Examining Students' Mathematical Evidence in CER Explanations during Science Inquiry Contexts

Using mathematics and constructing explanations are practices outlined by the NGSS (2013). However, students struggle with mathematizing in science: determining the mathematical relationships between data (linear; Lai et al., 2016; Shah & Hoeffner, 2002), understanding the components of equations (slope; Nixon et al., 2016, Planinic et al., 2012), applying best-fit lines to data (Casey, 2015), and generating explanations about covariational relationships (McDermott et al., 1987; Sokolowski, 2019). These are critical barriers to high school science (Basson, 2002; Sadler & Tai, 2001), and "plugging and chugging" rote formulas fails to develop deep understandings of the DCIs expected by the NGSS (Brandiet et al., 2018). Thus, students need to be supported at using mathematics in science inquiry contexts so that they can develop *deep* understandings of phenomena.

Methods

84 students (Table 1) completed an Inq-ITS virtual lab that involved mathematical modeling (Dickler, 2021) to determine how the mass of a sled going down a ramp affects the momentum of the sled when it reaches the end of the ramp (Figure 1, Table 2).

We developed and refined a fine-grained rubric for Claim, Evidence, Reasoning (CER; McNeill et al., 2006) statements that assesses the mathematical evidence provided in students' science explanations in a virtual lab in Inq-ITS (Gobert et al., 2013; Figure 2; Table 3) and substantiated the rubric elements using prior literature on explanations (Li et al., 2017; Sokolowski, 2019). Four of the authors independently coded the first 10 sets of responses for each rubric element; 91% agreement was reached, disagreements were discussed, and the agreedupon codes were used for analyses. Each rater then coded their own assigned portion of the remaining sets of responses.

Results

In the students' claims (Table 4), most (45%) described a covariational relationship between the variables; 27% described a mathematical relationship; only 18% described both the covariational relationship *and* the mathematical relationship (the ideal answer); and 10% did not describe either. When asked to substantiate claims with evidence, only 5% included the *specific* equation of the model they built in the lab, and 7% mentioned the fit of the model to their data. Regarding students' reasoning statements, only 8% mentioned a scientific theory or concept that explains why their evidence supports their claim, which aligns with earlier CER research (McNeill et al., 2006). Overall, these results indicate that students struggle with understanding how to generate CER statements using mathematical evidence.

Scholarly Significance

The NGSS (2013) emphasize the importance of students using mathematics and constructing explanations to develop deep understandings in science. However, we found that students struggled with using mathematical evidence in the CER responses, which aligns with previous findings about students' difficulties with mathematizing in science (e.g., McDermott et al., 1987). To be able to assess and support students with these intersecting practices, we must be able to operationalize these practices in a fine-grained and rigorous way. Our present work contributes to this goal and will ultimately inform future development of automated scoring and scaffolds of students' scientific explanations involving mathematical evidence for Inq-ITS.

References

Basson, I. (2002). Physics and mathematics as interrelated fields of thought development using acceleration as an example. *International Journal of Mathematical Education in Science and Technology*, *33*(5), 679-690.

Brandiet, A., Rupp, C.A., Lazenby, K., & Becker, N.M. (2018). Evaluating students' abilities to construct mathematical models from data using latent class analysis. *Chemistry Education Research and Practice, 19*(1), 375-391.

Casey, S. A. (2015). Examining student conceptions of covariation: A focus on the line of best fit. *Journal of Statistics Education*, *23*(1), 1-33.

Dickler, R. (2021). *An intelligent tutoring system and teacher dashboard to support students on mathematics in science inquiry* [Doctoral dissertation, Rutgers University]. ProQuest Dissertations Publishing.

Fretz, E.B., Wu, H.K., Zhang, B., Davis, E.A., Krajcik, J.S., & Soloway, E. (2002). An investigation of software scaffolds supporting modeling practices. *Research in Science Education*, *32*(4), 567-589.

Gobert, J. D., Sao Pedro, M., Raziuddin, J., & Baker, R. S. (2013). From log files to assessment metrics: Measuring students' science inquiry skills using educational data mining. *Journal of the Learning Sciences*, *22*(4), 521-563.

Li, H., Gobert, J., & Dickler, R. (2017). Automated assessment for scientific explanations in on-line science inquiry. In A. Hershkovitz & L. Paquette (Eds.), *Proceedings of the 10th International Conference on Educational Data Mining* (pp. 214-219). EDM Society.

McDermott, L. C., Rosenquist, M. L., & Van Zee, E. H. (1987). Student difficulties in connecting graphs and physics: Examples from kinematics. *American Journal of Physics*, *55*(6), 503-513.

McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *The Journal of the Learning Sciences*, *15*(2), 153-191.

NGSS Lead States. (2013). *Next generation science standards: For states, by states.* National Academies Press.

Nixon, R. S., Godfrey, T. J., Mayhew, N. T., & Wiegert, C. C. (2016). Undergraduate student construction and interpretation of graphs in physics lab activities. *Physical Review Physics Education Research*, *12*(1), 101041-1010419.

Planinic, M., Milin-Sipus, Z., Katic, H., Susac, A., & Ivanjek, L. (2012). Comparison of student understanding of line graph slope in physics and mathematics. *International Journal of Science and Mathematics Education*, *10*(6), 1393–1414.

Sadler, P. M., & Tai, R. H. (2001). Success in introductory college physics: The role of high school preparation. *Science Education*, *85*(2), 111-136.

Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review*, *14*(1), 47-69.

Sokolowski, A. (2019). Developing mathematical reasoning using a STEM platform. In, *Interdisciplinary mathematics education* (pp. 93-111). Springer, Cham.

Appendix

Table 1. Demographics of Participating High Schools

Teacher (School)	STEM Subject	# of Students	Qualify for Free/ Reduced Lunch	Asian	Black	Hispanic	Native American	Pacific Islander	White	Two or More Races
Ms. A (1)	Engineering	32	73%	0.4%	45.8%	52.7%	0.1%	0%	0.9%	0.2%
Ms. B (2)	Chemistry	6	32%	22.5%	11%	21.6%	0.2%	0.2%	39.6%	4.9%
Mr. C (3)	Physics	46	8%	8.1%	2.1%	10.8%	0%	0.2%	78.6%	0.3%

Figure 1. Screenshots of Inq-ITS Virtual Lab with Mathematical Modeling

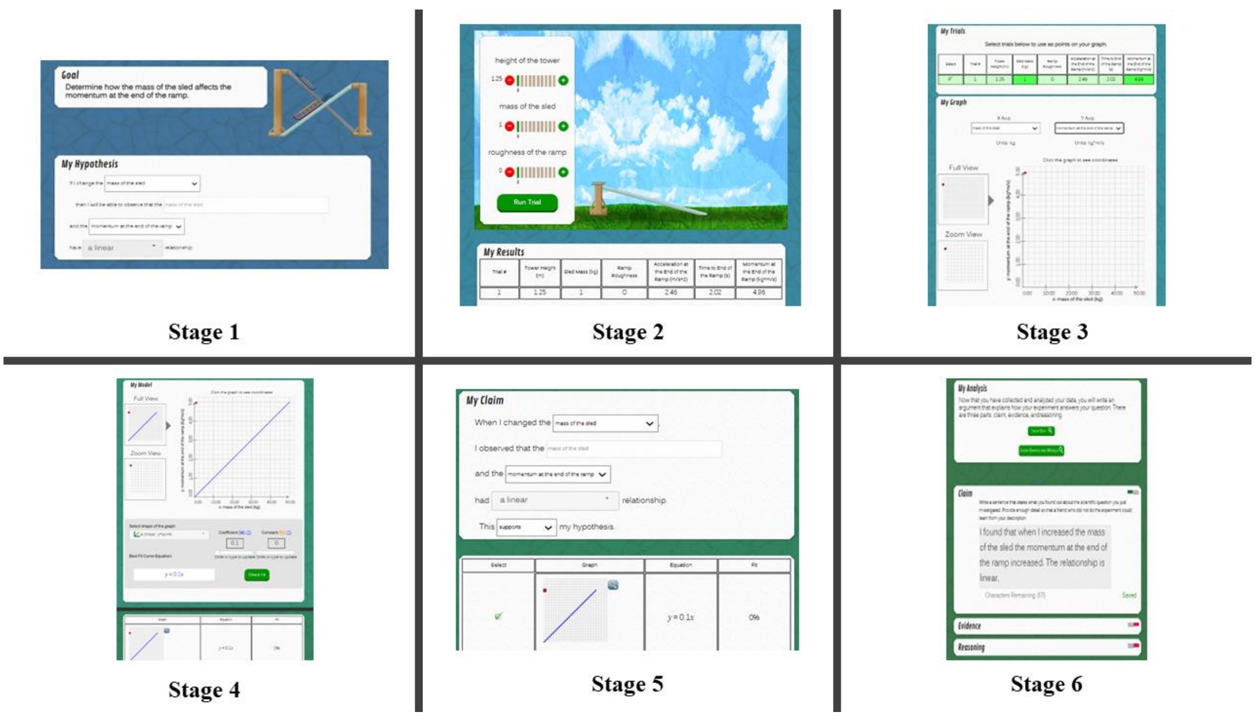

Table 2. Descriptions the Inq-ITS Virtual Lab Stages

Stage	**Related NGSS Practice(s)**	**Description of Stage**
1.Hypothesizing	Practice 1: Asking Questions & Defining Problems	Students form a hypothesis about the mathematical relationship between an independent and dependent variable based on a given goal (e.g., if I change the <u>mass of the sled</u>, then I will be able to observe that the <u>mass of the sled</u> and the <u>momentum of the sled at the end of the ramp</u> have a <u>linear</u> relationship).
2. Collecting Data	Practice 3: Planning & Carrying Out Investigations	Students run trials using a simulation to investigate the relationship between the variables that they outlined in their hypothesis (e.g., mass of the sled and momentum at the end of the ramp). The data that they collect are automatically stored in a data table.
3. Plotting Data	Practice 2: Developing & Using Models Practice 5: Using Mathematics & Computational Thinking	Students select trials from the data they had collected in the previous stage to plot in a graph. Students select the variable to place on the x-axis of their graph and the y-axis of their graph. Ideally, students should place their independent variable (e.g., mass of the sled) on the x-axis and their dependent variable (e.g., momentum of the sled at the end of the ramp) on the y-axis, and students should only plot *controlled* data.
4. Building Models	Practice 2: Developing & Using Models Practice 5: Using Mathematics & Computational Thinking	Students select the type of mathematical relationship that best fits the shape of the plotted data (e.g., linear, inverse, square, inverse square, or horizontal). Students also determine the coefficient and constant for the equation of the best-fit curve/line as well as check the fit (i.e., coefficient of determination, R^2), which is automatically calculated and stored in their table along with a snapshot of their graph and the equation that they built. Ideally, students' model should align with the underlying mathematical formula for the phenomenon (e.g., their model for mass of the sled vs. momentum at the end of ramp should be $y = a \cdot x$, where a is the velocity of the sled, which they

		should hold constant in their experiment. This aligns with the canonical formula for momentum; i.e., momentum = velocity · mass).
5. Analyzing Data	Practice 4: Analyzing & Interpreting Data	Students interpret the results of their graphs by making a claim about the relationship between the variables, identifying if it was the relationship that they had initially hypothesized, and selecting the graphs and corresponding equations that best demonstrated this relationship.
6. Explaining Findings	Practice 6: Constructing Explanations & Designing Solutions	Students write an explanation of their findings in the claim, evidence, and reasoning (CER) format.

Figure 2. Close-up Screenshot of the CER Prompts in the Explain Findings Stage

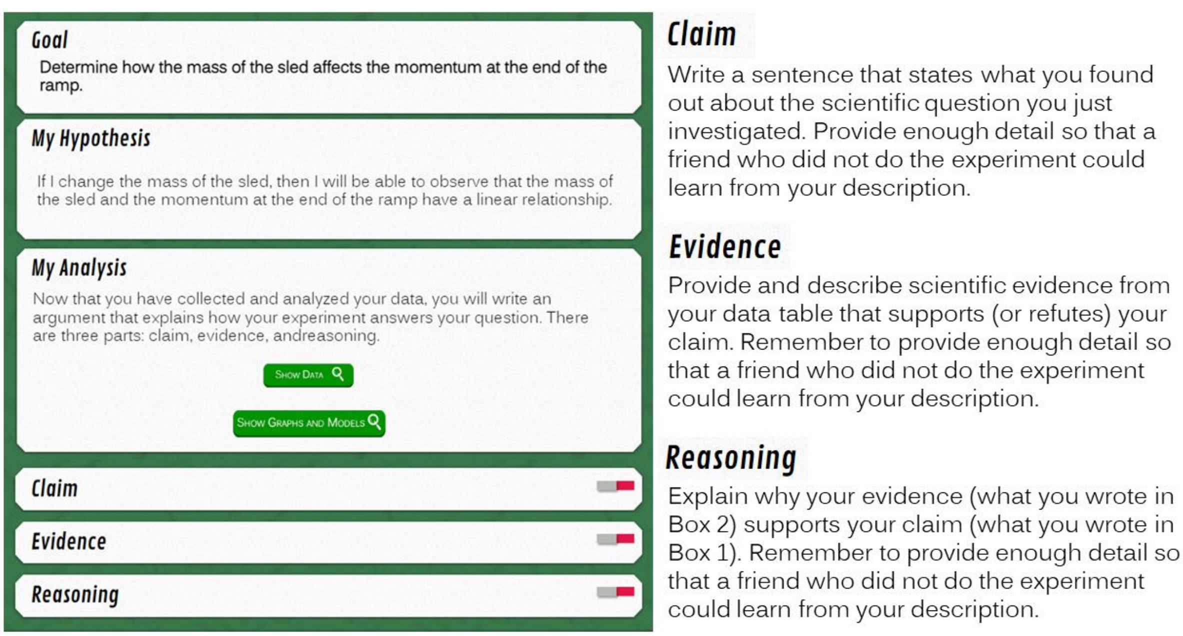

Table 3. Coding Rubric for CER Responses Using Mathematics in Science Inquiry

CLAIM	
Independent Variable	Does the student mention an independent variable (IV)?
	Does the student mention the goal-aligned independent variable (IV) for this lab (i.e., mass of the sled)?
Dependent Variable	Does the student mention a dependent variable (DV)?
	Does the student mention the goal-aligned dependent variable (DV) for this lab (i.e., momentum at the end of the ramp)?
Mathematical Relationship	Does the student mention any mathematical relationship?
	Does the student mention the *CORRECT* mathematical relationship for this lab (i.e., linear for mass of sled vs. momentum at end of ramp)?
Covariational Relationship	Does the student mention a *covariational* (change/increase/decrease) relationship between variables (e.g., "when I increased the mass, the momentum increased")?
	Does the student mention the CORRECT *covariational* (change/increase/decrease) relationship between variables (e.g., "when I increased the mass, the momentum increased")?
EVIDENCE	
Quantitative Data	Does student mention data (exact numbers) from their trials as evidence of the claim?
	Does student mention *at least 2 points of* data (exact numbers) from their trials as evidence of the claim?
	Does student mention *relevant/appropriate* data from their trials as evidence of the claim?
Model Equation	Does the student mention the equation for the chosen best fit model as evidence of the claim?
	Does the student mention the *exact* equation for the chosen best fit model as evidence of the claim?
Model Graph	Does the student mention their graph (including shape of the graph in their own words; e.g. "straight line")?
	Does the student mention the shape of the graph in their own words; e.g. "straight line"?
Model Fit	Does the student mention the fit score or how well the graph fits the data of the best fit model as evidence of the claim?
	Does the student mention the *exact* fit score of their chosen best fit model as evidence of their claim?
Mathematical Relationship	Does the student mention any mathematical relationship?
	Does the student mention the *CORRECT* mathematical relationship for this lab (i.e., linear for mass of sled vs. momentum at end of ramp)?
Covariational Relationship	Does the student mention a covariational (change/increase/decrease) relationship between variables in their stated claim (e.g., "when I increased the mass, the momentum increased")?

	Does the student mention the CORRECT *covariational* (change/increase/decrease) relationship between variables in their stated claim (e.g., "when I increased the mass, the momentum increased")?
	Does the student describe the *rate* at which the DV changes with respect to the IV (e.g., "The momentum increases as at a constant rate when the mass changes")?
REASONING	
Connection	Does the student employ language that draws a connection between the claim and the evidence (e.g., "the evidence supports my claim because...")?
Scientific Theory	Does the student draw connections between their reasoning and a broader scientific theory, principle, or concept?
Quantitative Data	Does student mention data (exact numbers) from their trials?
	Does student mention *at least 2 points of* data (exact numbers) from their trials?
	Does student mention *relevant/appropriate* data from their trials?
Model Equation	Does the student mention the equation for the chosen best fit model?
	Does the student mention the *exact* equation for the chosen best fit model?
Model Graph	Does the student mention their graph (including shape of the graph in their own words; e.g. "straight line")?
	Does the student mention the shape of the graph in their own words; e.g. "straight line"?
Model Fit	Does the student mention the fit score or how well the graph fits the data of the best fit model?
	Does the student mention the *exact* fit score of their chosen best fit model?
Mathematical Relationship	Does the student mention any mathematical relationship?
	Does the student mention the *CORRECT* mathematical relationship for the variables in their stated claim(i.e., linear for mass of sled vs. momentum at end of ramp)?
Covariational Relationship	Does the student mention a covariational (change/increase/decrease) relationship between variables in their stated claim (e.g., "when I increased the mass, the momentum increased")?
	Does the student mention the CORRECT *covariational* (change/increase/decrease) relationship between variables in their stated claim (e.g., "when I increased the mass, the momentum increased")?
	Does the student describe the *rate* at which the DV changes with respect to the IV (e.g., "The momentum increases as at a constant rate when the mass changes")?

Table 4. Student Examples of CER Responses for the Inq-ITS Virtual Lab

<table>
<tr><th colspan="3">Student A - Uses the equation and graph as mathematical evidence to support claim about mathematical relationship</th></tr>
<tr><td rowspan="2">Claim</td><td>Student Response</td><td>If I change the mass of the sled, then I will be able to observe that the mass of the sled and the momentum at the end of the ramp have a linear relationship.</td></tr>
<tr><td>Rubric Elements</td><td>- the correct independent variable ("mass of the sled")
- the correct dependent variable ("momentum at the end of the ramp")
- the correct mathematical relationship ("linear relationship")</td></tr>
<tr><td rowspan="2">Evidence</td><td>Student Response</td><td>When my data is graphed, it shows a linear graph. Also, it follows the equation of y=5x which is a linear equation.</td></tr>
<tr><td>Rubric Elements</td><td>- the graph ("when my data is graphed")
- a specific model equation ("y=5x")
- the correct mathematical relationship ("shows a linear graph")</td></tr>
<tr><td rowspan="2">Reasoning</td><td>Student Response</td><td>My claim was that the mass of the sled and the momentum at the end of the ramp have a linear relationship. My evidence proves this because the graph of the relationship between the mass of the sled and the momentum of the sled is linear.</td></tr>
<tr><td>Rubric Elements</td><td>- a connection between claim and evidence ("My evidence supports this because…")
- the graph ("the graph of the relationship between…")
- the correct mathematical relationship ("linear")</td></tr>
<tr><th colspan="3">Student B – Does not use mathematical evidence (only quantitative data from trials) to support claim about covariational relationship</th></tr>
<tr><td rowspan="2">Claim</td><td>Student Response</td><td>When the mass of the sled was increased, the momentum at the end of the ramp was also increased.</td></tr>
<tr><td>Rubric Elements</td><td>- the correct independent variable ("mass of the sled")
the correct dependent variable ("momentum at the end of the ramp")
the correct covariational relationship ("When the mass of the sled was increased, the momentum at the end of the ramp was also increased. ")</td></tr>
<tr><td rowspan="2">Evidence</td><td>Student Response</td><td>When the sled was 4kg the momentum was 19.82kg*m/s. When the sled was 7kg the momentum was 34.55kg*m/s.</td></tr>
<tr><td>Rubric Elements</td><td>- sufficient & appropriate data (i.e., exact numerical data from at least 2 trials)</td></tr>
<tr><td rowspan="2">Reasoning</td><td>Student Response</td><td>The evidence supports my claim because it shows the mass of the sled increasing and causing the momentum to increase aswell.</td></tr>
<tr><td>Rubric Elements</td><td>- a connection between claim and evidence ("My evidence supports my claim because…")
the covariational relationship ("the mass of the sled increasing and causing the momentum to increase aswell.")</td></tr>
</table>

Student C - *Uses a variety of mathematical evidence (quantitative data, equation, graph, model fit) across both "Evidence" and "Reasoning" responses to support a claim that includes both the covariational relationship and mathematical relationship between the variables*		
Claim	***Student Response***	When I increased the mass of the sled, the momentum of the sled at the end of the ramp also increased. This shows a linear relationship between the mass of the sled and the momentum at the end of the ramp.
	Rubric Elements	- the correct independent variable ("mass of the sled") - the correct dependent variable ("momentum of the sled at the end of the ramp") - the correct covariational relationship ("When I increased the mass of the sled, the momentum of the sled at the end of the ramp also increased.") - the correct mathematical relationship ("linear relationship")
Evidence	***Student Response***	Every time the mass of the sled increased by 3 kilograms, the momentum at the end of the ramp also increased. In trial 1, the mass of the sled was 1 kilogram, and the momentum at the end of the ramp was 4.91 kg*m/s. In trial 10, when the sled mass was 28 kg, the momentum had increased to 138.18 kg*m/s. When this data is graphed, the line of best fit(around 99.96% fit) is linear.
	Rubric Elements	- the correct covariational relationship ("Every time the mass of the sled increased...the momentum at the end of the ramp also increased.") - sufficient & appropriate data (i.e., exact numerical data from at least 2 trials) - the graph ("when this data is graphed, the line of best fit...") - the correct mathematical relationship ("linear") **NOTE: The student would have also gotten a point for the rate of the covariational relationship if the student had specified how much the momentum increased each time the mass of the sled increased by 3.
Reasoning	***Student Response***	Even without graphing, we can see that as the mass of the sled increases, the momentum at the end of the ramp also increases. Then, when the data is graphed, the line of best fit is linear. It is the equation y=5x with a fit of 99.96%. This is a very high fit, showing that the data is linear.
	Rubric Elements	- the correct covariational relationship ("as the mass of the sled increases, the momentum at the end of the ramp also increases") - the graph ("when this data is graphed, the line of best fit...") - the correct mathematical relationship ("linear") - the specific equation of a model ("y=5x") - the specific fit of the model to the data ("a fit of 99.96%)
Student D - *Uses quantitative data as evidence to support claim about mathematical relationship, but the evidence does not support the claim that was made*		
Claim	***Student Response***	The mass of sled had an inverse relationship with the momentum t the end of the ramp.

	Rubric Elements	- the correct independent variable ("mass of sled") - the correct dependent variable ("momentum t the end of the ramp") **NOTE: although the student does mention a mathematical relationship ("inverse relationship"), this is not the correct mathematical relationship. This student may be struggling to determining the type of mathematical relationship (i.e., linear, inverse, etc.) that exist between the data.
Evidence	***Student Response***	When the sleds mass was at 1 the momentum was 4.89 but when the mass was at 28 the momentum was 138.01.
	Rubric Elements	- sufficient & appropriate data (i.e., exact numerical data from at least 2 trials) **NOTE: The "data" are considered appropriate because they refer to correct independent and dependent variables for this experiment. However, this data does not support the student's claim that the relationship between these variables are "inverse".
Reasoning	***Student Response***	As the mass increased the momentum increased
	Rubric Elements	- the correct covariational relationship ("As the mass increased the momentum increased")