



Impact of Modes of Instruction and Classroom Activities on the 2019 NAEP Eighth-Grade Science Assessment

Susan Sobehrad¹, Mingyuan Zhang^{2*}

Central Michigan University, USA.

*Corresponding Author
Mingyuan Zhang

Central Michigan
University, USA.

Article History

Received: 14.08.2024

Accepted: 06.09.2024

Published: 17.09.2024

Abstract:

Objective: This investigation focuses on the impact of classroom practices that may influence a student's ability to conduct research and thus affect their ability to demonstrate increases in content knowledge through inquiry.

Methods: Data for this study was selected from The National Assessment of Education Progress (NAEP). The NAEP Data Explorer was used for data analysis. Overall science scale scores from the 2019 eighth-grade science assessment for specific questions related to instructional methods for scientific research are examined. Because more than one variable is examined, the NDE is used to create descriptive cross-tabulated tables, and t-tests are run to determine significance.

Results: The findings include: (1) The frequency with which students used technology in the classroom has no effect on science achievement scores. (2) Intermittent use of scientific inquiry shows bigger gains in science achievement than consistent use. (3) A high frequency of opportunities to combine information from multiple sources has no significant impact on science achievement.

Conclusion: High frequency use of technology, inquiry methods, or online searches does not correlate with high gains in science achievement. Adequate prior knowledge, instruction in query development, and appropriate scaffolding are necessary to ensure that students can construct meaning and build knowledge while using technology for online tasks.

Keywords: science inquiry, knowledge synthesis, NAEP, science achievement.

Cite this article:

Sobehrad, S, Zhang, M., (2024). Impact of Modes of Instruction and Classroom Activities on the 2019 NAEP Eighth-Grade Science Assessment. *ISAR Journal of Arts, Humanities and Social Sciences*, 2(9), 64-76.

Introduction

Background

Inquiry learning is a student-centered approach to knowledge-building where students construct their own knowledge by exploring open-ended tasks. Increasingly, inquiry learning is being incorporated into STEM (science, technology, engineering, and math) practices to actively engage students in critical thinking, including answering scientific questions, investigating natural phenomena, and creating solutions to real-world problems (Duncan et al., 2021; Chen & Chen, 2021). For students to conduct such explorations, it is important for them to be able to define the nature of the task or problem they have been given, which means that they need to be able to think critically, ask relevant questions, and discern key attributes related to the task as well as to potential answers and solutions. This requires guidance by an expert—and many times that expert guidance is handed off to online search engines.

The Research Problem

Students have a difficult time finding good answers to ill-structured questions using online resources (Sobehrad, 2023). The ability to think critically and ask relevant scientific questions are not innate skills; students need opportunities in the classroom to learn and practice these skills as they are learning new content. The purpose of this study is to examine both teacher and student views regarding the emphasis placed on the development of online research skills in the eighth-grade science classroom.

Justifying the Research

Online research involves using the internet and digital resources to gather information and is often used as a tool within various educational approaches, including inquiry learning, to find and evaluate information. However, the results of online searches are only as good as the queries generated by the researcher. If a student doesn't have adequate prior knowledge of the associated content, their ability to successfully navigate the internet in order to find the information they need to complete the assigned inquiry tasks is diminished. Less successful inquiries have the potential to

generate or promote misconceptions related to science content. At the same time, algorithmic influences may result in a narrowing of results toward wrong information as students select and click on sites that they may choose erroneously due to their lack of domain knowledge (Gruber & Hargettai, 2023). Furthermore, much of the information provided via common search engines such as Google is at a level above a middle school student's ability to understand (Sobehrad, 2023; Anuyah et al., 2020). Because of the vast amount of information available on the internet, searches have the potential to generate so many options that students either suffer from information overload or lack the experience to separate relevant and irrelevant information.

If by chance students can find pertinent information at an appropriate level of cognition, they may or may not know how to combine information from a variety of sources into a coherent picture of the explanatory science that will help them move beyond comprehension into higher levels of functioning such as evaluation or creation. Again, the ability to synthesize knowledge from multiple sources requires intervention by an expert—hopefully the teacher—although it seems likely that this role will eventually be handed off to a digital partner in the form of AI.

Deficiencies in Evidence

Searching for information online is a crucial skill for students to master in today's classrooms. Previous studies have indicated that students across various age groups often struggle to find and efficiently gather information online (Zhou & Lam, 2019). The fact that search engines fulfill a role as surrogate expert means that students must be able to discern relevant search results, evaluate the validity and reliability of those results, and be aware of potential biases that arise from the filter bubble (Simpson, 2013; Martin et al., 2022). Yet, there is a lack of information specific to how much time is spent in the classroom on activities designed to help students navigate these challenges to acquire new knowledge.

The Audience

Understanding student and teacher perceptions about how much time is spent in the classroom to improve students' ability to solve problems using online resources is an important next step. Before curriculum experts and instructional designers can design methods for use in the classroom and provide relevant and timely professional development offerings to teachers, they must first know how much time is spent already doing so. If change is required, the information would be helpful to administrators who may need to adjust policy to accommodate a major pedagogical shift in the way research is taught in the classroom.

Research Questions

Inquiry learning combined with STEM practices and online research complement one another, with each contributing uniquely to the development of informed, critical, and engaged learners who are able to synthesize web-mediated content to devise solutions to complex problems. This investigation focuses on the impact of classroom practices that may influence a student's ability to conduct research and thus affect their ability to demonstrate increases in content knowledge through inquiry. Specifically, 1) How often do teachers in an eighth-grade science class ask students to use a computer or other technological resources to conduct a search for science information? 2) How often do students participate in inquiry activities in an eighth-grade science class? and 3) How often are eighth grade science students asked to

combine information about science from multiple sources for an assignment?

Literature Review

Introduction

Inquiry learning is a student-centered approach to knowledge-building where students construct their own knowledge by answering scientific questions, investigating natural phenomena, and creating solutions to authentic issues (Duncan et al., 2021; Chen & Chen, 2021). For students to successfully navigate these processes, they must obtain enough relevant domain knowledge to enable them to define the task or problem they have been given and to evaluate potential solutions for their utility. This requires guidance by an expert. That expert may be the teacher, who could potentially share the responsibility with online search engines or AI tools.

Online research involves using digital resources to gather information. However, online searches and AI queries will generate just what the researcher asks for, and sometimes unintended results will be included. Students without adequate prior knowledge of required content will less successfully find the information they need. Additionally, unintended or unasked for results have the potential to promote science misconceptions. For example, misinformation spread about the erroneous dangers of vaccination has resulted in the re-emergence of preventable diseases. At the same time, algorithmic influences may result in a narrowing of results toward wrong information as search results are skewed toward what students most often search for rather than toward specific answers to the queries; important information may be left out, its omission undetected by students who may have limited domain knowledge (Gruber & Hargettai, 2023). Furthermore, much of the information provided by common search engines such as Google is at a much higher level of comprehension than a middle school student's ability to understand (Sobehrad, 2023; Anuyah et al., 2020). Finally, online searches have the potential to generate so many options from so many perspectives that students will experience cognitive overload which can impact memory and the ability to make content-related decisions (Rutkowski & Saunders, 2018).

Even finding relevant, digestible data at an appropriate level of cognition is no guarantee that the student will then be able to build new knowledge. The ability to combine information from multiple, multimodal sources is not an innate talent; synthesizing knowledge from multiple sources requires intervention by an expert—hopefully the teacher, perhaps in the form of a digital partner powered by AI.

Inquiry learning combined with STEM practices and online research complement one another and make it possible for students to synthesize web-mediated content and devise solutions authentic problems. This investigation focuses on the impact of classroom practices that may influence a student's ability to conduct research and thus affect their ability to demonstrate increases in content knowledge. Specifically, 1) How often do teachers in an eighth-grade science class ask students to use a computer or other technological resources to conduct a search for science information? 2) How often did students do inquiry activities in science this year? and 3) How often are eighth grade science students asked to combine information about science from multiple sources for an assignment?

Using Technology for Science Research in the Middle Grades

The ubiquity of 1:1 technology initiative and the continued emphasis on integrating technology in the classroom has transformed searching the internet for information from an occasional, intriguing classroom event into a less intriguing routine task, the complexity of which is hidden by sophisticated algorithmic frameworks which presume to know the intent of the researcher. The internet has evolved into a dynamic space where educators and learners can gain new knowledge on a daily basis (Heersmink, 2016; Marsh & Rajaram, 2019). The internet is an integral component of a classroom's cognitive ecology; it emphasizes an exchange of information between search engine and student during internet searches, providing students with access to a wealth of new knowledge enriched by a variety of perspectives, experiences, and insights (Martin et al., 2022). However, less experienced researchers, such as middle grade students, have great potential to be seduced by the ease with which their technology can answer their questions. Giving students the opportunity to conduct online research in no way ensures that those same students will achieve the specific learning outcomes intended by the educator. Implementing targeted instruction on subject-specific content before permitting students to engage in web-based research could significantly enhance their ability to locate relevant information effectively (Kirschner, Sweller, & Clark, 2006). Teaching students effective ways to structure online searches has also been found to increase achievement of learning outcomes, underscoring the importance of structured guidance in the learning process, particularly in the context of utilizing the internet as an educational tool.

Some studies have shown that teaching practices requiring students' frequent active involvement in technology use have a negative association with learning outcomes (OECD, 2011; Biagi and Loi, 2013). Cairns & Areepattamannil (2019, p. 16) showed similar findings, stating that their study "found a significant negative relationship between inquiry-based science teaching and science achievement" when there was a focus on use of technology in the classroom. Cairns & Areepattamannil (2019) noted that the efficacy of both teacher and student in the inquiry process might have contributed to this negative effect on achievement.

Scaffolds and Inquiry-Based Learning

In several studies (Zhang, 2013; Teig et al., 2018; Jerrim et al., 2019; Cheng et al., 2017) inquiry-based learning led to greater achievement in science scores only when teachers provided a significant level of guidance during the inquiry process. Teig et al. (2018, p. 22) reported that "empirical research that concentrated on the instructional guidance given to students found that [inquiry-based] activities involving teacher-guided instruction were more effective than independent or open-inquiry based approaches." Jerrim et al. (2019, p. 42) found that "inquiry-based teaching has a very weak relationship with attainment in science—and that any positive effects are confined to moderate levels of inquiry combined with high levels of guidance." Cheng et al. (2017) found that when a structured approach was used to teach problem-solving during inquiry learning, student outcomes were higher than students who were not provided with problem-solving methods. Zhou & Lam (2019, p. 1359) explored the use of metacognitive prompts to help learners identify what they needed to learn during the online search process; metacognitive prompts "contain

questioning words that offer learners explanations, reminders, or hints of what they are doing and how and why they are doing it." They found multiple studies that showed this method of guided research increased student achievement, and cited another source that suggested technology-enhanced prompts (on-screen prompts) were most beneficial in increasing students' metacognitive awareness during online searches when compared to teacher-enhanced prompts (Raes et al., 2012). Koyunlu Unlu & Dokme (2022) found that the implementation of technology-supported, inquiry-based learning improved students' academic achievement and positively affected the students' science processing skills. They found that student skills related to observation, comparison, inference, hypothesizing, experimentation, and data analysis improved due in part to the scientific research undertaken during the inquiry task (Koyunlu Unlu & Dokme, 2022, p. 129).

Colwell et al. (2013) suggests scaffolding online student research using internet reciprocal teaching (IRT) has the potential to improve development of student digital literacy. Because IRT involves teacher and student modeling of online research and comprehension strategies such as questioning, locating, critically evaluating, synthesizing, and communicating information gained online, the potential for positive impact on student outcomes and increased digital literacy is high. However, Colwell et al. (2013) discovered that while students were able to use the strategy shortly after the scaffolds were made, long term behavioral changes did not result. "For the students in this study. . .previous strategies, often developed through activities on the Internet outside of school, appeared to be firmly entrenched and not easily altered beyond a few tasks for a relatively brief time (Colwell et al., 2013, p. 314)."

Creative Knowledge Synthesis and Inquiry Learning

Educators and students now have access to a vast array of resources to support new learning created by combining disparate information from a variety of online sources into new understanding (Lierat, & Glăveanu, 2018). This knowledge-building process exemplifies creative synthesis, which has been defined by the American Psychological Association as "the combination of several ideas, images, or associations into a new whole that differs fundamentally from any of its components." Also involved is synthesis of meaning that results when information from multiple, diverse sources is integrated to develop a coherent interpretation content with few changes to the information (DeSchryver, 2014). Research has shown that knowledge building is an important requirement for comprehension (Elleman & Compton, 2017; Catts & Kahmi, 2017). However, the influx of technology and the exponential growth of information now available on the internet can be overwhelming for students.

Cognitive load theory suggests that working memory can process only a limited amount of information at any one time (MCW, 2022). Thus, there is a danger in utilizing the internet to build knowledge in that students' working memory can easily be overloaded due to the large amount of information available, especially if a student does not have enough domain knowledge to be able to either classify information into smaller chunks, or to eliminate unnecessary information. Some critics of inquiry-based learning also use working-memory overload to argue against the practice (Kirschner et al., 2006; Rosenshine, 2012; Zhang, 2013). Zhang (2013) proposes that using specific strategies to guide student inquiry using internet resources can help students focus on

purposes for reading scientific information online and avoid a fragmented and disjointed collection of knowledge. Providing guidance can potentially reduce cognitive overload.

Methods

Data for this study was selected from The National Assessment of Education Progress (NAEP), a congressionally mandated program organized by the National Center for Education Statistics (NCES). As part of the Nation’s Report Card, NAEP has collected information about academic achievement and learning for students in grades four, eight, and twelve in a variety of subjects since 1969 (National Center for Education Statistics, 2024).

The NAEP science assessments for grades four, eight, and twelve transitioned from paper-based to digitally based formats in 2019. The process was multi-step, which entailed offering the assessments in both formats (NCES, 2024). In this study, data from the 2019 NAEP results was used to investigate the impact of classroom practices that may influence a student’s ability to conduct scientific research and thus affect their ability to demonstrate increases in content knowledge through inquiry.

Participants and Sampling

Data is collected from public schools across the nation, including public charter schools. Schools are selected by probability sampling proportional to size. According to NCES (2024) this method ensures that “demographic and achievement characteristics of the sample are consistent and representative of the nation (n.p).” All states have participated in NAEP assessments since 2002.

Prior to 2022, the NAEP science assessment was given every other year. Currently it is given in even numbered years to a sample of students in grades four and eight, and to twelfth grade students every four years. For this study, results from the NAEP science assessment for eighth grade from the year 2019 were investigated.

Public School Selection in State Assessment Years

Schools who will potentially be selected to participate in the NAEP are first identified using the U.S. Department of Education's public school system database. Then, a representative sampling of public schools from across the nation are classified into groups called strata first by location (e.g., urban, suburban, rural) and then by racial/ethnic composition. Subsequently, schools are sorted using student achievement to ensure that the NAEP represents

schools from all levels of performance ability. Next, a sample of schools from each stratum, with probability proportional to school size, is chosen. Small schools, high minority schools, and private schools are included in the sample to ensure adequate representation in each category (*Focus on NAEP*, 2024). Finally, the completed list is sent to each state’s department of education to ensure that the school is eligible to participate in the NAEP (*Focus on NAEP*, 2024).

NAEP Sampling and Data Collection

Students in the sample schools are randomly chosen to answer questions. Generally, 30 students per grade per subject from each school are selected. Only schools reporting 85 percent student participation are included in the final NAEP statistics (*Focus on NAEP*, 2024). Each student only completes a portion of the NAEP items. Balanced Incomplete Block (BIB) spiraling systematically groups items within the test booklets to ensure that all content domain questions are asked, and that all items are completed by a representative sample of students.

Data Analysis

NAEP Data Explorer

The NAEP Data Explorer (NDE) was used for data analysis. NDE is a web-based application that enables users to create customizable tables and graphs to display the results of statistical information gleaned from the NAEP assessment data. Data can be filtered by testing year, grade level, and subject area. Additional filters include demographic information about students, schools, and the community, instructional methods, and factors from outside of school that may impact assessment results.

For this study, overall science scale scores from the 2019 eighth-grade science assessment for specific questions related to instructional methods for scientific research are examined. (See Table 1 for the specific coded questions examined in this study.) Because more than one variable is examined, the NDE is used to create descriptive cross-tabulated tables, and t-tests are run to determine significance (*NDE Core Web*, 2024). Cohen’s *d* effect size was calculated through the University of Colorado’s *Effect Size Calculator* (Becker, 2000). Effect sizes describe the practical significance of a statistical test result, independent of the sample size and the measurement scale (Gliner et al., 2001). General rules for describing the magnitude of effect sizes can be attributed to Jacob Cohen (1969). Cohen classified effect sizes as *small* ($d = 0.2$), *medium* ($d = 0.5$), and *large* ($d \geq 0.8$).

Table 1. Coded Questions from the 2019 NAEP Eighth-Grade Science Assessment Used for Analysis

In your eighth-grade class, how often do your students use a computer or other technological resources to conduct a search for science information? (teacher-reported) ID: T113301	Never or hardly ever Once or twice a month Once or twice a week Every day or almost every day
How often science students did inquiry activities (average response) ID: TSINQR	Never Sometimes Often to Always
In this school year, how often have you done the following activities in your science class? Combined information about science from multiple sources (for example, books, websites, or articles) for an assignment (student-reported) ID: K824208	Never or hardly ever Once in a while Sometimes Often Always/almost always

Results

Information from the NAEP 2019 science assessment was examined in this study. Statistical information was calculated on three variables in order to determine whether activities related to solving ill-defined problems affected eighth grade students’ science assessment scores. This section discusses the statistical analysis of scale scores and standard deviations for all three variables

- 1. How often students use computers in science
- 2. How often students participate in scientific inquiry
- 3. How often students’ information from multiple sources

The results of independent t-tests with an alpha level of .05 are also reported. Cohen’s d effect size was calculated to examine significance. Cohen’s d effect size was calculated through the University of Colorado’s *Effect Size Calculator* (Becker, 2000). Effect sizes describe the practical significance of a statistical test result, independent of the sample size and the measurement scale (Gliner et al, 2001).

Data Table 1 presents the average overall scale score for the 2019 science assessment for all students in national public schools. The total number of students is not reported by the NAEP Data Explorer.

Data Table 1: Average scale scores for grade 8 science, by all students [TOTAL], national public, 2019			
Year	Jurisdiction	All students	Average scale score
2019	National public	All students	153

NOTE: Some apparent differences between estimates may not be statistically significant.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2019 Science Assessment.

Research Question #1: In your eighth-grade class, how often do your students use a computer or other technological resources to conduct a search for science information?

The 2019 NAEP eighth grade overall science scores and standard deviations were analyzed based on the variable of how often teachers asked students to use the computer to search for science information (Data Table 2).

Data Table 2: Average scale scores and standard deviations for grade 8 science, by Students use computer to search for science information [T113301], national public 2019

Year	Jurisdiction	Students use computer to search for science information	Average scale score	Standard deviation
2019	National public	Never or hardly ever	151	34
		Once or twice a month	154	35
		Once or twice a week	154	35
		Every day or almost every day	154	35

NOTE: Some apparent differences between estimates may not be statistically significant.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2019 Science Assessment.

Students who used the computer to search for science information “never or hardly ever” had an average scale score of 151 (SD=34). Students who used the computer to search for science information “once or twice a month” had an average scale score of 154 (SD=35), those who used the computer to search for science information “once or twice a week” had an average scale score of 154 (SD=35), and those who used the computer to search for science information “everyday or almost every day” had an average scale score of 154 (SD=35).

An independent t-test was used to determine the significance of average scale score differences between scores in classes where teachers used the computer hardly ever or almost every day. The t-test results determined there is no significant difference (Data Table 3).

Data Table 3: Science Grade 8, Difference in average scale scores between variables, for students use computer to search for science information [T113301], national public, 2019

	Never or hardly ever	Once or twice a month	Once or twice a week	Every day or almost every day
Never or hardly ever				
Once or twice a month	x Diff = 3 P-value = 0.0965 Family size = 6			
Once or twice a week	x Diff = 3 P-value = 0.1572 Family size = 6	x Diff = 0 P-value = 0.7884 Family size = 6		
Every day or almost every day	x Diff = 2 P-value = 0.2112 Family size = 6	x Diff = -1 P-value = 0.7057 Family size = 6	x Diff = 0 P-value = 0.8998 Family size = 6	
LEGEND:				
< Significantly lower.				
> Significantly higher.				
x No significant difference.				

NOTE: Within jurisdiction comparisons on any given year are dependent with an alpha level of 0.05.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2019 Science Assessment

Research Question #2: How often did students do inquiry activities in science this year?

The 2019 NAEP eighth grade average scale scores and standard deviations were analyzed based on the variable of how often teachers provided students with opportunities to participate in science inquiry activities (Data Table 4).

Data Table 4: Average scale scores and standard deviations for grade 8 science, by How often student did inquiry activities in science class this year (average response) [SINQ08], national public 2019

Year	Jurisdiction	How often student did inquiry activities in science class this year (average response)	Average scale score	Standard deviation
2019	National public	Never	150	34
2019	National public	Sometimes	157	34
2019	National public	Often to always	156	35

NOTE: Some apparent differences between estimates may not be statistically significant.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2019 Science Assessment.

An independent t-test was used to determine the significance of average scale score differences between scores in classes where teachers provided differing amounts of opportunities for students to participate in inquiry activities, from never to always. The t-test results are shown in Data Table 5.

Data Table 5: Science, grade 8, Difference in average scale scores between variables, for How often student did inquiry activities in science class this year (average response) [SINQ08], national public, 2019			
	Never	Sometimes	Often to always
Never			
Sometimes	> Diff = 7 P-value = 0.0000 Family size = 3		
Often to always	> Diff = 6 P-value = 0.0000 Family size = 3	x Diff = -1 P-value = 0.6192 Family size = 3	
LEGEND:			
<	Significantly lower.		
>	Significantly higher.		
x	No significant difference.		

NOTE: Within jurisdiction comparisons on any given year are dependent with an alpha level of 0.05.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2019 Science Assessment

Data Table 5 presents differences in means and independent t-test results. Alpha was set at 0.05 rather than 0.001 as set a priori by the researcher. The average scale score (M=157, SD=34) of students who did inquiry activities sometimes in science class was significantly higher than students who never did science inquiry activities. The average scale score (M=156, SD=35) of students who did inquiry activities from often to always was significantly higher than students who never did science inquiry activities.

Cohen’s *d* effect size was calculated on the independent t-tests that indicated significance in order to determine the magnitude of the effect. Cohen (1969) classified effect sizes as small (*d* = 0.2), medium (*d* = 0.5), and large (*d* ≥ 0.8).

Data Table 6: Effect sizes of significant mean score differences when asking how often student did inquiry activities in science class this year (average response) [SINQ08]		
		Cohen’s <i>d</i>
Sometimes	Never	0.21
Often to always	Never	0.17

The effect size between teachers reporting that they provided opportunities for students to participate in inquiry activities in science sometimes and those who reported never providing opportunities for students to participate in inquiry activities in science was 0.21. The effect size between teachers reporting that they provided opportunities for students to participate in inquiry activities in science from often to always and those who reported never providing opportunities for students to participate in inquiry activities in science was 0.17.

Research Question 3: In this school year, how often have you combined information about science from multiple sources for an assignment?

The 2019 NAEP eighth grade average scale scores and standard deviations were analyzed based on the variable of how often students reported combining information about science from multiple sources for an assignment (Data Table 7).

Data Table 7: Average scale scores and standard deviations for grade 8 science, by Emphasis on teaching students how to develop good research questions [T140701], national public 2019

Year	Jurisdiction	How often combined information about science from multiple sources	Average scale score	Standard deviation
2019	National public	Never or hardly ever	150	35
2019	National public	Once in a while	153	35
2019	National public	Sometimes	152	35
2019	National public	Often	158	34
2019	National public	Always/almost always	155	35

NOTE: Some apparent differences between estimates may not be statistically significant.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2019 Science Assessment.

An independent t-test was used to determine the significance of average scale score differences between scores in classes where students reported having different amounts of opportunity to combine information about science from multiple sources, from “never or hardly ever” to “always/almost always.” The t-test results are shown in Data Table 8.

Data Table 8: Science Grade 8, differences in means and independent t-test results for how often teachers asked student to combine information about science from multiple sources [K824208], national public 2019

	Never or hardly ever	Once in a while	Sometimes	Often	Always/almost always
Never or hardly ever					
Once in a while	> Diff = 3 P-value = 0.0112 Family size = 10				
Sometimes	x Diff = 2 P-value = 0.1512 Family size = 10	x Diff = -1 P-value = 0.1434 Family size = 10			
Often	> Diff = 8 P-value = 0.0000 Family size = 10	> Diff = 5 P-value = 0.0000 Family size = 10	> Diff = 6 P-value = 0.0000 Family size = 10		
Always/ almost always	> Diff = 5 P-value = 0.0006 Family size = 10	x Diff = 2 P-value = 0.1557 Family size = 10	> Diff = 3 P-value = 0.0077 Family size = 10	< Diff = -3 P-value = 0.0056 Family size = 10	
LEGEND:					
<	Significantly lower.				
>	Significantly higher.				
x	No significant difference.				

NOTE: Within jurisdiction comparisons on any given year are dependent with an alpha level of 0.05.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2019 Science Assessment

Data Table 8 presents differences in means and independent t-test results. Alpha was set at 0.05 rather than 0.001 as set a priori by the researcher. The average scale score (M=153, SD=35) of students who combined information from multiple sources once in a while in science class was significantly higher than students who never combined information from multiple sources. The average scale score (M=158, SD=34) of students who combined information from multiple sources often while in science class was significantly higher than students who never, once in a while, or sometimes combined information from multiple sources while in science class. The average scale score (M=155, SD=35) of students who combined information from multiple sources always to almost always while in science class was significantly higher than students who never or sometimes combined information from multiple sources while in science class. The average scale score (M=155, SD=35) of students who combined information from multiple sources always to almost always while in science class was significantly lower than students who often combined information from multiple sources while in science class.

Cohen’s *d* effect size was calculated on the independent t-tests that indicated significance in order to determine the magnitude of the effect. Cohen (1969) classified effect sizes as small ($d = 0.2$), medium ($d = 0.5$), and large ($d \geq 0.8$).

Data Table 9: Effect sizes of significant mean score differences when students reported how often teachers asked them to combine information about science from multiple sources [K824208], national public 2019

		Cohen’s <i>d</i>
Once in a while	Never or hardly ever	0.09
Often	Never or hardly ever	0.23
Often	Once in a while	0.14
Often	Sometimes	0.17
Always/almost always	Never or hardly ever	0.14
Always/almost always	Sometimes	0.09
Always/almost always	Often	0.09

The effect size between students reporting that they had opportunities to combine information from multiple sources while in science class once in a while and those who reported never or hardly ever having opportunities to combine information from multiple sources while in science class was 0.09. The effect size between students reporting that they often had opportunities to combine information from multiple sources while in science class and those who reported never or hardly ever having opportunities to combine information from multiple sources while in science class was 0.23. The effect size between students reporting that they often had opportunities to combine information from multiple sources while in science class and those who reported having opportunities to combine information from multiple sources once in a while in science class was 0.14. The effect size between students reporting that they often had opportunities to combine information from multiple sources while in science class and those who reported sometimes having opportunities to combine information from multiple sources while in science class was 0.17. The effect size between students reporting that they always or almost always had opportunities to combine information from multiple sources while in science class and those who reported never or hardly ever having opportunities to combine information from multiple sources while in science class was 0.14. The effect size between students reporting that they always or almost always had opportunities to combine information from multiple sources while in science class and those who reported sometimes having

opportunities to combine information from multiple sources while in science class was 0.09. The effect size between students reporting that they always or almost always had opportunities to combine information from multiple sources while in science class and those who reported often having opportunities to combine information from multiple sources while in science class was 0.09.

Discussion

In this study, data was derived from the average scale scores of the 2019 NAEP science assessment administered to eighth grade students enrolled in public schools in the United States. The data is used to investigate the impact of classroom practices that may influence a student’s ability to conduct scientific research and thus affect their ability to demonstrate increases in content knowledge through inquiry. Because students have a difficult time finding good answers to ill-structured questions using online resources, this study examined both teacher and student views regarding the emphasis placed on the development of research skills in the eighth-grade science classroom.

Research Question #1: Frequency of Using Technology to Search for Science Information

The lack of significance of computer time to higher achievement opposes findings by Martin et al. (2022) who suggest that liberal access to technology creates opportunities for an exchange of information between search engines and students

during internet searches that give students access to a wealth of new knowledge. The non-significance of results supports other findings that suggest access to information doesn't ensure that new learning occurs (Fisher et al., 2022; Willoughby et al., 2009; Wood et al., 2016). Evidence indicates that the vast scope of information available on the internet makes it difficult for students to locate relevant new information and apply it to learning goals, as does the non-traditional composition of online information. Online information is not linear nor hierarchical and may often lack the headings and subheadings that students use to find relevant information quickly (Wood et al., 2016).

Other research indicates that successful searches on the internet can erroneously result in the belief that because the information has been located, it has been learned; this belief discourages storing new information in internal memory in a way that it can be retrieved and applied at a later time (Fisher et al., 2022). Fisher et al. (2022) further explain that this occurs because "people do not accurately distinguish accessing information through Internet search from accessing information from internal memory, they think they mastered new knowledge, when in fact, they have not (p. 1)."

Lack of domain knowledge may also impact learning achieved through use of online sources. Willoughby et al. (2009) suggest that higher content knowledge can lead to more efficient and effective searches. Learners with less domain knowledge are not as capable of identifying and organizing relevant information found online (Shapiro, 2004). At the same time, ineffective search strategies can also impact the success of online searching and the learning that results. Learners with greater efficacy at search strategies were able to find credible sites; however, they were less able to discern those sites that could provide detailed, complete information about a specific topic (Wood et al., 2016).

In short, studies have shown that students' frequent active involvement in technology may actually have a negative association with learning outcomes (OECD, 2011; Biagi and Loi, 2013). Researchers have corroborated these findings and noted that the skill levels of both teacher and student in the information search and inquiry process might have contributed to this negative effect on achievement (Wood et al., 2016; Fisher et al., 2022; Cairns & Areepattamannil, 2019). Thus, it is not a surprise that simply giving students more computer time does not equate with more learning.

Research Question #2: Frequency of Science Inquiry Activities

Scale scores of students who sometimes or often participated in science inquiry activities were significantly higher than those of students who never participated in science inquiry (Data Table 4). Interestingly, scale scores of students who only participated in science inquiry activities sometimes were higher than those of students who participated often or always. Independent t-test scores validated these scores, showing Cohen's-*d* effect sizes of 0.21 and 0.17, respectively (Data Table 6).

This information confirms the research of Jerrim et al. (2019) who found that just providing inquiry-based teaching was not enough to ensure student achievement gains without high levels of guidance. This claim is supported by the higher effect size of the findings that students had higher scores when they only sometimes participated in inquiry. The implication is that more time was spent in class on improving domain knowledge and

providing guidance, so that less time was available to spend on the actual inquiry process. The lower effect size for students who often or always engaged in inquiry-based learning is also supported by substantial research that indicate greater achievement in science scores during inquiry only improves when a significant amount of guidance is provided in advance (Zhang, 2013; Tieg et al., 2018; Jerrim et al., 2019; Cheng et al., 2017).

The NAEP data does not indicate whether or not the inquiry process was completed online; however, evidence suggests that scaffolding and support are needed for improved performance regardless of online or offline inquiry. The use of metacognitive prompts—hints, reminders, explanations—whether offline or on-screen, is an effective way to improve student outcomes during the inquiry process (Zhou & Lam, 2019). Koyunlu Unlu & Dokme (2022) found that students' science process skills improved as a result of online research done during inquiry.

Colwell et al. (2013) found that teacher and student modeling of online research and comprehension strategies improved student outcomes and increased digital literacy. However, these gains were short lived, and students were unable to retain these research strategies over time. Opportunities for inquiry are not enough to boost achievement; improving achievement through inquiry requires thorough, repeated scaffolding throughout a student's education.

Research Question 3: Frequency of Combining Science Information from Multiple Sources

Scale scores related to the frequency with which students combined science information from multiple sources were highest for students who often completed such activities, as opposed to those who completed such activities less often, or almost always (Data Table 7). Cohen's-*d* effect sizes showed the greatest magnitude in effect size when students who sometimes combined science information from multiple sources were compared to students who never or hardly ever practiced such skills (Data Table 9). Lacking specific information from the NAEP survey, the assumption is made here that multiple sources include multimodal sources, including sources found on the internet.

Knowledge synthesis results from combining information from multiple sources to develop a coherent interpretation of the content (DeSchryver, 2014). Creative knowledge synthesis also requires combination of multiple ideas, images, or associations, with the additional caveat that the resulting knowledge is fundamentally changed from any of the sources (American Psychology Association, n.d.). In other words, creative knowledge synthesis "transcends" the original information, and "transforms" it in new ways that lead to new understandings (DeSchryver, 2014; Henriksen et al., 2017, p.8). Regardless of whether knowledge is created or interpreted, it is crucial to comprehension (Elleman & Compton, 2017; Catts & Kahmi, 2017). Thus student scores for students who often combined science information from multiple sources had higher scale scores than students who only sometimes or never did so.

However, the vast array of resources on the internet can easily overwhelm students and impede their knowledge building abilities due to the potential for cognitive overload. There are limits to a student's working memory; if there is too much information to process, or if the student does not have enough domain knowledge to make sense of what they find online, their

success in knowledge building may suffer (MCW, 2022; Kirschner et al., 2006; Rosenshine, 2012; Zhang, 2013). This may help to explain why students who always or almost always use the internet have a somewhat lower effect size; too much information can be detrimental to knowledge building as can too little.

Zhang (2013) also suggests that scaffolding can enhance the amount of knowledge students can gather during inquiry using internet sources, supporting the possibility that students who often combined information from multiple sources spent additional time on scaffolding activities to improve their success in knowledge building. Conversely, students who always or almost always combine information from multiple sources would logically have had less time to develop knowledge building skills from online sources.

Conclusion

More and more often, students are participating in inquiry learning, a student-centered approach to learning where students construct their own knowledge by exploring open-ended tasks. Often, students will use online resources to look for solutions and build their knowledge base. However, students are not always well equipped to do so. This investigation focuses on the impact of classroom practices that may influence a student's ability to conduct this type of research and demonstrate increases in content knowledge through inquiry.

Using Technology to Search for Science Information

Most students today have access to their own device to use in their classrooms. As a result, they are expected to be able to use those devices to perform routine tasks such as using the internet and digital resources to gather information. However, results from this study show that—for better or worse—the frequency with which students used technology in their classroom had no effect on their achievement scores. This reinforces the idea that online searches are only as good as the queries generated by the researcher. If a student doesn't have adequate prior knowledge, or instruction in query development, their ability to use the internet to improve their learning is impaired. Additionally, accessing a valid online resource doesn't mean that the resource is at an appropriate level for meaning making.

Frequency of Science Inquiry Activities

Use of inquiry learning has been growing as a method designed to help students construct their own knowledge. However, simply using the inquiry method does not guarantee success. Without sufficient prior knowledge, instruction in methods for asking the right questions, or hints and reminders about key ideas, the inquiry process may miss the mark on increased learning potential. The current study suggests that too much inquiry and not enough scaffolding has little to no effect on achievement scores. However, intermittent use of inquiry shows bigger gains, presumably because more time was spent in preparing students for the inquiry tasks.

Frequency of Combining Science Information from Multiple Sources

The ubiquity of technology and its consistent use in classrooms as a pathway to learning indicates it is imperative that students are able to synthesize information from multiple sources in order to construct new knowledge and find solutions to authentic

problems. However, as with the inquiry process, use does not guarantee success. The fact that a student is able to identify reliable information does not mean they are able discern what information is most relevant to the learning targets, or that they know how to combine disparate pieces of information into a cohesive whole. Over reliance on collecting information from multiple sources without the supporting instruction on sense-making has no significant impact on learning. Additional scaffolding is the key to higher achievement.

Implications and Recommendations

- Frequent use of technology doesn't ensure increased achievement. Technology applications need to be intentionally chosen, and scaffolds must be in place to ensure successful implementation of the tools.
- Effective use of technology is not an intuitive process for students. Modeling and scaffolds should be used to help students master skills appropriate to the task.
- The inquiry process and knowledge synthesis both require that a student is able to handle a significant cognitive load. To offset these demands, pedagogy needs to include strategies designed to help teachers and students organize information in a way that streamlines thought processes and makes connections clear.

Limitations

Missing details about the concurrent work that is taking place in the classroom dilutes the impact of these results. The idea that less frequent inquiry, or the fact that students are asked to combine information from multiple sources less often, may be attributed to an increased amount of scaffolding surrounding the work is merely surmised. These results could mean that a teacher has less time to spend on this type of work, although that would suggest more time spent on traditional teaching methods—which could actually serve as necessary scaffolding. In addition, it is not clear whether students were given any direct instruction on how to effectively conduct internet searches or how to combine information from multiple sources.

Recommendations for Future Research

Additional research is needed to determine effective methods for teaching students how to write successful queries and how to combine information from multiple sources into new knowledge. Knowing how to make inferences based on incomplete information would increase student proficiency for both types of tasks. For inquiry, inferencing can help students ask more relevant questions. For knowledge synthesis, inference can make it possible to see connections between multimodal information—audio, video, images, charts, and graphs—to determine causality. Therefore, future research on effective methods for the use of inferencing to improve inquiry and knowledge synthesis would lead toward students better equipped to tackle solving ill-structured problems using online resources.

References

1. Anuyah, O., Milton, A., Green, M., & Pera, M. S. (2020). An empirical analysis of search engines' response to web search queries associated with the classroom setting. *Aslib Journal of Information Management*, 72(1), 88-111.

2. Becker, L. (n.d.). *Calculate d and r using means and standard deviations*. Effect Size Calculators. Retrieved April 14, 2024, from <https://lbecker.uccs.edu/>
3. Biagi, Federico, and Massimo Loi. 2013. Measuring ICT Use and Learning Outcomes: evidence from recent econometric studies. *European Journal of Education*, 48(1), 28–42.
4. Cairns, D., Areepattamannil, S. (2019). Exploring the relations of inquiry-based teaching to science achievement and dispositions in 54 countries. *Res Sci Educ* 49, 1–23.
5. Catts, H., & Kamhi, A. (2017). Prologue: Reading comprehension is not a single ability. *Language, Speech, and Hearing Services in Schools*, 48, 73–76.
6. Chen, K., & Chen, C. (2021). Effects of STEM Inquiry Method on Learning Attitude and Creativity. *Eurasia Journal of Mathematics, Science and Technology Education*, 17(11).
7. Cheng, S. C., She, H. C., & Huang, L. Y. (2017). The impact of problem-solving instruction on middle school students' physical science learning: Interplays of knowledge, reasoning, and problem solving. *Eurasia Journal of Mathematics, Science and Technology Education*, 14(3), 731-743.
8. Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.
9. Colwell, J., Hunt-Barron, S., & Reinking, D. (2013). Obstacles to developing digital literacy on the Internet in middle school science instruction. *Journal of Literacy Research*, 45(3), 295-324.
10. Deschryver, M. (2014). Higher Order Thinking in an Online World: Toward a Theory of Web-Mediated Knowledge Synthesis. *Teachers College Record* (1970), 116(12), 1–44. <https://doi.org/10.1177/016146811411601202>
11. Duncan, R. G., Na'ama, Y., & Chinn, C. A. (2021). Inquiry and learning in science. In *International handbook of inquiry and learning*. Routledge, 325-244.
12. Elleman, A. M., & Compton, D. L. (2017). Beyond comprehension strategy instruction: What's next?. *Language, Speech, and Hearing Services in Schools*, 48(2), 84-91.
13. Fisher, M., Smiley, A. H., & Grillo, T. L. (2022). Information without knowledge: The effects of Internet search on learning. *Memory*, 30(4), 375-387.
14. *Focus on NAEP*. (n.d.). The Nation's Report Card. Retrieved April 14, 2024, from https://www.nationsreportcard.gov/focus_on_naep/
15. Gliner, J. A., Vaske, J. J., & Morgan, G. A. (2001). Null hypothesis significance testing: Effect size matters. *Human Dimensions of Wildlife*, 6(4), 291-301.
16. Gruber, J. & Hargittai, E. (2023). The importance of algorithm skills for informed Internet use. *Big Data & Society*, 10(1), 1-14.
17. Heersmink, R. (2016). The internet, cognitive enhancement, and the values of cognition. *Minds and Machines*, 26, 389-407.
18. Henriksen, D., Mishra, P., & Fisser, P. (2016). Infusing Creativity and Technology in 21st Century Education: A Systemic View for Change. *Educational Technology & Society*, 19 (3), 27–37.
19. Jerrim, J., Oliver, M., & Sims, S. (2019). The relationship between inquiry-based teaching and students' achievement. New evidence from a longitudinal PISA study in England. *Learning and Instruction*, 61, 35–44.
20. Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational psychologist*, 41(2), 75-86.
21. Koyunlu Unlü, Z., & Dökme, I. (2020). The Effect of Technology-Supported Inquiry-Based Learning in Science Education: Action Research. *Journal of Education in Science Environment and Health*, 6(2), 120-133.
22. Marsh, E. J., & Rajaram, S. (2019). The digital expansion of the mind: Implications of internet usage for memory and cognition. *Journal of Applied Research in Memory and Cognition*, 8(1), 1-14.
23. Martin, C., Sowden, P. T., & Warren, F. (2022). Connected Creativity: The Impact of Web Search on Everyday Creative Thinking. *European Psychologist*, 27(3), 227–236.
24. MCW. (2022). *Cognitive load theory: A guide to applying cognitive load theory to your teaching*. Office of Educational Improvement, Medical College of Wisconsin. <https://www.mcw.edu/-/media/MCW/Education/Academic-Affairs/OEI/Faculty-Quick-Guides/Cognitive-Load-Theory.pdf>
25. NCES. (2024, January 29). *The Nation's Report Card | NAEP*. National Center for Education Statistics. Retrieved April 14, 2024, from <https://nces.ed.gov/nationsreportcard/>
26. *NDE Core Web*. (n.d.). The Nation's Report Card. Retrieved April 14, 2024, from <https://www.nationsreportcard.gov/ndecore/xplore/NDE>
27. OECD. (2011). *PISA 2009 Results: Students online: Digital technologies and performance*. (Volume VI). Paris: OECD publishing.
28. Porat, E., Blau, I., & Barak, A. (2018). Measuring digital literacies: Junior high-school students' perceived competencies versus actual performance. *Computers & Education*, 126, 23-36.
29. Raes, A., Schellens, T., De Wever, B., & Vanderhoven, E. (2012). Scaffolding information problem solving in web-based collaborative inquiry learning. *Computers & Education*, 59(1), 82-94.
30. Rosenshine, B. (2012). Principles of instruction: Research-based strategies that all teachers should know. *American educator*, 36(1), 12.

31. Rutkowski, A. F., & Saunders, C. (2018). *Emotional and cognitive overload: the dark side of information technology*. Routledge.
32. Shapiro, A. (2004). How including prior knowledge as a subject variable may change outcomes of learning research. *American Education Research Journal*, 41(1), 159-189.
33. Simpson, T. W. (2013). Evaluating Google as an epistemic tool. *Philosophical Engineering: Toward a Philosophy of the Web*, 97-115.
34. Sobehrad, S. (2023). Analyzing the Ability of Students to Creatively Solve Ill-Defined Problems Using Web-Mediated Content. Department of Education and Human Services, Central Michigan University, Mt. Pleasant, MI.
35. Teig, N., Scherer, R., & Nilsen, T. (2018). More isn't always better: The curvilinear relationship between inquiry-based teaching and student achievement in science. *Learning and Instruction*, 56, 20–29.
36. Willoughby, T., Anderson, S. A., Wood, E., Mueller, J., & Ross, C. (2009). Fast searching for information on the Internet to use in a learning context: The impact of domain knowledge. *Computers & Education*, 52(3), 640-648.
37. Wood, E., De Pasquale, D., Mueller, J. L., Archer, K., Zivcakova, L., Walkey, K., & Willoughby, T. (2016). Exploration of the relative contributions of domain knowledge and search expertise for conducting Internet searches. *The Reference Librarian*, 57(3), 182-204.
38. Zhang, M., & Quintana, C. (2012). Scaffolding strategies for supporting middle school students' online inquiry processes. *Computers and Education*, 58(1), 181–196.
39. Zhang, M. (2013). Supporting middle school students' online reading of scientific resources: moving beyond cursory, fragmented, and opportunistic reading. *Journal of Computer Assisted Learning*, 29(2), 138–152.
40. Zhou, M., & Lam, K. K. L. (2019). Metacognitive scaffolding for online information search in K-12 and higher education settings: a systematic review. *Educational technology research and development*, 67(6), 1353-1384.