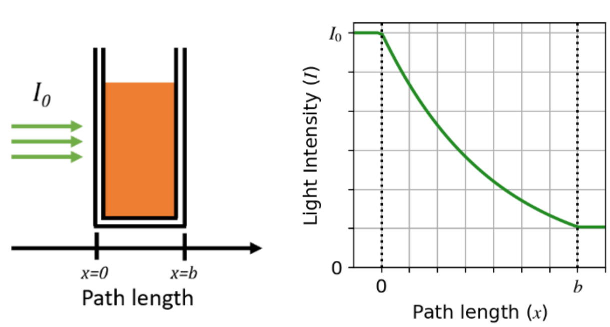
Scientific Practice elicited by each question (color-coded):

- Developing and Using Models
- Analyzing and Interpreting Data
- Using Mathematics and Computational Thinking
- Constructing Explanations and Engaging in Argument from Evidence

Challenge Questions on S*pectrophotometric Determination of Iron* lab.

Preface: This lab highlights the usefulness of Beer's law in accurately determining the concentration of iron in solution. The following Challenge Questions ask you to delve a little deeper into the individual "components" of Beer's law (such as light intensity, concentration, and path length) and consider in more detail how light interacts with a solution. While some of these questions may seem challenging, please answer them to your best ability and note that there are multiple ways of approaching them. As long as you show your reasoning, any honest attempt to answer each of these questions will be awarded full points.

Exercise 1: The intensity profile of a beam of green light, as it passes through a cuvette containing a 0.01-M solution of an iron (II) - 2,2-bipyridyl complex, is shown in the plot below:

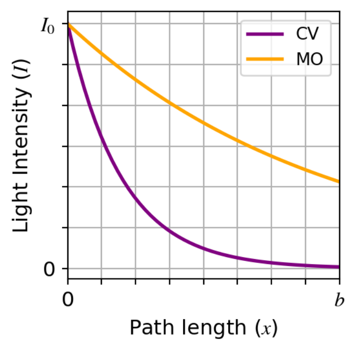


a. What information does the graph tell you about the evolution of the intensity profile of the beam through the solution? Specifically commenting on the sign(s) and magnitude(s) of the slope of the graph, describe how the intensity of light changes before, through, and past the solution.

b. Draw a picture in your lab notebook that depicts how you think light interacts with the iron complex solution. Use that illustration to explain why the intensity profile follows the shape shown above. You can ignore any effects of the solvent on the intensity of light.

c. Explain or show how your illustration in part (b) changes when a 0.02 M iron complex solution is instead illuminated with the same beam of light. Then, in your lab notebook, reproduce the intensity plot shown for the 0.01 M solution and, on the same graph, predict and sketch the intensity profile through the 0.02 M solution.

Exercise 2: Two equimolar dye solutions are excited by beams of light of the same energy. The first solution contains Crystal Violet (CV), while the other one contains Methyl Orange (MO). The intensity profile of the beam through each of the solutions is plotted below.

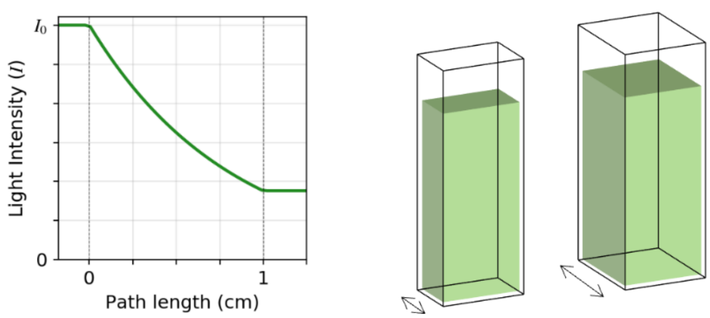


How is it possible that two solutions of the same concentration exhibit different intensity profiles under the same excitation conditions? Provide:

a. A mathematical explanation based on Beer's law.

b. An explanation based on the molecular behavior of CV and MO. To support your explanation, make sure to draw an illustration of how you think each solution interacts with light.

Exercise 3. Monochromatic light, passing through a 1-cm cuvette containing a 100 μM iron solution with molar absorptivity of 8.5×10^3 cm^{-1} M^{-1}, exhibits the intensity profile shown on the left:



In general and analytical chemistry labs, it is most common to use 1-cm cuvettes to run spectroscopy experiments. However, in experiments involving precious reagents, it is sometimes common to use thinner cuvettes (see right figure), which require a smaller volume of analyte.

a. Reproduce the graph above in your lab notebook and sketch on it the intensity profile through an identical solution placed instead in a cuvette with length b = 0.5 cm. You can assume the cuvette walls are infinitely thin and transparent.

b. Explain your reasoning for the curve you drew in part (a).

c. Calculate the fraction of incident light intensity (I/I_0) that exits the 1.0-cm-thick and the 0.5-cm-thick cuvette containing the same 100 µM iron complex solution. How do the values of I/I_0 you obtain for each cuvette thickness compare to your graphical estimate from part (a)?

d. If we want to conserve the volume of analyte used in spectroscopy experiments, why not go all the way? Calculate the minimum cuvette thickness that would still allow you to perform a reliable absorbance measurement of the 100 µm iron solution above with your spectrometer. You can use the following information and assumptions:

- Your light source produces 1.000 W of light at the wavelength of interest.
- You have calibrated the spectrometer with a clean reagent blank that let ~100% of the incident light intensity (I_0) be transmitted (consider this to apply to measurements with any cuvette thickness).
- The detector can reliably measure changes in intensity down to 0.005 W.
- The cuvette walls are infinitely thin and transparent.

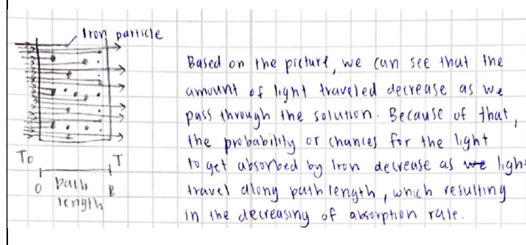
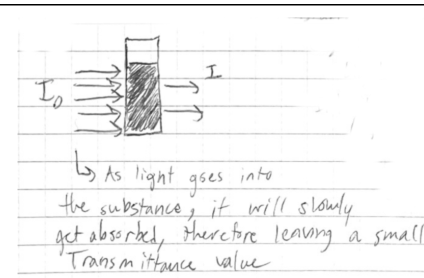
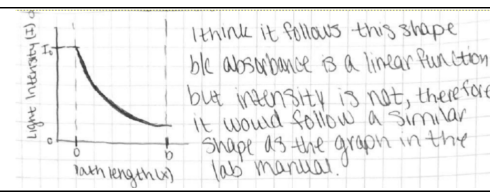
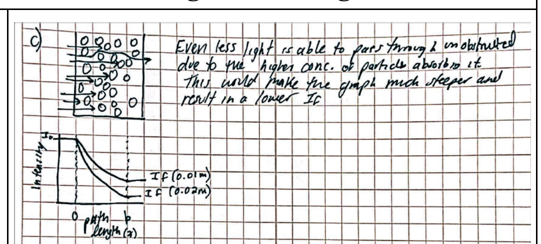
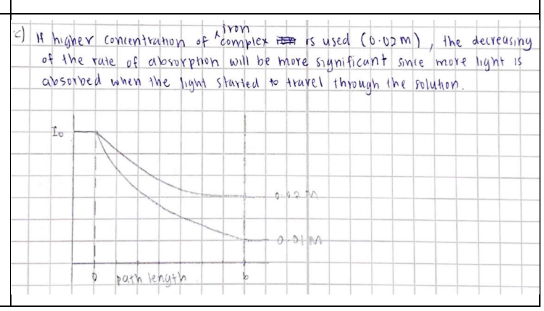
To construct your answer, you first need to establish what criteria define a “still-reliable measurement under minimum cuvette thickness.” In your answer, please address the following questions along with any calculations:

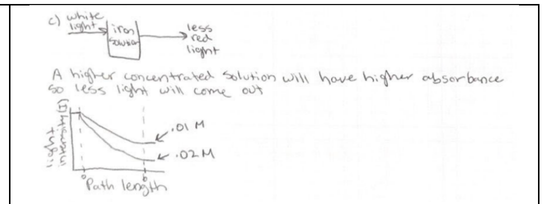
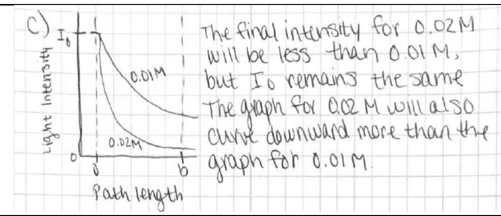
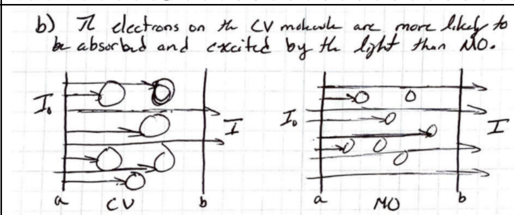
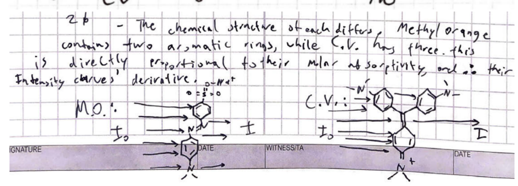
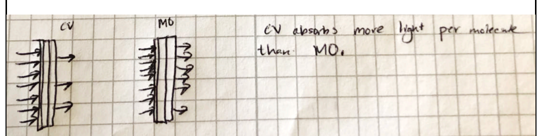
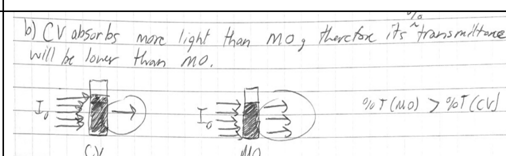
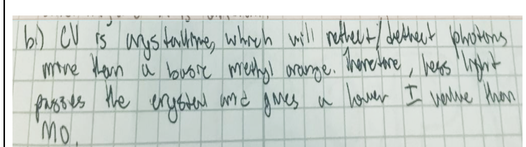
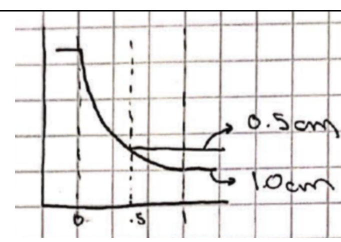
i. Which quantity in your experimental measurement is maximized or minimized when the cuvette thickness is minimized?
ii. What condition must this quantity meet to ensure the absorbance measurement is reliable?
iii. What is the cuvette thickness under this condition?

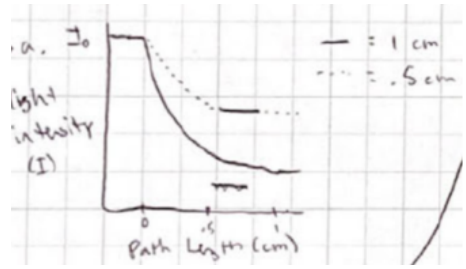
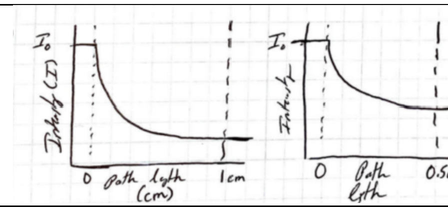
Section 4: Codebook, Flowcharts, and transcriptions

S4.1. Codebook with examples

Table S2. Codebook with examples. Transcriptions are shown below.

Task	Code abbr.	Type of reasoning represented	Example coded segment(s)
Q1	*submicro*	submicroscopic reasoning such as depiction of light rays hitting particulate representations of molecules as they pass through the cuvette (rays hitting *individual* molecules should be shown)	
	nonSubmicro	macroscopic depiction of light entering/exiting the solution without showing an interaction between light rays and molecules/ particles, or purely mathematical reasoning combined with mentioning of a phenomenon (e.g. absorption)	
	other	reasoning without a drawing of light and solution, or reasoning in neither of the above categories, or no reasoning (even in the presence of a drawing)	
Refer to Figure S4 for coding flowchart for Q2. Below are the general categories.			
Q2	*clearSub*	clearly submicroscopic reasoning: drawing more molecules in a picture with rays hitting molecules, or connecting molecule-related wording to previous submicroscopic-level picture	
	unclear	mentions increased interaction between the light and more-concentrated solution, without providing an illustration that shows whether the student approached the question from a submicroscopic or non-submicroscopic point	

	clearNonSub	clearly macroscopic or mathematical: algorithmic reasoning in terms of Beer's law, % transmittance, concentration, etc.	
	noReas	only shows illustration without explanation, or no explanation that goes beyond a description of the graph (no phenomena addressed)	
Q3	*submicro*	light is shown interacting either with a single molecule or an ensemble of molecules in each solution (rays hitting *individual* molecules should be shown)	
		OR, if not, the explanation addresses how much light CV and MO absorb "per molecule" (e.g. "CV absorbs more per molecule than MO")	
	nonSubmicro	the different final intensities of light for the two solutions are explained/depicted, without light-molecule interaction shown	
	other	no visualization is used in the explanation, or reasoning in neither of the above categories, or no reasoning (even in the presence of a drawing)	
Refer to Figure S5 for coding flowchart for Q4 cor. Below are the general categories			
Q4 cor.	*accurate*	the second intensity curve is drawn or described as completely overlapping with the first graph, but stops decreasing earlier (when x = 0.5 cm)	

	inaccurate	the two intensity curves are shown to decrease with different slopes within the solution (e.g. the intensity curve for 0.5 cm is decreasing slower), or are otherwise incorrect (e.g. higher transmittance for longer path length)	
	unclear	the graphs are drawn in a way that makes it impossible to determine if the student is comparing the two curves correctly	
If the student makes an *accurate* or *unclear* comparison in previous part, assign the codes below			
Q4 expl.	*sameCurve*	addresses the unchanged interaction between light and the solution	(b) Because the path length decrease, the light passes shorter length of molecular layers and exit the sample. The ability of molecule to absorb / the amount of molecule per layer does not change so the slope is the same.
		OR an explicit connection is made between "same solution" and "same shape"	b) The molar absortivity and concentration are the same so the intensity will decrease at the same rate for both, but the 0.5cm cuvet will have a higher %T because there is less solution to pass through.
	onlyFinal	the different final light intensities of the two solutions are addressed in terms of phenomena or algorithmic reasoning (i.e. mentioning of absorption or transmittance), without further discussion of light-solution interaction	b. since the path length was cut in half, there is less volume of solution so the light interacts with less molecules. Absorbance decreases in the 0.5 cm curvet due to its linear relationship with path length. Therefore, the solution in the 0.5 curvet has a higher light intensity than the 1.0 cm curvet because of the decrease in absorbance. b. I drew the 0.5cm graph because once the light that ~~is left~~ has not been absorbed leaves the curvet, it is constant. So, at the spot of 0.5cm, the light stops getting absorbed and is constant.
	noReas	accurate graph with no explanation that goes beyond a description of the graph (mentioning only intensity or path length constitutes a description of the graph)	b) The Intensity of light would not change, so at 0.5cm, the light would exit and stay constant.

S4.2. Transcriptions of the example text answers and reasoning for applied codes:

Drawings are described in brackets. Any grammatical or spelling errors of the students have been left as-is.

Q1

- *"Based on the picture, we can see that the amount of light traveled decrease as we pass through the solution. Because of that, the probability or chances for the light to get absorbed by iron decrease as light travel along path length, which result in the decreasing of absorption rate." {Shows a picture of light rays hitting individual molecules inside a cuvette and being primarily stopped by the first few layers of molecules they encounter}*

  This answer was coded as **submicro** because the student drew a molecular-level representation and implicitly related it to the rate of decrease of intensity in the graph.

- *"As light goes into the substance, it will slowly get absorbed, therefore leaving a small transmittance value." {Draws a cuvette with incoming and outgoing rays of light}*

  This answer was coded as **nonSubmicro** (which includes mathematical reasoning) because the student mentioned that absorption affects the transmittance of light, but did not use a molecular-level representation to explain light intensity changes inside the solution.

- *"I think it follows this shape because absorbance is a linear function but intensity is not, therefore it would follow a similar shape as the graph [shown in the question]. {shows a copied drawing of the graph in the question}"*

  This answer was coded as **other** because the student did not address any submicro or macro phenomena in their answer, nor did they provide a drawing. While some mathematical concepts are addressed, they are not connected to light-solution phenomena.

Q2

- *"Even less light is able to pass through unobstructed due to the higher conc. of particles absorbing it. This would make the graph much steeper and result in lower I_f" {shows picture of light rays hitting a large number of molecules, and draws the graphs for 0.01 M and 0.02 M superimposed}*

  This answer was coded as **clearSub** according to the flowchart for this question (Fig. S2); a drawing of the solution is provided, it is framed within an explanation, it depicts rays hitting molecules, and the molecules could not reasonably be interpreted macroscopically.

- *"A higher concentrated solution will have higher absorbance so less light will come out" {Shows a picture with a cuvette and incoming and outgoing rays, as well as a superposition of the two intensity curves}*

  This answer was coded as **clearNonSub** according to the flowchart: a drawing is provided, it is framed within an explanation, and it is not depicting light rays hitting molecules; the reasoning can be categorized as mathematical (higher concentration leads to higher absorbance).

- *"If higher concentration of iron complex is used (0.02 M), the decreasing of the rate of absorption will be more significant since more light is absorbed when the light started to travel through the solution."{no picture of the solution; only the intensity graph is present}*

  This answer was coded as **unclear** according to the flowchart: no new drawing is provided in this prompt, so we have to take into account the drawing the student provided in the previous prompt (i.e. the "submicro" example for Q1). There is no wording directly referencing changes that the student would make to the drawing, but the answer contains an explanatory component (higher concentration → more light absorbed → more significant rate of absorption). Even though the drawing in the previous question is submicroscopic, this explanation does not address particle behavior, so we cannot determine if the student is approaching the question from a macroscopic or submicroscopic standpoint.

- *{Shows a drawing of the two intensity curves superimposed (correctly)} "The final intensity for 0.02 M will be less than 0.01 M, but I_0 remains the same. The graph for 0.02 M will also curve downward more than the graph for 0.01 M."*

  This answer was coded as **noReas** because it only includes a description of the intensity curves, not an explanation.

<u>Q3</u>

- *{Shows two side-by-side pictures of light rays interacting with circles-molecules in the CV and MO solutions; the CV molecules are shown as larger and therefore "blocking" more rays} "The electrons on the CV molecule are more likely to be absorbed and excited by the light than MO."*

  This answer was coded as **submicro** because the student drew out ensemble of molecules interacting with light inside the cuvette and provided an explanation for the graph shape. The inaccuracy of the description of what occurs between the light and each molecule ("electrons being absorbed") did not affect the coding assignment.

- *{Shows the molecular structures of CV and MO side-by side along with an equal amount of light rays hitting each; less light comes through the CV molecule} "The chemical structure of each differs. Methyl orange contains two aromatic rings, while C.V. has three. This is directly proportional to their molar absorptivity, and therefore their intensity curve's derivative."*

  This answer was also coded as **submicro** because the student used molecular structures to explain the light-matter interaction difference between CV and MO.

- *{ Draws pictures of cuvets labelled "CV" and "MO" with different amounts of light transmitted through } "CV absorbs more light per molecule than MO."*

  This third example was also coded as **submicro** due to explicitly acknowledging that an individual molecule of CV absorbs more light than one of MO.

- *{Draws pictures of cuvets labelled "CV" and "MO" with different amounts of light transmitted through} "CV absorbs more light than MO, therefore its % transmittance will be lower than MO."*

  This answer was coded as **nonSubmicro** because the student only addressed the absorbance of the solution as a whole and used the Beer's Law equation to explain the graphs. There is no evidence that the student is referring to molecules when mentioning "CV" and "MO."

- *"CV is crystalline, which will reflect/deflect photons more than a basic methyl orange. Therefore, less light passes through the crystal and gives a lower [final] I value than MO."*

  This answer was coded as **other** because it demonstrates an alternative interpretation of the question (both CV and MO are solutes).

<u>Q4 accuracy</u>

The assignment of the codes "**accurate**" and "**inaccurate**" for the examples shown is self-explanatory. For the case coded as **unclear**, we note that there were no additional elements in the explanation portion that hinted at either an inaccurate or accurate comparison. If there had been elements in the explanation portion that suggested either an accurate (e.g. *"the graphs would look the same until 0.5 cm"*) or inaccurate (e.g. *"the intensity would decrease more slowly throughout in the 0.5 cm cuvette"*), comparison, this example would have been coded as *accurate* or *inaccurate*, respectively.

<u>Q4 explanation</u>

- *"Because the path length decreases, the light passes a shorter length of molecular layers and exits the sample. The ability of molecules to absorb/the number of molecules per layer does not change, so the slope is the same."*

  This answer was coded as **sameCurve** because it addresses that the number of molecules per layer stays the same and the interaction between light and the solution does not change.

- *"The molar absorptivity and concentration are the same, so the intensity will decrease at the same rate for both, but the 0.5 cm cuvet will have a higher %T because there is less solution to pass through."*

  This answer was also coded as **sameCurve** because it makes specific connection between "same solution" and "same decreasing rate (same shape)".

- *"Since the path length was cut in half, there is less volume of solution, so the light interacts with less molecules. Absorbance decreases in the 0.5 cm cuvet due to its linear relationship with path length. Therefore, the solution in the 0.5 cm cuvet has a higher light intensity than the 1.0 cm cuvet because of the decrease in absorbance."*

  This answer was coded **onlyFinal** because the student only explained why the final intensities were different, but no explanation related to unchanged interaction within the solution.

- *“I drew the 0.5 cm graph because once the light that has not been absorbed leaves the cuvet, it is constant. So, at the spot of 0.5 cm, the light stops getting absorbed and is constant.”*

  This answer lies close to the boundary of “simply describing the graph” and providing an explanation. It was coded **onlyFinal** (as opposed to *noReas*), because the student addresses absorption (which is a light-matter interaction) and the fact that it stops occurring outside of the cuvet. This interpretation clearly goes beyond a description of the graph.

- *“The intensity of light would not change, so at 0.5cm, the light would exit and stay constant.”*

  This answer was coded as **noReas** because it only provides a description of the graph.

S4.3. Coding Flowchart for prompt Q2: doubled concentration

Note: Prompts Q1 and Q2 correspond to "Exercises" 1b and 1c in the original assignment, respectively.

Reason for this flowchart

Because 1c is continuation of 1b, and because not all students drew a new drawing in 1c, we may need to consider students' responses in 1b when judging if they approached 1c from a submicroscopic or non-submicroscopic/other perspective.

Terminology used in this flowchart:

Graph: the intensity vs. path length graph that students are asked to predict. Does not count as a drawing.

Drawing: an illustration of how the student thinks the light interacts with the solution. The students were asked to draw a picture of how light interacts with the solution in 1b, and then asked how the picture would change in 1c. Some students gave a new drawing for 1c, some described how they would change their picture in 1b, and some did not provide either.

Explanation: Any text accompanying the students' answer, even if it does not constitute an explanation on strict terms (sometimes students describe the graph despite being asked to explain it).

General Categories

The general descriptions of the four categories that answers should fit within are listed below. The flowchart (not these definitions) should be used for coding.

clearSub: *submicroscopic resources clearly activated*

clearNonSub: clearly only non-submicro resources (macro + mathematical) are activated, including Beer's Law

unclear: cannot tell if question is approached in submicroscopic or macroscopic way

noReas: no explanation for the graph; if there is an illustration, it is not used further

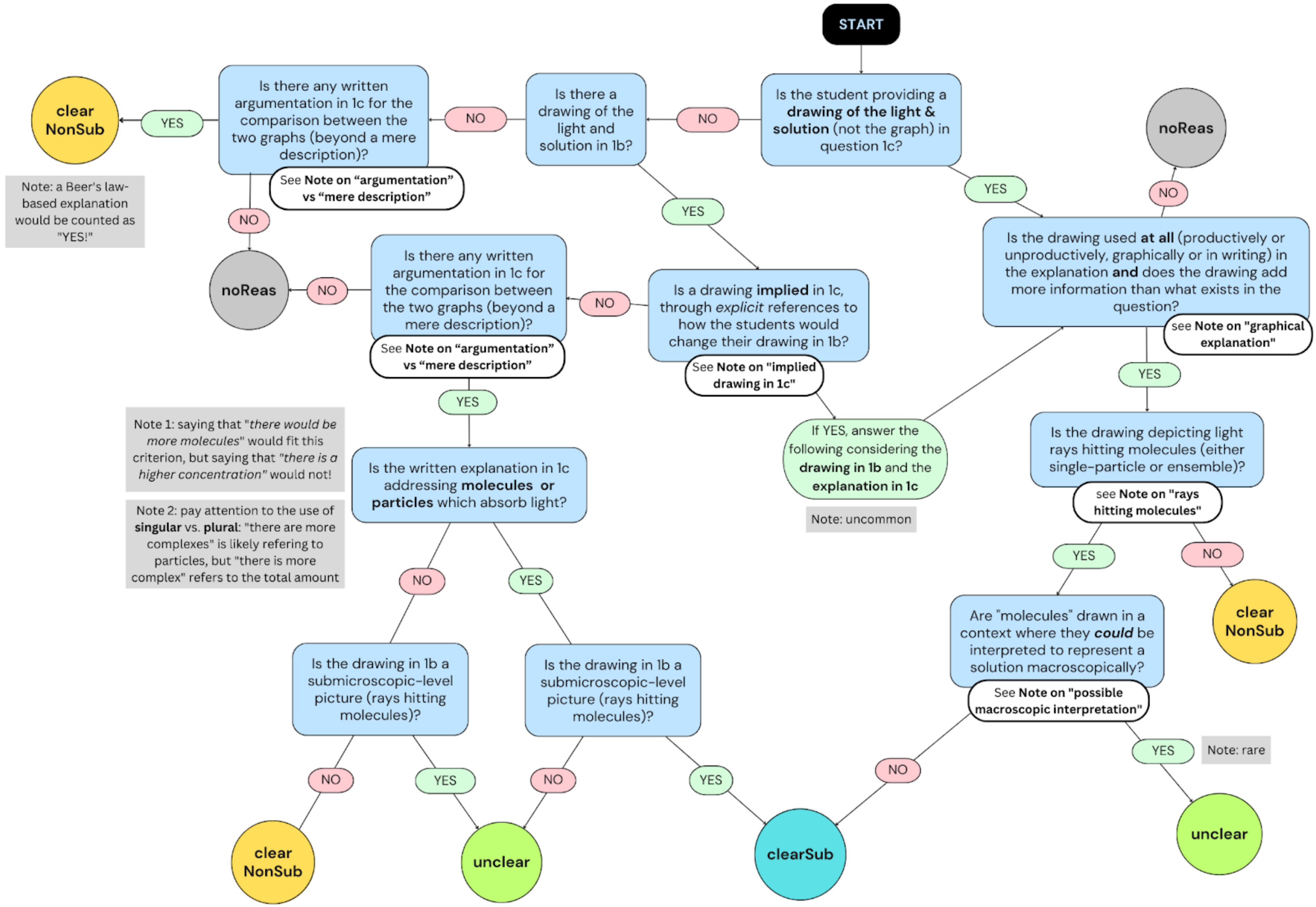


Figure S4. Coding flowchart for prompt Q2: doubled concentration.

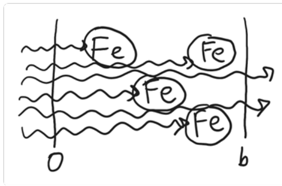
Note on “implied drawing in 1c”

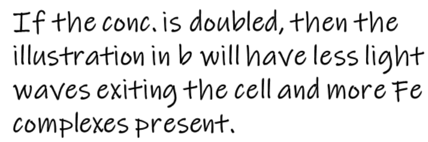
Consider case 1. The student has not provided a new drawing in 1c, but clearly explains how the illustration in 1b would change.

1b drawing:

Part of 1c explanation:

Case 1





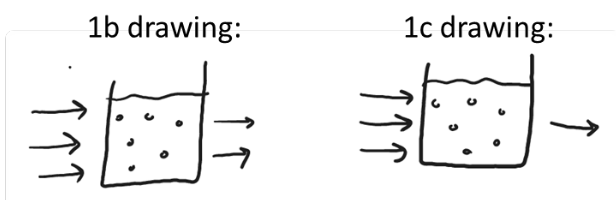
As a result, we can consider an implied drawing for 1c (one that looks like 1b but with more molecules and less outcoming rays), to which we can apply the same criteria in the coding flowchart that we would apply if a picture was actually drawn in 1c.

Note on “graphical explanation”

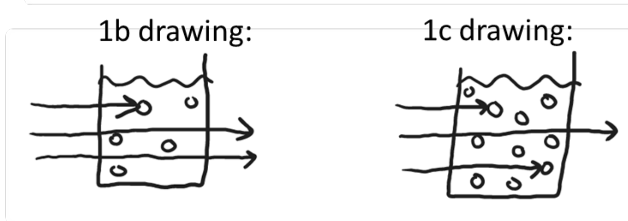
A "graphical" explanation could arise if the drawings in 1b and 1c can be compared to **justify** why less light comes out in 1c.

Let us consider two cases of drawings provided in 1c with no accompanying written explanation of how they influence the graph. We examine if the drawings themselves include an implicit justification of why there is a lower intensity exiting the solution with the doubled concentration:

Case 2



Case 3



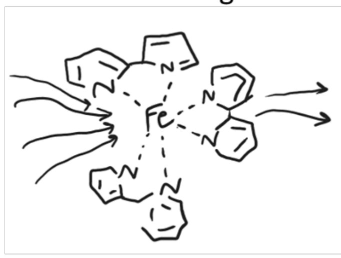
In case 2, there is no logical explanation of why there are fewer light rays exiting the solution in 1c, since the solutions look identical. For this case, NO should be selected. In case 3, it is clear that more rays are intercepted by molecules in 1c, and therefore fewer rays exit the solution. In this case, YES should be selected.

Note on “possible macroscopic interpretation”

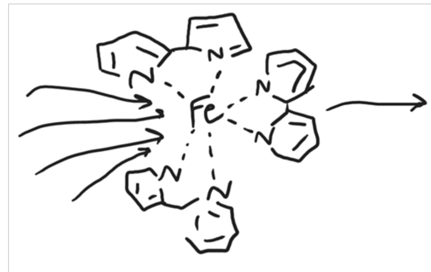
Consider case 4. At first glance, the drawing in 1c is depicting the submicroscopic level, since a complex molecule is drawn interacting with light rays. Four light rays are incident on the molecule and one light ray comes out of it.

Case 4

1b drawing:



1c drawing:

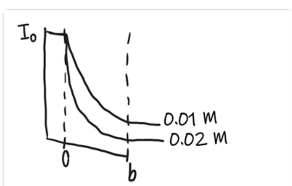


However, if one looks at the same student's drawing for 1b, the same, single molecule is drawn, though, inexplicably, two light rays are exiting it instead of one. This is more consistent with what one would observe with a cuvette full of complex molecules, than with a single molecule. Therefore, it is **unclear** whether the student is actually thinking of photon-molecule interactions at the submicroscopic level in 1c.

<u>Note on "argumentation" vs "mere description"</u>

Written argumentation should mention a *phenomenon*, e.g. "absorption", "deflection", "scattering", etc. or, in the case of Beer's law, "absorbance." Merely stating that "intensity decreases" or that "the concentration is doubled" counts as describing the graph and/or the information provided in the question. Consider Cases 5 and 6, which provide two possible "explanations" to the same intensity graph (without a drawing):

Graph
(same for **Cases 5 and 6**)



Case 5 explanation

More of the light will be absorbed at the beginning than in the 0.01 M [solution]

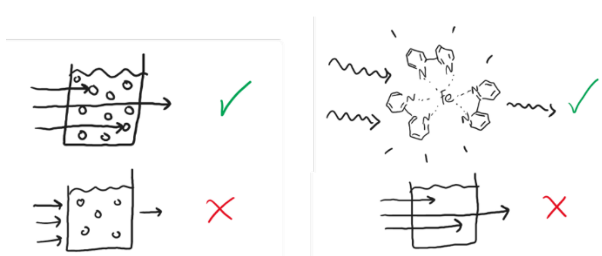
Case 6 explanation

Concentration is doubled, so intensity should decrease twice as fast.

In case 5, the phenomenon of absorption is mentioned, so (rudimentary or not!) this counts as an argumentation (more light absorbed = lower intensity seen in the graph). In case 6, the "explanation" does not provide any new information not seen in the graph or the question: the faster decrease of intensity is just what we see in the graph, and the "doubled concentration" is the information provided by the question. So, we consider argumentation to exist in Case 5 (YES) and *not* to exist in Case 6 (NO).

<u>Note on "rays hitting molecules"</u>

We consider an answer to depict "rays hitting molecules" if light rays are shown to interact with individual molecules, either on a single-molecule level or at the ensemble level. The answers below marked with a checkmark would qualify for "rays hitting molecules."



S4.4. Coding Flowchart for prompt Q4 “correctness”: halved path length

Note: The “correctness” aspect of Prompt Q4 corresponds to Exercise 3a in the assignment (*sketch the intensity profile*) and the “explanation” aspect of Prompt Q4 corresponds to Exercise 3b in the assignment (*explain your reasoning*).

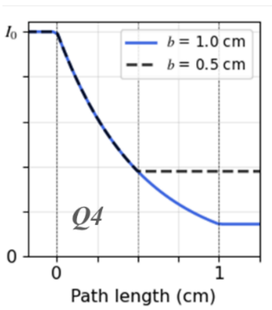
Reason for this flowchart

In question 3a, not all students drew their curves for the b = 0.5 cm and b = 1.0 cm in a way that makes them clearly comparable to each other. In attempt to better classify the cases where the graphical comparison was unclear, we attempt to draw on information that the students gave in their explanation (question 3b) to determine if they are thinking correctly about the comparison between the two curves. In some cases, the comparison will remain “unclear” even after consulting the students’ answers in 3b, but, in other cases, the students provide clear evidence that they are comparing the two curves correctly or incorrectly.

General categories:

The general descriptions of the four categories that answers should fit within are listed below. The flowchart (not these definitions) should be used for coding.

Correct*:* The intensity curve for the solution with 0.5 cm path length should completely overlap the curve for the solution with 1 cm path length, until the point where x = 0.5 cm. Beyond that point, the intensity should stay constant.



Incorrect: No answer is provided, or a different comparison than the above is provided. The most common mistake would be for the curves *not* to overlap where they should.

Unclear comparison: it is impossible to determine if the student is comparing the two curves correctly.

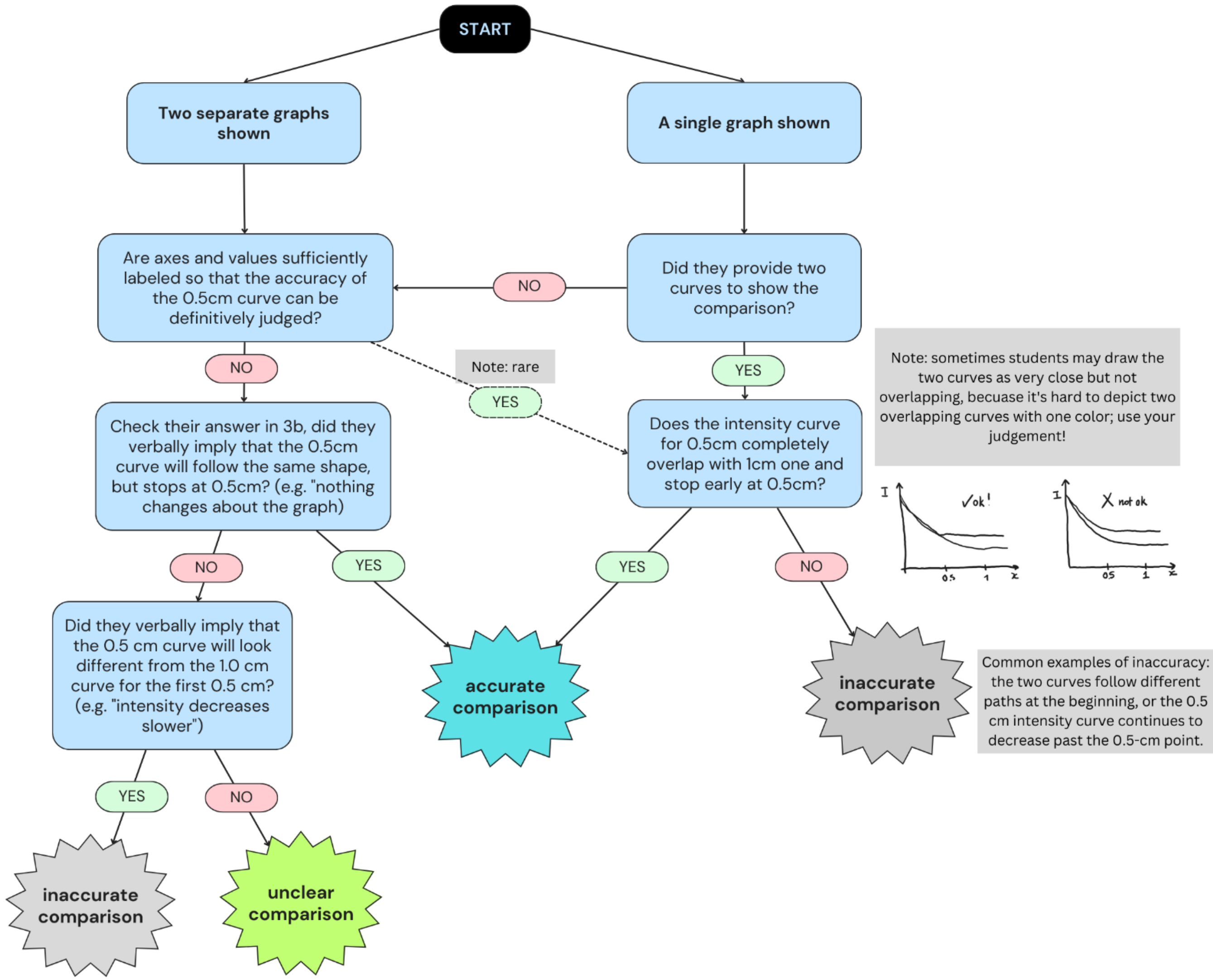


Figure S5. Coding Flowchart for prompt Q4 "correctness": halved path length.

S4.5. Transcriptions of answers displayed in Fig. 12 and Fig. 13

Drawings are described in brackets. Any grammatical or spelling errors of the students have been left as-is.

Figure 12:

a) *"When light begin to pass the solution, there are more particles that could block the light. When the light is passing the solution, the possibility for a single photon to encounter a particle stay the same but the overall light intensity is decreasing, so that the decreasing rate of the intensity is decreasing." {Shows a picture of light rays (arrows) hitting individual molecules (circles) and some exiting at the end}*

b) *"As the light goes through the beam it is deflected by things in the solution (Fe) so therefore the intensity recorded decreases as it goes through." {Draws a cuvette with incoming and outcoming rays of light, as well as rays going to different directions}*

c) *Molar absorptivity, ε of CV is higher than MO. Particle [sic] in CV absorbs more light compared to MO. CV has bigger molecules that absorbs light compared to molecules in MO. {Shows two rectangular cuvettes labeled "MO" and "MO", each containing the same number of circular particles hit by light rays (arrows). The CV cuvette is shown to contain bigger particles than MO, but the same number of particles.*

c) *"The absorbance of light increase energy (exitation [sic])" {Shows a drawing of light source, cuvet and detector, representing a spectrometer set-up (no clear indication of light rays), as well as a graph resembling an absorption spectrum and a drawing of excitation of a system from ground level to higher energy level}*

d) "light of the same energy have [sic] the same wavelength according to E = hc/λ . Crystal violet have maximum absorptivity at 420 and 620 nm while methyl orange have maximum absorptivity at 464 nm." {draws a visible light spectrum showing the (labeled) maximum absorptivity curve for each solute}

e) "CV has a much higher molar absorptivity, so it will absorb all colors shown through as where the methyl orange has less absorptivity so other colors could show through" {shows two cuvettes with yellow light coming in, and different color of light coming out}

Figure 13a:

"CV absorbs more photons than MO per molecule. This would be like last week's law [sic] when we increased the radius of the disks." {Draws two "stacks" of particles (one labeled "CV," one "MO"), with four light intensity arrows "entering" each. The CV particles are depicted as bigger, and only let one arrow pass through, whereas the MO particles are smaller and let two arrows pass through.}

Section 5: Interrater reliability results

Table S3 summarizes the interrater agreement between the two original coders (NS and YZ, treated as one rater due to 100% internal agreement) and the fourth coder (SAF) for the subset of files coded collectively. Krippendorff's alpha (α) was chosen as a metric of interrater agreement (as opposed to Cohen's kappa), because certain codes (e.g. the "submicro" code in Q2) were very dominant in the subset of files coded while others (e.g. the "other" code in Q2) were rarely represented.

Table S3. Negotiated agreement vs. raw agreement

	Prompt Q1	*Prompt Q2*	*Prompt Q3*	*Prompt Q4 correctness*	*Prompt Q4 explanation*
Agree	24	14	18	25	16
Negotiated agreement	2	11	7	1	4
Disagreement	0	1	1	0	0
Total files	26	26	26	26	20*
% Agreement (pre-negotiation)	92%	54%	69%	96%	80%
Negotiated % Agreement	100%	96%	96%	100%	100%
Krippendorff's α (pre-negotiation)	.84	.36	.31	.92	.70
Krippendorff's α (post-negotiation)	1	.95	.91	1	1

** Files with inaccurate comparison in prompt Q4 corr. were excluded.*

Section 6: The Assumptions of the Chi-square Test of Independence

- *Assumption 1*. Data used in analyses are frequencies; this assumption was met by using raw code counts rather than percentages.
- *Assumption 2*. There are two variables, and both are categorical; this assumption was met by the design of our study, i.e., examining the association between the classroom condition (simulation vs. comparison groups) and code categories associated with student explanations.
- *Assumption 3*. The levels of the categorical variables are mutually exclusive; this assumption is met by the design of the study, i.e., a student's prompt response could only be associated with one classroom condition level and one response code category.
- *Assumption 4*. Each response could contribute data to only one cell in the chi-square contingency table; this assumption is met by the design of the study, i.e. a student's response to a prompt can only belong to (a) either the simulation or comparison group, and (b) one response code category.
- *Assumption 5*. All observations are independent; this assumption is met by the structure of the course assignment from which data were collected—all students completed the spectroscopy assessment prompts individually.
- *Assumption 6*. Expected values of cells should be five or greater in at least 80% of the cells, and no cell should have an expected value less than one; this assumption was met by inspecting contingency tables, which are reported below for each chi-square test.

Section 7: Connection of students' models with graph shape in Q1

Students responses to prompt Q1 (explain graph shape) were coded in terms of whether the students connected their drawing of the light-solution interaction to the shape of the graph. Answers were categorized as *noGraphRef* when no reference was made to any aspect of the graph, *shallowGraphRef* when only the decrease (not the leveling-off) of the graph was explained, or *graphShapeRef* when both the decrease and the leveling-off (decreasing slope magnitude) of the intensity profile were explained.

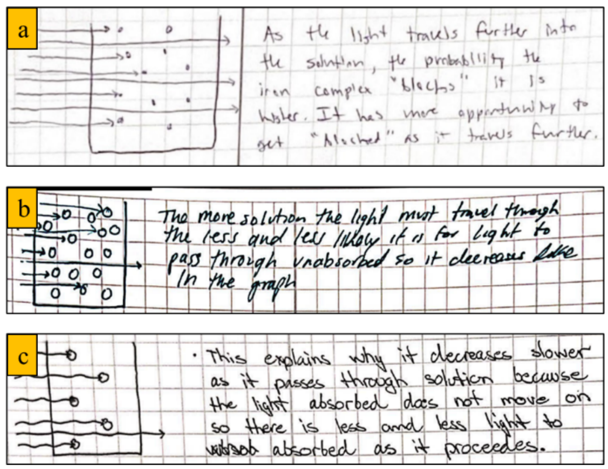


Figure S6. Examples of student answers to Q1 (explain graph shape) that explained the shape of the graph to various degrees: (a) no reference to the graph, (b) a shallow reference to the graph and explanation of its decrease, and (c) a reference to the graph shape and an explanation of the decreasing decrease rate of the intensity profile.

While a slightly higher proportion of students in the simulation group (9 out of 53 compared to 8 out of 71 in the comparison group) were able to successfully explain the decreasing magnitude of the slope of the intensity profile, there is no statistically significant effect of the simulation on students' ability to connect their drawings with the graph, χ^2 (2, n = 124) = 1.04, p = .59, Cramér's V = .09.

Table S4. Contingency table for the χ^2 test examining the association between completing BLSim activity and reasoning code distribution in terms of explanations of the intensity profile shape.

	Comparison Group (n = 71)	**Simulation Group (n = 53)**
noGraphRef	SR: 0.0 expected: 33.2 observed: 33	SR: 0.0 expected: 24.8 observed: 25
shallowGraphRef	SR: .4 expected: 28.1 observed: 30	SR: –.4 expected: 20.9 observed: 19
graphShapeRef	SR: –.6 expected: 9.7 observed: 8	SR: .6 expected: 7.3 observed: 9

References cited in the ESI

Laverty, J. T., Underwood, S. M., Matz, R. L., Posey, L. A., Carmel, J. H., Caballero, M. D., Fata-Hartley, C. L., Ebert-May, D., Jardeleza, S. E., & Cooper, M. M. (2016). Characterizing College Science Assessments: The Three-Dimensional Learning Assessment Protocol. *PLOS ONE*, *11*(9), e0162333. https://doi.org/10.1371/journal.pone.0162333

Spitha, N., Doolittle, P. S., Buchberger, A. R., & Pazicni, S. (2021). Simulation-Based Guided Inquiry Activity for Deriving the Beer–Lambert Law. *Journal of Chemical Education*, *98*(5), 1705–1711. https://doi.org/10.1021/acs.jchemed.0c01433