Electronic Supplementary Material (ESI) for Chemistry Education Research and Practice.
This journal is © The Royal Society of Chemistry 2023

Part A: Evidence-Centered Design Application

Claim Space: The objective of this assessment is to characterize the ways n which students reason about conflicting data in order to generate an argument. Based on *The Framework for K-12 Science Education* students should be able to handle uncertainty in the following ways:

(a) Recognize when data conflicts with an initially held idea or expectation
(b) Recognize the limitations of the evidence in an argument
(c) Make the necessary changes to an existing model or explanation to incorporate new evidence

To assess these competencies, we target the following KSAs of argumentation in this assessment:

(a) How students use data as evidence in their argument
(b) How students coordinate multiple pieces of data to generate a claim
(c) The types of reasoning students utilize to manage uncertainty

Evidence Space:

(a) What evidence will be collected to draw conclusions about the targeted competency?
We will collect students written arguments from engaging in a task that contained conflicting data to draw our conclusions. We also will collect data from semi-structured follow-up interviews during which students will provide more explanation to their argument.

(b) How will the collected evidence be analyzed and interpreted?
We anticipate utilizing qualitative methodologies, such as open coding, to analyze student responses.

(c) How will we interpret our evidence to align with the desired competencies?
We will consider the following when analyzing the student responses:
i. What evidence was/was not incorporated into the student's argument?
ii. What reasoning did the student use to justify their condition choice?

Task Space:

(a) What features does the task need to elicit evidence that valid claims can be drawn from?
The task must contain some elements that create uncertainty for the students. These elements including things such as conflicting data, several variables that must be weighed separately, or asking questions that have no correct answers.

(b) What features of the task can be varied based on the needs of the instructor?
The level of conceptual knowledge required to engage in the task can be varied based on the student population.

(c) What scaffolding will be utilized in the task to ensure the desired construct is measured?
We will show students each piece of data individually to interpret before asking them to coordinate multiple pieces of evidence to construct their claim. We will also provide context on what the optimal results for each piece of data is. This is done to help students focus on coordinating multiple pieces of evidence and constructing a thorough argument rather than trying to understand the goals of the assignment.

Part B: The Onion Task

Vegetables are a key component to diets as they provide the body with key nutrients it needs in order to sustain biological processes. One commonly used vegetable is the onion, which provides numerous health benefits to its consumer. Dehydrated onions are in high demand globally, as they are often used in soups, stews, and other food production. As the recently appointed supervisor in a food company, you have decided to expand your business to begin producing dehydrated onions.

You've hired a research team to go out and investigate different variables that affect the quality of onions after they've been dehydrated. This research team started their investigation by collecting onions and splitting them into two different treatment groups. The first group received a salt treatment, while the second group received no treatment before being dehydrated. The researchers then took each of the treatment groups and split them into three different subgroups. Each subgroup's onions were cut into slices at different thicknesses. A visual for this is depicted in Figure 1a.

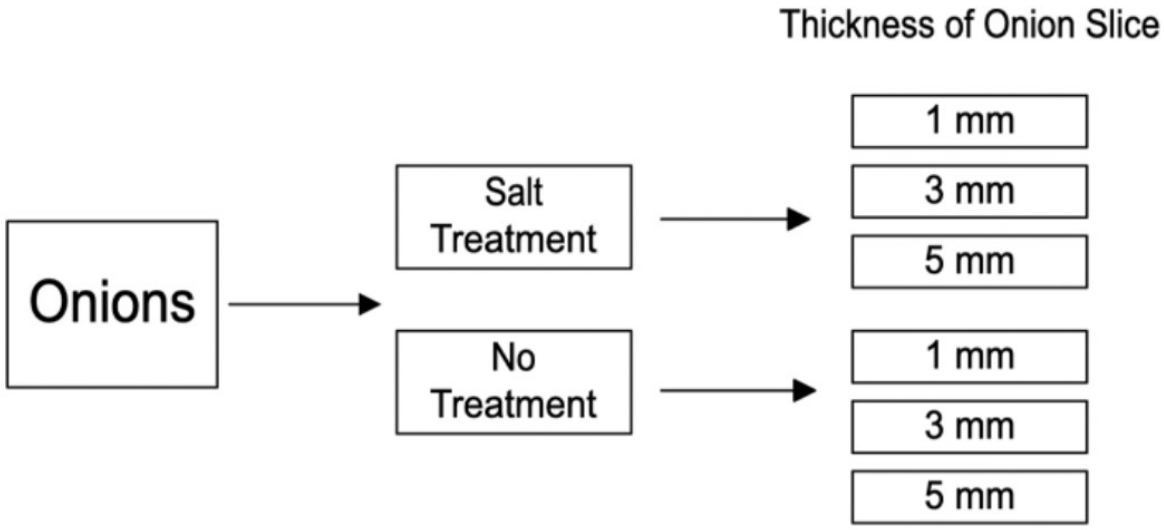


Figure 1a: Process of splitting onions up into treatment group and slice thickness

The researchers then took each of the different thickness groups and split them into three more groups. Each group was then dehydrated at a different temperature. This process is shown below in Figure 1b.

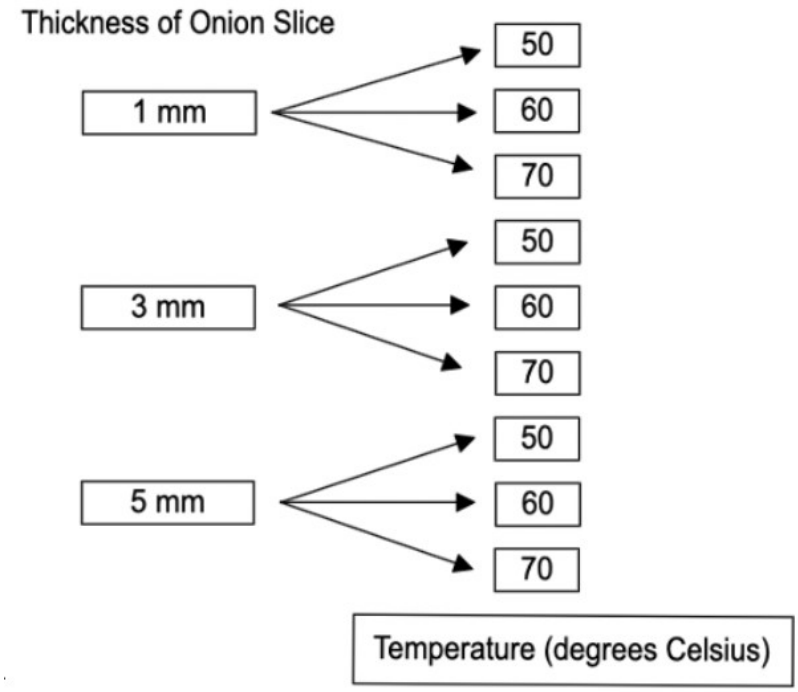


Figure 1b: Process of splitting each thickness group into three different dehydration temperatures.

These steps combine to give the experimental setup depicted in Figure 1c.

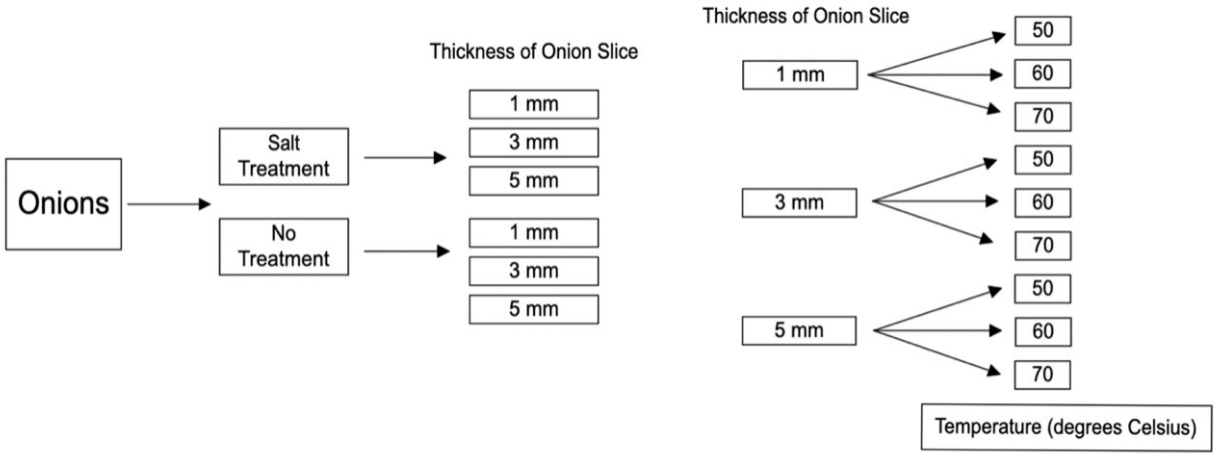
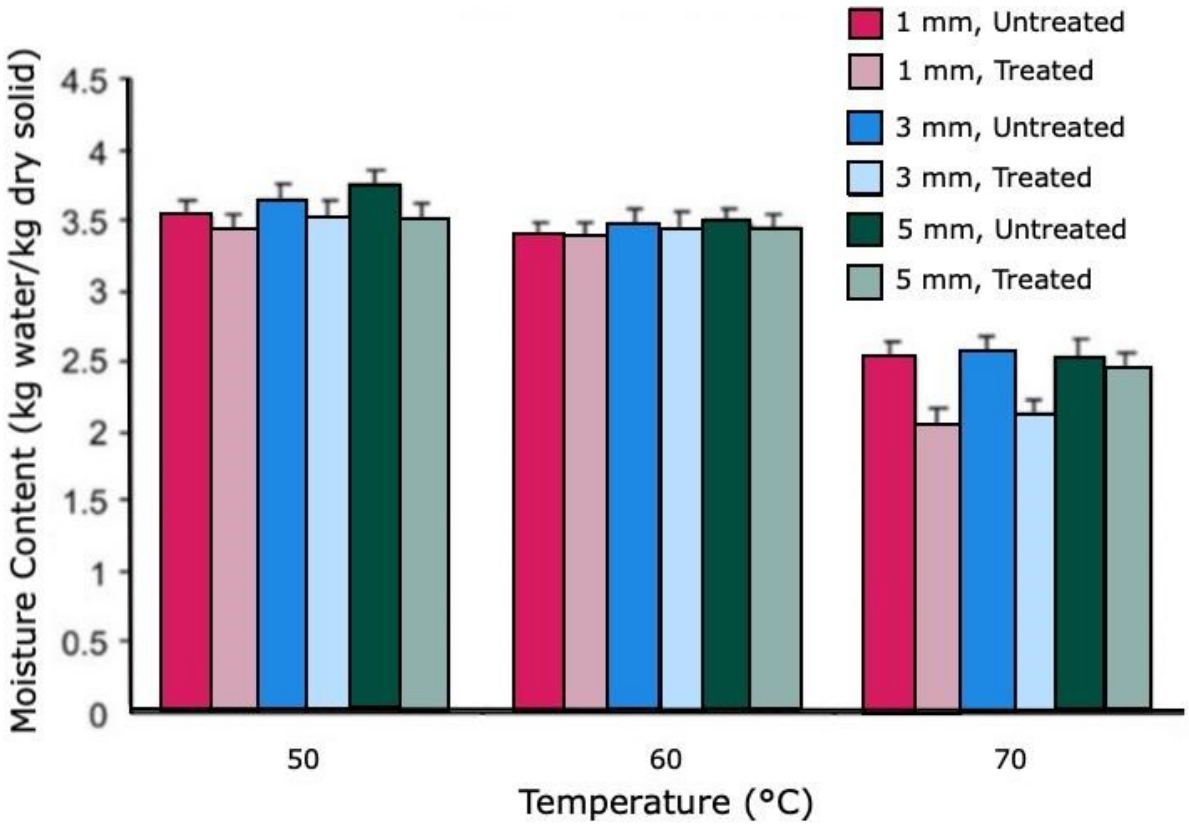


Figure 1c: Total experimental set up for testing variables affecting onion dehydration.

After the researchers dehydrated all the onions, they investigated three different aspects to determine which dehydration conditions yielded the best onions overall. The first quality they investigated was the onion's moisture content. Moisture content is a measure of how much water remains in the onion after the slice is dehydrated. Because most chemical reactions that result in molding require water to occur, the higher the moisture content, the lower the quality of the dehydrated onion. The measured moisture content for each of the dehydrated onion slices are depicted here in Figure 2:



What is your interpretation of the data in Figure 2? What might account for these differences between samples?

Once the investigators had finished collecting data about moisture content, they moved on to considering a new quality in the onion: browning. Browning in this context represents how fast an onion browns after it is dehydrated. As the onion browns, it begins to spoil and become rancid. Table 1 below depicts the value of the rate constants for the chemical reaction that causes the browning in the onions to occur.

Thickness (mm)	Temperature (°C)	Untreated Sample's k	Treated Sample's k
1	50	0.252	0.224
	60	0.327	0.298
	70	0.388	0.346
3	50	0.089	0.075
	60	0.123	0.116
	70	0.296	0.217
5	50	0.073	0.074
	60	0.098	0.094
	70	0.170	0.133

Table 1: Rate constants of browning reaction in dehydrated onions

What is your interpretation of the data in Table 1? In what ways do you think the rate of browning is related to treatment, temperature, and thickness?

The final quality that the research team investigated was the flavor of the onion after it was dehydrated. The chemical species responsible for giving onions their flavor is thiosulphiante. The investigators were interested to see how this chemical was broken down once the onions were dehydrated. Table 2 depicts the value of the rate constant for the chemical reaction that is responsible for causing the loss of thiosulphinate over time.

Thickness (mm)	Temperature (°C)	Untreated Sample's k	Treated Sample's k
1	50	13	5.0
	60	14	6.0
	70	15	7.0
3	50	3.0	0.6
	60	4.0	2.0
	70	8.0	3.0
5	50	0.9	0.2
	60	1.0	0.3
	70	2.0	0.7

What is your interpretation of the data in Table 2? In what ways do you think the rate of browning is related to treatment, temperature, and thickness?

Based on the information depicted in the previous figures and your understanding of chemical kinetics, what are the best conditions to dehydrate the onions at? **Be sure to include a temperature, thickness, and treatment for the onion slice.**

Why do you believe these are the best conditions to dehydrate the onions at? **Be sure to consider all three qualities (moisture content, browning, and flavoring) and cite evidence and reasoning to support your claim.**

Part C: Rubric Iterations and Inter-Rater Reliability Process

Initially, the reasoning codes were separated by what the student selected as their best conditions. An excerpt of two common condition sets and their respective rubrics are shown below:

Rubric separated out by condition codes:

50 °C, 5 mm, Treated	
Code	**Description**
Misinterpret Moisture Content	Student misinterprets some component of the moisture content figure, and that misinterpretation carries through to their reasoning for selecting the best conditions.
Neglected Moisture Content	Student omits any discussion of moisture content from their response.
2 out of 3	Student picked conditions that favor browning and flavoring and acknowledge that they are not the optimal conditions for moisture content.

70 °C, 5 mm, Treated	
Code	**Description**
Compartmentalize	The student picks one condition to favor one quality, but another condition to favor another quality. The student does not acknowledge the conflict in their response.
Balancing Positives and Negatives	Student acknowledges that the data is conflicting on what constitutes the best conditions. The student will select a condition and recognize its negative impact on a quality. They will then try and benefit that neglected quality in the selection of their next condition.
Variance	Student recognizes that the data conflicts on what the optimal conditions are. The student will recognize both the positive and negative impacts of selecting a condition. Student considers the difference between data points within one condition (treatment, thickness, or temperature) for one or more quality (moisture content, browning, or flavor loss) and uses those differences as justification for selecting a certain condition.

As we continued to test the rubric, we determined that only coding within the condition rubric did not always capture the type of reasoning the student was employing. To address this, we decided to combine all the individual condition rubrics into one comprehensive rubric. This rubric is shown below:

Combined Rubric (no condition codes):

Code	**Description**
Vague	Student does not assert a set of conditions or does not provide reasoning to justify why they selected their condition. This also includes students who asserted that a quality or variable was the best to dehydrate the onion.

No Justification of Thickness or Temperature	Student does include any reasoning to justify their choice in temperature and/or thickness.
Misinterpret Moisture Content	Student misinterprets some component of the moisture content figure and that misinterpretation carries through to their reasoning for selecting the best conditions.
Misinterpret Rate Constants	Student misinterprets the meaning of the rate constant in the task and asserts that a larger rate constant means a slower rate and vice versa.
Misinterpret Goal	Student misinterprets what constitutes the best onion slice. For example, a student may assert that they want to maximize moisture content, browning, or flavor loss.
Neglected Moisture Content	Student omits any discussion of moisture content from their response.
Neglected Browning and Flavoring	Student omits any discussion of browning or flavoring from their response.
Compartmentalize	The student picks one condition to favor one quality, but another condition to favor another quality. The student does not acknowledge the conflict in their response.
Middle Ground	Student explicitly states that there is no optimal condition set due to the conflicting data, so they opt to select the middle value of the conditions to not favor one quality over another.
Prioritize Moisture Content	Student acknowledges that their conditions will not be optimal for browning and flavoring, but ultimately decides that moisture content is the most important quality to benefit.
2 out of 3	Student picked conditions that favor browning and flavoring and acknowledge that they are not the optimal conditions for moisture content.
Balancing Positives and Negatives	Student acknowledges that the data is conflicting on what constitutes the best conditions. The student will select a condition and recognize its negative impact on a quality. They will then try and benefit that neglected quality in the selection of their next condition.
Variance	Student recognizes that the data conflicts on what the optimal conditions are. The student will recognize both the positive and negative impacts of selecting a condition. Student considers the difference between data points within one condition (treatment, thickness, or temperature) for one or more quality (moisture content, browning, or flavor loss) and uses those differences as justification for selecting a certain condition.
2 out of 3 & Middle Ground	Student utilizes 2 out of 3 reasoning to justify their choice in selecting one condition, but then middle ground reasoning to justify their choice in selecting another condition.
Prioritize Moisture Content & Middle Ground	Student prioritizes moisture content reasoning to justify their choice in selecting on condition, but then middle ground reasoning to justify their choice in selecting another condition.

Once the rubrics had all been combined into one, we moved on to establishing IRR. After our first round of coding, we calculated a Cohen's kappa value of 0.558, which indicates weak

agreement. At this point, we further clarified how the conditions the student selected could help inform the type of reasoning they utilized, but it did not dictate what codes the response could earn. This clarification process involved going through several student responses together and discussing how we are thinking about the student's conditions and reasoning working together. Then we completed a new round of coding and calculated a new Cohen's kappa value of 0.631.

At this point, we determined that separating the different types of misinterpretations out was too nuanced. It was difficult to tell when a student misinterpreted data and when they misunderstood the goal of the assignment. In some cases, the students also had multiple misinterpretations in their response. We did not feel it would be appropriate to assign a hierarchy to the different misinterpretations, as they capture different ideas that are equally important. Therefore, we decided it would be appropriate to collapse all the misinterpretation codes into one to help with the coding process.

Finally, we decided to remove the "No Justification of Thickness or Temperature" code from the rubric. We felt like these responses were better categorized under the "Vague" category, as they did not offer clear insight as to why the student selected the conditions that they did.

After discussing these changes, we completed a new round of IRR. We calculated a Cohen's Kappa value of 0.799 which indicates a strong/moderate agreement.

The final rubric can be seen below:

Final Rubric:

	Code	**Description**
0	Vague	Student does not assert a set of conditions or does not provide reasoning to justify why they selected their condition. This also includes students who asserted that a quality or variable was the best to dehydrate the onion.
1	Misinterpretation	Student's response has some misinterpretation in their reasoning response. This includes misinterpreting the assignment instructions or misinterpreting the data.
2	Neglected Moisture Content	Student omits any discussion of moisture content from their response.
3	Neglected Browning and Flavoring	Student omits any discussion of browning or flavoring from their response.
4	Compartmentalize	The student picks one condition to favor one quality, but another condition to favor another quality. The student does not acknowledge the conflict in their response.
5	Middle Ground	Student explicitly states that there is no optimal condition set due to the conflicting data, so they opt to select the middle value of the conditions to not favor one quality over another.
6	Prioritize Moisture Content	Student acknowledges that their conditions will not be optimal for browning and flavoring, but ultimately decides that moisture content is the most important quality to benefit.

7	2 out of 3	Student picked conditions that favor browning and flavoring and acknowledge that they are not the optimal conditions for moisture content.
8	Balancing Positives and Negatives	Student acknowledges that the data is conflicting on what constitutes the best conditions. The student will select a condition and recognize its negative impact on a quality. They will then try and benefit that neglected quality in the selection of their next condition.
9	Variance	Student recognizes that the data conflicts on what the optimal conditions are. The student will recognize both the positive and negative impacts of selecting a condition. Student considers the difference between data points within one condition (treatment, thickness, or temperature) for one or more quality (moisture content, browning, or flavor loss) and uses those differences as justification for selecting a certain condition.
10	2 out of 3 & Middle Ground	Student utilizes 2 out of 3 reasoning to justify their choice in selecting one condition, but then middle ground reasoning to justify their choice in selecting another condition.
11	Prioritize Moisture Content & Middle Ground	Student prioritizes moisture content reasoning to justify their choice in selecting on condition, but then middle ground reasoning to justify their choice in selecting another condition.