Original Research Article

Journal of Educational Computing Research
2022, Vol. 0(0) 1–22

DOI: 10.1177/07356331221081484
journals.sagepub.com/home/jec
SAGE

Assessing the Development of Digital Scientific Literacy With a Computational Evidence-Based Reasoning Tool

Nancy Holincheck[1], Terrie M. Galanti[2], and James Trefil[3]

Abstract

The evolving digital world requires scientifically literate citizens who are able to critically evaluate Internet sources of varying credibility. Instruction on evidence evaluation in postsecondary education often focuses on peer-review as a singular indicator of credibility. With increased access to web-based scientific information, students must also learn to think critically in real-time about the dimensions of credibility. This study describes the integration of *sInvestigator*, a computational evidence-based scientific reasoning tool, with a class of 32 students in an undergraduate honors course focused on socio-scientific issues. A cross-disciplinary team of researchers with expertise in science education, scientific literacy, and evidence evaluation developed and implemented an online questionnaire to measure students' development of digital scientific literacy. After using *sInvestigator* to evaluate sources of scientific evidence based on publisher reputation, author competence, and author objectivity, students were better able to assess the credibility of online information. Results of this study also confirm the potential to authentically assess students' use of author and publisher information to evaluate digital scientific sources. The need for further research on the operationalization and measurement of digital scientific literacy is discussed.

[1]College of Education and Human Development,George Mason University, Fairfax, VA, USA
[2]College of Education & Human Services, University of North Florida, Jacksonville, FL, USA
[3]College of Science, George Mason University, Fairfax, VA, USA

Corresponding Author:
Nancy Holincheck, George Mason University, 4400 University Drive, MS 1E8, Fairfax, VA 22030, USA.
Email: nholinch@gmu.edu

Keywords
digital literacy, scientific literacy, evidence credibility, science, internet sources, sInvestigator

Introduction

A scientifically literate populace which relies on evidence to evaluate the quality of information has been the shared pursuit of scholars and national organizations alike for over 50 years (American Association for the Advancement of Science, 1989; NGSS Lead States, 2013; National Research Council, 2012). The many different conceptions of scientific literacy in the literature have been driven by the varied goals of their authors (Norris et al., 2014). Trefil and Hazen (2010) define scientific literacy as “the matrix of knowledge needed to understand enough about the physical universe to deal with issues in the news and elsewhere” (p. 58). This knowledge-based classification of scientific literacy is motivated by the need to prepare an educated citizenry to participate fully in political discourses. Yet scientific literacy in the 21st century cannot be separated from an individual’s ability to assess the quality of increasingly accessible digital scientific information. The urgent need for flexible, student-centered distance learning opportunities during the global health crisis has meant that instructors are increasingly reliant on Internet-based information in K-16 science education. Understanding the relationship between science and culture (Holbrook & Rannikmae, 2007) within the broader construct of scientific literacy is dependent upon an individual’s ability to think critically about information from a variety of sources on the Internet.

Digital literacy has been broadly conceptualized in education literature as the knowledge and application of technological, cognitive, and sociological skills in the use of electronic information. Digital and scientific literacies “are constructed simultaneously and socially” (Dias da Silva & Heaton, 2017) as students learn to evaluate and synthesize information. Students apply these skills as they process scientific online information, and they must critically consider the credibility of freely-accessible articles to engage in rational civic debate (Trefil, 2008). Anelli (2011) argues that insufficient attention has been given to teaching undergraduate students to evaluate the credibility and expertise of scientists in their pursuit of scientific knowledge. When instructors leverage online scientific information in virtual teaching and learning, scientific literacy skills become increasingly interconnected with digital literacy skills. As the science education community strives to understand scientific literacy in relation to digital literacy, there is a need to measure the students’ abilities to evaluate the credibility of digital sources of information. Therefore, this research aims to work toward the operationalization and measurement of digital scientific literacy in K-16 science education.

This article describes the development of an online Digital Scientific Literacy Questionnaire (DSLQ) by which students assess the credibility of scientific information from a set of purposefully-selected websites. The DSLQ was developed as an evaluation resource for the NSF-funded Improving Undergraduate Science Education project titled "Teaching Critical Thinking Skills in Science with *sInvestigator*." Faculty from science, engineering, and education worked with the instructor of an honors undergraduate science course to use *sInvestigator*, a cognitive assistant designed to develop students' critical thinking skills for evidence-based scientific argumentation. The DSLQ was used throughout the project to evaluate *sInvestigator* and to inform instructional decision-making.

Throughout the development of the instrument, the research team recognized the additional potential of this tool for formative assessment of digital scientific literacy skills. The DSLQ provides an empirical window on the real-time decision-making of students as they evaluate the credibility of unfamiliar contemporary scientific sources. The findings from this exploratory study reveal the important potential for further development of the DSLQ for use as both a research instrument and as a formative assessment tool across multiple disciplines.

Research Questions

This study examined the following research questions:

1. How do students' ratings of publisher reputation, author competence, and author objectivity relate to their overall evaluation of the credibility of online scientific sources?
2. In what ways do students evaluate digital scientific information differently after using these credibility dimensions within *sInvestigator*?

Theoretical Framework

The Internet and its search engines have made scientific information more universally accessible both inside and outside the classroom, but these structures have "removed or at least enabled a bypass of traditional filters and interpreters" (Britt et al., 2014, p. 105) that would otherwise control the quality of information that students read and interpret. The removal of this filtering function complicates the development of scientific literacy skills as students independently filter digital information. Students are often not aware of the importance of evaluating sources (Scharrer et al., 2019; Hämäläinen et al., 2021; Kiili et al., 2021) and must learn to make real-time assessments about the quality of an article. Students engage in a predictive judgment of credibility as they filter an extensive list of search engine results before they engage in an evaluative judgment of credibility of scientific content (Hahnel et al., 2020; Hendriks et al., 2020).

Digital Literacy in Relation to Science Teaching and Learning

As science educators leverage the proliferation of electronic information to innovate classroom instruction, they must attend to digital literacy as a “prerequisite for learning in a student-centered educational culture” (Fieldhouse & Nicholas, 2008, p. 50). Gilster (1997) characterized digital literacy as the general ability to evaluate and interpret information from multiple online sources. More specifically, Eshet-Alkalai and Chajut (2009) offered a skills-based framework for digital literacy aligned with Gilster’s definition (Porat et al., 2018). Three of Eshet-Alkalai and Chajut’s (2009) six distinct literacies are especially critical for students who use the Internet to explore scientific phenomena. *Branching literacy* is the ability to construct knowledge by a non-linear navigation through electronic domains, *information literacy* addresses the critical evaluation of information for bias, and *real-time thinking* is how students process and evaluate large volumes of information. Fieldhouse and Nicholas (2008) further described information literacy as the ability to assess the “authority, relevance, currency, quality, coverage, and objectivity of information” (p. 55). They elaborated on the challenges of encouraging college students who consider themselves information-savvy to become information-wise in a world where they are as likely to rely on peers as they are on educators.

As students evaluate Internet evidence and develop argumentation based on this evidence, the relationship between digital and scientific literacy becomes more interdependent than the characterization of parallel competencies in recent literature (Bliss, 2019; Dias da Silva & Heaton, 2017). Because of this interdependence, the research team identified a singular construct called *digital scientific literacy* to acknowledge the inseparability of these competencies. This singular construct can move the field of science education forward in describing students’ abilities to predict and critically evaluate the credibility of online evidence. Digital scientific literacy encapsulates the skills associated with searching, evaluating, and understanding Internet sources as theorized by Wiley et al. (2009) and becomes a prerequisite for students to engage productively in complex scientific inquiry in the 21st century (Van Laar et al., 2017; Silber-Varod et al., 2019).

Building Digital Scientific Literacy in K-16 Classrooms

Today’s students have been broadly classified as digital natives (Prensky, 2001) who prefer to learn through random access and parallel processing of online information. Although students increasingly use web-based sources to obtain scientific information, prior research has shown that students do not critically consider information they find on the Internet (e.g., Brem et al., 2001; Gormally et al., 2012). They often focus on the first three information sources in an Internet search (Wineburg & McGrew, 2019) and determine trustworthiness based on superficial criteria (Bråten et al., 2015; Halverson et al., 2010). The relaxation of parameters for publishing and the lack of editorial control in the digital world (Kammerer et al., 2020) adds to the challenges of evaluating credibility of scientific information. As students seek to make sense of scientific information found on the Internet, they must evaluate the credibility, originality, and integrity of the information they discover (Eshet, 2004). Digital scientific literacy thus becomes essential in inquiry-based instruction which models the work of practicing scientists in corroborating evidence from multiple websites (Wiley et al., 2009).

Several researchers have designed classroom interventions to support their students in building digital literacy with a specific focus on critical thinking about Internet evidence and scientific argumentation (e.g., Brem et al., 2001; Halverson et al., 2010; Hoffman et al., 2003). While some attention has been paid to developing students' scientific literacy in digital contexts, a thorough search of the literature for existing tools failed to yield an appropriate measure of scientific and digital literacy as a singular concept. Although separate measures of scientific and digital literacy exist (e.g., Hahnel et al., 2020; Jin et al., 2020; Macedo-Roouet et al., 2019; Opitz et al., 2017), none of the extant measures assess a student's real-time ability to think critically about Internet-based scientific information. The research team hypothesizes that a student's overall credibility rating of online sources can be explained by ratings of publisher reputation, author competence, and author objectivity. There is a need for a measure that simultaneously emulates students' Internet filtering strategies and assesses their digital scientific literacy skills. This study describes the design of the DSLQ as a classroom tool that served two purposes for this study. It informed the development and implementation of *sInvestigator* as a critical thinking tool, and it served as a formative assessment of students' digital scientific literacy.

Methods

A mix of quantitative and qualitative methods were used to study students' development of digital scientific literacy while using *sInvestigator* in an undergraduate honors course focused on socio-scientific issues. The research team designed the DSLQ in the context of *sInvestigator* as an instructional intervention, but the measure was intentionally designed to be relevant in any digital scientific inquiry context.

sInvestigator as an Instructional Intervention

sInvestigator was developed for educational purposes as part of NSF grant #1611742 and builds upon *Cogent* (Tecuci et al., 2015), a cognitive assistant developed for evidence evaluation for intelligence analysis. *sInvestigator* constructs a real-time computational model, or knowledge base, of evidentiary reasoning based on student data entry and justifications (Boicu et al., 2016). Students can collaborate to evaluate Internet scientific sources and to develop probabilistic hypotheses and arguments. The initial version of the *sInvestigator* tool (Fall 2016) prompted students to rate the credibility of each item of evidence from a drop-down menu (e.g, certain, almost certain, likely certain, lacking support). To improve the educational potential of *sInvestigator*, the research team identified several distinct but related criteria for

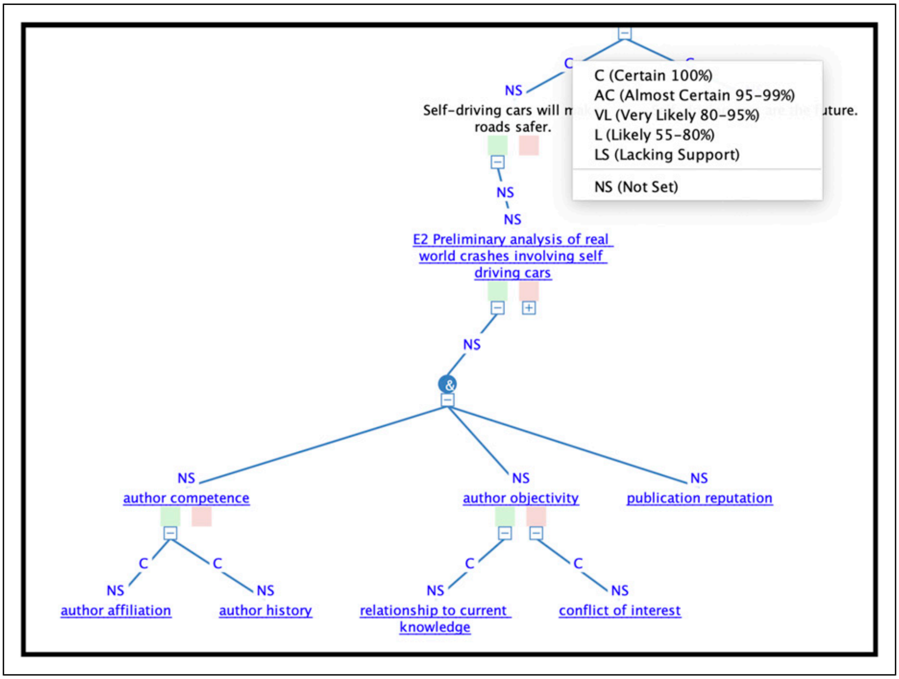

Figure 1. Screenshot of *sInvestigator* tool's credibility analysis. Note: The *sInvestigator* tool can be downloaded from the "Teaching Critical Thinking Skills in Science with *sInvestigator*" website (http://lac.gmu.edu/sInvestigator/index.html).

evaluating the credibility of digital scientific sources. Drawing on their prior conceptualizations of scientific literacy, the team hypothesized that students' ratings of publisher reputation, author competence, and author objectivity would explain their overall ratings of credibility.

The *sInvestigator* tool was modified to prompt students to evaluate the credibility of multiple items of digital evidence in relation to scientific hypotheses (see Figure 1). Students use drop-down menus to assign their ratings of author competence (including author affiliation and author history), author objectivity (including relationship to current knowledge and conflict of interest), and publisher reputation. *sInvestigator* uses a probabilistic algorithm to calculate an overall credibility rating for each item of evidence based on these assignments, but students can override that overall rating. Students are also prompted to write a short justification for their credibility ratings.

Design of the DSLQ

The DSLQ measure was developed in parallel with the modifications to *sInvestigator,* but the research team constructed the measure to be used for more than project

Article #1	Article #2	Article #3
Title	***Title***	***Title***
Energy Drinks' Effects on Student-Athletes and Implications for Athletic Departments	Improved time to exhaustion following ingestion of the energy drink Amino Impact™	Do Energy Drinks Improve Sports Performance?
Publisher	***Publisher***	***Publisher***
The Sports Journal	Journal of International Society of Sports Nutrition	TeamSnap

2 You need to write a paper about the effect of energy drinks on sports performance. A Google search leads you to the three articles listed in the table above. Which article are you *most likely* to use as a reference?

- ◯ *Energy Drinks' Effects on Student-Athletes and Implications for Athletic Departments*
- ◯ *Improved Time to Exhaustion Following Ingestion of the Energy Drink Amino Impact™*
- ◯ *Do Energy Drinks Improve Sports Performance?*

The following text is an excerpt from the article.

Does ingestion of an energy drink really boost athletic performance? Caffeine is the only ingredient in energy drinks that has been studied in depth and that shows proven effects; short- and long-term effects of high doses of taurine and glucuronolactone require additional study. Athletes have long used caffeine prior to training sessions and competitions, but most nevertheless do not well understand how the drug works, for example that, as a diuretic, caffeine is capable of aggravating the dehydration athletes may experience during competition. The scientific literature itself provides mixed messages about caffeine's performance-enhancing capability and its value prior to exercise. Fornicola (2007) stated that no real need exists to use energy drinks for performance advantage and that that quick caffeine fix is not a very intelligent strategy. In contrast, the website of the Sports Medicine Council of Manitoba reports that endurance athletes might gain some advantage by exploiting caffeine to derive energy from fat early in a competition, thereby leaving more muscle glycogen available to provide energy later on (Caffeine—Performance, n.d., p. 1). However, the website also advises athletes that "4% dehydration equals 20% of performance lost" (p. 1).

4 What is the reputation of the publisher of this article?

- ◯ **Certainly** reputable
- ◯ **Very likely** reputable
- ◯ **Likely** reputable
- ◯ **Likely not** reputable
- ◯ **Very likely not** reputable
- ◯ **Certainly not** reputable
- ◯ I have no basis by which to evaluate the publisher's reputation.

5 What is the competence of the author of this article?

- ◯ **Certainly** competent
- ◯ **Very likely** competent
- ◯ **Likely** competent
- ◯ **Likely not** competent
- ◯ **Very likely not** competent
- ◯ **Certainly not** competent
- ◯ I have no basis by which to evaluate the author's competence.

6 What is the objectivity of the author of this article?

- ◯ **Certainly** objective
- ◯ **Very likely** objective
- ◯ **Likely** objective
- ◯ **Likely not** objective
- ◯ **Very likely not** objective
- ◯ **Certainly not** objective
- ◯ I have no basis by which to evaluate the author's objectivity.

7 What is the credibility of this article?

- ◯ **Certainly** credible
- ◯ **Very likely** credible
- ◯ **Likely** credible
- ◯ **Likely not** credible
- ◯ **Very likely not** credible
- ◯ **Certainly not** credible
- ◯ I have no basis by which to evaluate the credibility of the article.

8 Would you use the article as a reference for your paper? Justify your response in 2 or 3 sentences.

Figure 2 . Excerpts from DSLQ Pre-Questionnaire—Medium-Credibility Article.

evaluation. While the dimensions of credibility aligned with the prompts in *sInvestigator*, the DSLQ was designed to be used by teachers and researchers to better understand how K-16 students evaluate scientific evidence in broader instructional contexts. The DSLQ prompts students to evaluate publisher reputation, author competence, author objectivity, and overall article credibility with a sequence of Likert-type items and an open-ended justification for each of three online articles (see Figure 2).

The DSLQ integrates evaluation of scientific information with digital literacy skills as defined by Eshet-Alkalai and Chajut (2009). It is hosted on an online survey platform to provide an authentic assessment of branching literacy, information literacy, and real-time thinking skills. The DSLQ emulates real-time information filtering in a naturalistic Web search and quantifies the complexity of evaluating digital scientific evidence. A Likert-type scale design provides a numerical spectrum of confidence for assessing credibility. Students are asked about their confidence in the credibility of three articles within the context of writing a research paper. They are also provided with the following definitions of the three dimensions of overall credibility.

- **Author competence** refers to the extent to which we can believe that the author of an article is an expert in the subject matter of that article.
- **Author objectivity** refers to the extent to which we can believe that the author's claims have a scientific basis and are not influenced by non-scientific factors.
- **Author credibility** refers to the extent to which we can believe what an article is telling us.

The DSLQ design was inspired by a multiple-choice item about confidence in article accuracy from the Test of Scientific Literacy Skills (Gormally et al., 2012). The DSLQ is unique in that it embeds active hyperlinks to each of three articles of varying credibility to allow students to use non-linear search strategies to evaluate digital scientific information. Because the DSLQ is web-based, students can open additional windows to visit other websites to investigate potential bias in the articles they are evaluating. Four rounds of pilot testing were conducted to refine the questions and format of the DSLQ before initial classroom implementation.

For the surveys administered in this study (pre- and post-DSLQ), the research team identified two socio-scientific topics (risk of medical X-rays and performance effects of energy drinks) that were not directly addressed in the course but were considered relevant to the participants. Once the topics were identified, the research team reviewed online sources and selected three articles of varying credibility. These included a high-credibility article that was peer-reviewed and published in a well-known journal, a medium-credibility article from a respected source that was not clearly peer-reviewed, and a low-credibility article from a blog on a commercial site. The DSLQ is structured so that the set of high-, medium-, and low-credibility articles can be easily changed when the measure is used in other contexts.

On the entry pages of the online survey, students view a table with the titles and publisher names of all three articles and definitions of key credibility terms. Students are then presented with screenshots and hyperlinks to the article web pages and asked to evaluate the publisher reputation, author competence, author objectivity, and overall article credibility (see Figure 2). Students choose from six different levels of confidence ranging from "certain" to "not certain" for each evaluation question. Students are also given the option of choosing "I have no basis by which to evaluate" to increase the likelihood their responses are carefully considered. After evaluating the hypothesized dimensions of credibility and overall credibility for each article, students are asked if they would use each article as a reference in a research paper.

Setting and Participants

The DSLQ was implemented in a 16-week undergraduate honors "History of Science" course at a large mid-Atlantic public university. The class was taught by a member of the research team and was traditionally a lecture-based course with tests, class presentations, and a final paper on a topic selected by the student. For the purposes of this project, the course was restructured to engage students in a series of four inquiry lessons using *sInvestigator.* In the *sInvestigator* lessons, a hybrid instructional model of 50% lecture and 50% collaborative inquiry with preselected digital sources of varying credibility was used to develop students' understanding of the evolution of scientific thought. The other class sessions included a mix of traditional lecture, tests related to the course readings and lectures, individual consultations with the instructor, and student presentations.

The participants were the 32 students (13 female and 19 male) enrolled in the course in Fall 2017. All students were sophomores and juniors in the university's honors college and had previously completed a course on inquiry and research as part of the honors program. Students worked collaboratively during the four *sInvestigator* lessons in class, but the DSLQ was administered to students individually.

Data Sources

This study uses quantitative and qualitative data from pre- and post-administrations of the DSLQ. This data was collected during the first and last class meetings of the 16-week course in which *sInvestigator* was used. Quantitative data includes the Likert responses (*1 = certainly not* to *6 = certainly*) to questions about the publisher reputation, author competence, author objectivity, and overall credibility of the article. The qualitative data used in this study are the students' open-ended responses in which they justified whether or not they would use each article in a research paper. Student justifications were analyzed to help the research team understand student reasoning behind the credibility ratings.

Data Analysis

To answer Research Question 1, we tested our hypothesis that an overall credibility rating can be explained by ratings of publisher reputation, author competence, and author objectivity. Student responses to these ratings across all six articles were compared including student responses to the pre- and post- administrations of the DSLQ. All quantitative analyses were performed using the statistics program SPSS (SPSS 26). Before modeling how the ratings of the three hypothesized dimensions of

Table 1. Mean Overall Credibility Ratings for Articles used on Pre- and Post-Questionnaires.

Researcher Rating of Credibility	Pre-Mean Overall Credibility				Post-Mean Overall Credibility			
	Score	N	SD	Equivalent Rating	Score	n	SD	Equivalent Rating
Low-credibility	2.17	29	1.05	Very likely not credible	3.08	26	1.02	Likely not credible
Medium-credibility	3.64	25	1.00	Likely credible	4.80	25	0.76	Likely to very likely credible
High-credibility	5.43	30	0.68	Very likely to certainly credible	5.67	31	0.61	Very likely to certainly credible

Relating Dimensions of Credibility to Overall Credibility.

credibility contributed to a students' rating of overall credibility, the data was disaggregated by article source and correlation analyses were performed. The correlational analyses of these ratings were necessary to confirm sufficient variability to fit a model.

Analysis of the descriptive statistics from the Fall 2017 pre- and post-administration of the DSLQ provided insight into students' consideration of the three dimensions of article credibility in relation to one another. This analysis included examination of the means and standard deviations of students' overall credibility ratings of each article.

To answer Research Question 2, we evaluated the effectiveness of *sInvestigator* using paired samples t-tests to compare participants' pre- and post-ratings of the low, medium, and high-credibility articles on the DSLQ. Content analysis (Neuendorf, 2016) of participants' responses to open-ended questions on the DSLQ was completed to provide additional context for both research questions. Students were asked to justify their ratings of each dimension and the overall credibility; their justifications varied in length from a few words to short paragraphs. Two members of the research team separately analyzed student responses and compared their results to inform their interpretation of the quantitative findings.

Results and Discussion

Before proceeding with inferential statistical analysis, the research team used descriptive statistics to confirm that participant ratings of overall credibility aligned with researcher ratings. Participants' mean overall credibility ratings of each source on both the pre- and post-questionnaires (Table 1) were within the specified range of ratings from the research team for low quality (1–3.5), medium quality (3.5–5), and high-quality articles (5–6). These results indicate that articles selected for the DSLQ provided an appropriate range of credibility for student consideration.

The research team calculated correlation coefficients to better understand potential relationships between publisher reputation, author competence, author objectivity, and overall credibility. Analysis of a combined correlation matrix for all four of these DSLQ

Table 2. Correlation Matrix for Overall Credibility Versus Predictor Variables Across All Six Articles.

	n	*M*	*SD*	Publisher Reputation	Author Competence	Author Objectivity
Publisher Reputation	*168*	*4.37*	*1.49*	--		
Author Competence	*163*	*4.52*	*1.24*	*0.879***	--	
Author Objectivity	*165*	*4.15*	*1.45*	*0.656***	*0.619***	--
Overall Credibility	*171*	*4.21*	*1.52*	*0.933***	*0.895***	*0.655***

***p < .01

variables (see Table 2) indicated that each predictor was highly correlated with each of the other predictors and with the overall credibility outcome. The strong correlations between the predictor variables and overall credibility for all six articles are consistent with the research team's hypothesis that each of these dimensions is related to assessment of overall credibility.

To further explore the hypothesis that student ratings of publisher reputation, author competence, and author objectivity could serve as predictors of the student rating evaluation of the overall credibility of a scientific article, separate correlation matrices were calculated for each of the low, medium, and high-credibility articles. This analysis revealed a moderate to high correlation ($r > 0.5$) between author competence and overall credibility for all six articles, and between publisher reputation and overall credibility for five of the six articles (see Table 3). Further examination of these article-specific correlation matrices revealed differential relationships across the three levels of credibility. Unlike the other hypothesized predictors, author objectivity was moderately correlated with overall credibility for only two of the six articles (Johnson & Christensen, 2019).

Regression Analysis

The post–high-credibility article met the criteria for standard regression analysis because none of the variables were highly correlated. Assumptions of linearity, independence, homogeneity and normality were confirmed by examination of residual plots. A multiple linear regression was conducted to examine how the three ratings of publisher reputation, author competence, and author objectivity contributed to students' overall rating of credibility for the post–high-credibility article. Results of the multiple linear regression indicated that there was a collective significant effect between publisher reputation, author competence, and author objectivity ($F(3, 25) = 11.416$, $p < .001$, $R^2 = 0.578$). According to Cohen (1988), this R^2 value suggests a large effect. The overall credibility is equal to −0.430 + 0.678 (publisher reputation) + 0.323

Table 3. Correlations for Overall Credibility of Each Article with Predictor Variables.

	Publisher Reputation	Author Competence	Author Objectivity
Overall Credibility (Pre-Low)	0.90**	0.71**	0.41
Overall Credibility (Pre-Medium)	0.77**	0.61**	0.56**
Overall Credibility (Pre-High)	0.68**	0.56**	0.36
Overall Credibility (Post-Low)	0.86**	0.65**	0.54**
Overall Credibility (Post-Medium)	0.75**	0.70**	0.34
Overall Credibility (Post-High)	0.42	0.85**	0.16

**$p < .01$.

(author competence) + 0.057 (author objectivity), but only publisher reputation (t = 3.814, p = .001) was a significant predictor in the model. For every one point increase in student rating of publisher reputation of this article, the overall credibility rating of the article increased by 0.678 points.

This result is consistent with earlier correlation analyses of all six articles as it suggests that students are relying on first impressions of publisher information as displayed and linked in the DSLQ to make a determination of overall credibility. Although students' use of a peer-reviewed designation is an important aspect of assessing digital scientific literacy, the lack of significant contribution of author competence and author objectivity is worthy of additional exploration. It indicates that students need additional instruction on evaluating these other dimensions of article credibility. Although only one of the six articles met the criteria for this standard regression analysis and model fitting, there is potential for future applications of this analysis using different sets of scientific articles and a student population with more varied experience with academic research.

Examining the Impact of *sInvestigator*: Differences in Student Evaluation of Digital Scientific Information

To understand the role of *sInvestigator* as a classroom intervention in building digital scientific literacy skills, paired samples t-tests were used to examine differences in student performance on the pre- and post-administration of the DSLQ. The research team compared means for publisher reputation, author competence, author objectivity, and overall credibility for each pair of articles (low, medium, and high credibility from the pre- and post-questionnaires). The mean scores and standard deviations for each variable are provided in Tables 4, 5, and 6. Because of the variability in the number of students who responded "I have no basis by which to evaluate (criterion)" across the six articles, these values are also reported in Tables 4, 5, and 6 in italics and parentheses after the number of responses.

Table 4. Mean Ratings for High-Credibility Articles.

		Publisher Reputation	Author Competence	Author Objectivity	Overall Credibility
Pre	Mean rating	5.52	5.60	5.18	5.43
	n *(no basis)*	31	31	30 *(1)*	30 *(1)*
	Standard deviation	0.63	0.56	1.09	0.68
Post	Mean rating	5.81	5.67	4.89	5.67
	n *(no basis)*	31	31	30 *(1)*	31
	Standard deviation	0.40	0.54	1.37	0.61

After using *sInvestigator*, students were better able to consider potential nuances of publisher reputation and author competence when evaluating overall credibility of medium and low-credibility articles. Students in this study had previously learned of the importance of using peer-reviewed sources; use of *sInvestigator* equipped students to consider multiple aspects of the credibility of digital scientific information.

High-Credibility Articles

Students rated the publisher reputation of high-credibility articles as "very likely to certainly reputable" on both the pre- and post-questionnaires, yet assigned a significantly higher rating for publisher reputation ($t(30) = -2.19$, $p = .037$) with less variance on the post- than on the pre-questionnaire (see Table 4). After using *sInvestigator*, students were more confident of the overall credibility of the peer-reviewed, high-credibility article on the post-questionnaire than on the pre-questionnaire. Students felt confident in judging these dimensions for both peer-reviewed highly-credible articles, as 30 of 31 students (97%) were able to assign these ratings. Only one of the 32 students (3%) responded "I have no basis by which to evaluate" for any of the dimensions of credibility. Analysis of students' open-ended justifications indicated that peer-review and the author's level of education were important factors in their evaluation of article credibility.

Medium-Credibility Articles

For the medium-credibility articles (see Table 5), students assigned higher ratings for publisher reputation ($t(21) = -3.74$, $p < .001$)), author competence ($t(21) = -4.71$, $p < .001$), and overall credibility ($t(24) = -4.418$, $p < .001$) on the post- than on the pre-questionnaire. More than half of the students (56%) noted that the website of the post-test article included a statement that the journal was peer-reviewed. This result provides additional evidence that peer-review is an important consideration in credibility evaluation. The reduction in the number of students who selected "I have no basis by

Table 5. Mean Ratings for Medium-Credibility Articles.

		Publisher Reputation	Author Competence	Author Objectivity	Overall Credibility
Pre	Mean rating	3.77	3.91	3.87	3.64
	n *(no basis)*	23 *(7)*	23 *(7)*	26 *(3)*	27 *(3)*
	Standard deviation	1.15	0.87	1.29	1.00
Post	Mean rating	4.91	4.91	4.48	4.80
	n *(no basis)*	30 *(1)*	29 *(2)*	27 *(3)*	29 *(2)*
	Standard deviation	0.68	0.61	0.99	0.76

which to evaluate (criterion)" also indicates that after using *sInvestigator*, students were more confident of the overall credibility of the medium-credibility article.

Low-Credibility Articles

For the low-credibility articles, students assigned a higher overall credibility rating on the post- than on the pre-questionnaire, ($t(23) = -3.82$, $p = .001$ (see Table 6). After using *sInvestigator*, students' rating of mean overall credibility changed from "very likely not credible" to "likely not credible." The increase in the number of students who selected "I have no basis by which to evaluate (criterion)" also suggests that students are more aware of the need to consider multiple dimensions of article credibility. Analysis of the qualitative response to the question, "Would you use this article as a reference for your paper" indicated that students looked beyond the commercial indicators on the website and the blog format of the article. They considered both information about the author and the use of journal citations within the article in their credibility assessments.

Discussion

The DSLQ was initially developed to evaluate the effectiveness of *sInvestigator* as a curricular intervention. As the research team developed and refined the instrument, they recognized the additional potential of this tool for formative assessment of digital scientific literacy skills. The DSLQ provided an empirical window on the real-time decision-making of students as they evaluated the credibility of unfamiliar contemporary scientific sources. Student ratings of publisher reputation, author competence, and author objectivity as dimensions of overall credibility have specific implications for how students evaluate digital scientific information.

Table 6. Mean Ratings for Low-Credibility Articles.

		Publisher Reputation	Author Competence	Author Objectivity	Overall Credibility
Pre	Mean rating	2.57	2.95	2.86	2.17
	n (no basis)	28 (3)	27 (4)	30 (1)	29 (2)
	Standard deviation	1.12	1.07	1.58	1.05
Post	Mean rating	3.13	3.57	3.59	3.08
	n *(no basis)*	26 *(2)*	23 *(5)*	23 *(5)*	26 *(2)*
	Standard deviation	1.06	0.60	0.80	1.02

Publisher Reputation

The designation of an article as "peer reviewed" was both a necessary and sufficient filter for publisher reputation, and by extension overall credibility, of the high-quality articles. Publisher reputation explained over 60% of the variance in overall credibility for five of the six articles (see Table 3), suggesting that publisher reputation is the primary consideration for honors undergraduate students in assessing overall credibility. Three of the six articles in the DSLQ had an easily-discernible peer-reviewed designation at the top of the webpage, and nearly all (97%) of the students assigned a rating for publisher reputation on these articles. There were 12 "I have no basis by which to evaluate publisher reputation" responses for the three non-peer-reviewed articles (13%) compared to one "no basis" response for the peer-reviewed articles (1%). It is important to note that the honors students in this study had taken a research inquiry course at the beginning of their undergraduate studies which emphasized the importance of using only peer-reviewed references for academic papers. In their qualitative responses to the question of whether they would use the article as a reference, students often framed peer-review as a singular criterion for credibility.

Author Competence

Author competence explained over 70% of the variance *(r^2)* in overall credibility for the post–high-credibility article (see correlation coefficients *r* in Table 3). The medium to high correlations between author competence and overall credibility for the other articles indicates that author competence is an important consideration for participants in this study in assessing overall credibility. After using *sInvestigator*, students' ratings of author competence were more clustered, particularly on the low-credibility article. The reduced variability in student responses suggests that use of *sInvestigator* led to students engaging in careful consideration of author competence as a dimension of overall credibility.

Author Objectivity

Student ratings of author objectivity did not influence their ratings of overall article credibility for four of the six articles. The standard deviations of ratings of author objectivity at the individual article level had higher variability than the other dimensions, suggesting that students need additional instruction on how to assign these ratings when author information is not clearly displayed on the website. Furthermore, author objectivity was not a significant predictor in the model for the pre–high-credibility article. Although a high percentage of the students entered reasonable credibility ratings on both the pre- and post-questionnaire, the high variability in author objectivity ratings indicate that the undergraduate honors students who participated in this study were ill-prepared to judge author objectivity for the articles used in the DSLQ. Students would benefit from additional instruction on evaluating whether an author's claims have a scientific basis or if authors are likely influenced by non-scientific factors. Further research into digital scientific literacy can help educators better understand how author objectivity contributes to student understanding of the credibility of scientific evidence, and whether evaluation of this variable depends on individual student knowledge of a field.

Overall Credibility in Relation to Digital Scientific Literacy

The DSLQ is a tool to support researchers and instructors in understanding how students learn to evaluate digital scientific information. In the DSLQ, students rate the credibility of a digital source by examining a screenshot of the website and the first few sentences of the article with an active link to the full article. The design of this questionnaire simulates the search experience that students use when they are searching for digital evidence. The purposeful sequencing of questions about publisher reputation, author competence, and author objectivity within an online survey platform allows students to use digital information beyond the website itself to assign an overall credibility rating.

There is a need for additional research on the interaction of branching literacy and real-time literacy (Eshet-Alkalai & Chajut, 2009) within the broader concept of digital scientific literacy. A student who is completing the DSLQ has the explicit option to visit the webpage with the full article and the implicit option to visit additional websites to explore author and publisher information on other websites. Student decision-making about branching to visit other websites and time spent reading websites are important aspects of credibility evaluation, but data was not collected on these behaviors. Future iterations of the DSLQ should include a mechanism to track student search behaviors, including websites visited and time spent on websites, as part of the overall credibility determination.

The findings from this exploratory study reveal the important potential for further development of the DSLQ for use as both a research instrument and as a formative assessment tool in science. The reliance on peer-review as a primary measure of quality is problematic in the increasingly complex world of digital science information. There is a need for targeted instruction using both peer- and non–peer-reviewed sources of scientific information in the pursuit of a more nuanced consideration of overall credibility. Articles which are not peer-reviewed yet reference and synthesize peer-reviewed information are increasingly accessible and relevant in academic writing. Cultivating a more flexible idea of the credibility of scientific sour in digital contexts can create new opportunities for science learning.

The use of *sInvestigator* as a computational tool improved students' ability to consider multiple dimensions of credibility in evaluating digital scientific information. Yet, the study's findings with respect to author objectivity suggest that students need more explicit instruction in considering this dimension of overall credibility. It is important for students to consider funding sources and the potential political agenda of both authors and publishers (Anelli, 2011) in assessing credibility.

Limitations

There are several limitations to consider in further work on the DSLQ. The number of participants limits the generalizability of this study to a larger student population. Participants included all honors students enrolled in a science class at a large public university; the sample may not be representative of the college population in the United States. All participants in the study had previously taken a course in which they were taught to use peer-referenced journal articles, and our findings indicate that students recognized the importance of this even before using *sInvestigator*. Further research using the DSLQ should examine its use with a student sample that is more representative of the average college student for the purpose of establishing reliability and validity.

The use of different articles on the pre- and post-assessments may have been a confounding factor in this study. Future iterations of the DSLQ could counterbalance the two versions of the DSLQ to guard against order effects.

Conclusions and Implications

Digital scientific literacy is a skill that depends on students' ability to evaluate the credibility of scientific sources within an uncontrolled and ever-changing online environment. It is distinct from digital literacy in science classrooms (e.g., Leu et al., 2015) in that digital literacy is inseparable from the evaluation of scientific content. Students at both the secondary and postsecondary levels need to learn how to evaluate the credibility of sources (Hämäläinen et al., 2021) as a multidimensional critical thinking skill within their development of digital scientific literacy.

By quantifying students' consideration of author and publisher information as predictive of overall credibility, the DSLQ extends the existing literature basis on how students' process the uncertainty and complexity of scientific knowledge in the online information environment (Hendricks et al., 2020). While the DSLQ in its current form is a starting point for quantitative assessment of digital scientific literacy, there is a need for further research on the design of the measure. It is important to consider whether additional dimensions should be incorporated, as well as if further refinement of the dimensions used in the DSLQ is warranted. It is especially important to test and evaluate this measure with students with varied backgrounds in science and literacy.

Future investigations into the development of digital scientific literacy should interrogate the simplistic view that peer-review is a sufficient indicator of quality. Today's students consume digital information from an ever-changing variety of sources on a variety of computing platforms. Students need explicit opportunities to learn to consider the positioning of authors and publishers of scientific information. As non–peer-reviewed sources of information proliferate, students need opportunities to integrate their informal consumption of digital information and their formal use of scientific information in academic settings. The role that these varied sources play in students' emerging understanding of the world cannot be separated from their formal development of scientific literacy.

As the field seeks to equip students to be critical consumers of scientific information in today's age, novel authentic assessments of emerging computational tools are necessary. Use of tools like *sInvestigator* in educational computing can scaffold students' ability to carefully interpret and apply scientific information, while measures like the DSLQ can be used in the development and evaluation of these tools. With the pervasive shift to online distance learning in K-16 education as a result of the global pandemic, building and evaluating both digital literacy and scientific literacy has never been more critical. Although the preponderance of prior research has looked at these literacies in isolation, it is imperative that digital scientific literacy be conceptualized as one synergistic literacy. Digital scientific literacy is a prerequisite for students to productively construct their scientific understandings in online learning. Students who can identify and interpret credible Internet sources will thrive in these changing academic settings and in a digital world where a scientifically literate populace has never been more important. Trefil and Hazen's (2010) goal of preparing an educated citizenry who are capable of engaging in political discourse cannot be realized unless the field builds a more robust empirical basis for how students learn to evaluate digital scientific information.

Acknowledgments

The authors wish to thank our collaborators, Dr. Gheorge Tecuci, Dr. Mihai Boicu, and Dr. Dorin Marcu of George Mason University's Learning Agents Center, as well as research assistants Mariama Bah, Manpriya Dua, and Chirag Uttamsingh.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was partially supported by the National Science Foundation under DUE-IUSE Grant #1611742.

Data Availability

Availability of software and measures the *sInvestigator* software and the DSLQ can be downloaded from the project website: http://lac.gmu.edu/sInvestigator/index.html.

ORCID iDs

Nancy Holincheck https://orcid.org/0000-0001-6999-4072
Terrie M. Galanti https://orcid.org/0000-0003-2948-7898

References

American Association for the Advancement of Science (1989). *Project 2061—Science for all Americans*. AAAS.

Anelli, C. (2011). Scientific literacy: What is it, are we teaching it, and does it matter. *American Entomologist*, *57*(4), 235–244. https://doi.org/10.1093/ae/57.4.235. http://entomology.wsu.edu/wp-content/uploads/2012/02/Anelli2011scientific-lit.pdf.

Bliss, A. C. (2019). Adult science-based learning: The intersection of digital, science, and information literacies. *Adult Learning*, *30*(3), 128–137. https://doi.org/10.1177/1045159519829042.

Boicu, M., Tecuci, G., Marcu, D., Trefil, J., & Holincheck, N. (2016). *Inquiry-Based Teaching and Learning of Science with Cognitive Assistants*, AAAI Fall Symposium Series. https://www.aaai.org/ocs/index.php/FSS/FSS16/paper/download/14116/13678

Bråten, I., Braasch, J. L., Strømsø, H. I., & Ferguson, L. E. (2015). Establishing trustworthiness when students read multiple documents containing conflicting scientific evidence. *Reading Psychology*, *36*(4), 315–349. https://doi.org/10.1080/02702711.2013.864362.

Brem, S. K., Russell, J., & Weems, L. (2001). Science on the web: Student evaluations of scientific arguments. *Discourse Processes*, *32*(2&3), 191–231. https://doi.org/10.1207/s15326950dp3202&3_06. https://www.researchgate.net/profile/Lisa_Weems/publication/228679787_Science_on_the_Web_Student_Evaluations_of_Scientific_Arguments/links/0f317535404d6b1f19000000.pdf.

Britt, M. A., Richter, T., & Rouet, J.-F. (2014). Scientific literacy: The role of goal-directed reading and evaluation in understanding scientific information. *Educational Psychologist*, *49*(2), 104–122. https://doi.org/10.1080/00461520.2014.916217.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Dias da Silva, P., & Heaton, L. (2017). Fostering digital and scientific literacy: Learning through practice. *First Monday*, *22*(6). https://doi.org/10.5210/fm.v22i6.7284.

Eshet, Y. (2004). Digital literacy: A conceptual framework for survival skills in the digital era. *Journal of Educational Multimedia and Hypermedia*, *13*(1), 93–106. https://www.learntechlib.org/primary/p/4793/.

Eshet-Alkalai, Y., & Chajut, E. (2009). Changes over time in digital literacy. *Cyber Psychology & Behavior*, *12*(6), 713–715. https://doi.org/10.1089/cpb.2008.0264.

Fieldhouse, N., & Nicholas, D. (2008). Digital literacy as information savvy: The road to information literacy. In C. Lankshear & M. Knobel (Eds.), *Digital literacies: Concepts, policies, and practices*, Peter Lang Publishing, (pp. 43–72).

Gilster, P. (1997). *Digital literacy*. Wiley Computer Publishing.

Gormally, C., Brickman, P., & Lutz, M. (2012). Developing a test of scientific literacy skills (TOSLS): Measuring undergraduates' evaluation of scientific information and arguments. *CBE-Life Sciences Education*, *11*(4), 364–377. https://doi.org/10.1187/cbe.12-03-0026.

Hahnel, C., Eichmann, B., & Goldhammer, F. (2020). Evaluation of online information in university students: Development and scaling of the screening instrument EVON. *Frontiers in Psychology*, *11* Article, 562128. https://doi.org/10.3389/fpsyg.2020.562128.

Halverson, K. L., Siegel, M. A., & Freyermuth, S. K. (2010). Non-science majors' critical evaluation of websites in a biotechnology course. *Journal of Science Education and Technology*, *19*(6), 612–620. https://doi.org/10.1007/s10956-010-9227-6.

Hämäläinen, E. K., Kiili, C., Marttunen, M., Räikkönen, E., González-Ibáñez, R., & Leppänen, P. H. (2020). Promoting sixth graders' credibility evaluation of web pages: an intervention study. *Computers in Human Behavior*, *110* Article, 106372. http://doi.org/10.1016/j.chb.2020.106372.

Hämäläinen, E. K., Kiili, C., Räikkönen, E., & Marttunen, M. (2021). Students' abilities to evaluate the credibility of online texts: The role of internet-specific epistemic justifications. *Journal of Computer Assisted Learning*, *37*(5), 1409–1422. http://doi.org/10.1111/jcal.12580.

Hendriks, F., Mayweg-Paus, E., Felton, M., Iordanou, K., Jucks, R., & Zimmermann, M. (2020). Constraints and affordances of online engagement with scientific information—A literature review. *Frontiers in Psychology*, *11* Article, 3458. http://doi.org/10.3389/fpsyg.2020.572744.

Hoffman, J. L., Wu, H.-K., Krajcik, J. S., & Soloway, E. (2003). The nature of middle school learners' science content understandings with the use of on-line resources. *Journal of Research in Science Teaching*, *40*(3), 323–346. https://doi.org/10.1002/tea.10079.

Holbrook, J., & Rannikmae, M. (2007). The nature of science education for enhancing scientific literacy. *International Journal of Science Education*, *29*(11), 1347–1362. https://doi.org/10.1080/09500690601007549.

Jin, K. Y., Reichert, F., Cagasan, L. P. Jr, de la Torre, J., & Law, N. (2020). Measuring digital literacy across three age cohorts: Exploring test dimensionality and performance differences. *Computers & Education*, *157*, Article 103968. http://doi.org/10.1016/j.compedu.2020.103968.

Johnson, R. B., & Christensen, L. (2019). *Educational research: Quantitative, qualitative, and mixed approaches*. Sage Publications.

Kammerer, Y., Gottschling, S., & Bråten, I. (2020). The role of internet-specific justification beliefs in source evaluation and corroboration during web search on an unsettled socio-scientific issue. *Journal of Educational Computing Research*, *59*(2), 073563312095273. https://doi.org/10.1177/0735633120952731.

Kiili, C., Forzani, E., Brante, E. W., Räikkönen, E., & Marttunen, M. (2021). Sourcing on the internet: Examining the relations among different phases of online inquiry. *Computers and Education Open*, *2*, Article 100037. http://doi.org/10.1016/j.caeo.2021.100037.

Leu, D. J., Forzani, E., Rhoads, C., Maykel, C., Kennedy, C., & Timbrell, N. (2015). The new literacies of online research and comprehension: Rethinking the reading achievement gap. *Reading Research Quarterly*, *50*(1), 37–59. http://doi.org/10.1002/rrq.85.

Macedo-Rouet, M., Potocki, A., Scharrer, L., Ros, C., Stadtler, M., Salmerón, L., & Rouet, J. F. (2019). How good is this page? Benefits and limits of prompting on adolescents' evaluation of web information quality. *Reading Research Quarterly*, *54*(3), 299–321. http://doi.org/10.1002/rrq.241.

National Research Council (2012). *A framework for K-12 science education: Practices, cross-cutting concepts, and core ideas*. National Academies Press. https://doi.org/10.17226/13165.

Neuendorf, K. A. (2016). *The content analysis guidebook*. Sage Publications.

NGSS Lead States (2013). *Next Generation Science Standards: For states, by states*. National Academies Press.

Norris, S. P., Phillips, L. M., & Burns, D. P. (2014). Conceptions of scientific literacy: Identifying and evaluating their programmatic elements, *International handbook of research in history, philosophy and science teaching* (pp. 1317–1344). Springer. https://doi.org/10.1007/978-94-007-7654-8_40.

Opitz, A., Heene, M., & Fischer, F. (2017). Measuring scientific reasoning–A review of test instruments. *Educational Research and Evaluation*, *23*(3–4), 78–101. http://doi.org/10.1080/13803611.2017.1338586.

Porat, E., Blau, I., & Barak, A. (2018). Measuring digital literacies: Junior high-school students' perceived competencies versus actual performance. *Computers & Education*, *126*, 23–36. http://doi.org/10.1016/j.compedu.2018.06.030.

Prensky, M. (2001). Digital natives, digital immigrants. *On the Horizon*, *9*(5), 1–6. https://doi.org/10.1108/10748120110424816. https://www.marcprensky.com/writing/Prensky%20-%20Digital%20Natives,%20Digital%20Immigrants%20-%20Part1.pdf.

Scharrer, L., Stadtler, M., & Bromme, R. (2019). Judging scientific information: Does source evaluation prevent the seductive effect of text easiness?. *Learning and Instruction*, *63*, Article 101215. http://doi.org/10.1016/j.learninstruc.2019.101215.

Silber-Varod, V., Eshet-Alkalai, Y., & Geri, N. (2019). Tracing research trends of 21st-century learning skills. *British Journal of Educational Technology*, *50*(6), 3099–3118. http://doi.org/10.1111/bjet.12753.

Tecuci, G., Marcu, D., Boicu, M., & Schum, D. A. (2015). *COGENT: Cognitive Agent for Cogent AnalysisAAAI Fall Symposium Series*. AAAI. https://www.aaai.org/ocs/index.php/FSS/FSS15/paper/download/11672/11478.

Trefil, J. (2008). Science education for everyone: Why and what? *Liberal Education*, *94*(2), 6–11. https://eric.ed.gov/?id=EJ822735.

Trefil, J., & Hazen, R. M. (2010). Scientific literacy: A modest proposal. In J. Meinwald, & J.G. Hildebrand (Eds.), *Science and the educated American: A core component of liberal education* (pp. 57–69). American Academy of Arts and Sciences.

Van Laar, E., Van Deursen, A. J., Van Dijk, J. A., & De Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior*, *72*, 577–588. http://doi.org/10.1016/j.chb.2017.03.010.

Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Source evaluation, comprehension, and learning in Internet science inquiry tasks. *American Educational Research Journal*, *46*(4), 1060–1106. http://doi.org/10.3102/0002831209333183.

Wineburg, S., & McGrew, S. (2019). Lateral reading and the nature of expertise: Reading less and learning more when evaluating digital information. *Teachers College Record*, *121*(11), n11. http://doi.org/10.1177/016146811912101102.