Multiple-Document Comprehension in High School Science: A Learning-Engineering Pilot Study

Andrew Potter, Tracy Arner, Kathryn S. McCarthy, & Danielle S. McNamara

integrated literacy
learning engineering
multiple-document comprehension
participatory research design
science learning

This study applies a learning-engineering process to refine and pilot an instructional framework designed to help high school science students read and write from multiple sources. Building on prior participatory design research with science teachers, the pilot examined how one biology teacher implemented a structured, five-lesson sequence and how her reflections and student work informed iterative improvement. Findings indicated that the lesson design was clear, manageable, and compatible with existing science routines, helping students build confidence before engaging with more rigorous texts. Paraphrasing provided an accessible entry point for writing, while students required additional modeling for source evaluation and elaboration. Teacher feedback led to targeted design refinements, including checklists and low-stakes grading to support accountability. The study illustrates how learning engineering can use classroom data to align

instructional design with authentic teaching contexts, promoting feasible and scalable approaches to integrated literacy instruction in science.

Introduction

Students today are expected to evaluate, synthesize, and communicate information drawn from multiple texts, which are skills that are essential for reasoning in science and emphasized in educational standards and international assessments (OECD, 2019). Yet high school science teachers report limited preparation and instructional time to engage students in integrated reading and writing tasks that require analysis across sources (Drew & Thomas, 2017). Addressing this gap requires approaches that make such instruction manageable within everyday classroom conditions. This study used a learning-engineering process to refine and pilot a structured lesson sequence for teaching these skills, which was originally designed through a participatory research study with high school science teachers (Potter et al., 2025). Serving as the implementation and data-instrumentation phase of a larger learning-engineering cycle (Goodell, 2023), the pilot examined teacher and student evidence to guide subsequent design improvements and identify practical conditions for classroom feasibility.

Related Work

Research on multiple-document (MD) comprehension (i.e., the process of reading, evaluating, and integrating information from more than one source) has shown that effective readers construct detailed understandings of individual texts as well as connections across ideas, sources, and perspectives (Perfetti et al., 1999). Recent frameworks emphasize that readers must evaluate source credibility, reconcile conflicting information, and synthesize evidence into coherent explanations, drawing on cognitive, metacognitive, and writing skills in tandem (List & Alexander, 2019). Students often need support to engage in these complex tasks. Instruction that explicitly scaffolds strategies such as paraphrasing, elaboration, and source evaluation can help students move from surface-level comprehension toward integration and argumentation (Brante et al., 2018; McNamara, 2017; Sonia et al., 2022).

Translating these insights into classroom practice depends on instructional structures that are adaptable, efficient, and easily incorporated into teachers' routines. Frameworks such as Before–During–After (BDA; Lewis & Strong, 2020) offer a simple sequence for guiding students through preparation, active reading, and synthesis. Moreover, lesson planning using a multiple text-set framework can help teachers organize thematically linked readings that build background knowledge and gradually increase in complexity (Lupo et al., 2019). Building on these foundations, the InSPECT framework was designed through a participatory study with high school science teachers (Potter et al., 2025) to help teachers integrate evidence-based practices for reading and writing from multiple texts. InSPECT provides a structured lesson sequence that supports both comprehension within texts and integration across them by embedding strategies (Investigate, Source Evaluate, Paraphrase, Explain, Connect Across Sources, Transform into Writing). Professional learning studies show that when teachers are provided with clear examples, adaptable templates, and time to tailor materials to their content areas, they are more likely to incorporate such approaches into everyday instruction (Goldman et al., 2019; Thomas & Drew, 2021). Nevertheless, sustained implementation of literacy instruction in science classes remains difficult due to limited planning time, assessment demands, and uneven student readiness (Drew et al., 2017).

Design processes that combine empirical evidence with iterative, user-centered development can improve implementation. Co-design and participatory design approaches invite teachers to shape materials and PD to ensure feasibility and contextual relevance (Cumbo & Selwyn, 2022; Roschelle & Penuel, 2006). Learning engineering builds on these approaches by

emphasizing collaborative, data-informed cycles of design, testing, and refinement conducted by multidisciplinary teams. These cycles typically include four interconnected stages: (a) identifying a learning challenge that guides the process, (b) creating a solution through design and instrumentation, (c) implementing the solution in an authentic context, and (d) investigating outcomes through mixed-method data analysis (Goodell, 2023). Learning engineering also emphasizes human-centered methods and tools for aligning designs with stakeholder goals and contextual constraints (Thai et al., 2023). Importantly, classroom pilots like the present study can be framed as nested cycles that generate decision-relevant evidence within a larger learning engineering effort (Craig et al., 2025). The present study represents the implementation and evaluation phase within this learning-engineering process, focused on improving the usability and classroom integration of integrated MD reading and writing instruction in high school science.

Present Study

Building on prior participatory design and pilot work with secondary teachers that informed the lesson sequence and professional development model (Christhilf et al., 2025; Potter et al., 2025), the present study reflects implementation and early investigation phases of the learning-engineering process, with evidence drawn from teacher reflections and student artifacts within a single biology unit on CRISPR gene editing. During this stage, a high school science teacher implemented the InSPECT framework in her classroom. As a learning-engineering pilot, this study focuses on classroom feasibility and use of instructional supports rather than measurement of cognitive processes or learning outcomes. The study was guided by the following research questions (RQs):

RQ1: How can a learning-engineering process be used to refine and implement a lesson sequence that supports reading and writing across multiple sources in science?

RQ2: What design insights from the teacher's reflections and student artifacts inform future iterations of the framework and professional development model?

Methods

Study Context & Implementation

This pilot study represented the implementation and data-instrumentation phase of a larger learning-engineering effort to iteratively design and refine instructional approaches for integrating reading and writing across multiple sources in science classrooms (Potter et al., 2025). The pilot was conducted in spring 2024 with one high school biology teacher, Dana Dana had previously participated in a focus group study, which served as the collaborative design-phase in which researchers and teachers collaborated on the lesson framework and materials. Dana implemented the instructional materials in her tenth-grade biology classes in Ohio.

The pilot consisted of a five-lesson unit on CRISPR gene editing, designed to engage students in reading, evaluating, and synthesizing information from multiple sources using the InSPECT framework. The instructional objective was for students to construct a source-based essay that describes how gene editing can repair faulty genes. To meet this objective, students worked with four sequenced sources that introduced the topic, built background knowledge, and culminated in a more rigorous scientific text. One source was an informational video designed to build background knowledge. For each source, students completed a structured reading guide aligned to InSPECT. Before reading, they evaluated the source for credibility and reflected on their prior knowledge of the topic. During reading, students took notes by paraphrasing key ideas, elaborated by connecting content to their prior knowledge, and wrote bridging statements by making connections to information within and across sources. After reading, students wrote a brief reflection on how each source could be used in their source-based essay. The unit concluded with a source-based essay in which students synthesized information from all four sources to construct their source-based essay.

Data Sources & Analysis

This single-case study, conducted within a broader learning-engineering process, drew primarily on a semi-structured interview with the pilot teacher. The interview captured perceptions of instructional usability, classroom feasibility, and alignment with existing teaching routines. To supplement these perspectives, we also examined student work. Of 51 enrolled students, 23 provided assent to share their work, and 9 provided both assent and parental consent, resulting in nine complete artifact sets for analysis. These students' structured notes and source-based essays were examined to provide contextual evidence of how the lesson sequence functioned in practice. Analyses followed a qualitative descriptive approach (Thomas, 2006), emphasizing design-relevant insights rather than generalizable claims. Two researchers independently reviewed the teacher and student data, coded for indicators of feasibility and application of key framework strategies, and consolidated codes through discussion. To enhance analytic efficiency, a GPT-4o–assisted interface (Potter et al., forthcoming) was used to locate illustrative excerpts, with all coding and interpretation completed by researchers following best practices for human–AI collaboration in qualitative analysis. All procedures were approved by the institutional review board, and informed consent and parental permissions were obtained.

Results

Analysis of the teacher interview and student work samples yielded three design-relevant findings that clarify (a) classroom feasibility, (b) student entry points for engaging with multiple documents, and (c) priorities for instructional refinement.

Theme 1: Structured Design Supports Feasibility

Dana described the sequence as straightforward to use and fitting into her routine. In the interview she said, "I found it really easy. The kids caught on really quick, and how to use it." She also explained the sequence's temporal logic: "it kept them moving in a chronological type, pattern or fashion to get to the end result." Dana noted the lessons started simple and built gradually: "I liked it, because it started with something easy ... and it moved up a little bit and kept adding more and more and more." These comments indicate the structure reduced perceived planning burden and that the materials were usable in regular class periods. The gradual progression from simpler to more demanding activities also appeared to help students build confidence before engaging with a more rigorous text at the end of the sequence, allowing them to practice strategies on accessible materials before applying them to a complex reading and writing task.

Theme 2: Paraphrasing as an Entry Point for Deeper Learning

Dana identified paraphrasing as a clear instructional win. She observed, "I think the paraphrasing helped." She described how paraphrasing supported essay writing: "once they [completed their reading guides] when they went to write their final essay, they were able to look at the paraphrasing from each of the four lessons and pull that together versus having to go back and look at the whole text, or the whole video, or whatever it may have been." By restating ideas in their own words as they read, students created a set of organized notes that they could later consult to assemble and compare information across sources. In other words, completing the reading guides served an aid for students to plan their source-based essay. She also expressed excitement about some student writing outcomes: " Their essays were fantastic." and "some of the kids who I did not expect some really good essays from were amazing." Dana viewed these successes as evidence that paraphrasing gave students an accessible way to engage with multiple documents, lowering the initial barrier to entry and providing a foundation for deeper synthesis in their writing.

Theme 3: Accountability and Source Evaluation as Areas for Refinement

Dana reported student difficulty with source-evaluation tasks and uneven completion of template sections. She said, "some of the students had a hard time verifying the publisher…they would just leave it blank…some of them would just skip that, or they

would say, interesting article or good." She described implementation choices to increase accountability: "The students for each of the 4 lessons. They got 6 points. So it was like a quiz grade...and then the final essay...was worth 26 points." Dana noted that these completion-based scores encouraged students to revisit and finish incomplete work, explaining that most "did go back and correct it." She attributed occasional omissions to limited motivation near the end of the school year but emphasized that accountability structures, such as visible checklists and low-stakes grading, helped sustain engagement. In this case, the structured reading guide served as a visual checklist for students, helping them track progress through each task. The guides were co-developed during the collaborative design phase, while the point-based completion grades were implemented at the teacher's discretion. This flexibility was intentional, as the framework was designed to accommodate teachers' existing assessment systems while maintaining consistent scaffolds for student accountability. These observations point to specific areas for refinement, including clearer modeling of how to verify sources and consistent reinforcement of accountability routines to ensure all students complete each component of the reading guide.

Discussion

The findings from this pilot study highlight how a structured instructional design can support classroom feasibility while lowering barriers to integrated reading and writing instruction in science. Science teachers frequently report limited time, preparation, and confidence to engage students in extended literacy tasks that require reading across sources and synthesizing evidence (Drew & Thomas, 2017). Specifically, the lesson sequence helped make complex literacy practices manageable for implementation, addressing this persistent challenge. Dana's reflections showed how a structured lesson sequence reduced the cognitive and logistical demands of implementation, as she explained that "I found it really easy. The kids caught on really quick, and how to use it," highlighting how a clear, sequenced structure supported both her planning and students' understanding. These observations underscore the importance of explicit organizational scaffolds for teacher usability when integrating reading and writing practices into content-based courses.

Dana also noted that paraphrasing was an especially effective entry point for students' engagement with multiple texts. She observed that "once they had all that paraphrasing…they were able to pull that together and actually write better essays," suggesting that structured note-taking served as a bridge to synthesis and writing. At the same time, she reported that students struggled with evaluating the credibility of sources and needed additional modeling to elaborate on and connect ideas across readings. These findings highlight the value of grounding complex literacy instruction in strategies that are familiar and accessible while gradually layering supports for higher-order reasoning.

From a learning-engineering perspective, this pilot represents a nested iteration within a broader design process rather than a single-pass evaluation. Earlier phases of the project focused on participatory design and professional development to establish the instructional framework and implementation supports (Christhilf et al., 2025; Potter et al., 2025). The present study extends that work by examining how those design decisions functioned in an authentic classroom context, using teacher reflections and student artifacts to surface feasibility constraints and strategy-specific needs. Evidence from this pilot directly informed targeted refinements, including clearer modeling of source evaluation, the use of reading guides as visible checklists, and low-stakes accountability structures to support sustained engagement. In this way, learner and teacher data functioned as design-relevant inputs that guided subsequent instructional refinement, consistent with an iterative learning-engineering process.

Although promising, this case represents a single classroom implementation with a small number of student artifacts, limiting the generalizability of the findings. The pilot occurred near the end of the school year, when motivation and attendance varied, potentially influencing student engagement and outcomes. Nevertheless, these contextual factors offer valuable insight into how instructional materials perform under realistic classroom conditions.

Future iterations will extend this work across multiple classrooms to evaluate the refined framework and professional learning model developed through this cycle. Subsequent learning-engineering cycles will continue to examine how instructional

supports are adapted, implemented, and refined over time, including how teachers support elaboration and source-evaluation strategies and how professional development structures can sustain use. More broadly, this work demonstrates the value of learning engineering as a bridge between research and practice, enabling iterative, data-driven improvement that aligns instructional design with the authentic contexts of teaching and learning.

Acknowledgments

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant [R305A180144] to Arizona State University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

References

Brante, E. W., & Strømsø, H. I. (2018). Sourcing in text comprehension: A review of interventions targeting sourcing skills. Educational Psychology Review, 30(3), 773-799.

Christhilf, K., Potter, A., Magliano, J. P., McCarthy, K. S., Allen, L. K., & McNamara, D. S. (2025). Constructed responses as a window into strategic processing: the role of prompts in multiple-document reading. Discourse Processes, 1-25. https://doi.org/10.1080/0163853X.2025.2578594

Craig, S. D., Avancha, K., Malhotra, P., C., J., Verma, V., Likamwa, R., Gary, K., Spain, R., & Goldberg, B. (2025). Using a Nested Learning Engineering Methodology to Develop a Team Dynamic Measurement Framework for a Virtual Training Environment. In International Consortium for Innovation and Collaboration in Learning Engineering (ICICLE) 2024 Conference Proceedings: Solving for Complexity at Scale (pp. 115-132). https://doi.org/10.59668/2109.21735

Cumbo, B., & Selwyn, N. (2022). Using participatory design approaches in educational research. International Journal of Research & Method in Education, 45, 60-72. https://doi.org/10.1080/1743727X.2021.1902981

Drew, S. V., Olinghouse, N. G., Faggella-Luby, M., & Welsh, M. E. (2017). Framework for disciplinary writing in science grades 6–12: A national survey. Journal of Educational Psychology, 109(7), 935–955. https://doi.org/10.1037/edu0000186

Drew, S. V., & Thomas, J. (2017). Secondary science teachers' implementation of CCSS and NGSS literacy practices: A survey study. Reading & Writing, 31(2), 267–291. https://doi.org/10.1007/s11145-017-9784-7

Goldman, S. R., Greenleaf, C., Yukhymenko-Lescroart, M., Brown, W., Ko, M.-L. M., Emig, J. M., George, M., Wallace, P., Blaum, D., & Britt, M. A. (2019). Explanatory modeling in science through text-based investigation: Testing the efficacy of the Project READI intervention approach. American Educational Research Journal, 56(4), 1148–1216. https://doi.org/10.3102/0002831219831041

Goodell, J. (2023). What is learning engineering? In J. Goodell & J. Kolodner (Eds.), Learning engineering toolkit (pp. 5–25). Routledge. https://doi.org/10.4324/9781003276579-3

Lewis, W. E., & Strong, J. Z. (2020). Literacy instruction with disciplinary texts. Guilford Publications.

List, A., & Alexander, P. A. (2019). Toward an integrated framework of multiple text use. Educational Psychologist, 54, 20-39. https://doi.org/10.1080/00461520.2018.1505514

Lupo, S. M., Tortorelli, L., Invernizzi, M., Ryoo, J. H., & Strong, J. Z. (2019). An exploration of text difficulty and knowledge support on adolescents' comprehension. Reading Research Quarterly, 54(4), 457-479. https://doi.org/10.1002/rrq.247

McNamara, D. S. (2017). Self-explanation and reading strategy training (SERT) improves low-knowledge students' science course performance. Discourse Processes, 54(7), 479-492. https://doi.org/10.1080/0163853X.2015.1101328

Organisation for Economic Co-operation and Development. (2019). PISA 2018 assessment and analytical framework. OECD Publishing. https://doi.org/10.1787/b25efab8-en

Perfetti, C. A., Rouet, J.-F., & Britt, M. A. (1999). Toward a theory of documents representation. In H. van Oostendorp & S. R. Goldman (Eds.), The construction of mental representations during reading (pp. 99–122). Lawrence Erlbaum Associates Publishers.

Potter, A. H., Arner, T., McCarthy, K. S., & McNamara, D. S. (2025). Integrating reading, writing, and digital tools in science: A participatory-design study of the InSPECT framework. Education Sciences, 16, Article 6. https://doi.org/10.3390/educsci16010006

Potter, A. H., Serhan, Z., Patne, N. A., Öncel, P., Ahmed, I., Arner, T., Islam, R., Allen, L. A., Roscoe, R. D., Crossley, S. A., & McNamara, D. S. (Forthcoming - b). Human-AI collaboration in participatory design: Refining the Writing Analytics Tool.

Roschelle, J., & Penuel, W. R. (2006, June). Co-design of innovations with teachers: definition and dynamics. In Proceedings of the 7th international conference on Learning Sciences (pp. 606-612). https://repository.isls.org/bitstream/1/3563/1/606-612.pdf

Sonia, A. N., Magliano, J. P., McCarthy, K. S., Creer, S. D., McNamara, D. S., & Allen, L. K. (2022). Integration in multiple-document comprehension: A natural language processing approach. Discourse Processes, 59(5-6), 417-438. https://doi.org/10.1080/0163853X.2022.2079320

Thai, K. P., Craig, S. D., Goodell, J., Lis., J., Schoenherr, J. R., & Kolodner J. (2023). Learning Engineering is Human-Centered. In J Goodell (Ed.), The Learning Engineering Toolkit (pp. 83-124). Routledge.

Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. American Journal of Evaluation, 27(2), 237-246.

Thomas, J. D., & Drew, S. V. (2021). Impact of A Practice-Based Professional Development on Secondary Science Teachers' Use of Disciplinary Literacy Practices: A Design Research Project. Journal of Science Teacher Education, 33, 1–31. https://doi.org/10.1080/1046560X.2021.1898763