Evidence of Accelerated Improvement of Claim-Evidence-Reasoning (CER) Skills with BrainPOP Science: A Case Study of Southeastern Region Districts

Melissa Hogan, EdD, Michelle Newstadt, PhD, Maneeza Dawood, PhD, Ilia Rushkin, PhD, Yigal Rosen, PhD

BrainPOP, 71 W. 23rd St., New York, NY 10010

Author Note

Correspondence concerning this article should be addressed to Melissa Hogan, EdD, BrainPOP, 71 W. 23rd St., New York, NY 10010. Email: melissah@brainpop.com

Abstract

BrainPOP Science is a research-based program that has been shown to be effective in improving and quickly developing students' scientific reasoning skills. This white paper presents a comprehensive case study conducted in school districts in the Southeastern United States. BrainPOP Science is one of the leading supplemental middle school programs. This study investigated the effectiveness of BrainPOP Science for enhancing students' proficiency in constructing scientific explanations and arguments by utilizing evidence and reasoning according to the Claim-Evidence-Reasoning (CER) framework. Data were collected from students' first and fourth CER submissions to explore changes in CER skills over time. A multivariate analysis of covariance (MANCOVA) was conducted to assess the joint significance of the difference between students' first and fourth claim, evidence, and reasoning subscores and total CER scores, as well as to examine the role of BrainPOP Science usage as a covariate. The results reveal significant learning gains on average between students' first and fourth CER scores. Multivariate analysis found the model was jointly significant for the total CER score, as well as each claim, evidence and reasoning subscores. A significant effect of BrainPOP usage was also observed, such that any usage (low, moderate, high) increased the total CER score, and evidence subscores. Additionally, moderate usage increased claim subscores, and moderate and high usage increased reasoning subscores.

Keywords: middle school science, claim-evidence-reasoning (CER) framework, scientific literacy, rapid skill development, BrainPOP, multidimensional learning, Edtech usage

Table of Contents

Evidence of Accelerated Improvement of Claim-Evidence-Reasoning (CER) Skills with BrainPOP Science: A Case Study of Southeastern Region Districts

The world is rapidly transforming, and with it, the need for scientific literacy and critical engagement with complex scientific concepts is more crucial now than ever before. The ability to carefully analyze information and draw grounded conclusions is a critical life skill for young people. Science education plays a significant role in cultivating these skills by testing ideas and theories against real-world evidence. In addition, 2019 data shows that the percent of students below proficiency on 8th grade science assessments such as the National Assessment of Educational Progress (NAEP) was 33% (National Center for Education Statistics, 2022). Developing critical thinking skills through multidimensional science learning in middle school can increase academic self-efficacy and achievement, preparing students for one of the most challenging experiences in secondary education: the transition from middle school to high school (Evans et al., 2018).

The Next Generation Science Standards (NGSS) recognize the importance of multidimensional science learning to foster critical thinking skills among students (National Research Council, 2013). NGSS emphasizes the development of students' abilities to analyze and interpret scientific information, construct evidence-based explanations and arguments, and communicate scientific ideas effectively. This approach recognizes that scientific knowledge is not limited to isolated facts; instead, it encompasses the integration of disciplinary core ideas (DCIs), science and engineering practices (SEPs), and crosscutting concepts (CCCs).

Scientific Literacy

Scientific literacy is a multifaceted concept that encompasses various dimensions and can be interpreted differently depending on the context. It is crucial to acknowledge the diverse perspectives and definitions of scientific literacy in order to understand its comprehensive nature. The Programme for International Student Assessment study conducted by the Organisation for Economic Co-operation and Development (OECD) considered scientific literacy an essential goal of science learning and teaching. The study defines scientific literacy as "the capacity to use scientific knowledge, to identify questions and to draw evidence-based conclusions in order to understand and help make decisions about the natural world and the changes made to it through human activity" (OECD, 2006). The National Academy of Sciences defines scientific literacy as "knowledge and comprehension of scientific principles and procedures necessary for making informed personal decisions, engaging in civic and cultural activities, and achieving economic productivity" (National Academy of Sciences, 1996).

Scientific Literacy and the Claim-Evidence-Reasoning (CER) Framework

One instructional framework that aligns with some components of scientific literacy and promotes multidimensional scientific knowledge development is the Claim-Evidence-Reasoning (CER) framework (McNeill & Krajcik, 2011). The framework finds that a student's explanation can be deconstructed into a *claim* that answers the proposed question, *evidence* based on relevant observations to support the claim, and *reasoning* that uses logic to link the cited evidence to the claim (Hardcastle et al., 2021). The CER employs this deconstruction to provide a structured approach for argumentation and communicating explanations. The framework encourages students to make clear and concise claims that address specific scientific questions. It engages them in scientific inquiry and helps students investigate questions through observations and gather relevant data to utilize as evidence. Constructing explanations and arguments is a higher-order thinking task that incorporates many science practices and skills.

While the CER framework embodies some aspects of scientific literacy, it is important to note that scientific literacy extends beyond this framework alone. Scientific literacy encompasses a broader range of skills, such as understanding the nature of science, critically evaluating scientific information, and recognizing the social and ethical implications of scientific discovery (Eisenhart et al., 1996). Still, the CER framework acts as a valuable tool for developing key elements of scientific literacy, including critical thinking, evidence-based reasoning and writing, and effective communication.

Effectiveness of CER Instruction

The CER framework has proven to be an effective approach to assessing students' multidimensional science knowledge (Hardcastle et al., 2021; Gotwals & Songer, 2013). A study conducted in eight middle school classrooms found that only 18.1% (*N*=72) of students were able to construct explanations encompassing all three components of claim, evidence, and reasoning. Surprisingly, 40% of students made claims without providing any supporting evidence or reasoning. The study did reveal a positive correlation between students' production of high-quality explanations and their overall classroom performance, unrelated to argumentative writing skills (Ruiz-Primo, Li, Tsai, & Schneider, 2008). This suggests that exposing students to the process of developing explanations can lead to a deeper understanding of scientific concepts. Concurrently, the study emphasizes the limited opportunities available for students to engage in such explanatory practices and the importance of providing adequate guidance and support during the construction of CER explanations. They suggested that science notebooks provide context into students' thinking and learning and can provide some evidence of teachers' communications with students about their progress. In this study, teachers who implemented notebooks as scaffolds for organizing claims, evidence, and reasoning struck a balance between structure and independent thinking, leading to more effective outcomes. Notebooks with less guidance lacked focus, while those with excessive guidance resulted in information copying without much interpretation from students. The process of constructing CER explanations through structured writing helps students to organize their thoughts, think critically, and develop their scientific literacy.

Another study aimed to understand students' intermediary knowledge as they progressed toward a more sophisticated knowledge of ecology. The findings of this study indicated significant gains in student learning following a curricular intervention. However, despite overall improvements, some students continued to struggle with explaining the potential impacts of disturbances on ecosystems (Gotwals & Songer, 2010). This research emphasizes the importance of having a learning progression framework to guide the design of assessment tasks and the interpretation of evidence. The assessment system employed in the study identified multiple types of middle knowledge that students may possess, highlighting the existence of "messy middles" as students navigate their development of reasoning abilities in complex scientific situations.

These studies highlight the necessity of high-quality educational resources and instruction for students to develop Claim-Evidence-Reasoning skills. While this research has demonstrated the positive impact of CER instruction, certain challenges and limitations remain. Implementing CER effectively requires teacher support, training, and scaffolding to ensure students grasp the underlying scientific concepts and reasoning processes (McNeill et al., 2006; Yao et al., 2016). Prior research has primarily focused on students' overall science performance, rather than the specific development of Claim-Evidence-Reasoning skills (Masters, 2020; Yao et al., 2016; McNeill & Krajcik, 2011). Previous research has also provided limited insight into student learning gains within the specific subcategories of CER, such as constructing claims, providing evidence, and developing reasoning skills. Prior research on CER assessment has found that students distribute themselves along a hierarchy of difficulty for each of the CER components, with writing a claim being the least difficult and providing reasoning being the most difficult for students to include in their writing (Gotwals & Songer, 2013). Understanding how students progress in each sub-category provides critical insight into changes in their ability to formulate and choose evidence and construct sound scientific reasoning. This is crucial for targeted instructional support. When investigation and design are at the center of learning, students can gather evidence and take ownership of the evidence they have gathered (Rosen et al., 2020). This process contributes to student agency as they make sense of phenomena and designs and extend their understanding of the natural and designed world.

BrainPOP Science

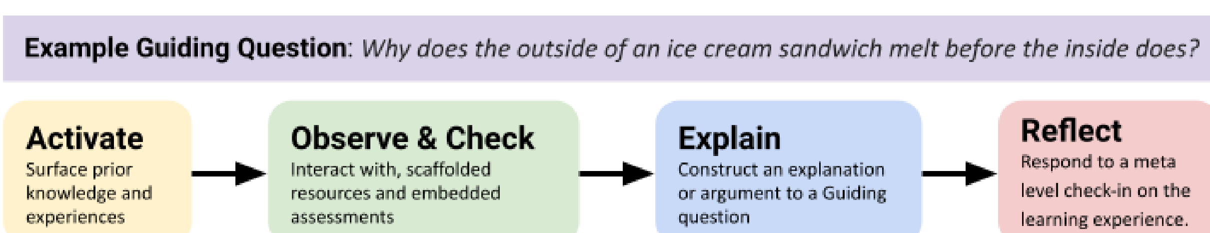
Education technology has revolutionized the way students engage in multidimensional science by blending the disciplinary core ideas, scientific practices, and crosscutting concepts through the use of a Claim-Evidence-Reasoning (CER) framework in the classroom (Rosen et al., 2021; Rosen, 2021). BrainPOP Science is a comprehensive, supplemental middle school science experience that embeds the CER process throughout the investigation (e.g., lesson) structure to support multidimensional learning.

BrainPOP Science Learning Experience

BrainPOP Science is designed around a learning progression approach in which students revisit concepts through multiple modalities and contexts. This allows students to build from concrete knowledge to more abstract and complex knowledge and practices. BrainPOP Science's ready-to-use, standards-aligned investigations center around relatable Guiding Questions and real-world phenomena that spark middle school learners' curiosity. Interactive resources—such as Data Manipulatives, simulations, and BrainPOP 3D Worlds™—and vocabulary-rich resources— like movies and Related Readings—encourage students to make observations, analyze real-world data sets, and model scientific concepts. Embedded formative assessments give students an opportunity to show what they know in a low-stakes environment. Additionally, multidimensional assessment items that mimic higher-stakes assessments can be assigned outside of the investigation structure.

Figure 1

BrainPOP Science Claim-Evidence-Reasoning (CER) Framework



The investigation structure embeds and scaffolds the CER process throughout the learning experience. Students are prompted and encouraged to make observations as they interact with resources, such as the 3D World, simulation, Data Manipulatives, Primary Sources, Movies, and Related Readings. At the end of each investigation, students synthesize the information presented in these resources by constructing an explanation or argument. Students review their observations and choose which observations become evidence to support their claim to the Guiding Question. In the reasoning section, they connect their claim and evidence using scientific principles and concepts. Each section (claim, evidence, and reasoning) includes scaffolds and examples that remind students how to make a concise claim, curate strong evidence, and construct reasoning with examples of each. Constructing a complete CER is a high cognitive demand task that encourages students to effectively communicate complex scientific concepts and practices.

Figure 2

Claim-Evidence-Reasoning Student Draft

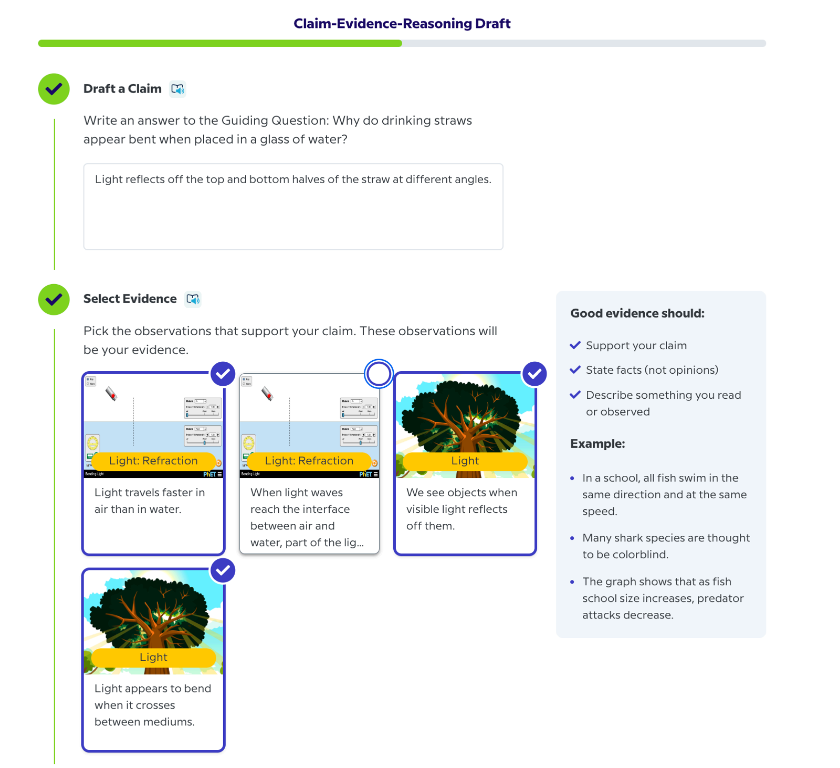
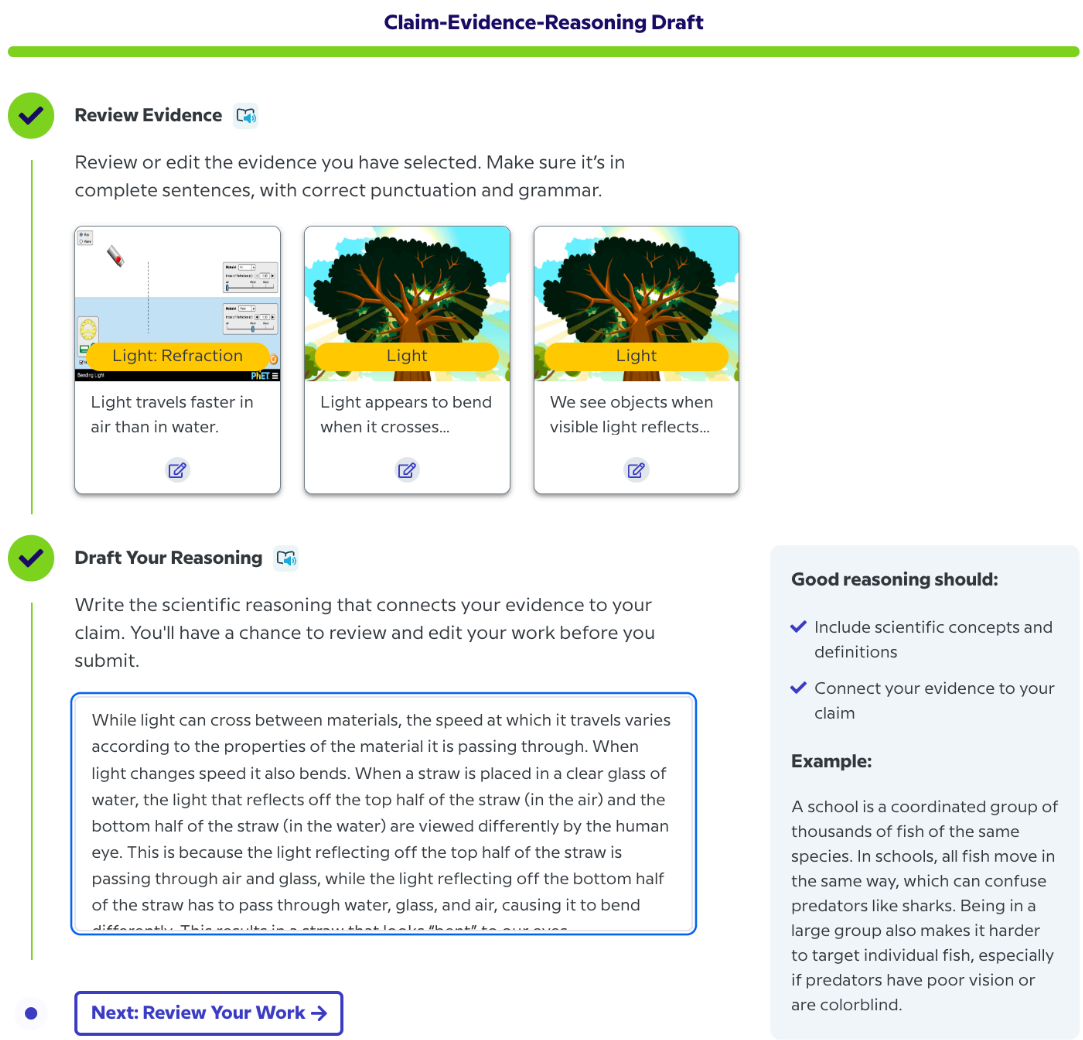


Figure 3

Claim-Evidence-Reasoning (CER) Student Submission



The Present Study

Given the aforementioned research, the present study aims to examine BrainPOP Science's approach to CER skill development (Rosen et al., 2021; Rosen, 2021). Specifically, this study examines students' knowledge development through evidence-based writing. Students interact with four BrainPOP Science investigations within a six-month timeframe and construct CER explanations at the end of each investigation. The goal is to assess students' progression in claim construction, evidence curation, and reasoning skills in a short period of time and after completing four CER submissions. As such, the present study asks two primary research questions:

1. What is the relationship between interactions with BrainPOP Science and skill progressions in developing multidimensional knowledge (e.g., CER construction)?
2. What is the relationship between consistent and sustained usage of BrainPOP Science on student learning gains in each CER subcategory (Claim, Evidence, Reasoning) between their first and fourth CER submissions?

Methodology

Data Collection and Participants

We began by finding the top twelve CER topics assigned to middle school science students on BrainPOP Science. By focusing on these specific topics, we aimed to ensure that the collected data would be representative of diverse concepts covered in middle school science curricula. Subsequently, we identified schools with at least 50 students who completed at least four CER submissions, indicating significant engagement with the CER framework. This criterion allowed for the selection of schools that had incorporated CER assignments as a substantial part of their instruction.

Seven schools in four districts of the Southeastern United States met this criteria. District demographic information for the four districts included in the study are shown in Table 1. By narrowing our study to these districts, we were able to explore the implementation and impact of the CER framework within a specific educational context, i.e., by examining the effectiveness and potential variations in CER implementation within a regional setting. Within these districts, each student's first and fourth CER submissions (approximately six months apart) were examined by content experts to assess potential changes in students' growth in multidimensional science knowledge over time. This paired approach enabled comparisons of participants' own performance as well as evaluations of learning gains in CER scores between their first and fourth submissions.

Table 1

District Demographics

	District A	District B	District C	District D
Total Students	161,784	2,794	270,978	49,769
Female	49%	48%	48%	49%
Male	51%	52%	52%	51%
Native American	<1%	2%	<1%	<1%
Asian/Pacific Islander	9%	6%	3%	1%
Hispanic	18%	37%	34%	9%
Black	22%	9%	38%	37%
White	45%	38%	19%	48%
Native Hawaiian/Other Pacific Islander	<1%	<1%	<1%	<1%
Two or more races	3%	8%	2%	2%
Low Income	33%	31%	60%	57%
ELL	9%	8%	12%	6%

Automated grading of open-ended CER responses

Teachers (and students) are provided with insights about “hard-to-measure” skills such as Claim-Evidence-Reasoning (CER) in Science and Literacy/ELA argumentation with evidence in Essentials and strategies for improvement.

Grading open-ended responses is done by means of Large Language Models (LLMs). In the prompts that we provided to the model, the grading rules are equivalent to the scoring rubrics (Figure 4) and request an outcome consisting of scores and score rationales. While using the foundational general-purpose LLMs is possible, we also deploy LLMs fine-tuned using a training set of CER responses hand-graded by content experts. Likewise, we have the option of including some of the hand-graded responses as primer examples in the LLM prompts.

Figure 4

BrainPOP's CER Scoring Rubric

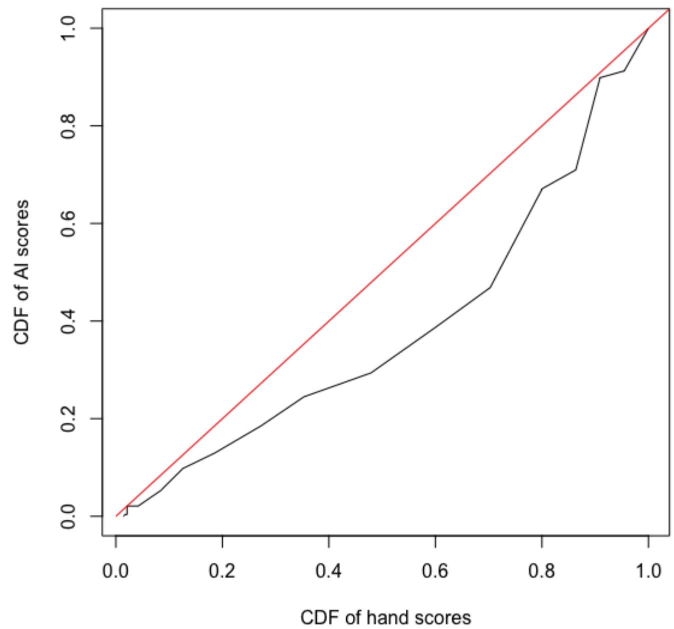
		0	1	2	3
Claim	**Completeness**	Not attempted	The claim: ☐ Is unrelated to the Guiding Question	The claim: ☐ Partially answers the Guiding Question	The claim: ☐ Completely answers the Guiding Question
	Accuracy and Relevance	Not attempted	Information is: ☐ Not relevant AND/OR accurate	Information is: ☐ Partially accurate and relevant scientific concepts	Information is: ☐ Fully scientifically accurate and relevant
Evidence	**Completeness**	Not attempted	Evidence statements: ☐ Are not complete AND/OR do not contain evidence	Evidence statements: ☐ Each have at least one piece of evidence that support the claim	Evidence statements: ☐ Have two or more pieces of evidence each that sufficiently support the claim
	Accuracy and Relevance	Not attempted	Evidence is: ☐ Not included or unrelated to the claim or inaccurate	Evidence is: ☐ Partially accurate and related to the claim	Evidence is: ☐ Fully accurate and supports the claim
Reasoning	**Connections**	Not attempted	Reasoning: ☐ Does not connect evidence to the claim	Reasoning: ☐ Connects some evidence to the claim	Reasoning: ☐ Completely connects all evidence to the claim
	Accuracy and Relevance	Not attempted	Scientific principles are: ☐ Not relevant AND/OR accurate	Scientific principles are: ☐ Partially accurate and include relevant scientific principles	Scientific principles are: ☐ Fully scientifically accurate and relevant
Writing	**Clarity**	Not clear	Claim, evidence, and reasoning are clearly communicated using complete sentences.		
	Organization	Not organized	Students use academic vocabulary and transitions to connect evidence and reasoning to claim.		

Interrater Reliability

Interrater Reliability is crucial for ensuring that data are consistent and accurate representations of the measured variables. The quadratic weighted kappa statistic was used to test interrater reliability as a measure to assess the level of agreement between hand-scored CERs and "Mantis," BrainPOP's automated scoring model (Nagaraj et al., 2018). The interrater reliability of hand-scored CERs against Mantis was 0.679, which is interpreted as a substantial agreement among scores (Landis & Koch, 1977). Overall, Mantis tended to give higher scores, acting as a more lenient scorer. This can be seen in Figure 5, where the curve goes less steeply than the x=y line in the region of lower scores, which means that Mantis assigns fewer of those.

Figure 5

Cumulative distribution function (CDF) of Mantis scores (y) vs CDF of hand-scores (x). The red line (x=y) represents the ideal agreement. The black line is the actual relationship found.



Results

Overall, 548 student participants who met the study criteria were identified and used in the analyses (Table 2). BrainPOP Science usage was examined as a covariate to explore the relationship between students' performance and growth in CER scores. High usage corresponded to districts that completed at least 55% or more of assigned BrainPOP Science investigations. Moderate usage corresponded to districts that completed between 30-54% of assigned BrainPOP Science investigations, and low usage corresponded to districts that completed less than 30% of assigned BrainPOP Science investigations.

Table 2

Study Descriptives

District	School	Usage	# of CERs	# of Students	CER 1	CER 4
A	A	Moderate	164	82	Solids, Liquids & Gases	Properties of Light
A	B	Moderate	122	61	Chemical Reactions	Process of Photosynthesis
A	C	Moderate	272	136	Atoms & Molecules	Properties of Light/Earth's Water Cycles
B	D	High	114	57	Properties of Matter	Human Use of Natural Resources
C	E	Moderate	124	62	Phase Changes	Refraction, Reflection, and Absorption
C	F	Moderate	146	73	Thermal Energy Transfer	Volcano Formation
D	G	Low	154	77	Phase Changes	Neurons and the Brain
	Total		1,096	548		

A multivariate analysis of covariance (MANCOVA) was conducted to assess the joint significance of the difference between students' first and fourth claim, evidence, reasoning and total CER scores. In addition to assessing the joint significance of subscores and total score, a multivariate analysis of covariance (MANCOVA) was conducted to examine districts' level of BrainPOP Science usage as a covariate. MANCOVA results found an overall joint significance of the average difference in scores between students' fourth and first CER for the total score, as well as each claim, evidence, and reasoning subscore (Table 3, Figure 6).

Table 3

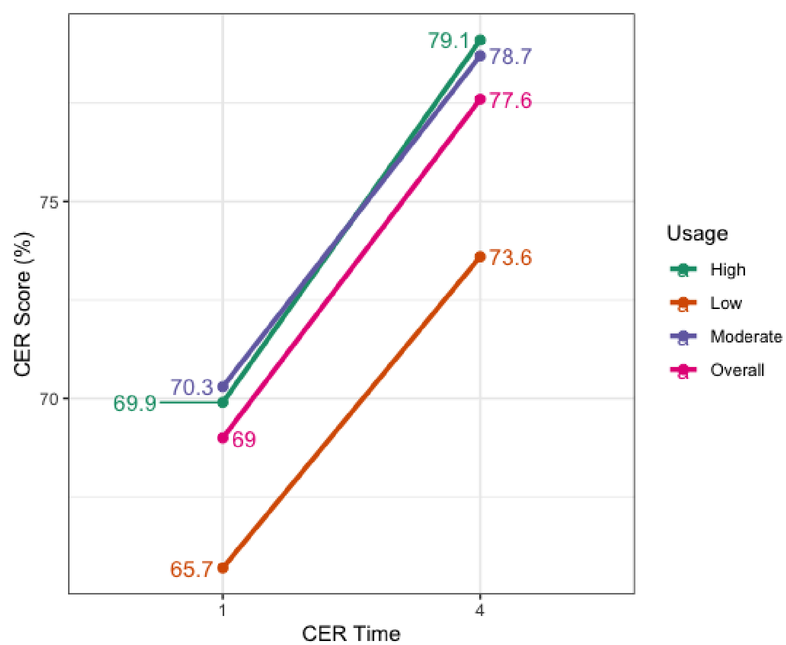
Joint Significance of Overall Model

Difference Score (T4-T1)	# of students	Parameters	RMSE	R-square value	F-value	P-value
Total	548	3	299.598	.11	23.212	<.001***
Claim	548	3	5.858	.02	3.391	.018*
Evidence	548	3	62.853	.11	21.579	<.001***
Reasoning	548	3	65.516	.10	18.988	<.001***

Note. *p < .05. **p < .01. ***p < .001.

Figure 6

Change in Total Score (%) Between First and Fourth CER



A significant multivariate effect of BrainPOP Science usage was observed on the joint CER scores (*Wilks' λ* = 0.811, *F*(12, 548) = 9.873, *p* < 0.001). Univariate analyses found a significant effect of BrainPOP usage, such that any usage (low, moderate, high) increased the total CER score, and evidence subscores. Additionally, moderate usage increased claim subscores, and moderate and high usage increased reasoning subscores.

Figure 7

Change in Subscores Between First and Fourth CER

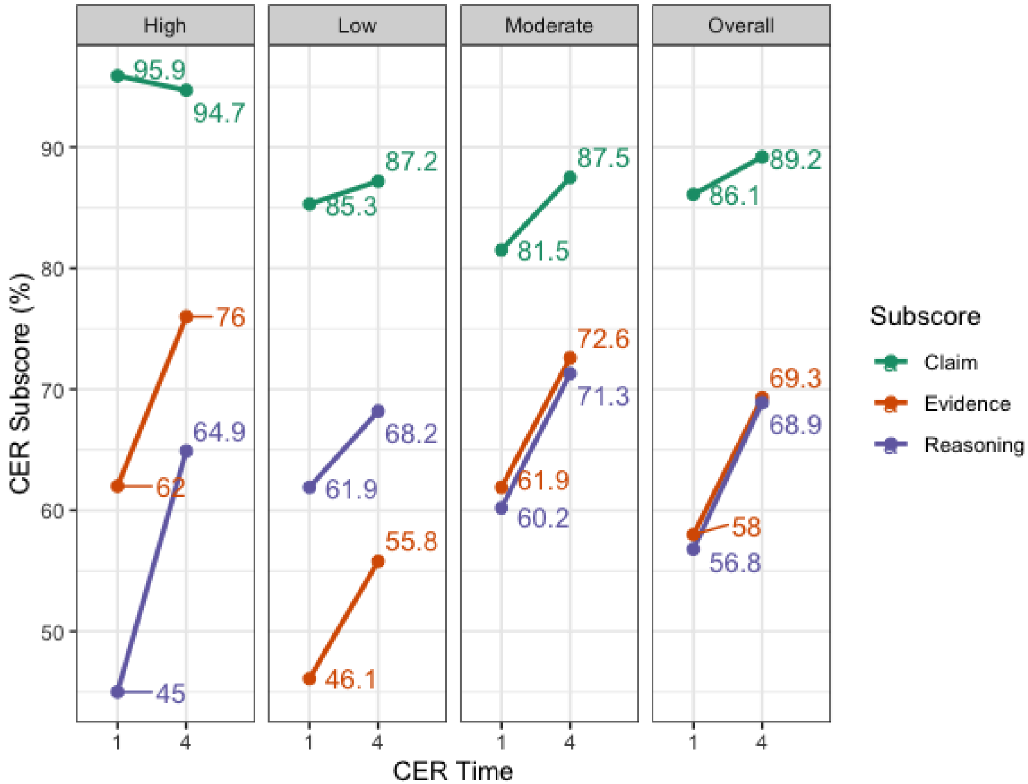


Figure 7 demonstrates that any level of BrainPOP usage (low, moderate, high) showed improvement in the total CER score and evidence subscore, as evidenced by the steep positive learning curve. Additionally, moderate BrainPOP Science usage showed improvement in the claim subscore, and both moderate and high BrainPOP Science usage showed improvement in the reasoning subscore.

Table 4

Univariate Tests of Usage on Each Subscore

	Coef.	Std. Error	T-value	P-value	Conf. Interval (95%)
Claim difference score					
Low Usage	.117	.150	.780	.436	[-.177,.411]
Moderate Usage	.198	.065	3.066	.002**	[.071,.325]
High Usage	-.070	.174	-.403	.687	[-.412,.272]
Evidence difference score					
Low Usage	.922	.194	4.741	<.001***	[.540,1.304]
Moderate Usage	.447	.084	5.327	<.001***	[.282,.612]
High Usage	.842	.226	3.725	<.001***	[.398,1.286]
Reasoning difference score					
Low Usage	.377	.212	1.779	.076	[.039,.792]
Moderate Usage	.502	.091	5.503	<.001***	[.323,.682]
High Usage	1.193	.246	4.849	<.001***	[.710,1.676]
Total difference score					
Low Usage	1.597	.409	3.902	<.001***	[.793,2.402]
Moderate Usage	1.109	.177	6.279	<.001***	[.762,1.456]
High Usage	1.842	.476	3.871	<.001***	[.907,2.777]

Note. $^*p < .05$. $^{**}p < .01$. $^{***}p < .001$.

BrainPOP Science Investigation Resources Impact on Growth

To contextualize and extend the findings from research questions one and two, as well as to provide further context of the use of BrainPOP Science on CER performance and growth, an additional exploratory research question was addressed.

Exploratory Research Question:

1. *To what extent did other (non-CER) BrainPOP Science resources impact Claim-Evidence-Reasoning learning gains?*

This additional question was designed to evaluate the effectiveness of the usage of the additional resources within investigations on students CER learning gains. Usage data was collected on the additional resources available to students within an investigation (Table 5).

Table 5

Usage of Other (non-CER) BrainPOP Science Investigation Resources

	District A	District B	District C	District D
Growth Percentage (Total CER Score)	29%	33%	27%	8%
Usage Group	Medium	High	Medium	Low
Investigations & CERs Assigned	1,266	206	708	882
CER Completion Rate	44%	55%	38%	17%
Checkpoint Completion Rate	68%	52%	64%	0%
Quiz Completion Rate	0%	66%	0%	0%

Table 5 provides additional evidence that students who engage with BrainPOP Science more frequently and more in depth by utilizing multiple resources within the investigations see greater learning gains in the development of multidimensional knowledge and CERs. As highlighted above, District B showed the greatest learning gains, had the highest usage, and utilized all other resources within the investigation at a high completion rate. District D, in the low usage group, showed minimal CER learning gains and did not show any usage of the additional resources. Both District A and District C, in the moderate usage group, demonstrated significant learning gains and both showed moderate usage for at least one additional resource. It is evident in the results that for the average student, general usage of BrainPOP Science Investigations and CERs are beneficial, but moderate to high usage of BrainPOP Science in addition to other supporting resources is all the more beneficial for students.

Discussion

The present study aimed to examine the relationship of consistent and sustained BrainPOP Science usage on students' learning gains in the development of multi-dimensional knowledge through evidence-based writing and in gathering evidence and constructing reasoning.

Results found that, on average, deeper learning with BrainPOP Science improved students' Claim-Evidence-Reasoning skills. These findings have important implications for educators and curriculum designers. Educators can leverage BrainPOP Science as an effective supplemental program to facilitate students' knowledge of scientific concepts, their ability to construct well-structured scientific explanations, and their growth in formulating claims, presenting evidence, and providing reasoning. Students' CER experience in BrainPOP Science prepares them for the multidimensional expectations on summative assessments. This test readiness experience gives students opportunities to demonstrate multidimensional knowledge that prepares them for state assessments in a low-stakes and engaging environment. This proficiency extends beyond test readiness; it cultivates analytical thinking, effective communication, and logical argumentation—which are not only relevant for assessments but also set the foundation for skills needed to be successful in high school and beyond..

Moreover, the study's findings contribute to the existing body of research on the effectiveness of educational resources in supporting students' development within the CER framework. The shifts observed in the total CER scores, as well as each claim, evidence and reasoning subscores suggest that students are actively engaging in the development of these challenging skills. It shows that they are capable of rapid skill development when provided with targeted instruction, scaffolding, and opportunities to practice constructing well-structured reasoning statements.

This substantial growth in students' evidence and reasoning skills highlights the potential of targeted instructional strategies, such as the CER framework employed in BrainPOP Science, to facilitate students' comprehension and application of scientific knowledge. By highlighting the positive outcomes associated with consistent engagement with BrainPOP Science, this study adds empirical evidence to the literature and reinforces the importance of incorporating CER instruction in middle school science education.

There are several limitations of this study that may be addressed in future research. First, this study only explored CER scores at two points in time. Future research can explore the longitudinal impacts of BrainPOP Science at additional time points to further explore the development of Claim-Evidence-Reasoning skills. Second, there is limited information regarding the context of the schools in the study that may have influenced the moderating role of school on the change in CER scores across the two time points. Future research can take a mixed-methods approach, connecting qualitative and observational data on the schools and classroom environments that may impact CER scores, in addition to quantitative data. Finally, this study explores change in CER scores using a within-subjects observational approach. Future research can expand on this work through a structured design in which students may be assigned CER instruction compared to a control group as well as using external measures such as Science state assessments.

The integration of research-based educational technology, exemplified by platforms like BrainPOP, has the potential to transform science education in middle schools, as demonstrated by previous studies (Rosen, 2009; Rosen at a., 2020). Research on CER has shown its effectiveness in promoting student learning outcomes and developing proficiency in scientific practices. When combined with educational technology, such as the interactive features and engaging content, science instruction becomes more immersive, accessible, and conducive to critical thinking and scientific inquiry. As the field of educational technology continues to evolve, it is essential to explore new ways to leverage these tools to enhance science education and prepare students for the demands of a rapidly changing world.

Implications

A key implication of this research is the expansion of teacher capacity. BrainPOP Science provides educators with ready-to-use, standards-aligned investigations that effectively support multidimensional science learning. The platform offers a scaffolded approach to evidence-based writing, incorporating exemplar answers and assessment rationales. This strategic integration of teaching resources empowers educators to confidently guide students in their development of Claim-Evidence-Reasoning skills, which is transferable to high-school readiness and preparation for state summative assessments. Through repeated engagement with the CER framework, students develop their scientific literacy and claim-evidence-reasoning skills through the iterative process of collecting additional evidence, writing and revising CER submissions (Masters & Docktor, 2022; Arias & Davis, 2017). The alignment between BrainPOP Science's CER instruction and the evolving landscape of educational assessment equips students with transferable scientific literacy skills to succeed in emerging summative assessment formats. By explicitly examining the effects of instructional interventions on scientific literacy skills within the CER framework, this study provides valuable insights that contribute to a more comprehensive understanding of students' development in evidence-based reasoning skills and scientific communication.

References

Arias, A. M., & Davis, E. A. (2017). Supporting children to construct evidence-based claims in science: Individual learning trajectories in a practice-based program. *Teaching and Teacher Education*, 66, 204–218.

Eisenhart, M., Finkel, E., & Marion, S. F. (1996). Creating the conditions for scientific literacy: A re-examination. *American Educational Research Journal*, *33*(2), 261-295.

Evans, D., Borriello, G. A., & Field, A. P. (2018). A review of the academic and psychological impact of the transition to secondary education. *Frontiers in Psychology*, *9*, 1482.

Gotwals, A. W., & Songer, N. B. (2010). Reasoning up and down a food chain: Using an assessment framework to investigate students' middle knowledge. *Science Education*, *94*(2), 259-281.

Gotwals, A. W., & Songer, N. B. (2013). Validity evidence for learning progression-based assessment items that fuse core disciplinary ideas and science practices. *Journal of Research in Science Teaching*, *50*(5), 597-626.

Hardcastle, J. M., Herrmann Abell, C. F., & DeBoer, G. E. (2021). Validating a Claim-Evidence-Science Idea-Reasoning (CESR) Framework for Use in NGSS Assessment Tasks. *Grantee Submission*.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159-174.

Masters, H. (2020). Using teaching rehearsals to prepare preservice teachers for explanation-driven science instruction. *Journal of Science Teacher Education*, *31*(4), 414-434.

Masters, H., & Docktor, J. (2022). Preservice teachers' abilities and confidence with constructing scientific explanations as scaffolds are faded in a physics course for educators. *Journal of Science Teacher Education*, *33*(7), 786-813.

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.

McNeill, K. L., & Krajcik, J. S. (2011). Supporting Grade 5-8 Students in Constructing Explanations in Science: The Claim, Evidence, and Reasoning Framework for Talk and Writing. *Pearson*.

Nagaraj, A., Sood, M., & Srinivasa, G. (2018, July). Real-time automated answer scoring. In *2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)* (pp. 231-232). IEEE.

National Academy of Sciences. 1996. *National science education standards*. Washington, DC: National Academy Press.

National Center for Education Statistics. (2022). Science Performance. *Condition of Education*. U.S. Department of Education, Institute of Education Sciences. Retrieved [date], from https://nces.ed.gov/programs/coe/indicator/cne.

National Research Council. (2013). Next generation science standards: For states, by states.

OECD. (2006). *Assessing scientific, reading and mathematical literacy: A framework for PISA 2006*. Paris: OECD.

Rosen, Y. (2021). Mind the Gap: Reimagining Science Education to Build Mastery for All. https://blog.brainpop.com/reimagining-science-education-for-all/

Rosen, Y. (2009). The effects of an animation-based on-line learning environment on transfer of knowledge and on motivation for science and technology learning. J*ournal of Educational Computing Research, 40*(4), 451-467.

Rosen, Y., Arieli-Attali, M., Ward, S., Seery, J., Simmering, V., & Ozersky, L. (2020). HERA: Exploring the power of adaptive scaffolding on scientific argumentation and modeling competencies in online learning systems. Paper presented at The International Conference of the Learning Sciences (ICLS), Nashville, TN.

Rosen, Y., Ozersky, L., Whitehead, M., Rushkin,I., Dawood,M. & O'Riordan, E. (2021). Building Science Mastery with BrainPOP: Three Dimensional Research-Based Learning Unlocked. https://content.brainpop.com/rs/567-ANS-609/images/BrainPOPScience-Mastery-Research-Report.pdf?utm_source=press&utm_medium=external&utm_campaign=scilaunch&_gl=1*12d54t6*_ga*NjgwNDEwNzE5LjE2MDkxNzk4MTk.*_ga_WC2T867EL3*MTY4NzI2NTkyNy4xLjAuMTY4NzI2NTkyNy4wLjAuMA..&_ga=2.56771946.1909940973.1687265928-680410719.1609179819

Ruiz-Primo, M. A., Li, M., Tsai, S. P., & Schneider, J. (2008). Testing One Premise of Scientific Inquiry in Science Classrooms: A Study That Examines Students' Scientific Explanations. CRESST Report 733. *National Center for Research on Evaluation, Standards, and Student Testing (CRESST)*.

Yao, J. X., Guo, Y. Y., & Neumann, K. (2016). Towards a hypothetical learning progression of scientific explanation. *Asia-Pacific Science Education*, *2*(1), 1-17.