



UvA-DARE (Digital Academic Repository)

Students evaluating the trustworthiness of historical sources and internet sources: A comparison

van der Eem, M.; van Drie, J.P.; Brand-Gruwel, Saskia; van Boxtel, C.A.M.

DOI

[10.1002/berj.4095](https://doi.org/10.1002/berj.4095)

Publication date

2025

Document Version

Final published version

Published in

British Educational Research Journal

License

CC BY

[Link to publication](#)

Citation for published version (APA):

van der Eem, M., van Drie, J. P., Brand-Gruwel, S., & van Boxtel, C. A. M. (2025). Students evaluating the trustworthiness of historical sources and internet sources: A comparison. *British Educational Research Journal*, 51(2), 665-686. <https://doi.org/10.1002/berj.4095>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)

Students evaluating the trustworthiness of historical sources and internet sources: A comparison

Maartje van der Eem¹ | Jannet van Drie¹ | Saskia Brand-Gruwel^{2,3} | Carla van Boxtel¹

¹Research Institute of Child Development and Education, University of Amsterdam, Amsterdam, The Netherlands

²Zuyd University of Applied Sciences, Heerlen, The Netherlands

³Maastricht University, Maastricht, The Netherlands

Correspondence

Maartje van der Eem, Research Institute of Child Development and Education, University of Amsterdam, Postbus 15776, Amsterdam 1001 NG, The Netherlands.
Email: m.vandereem@uva.nl

Funding information

Dudoc Alfa

Abstract

Fake news and disinformation are easily spread in today's digital society. Therefore, it is important that students learn how to evaluate the trustworthiness of online information, but this skill is often confined to a limited number of subjects in secondary education. History classes can potentially contribute to developing this skill. Therefore, the first aim of this study was to investigate whether there is a relationship between how Grade 9 students ($N=112$) perform while evaluating the trustworthiness of historical sources and internet sources. Average student performance on both kinds of sources was similar, but the distribution of the scores was significantly different. There was a moderate correlation between students' total scores on the historical task and the internet task. Two of the three criteria of trustworthiness that could be used on both kinds of sources were used by a great majority in both tasks. The second aim was to gain more insight into students' ($N=8$) and teachers' perspectives ($N=8$) on the usefulness of the evaluation skill learned in history class for other contexts, especially when searching on the internet. While most of the teachers mentioned the importance of the skill when using the internet, none of the students did so spontaneously. We suggest that history classrooms are an appropriate place to teach students not only about historical sources but also about internet sources,

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Author(s). *British Educational Research Journal* published by John Wiley & Sons Ltd on behalf of British Educational Research Association.

provided that more explicit attention is given to the relationship between both types of sources.

KEYWORDS

history education, internet sources, secondary education, trustworthiness

Key insights

What is the main issue the paper addresses?

Within history education, students learn to evaluate the trustworthiness of historical sources. Is there a relationship between how students evaluate historical sources and internet sources? How do students and teachers perceive the usefulness of the skill learned in history for other contexts?

What are the main insights that the paper addresses?

Students should realise that the historical evaluation skill is also valuable when reading online information. While most students used overlapping criteria when evaluating both types of sources, the interviewed students do not explicitly make the connection. Teachers do recognise this connection but could communicate that more clearly to their students.

INTRODUCTION

Since students rely heavily on the internet to find information, for school and for personal purposes, it is crucial that they know how to distinguish trustworthy information from false information (e.g. McGrew et al., 2019). The formal curriculum in secondary education often does not include a separate course or subject in which students learn how to evaluate the trustworthiness of online information (McGrew, 2020; Meneses, 2021; Nygren & Guath, 2019). Therefore, students are usually taught this skill in existing school subjects, frequently in history classrooms (McGrew, 2022). When education specialists address how to teach this evaluation skill in history classrooms, there are often two assumptions. First, there are common characteristics among the uses of this skill on different kinds of sources, for example, the criteria of trustworthiness that must be used when evaluating historical sources and internet sources. Second, positive transfer will occur: when students have learned to evaluate the trustworthiness of historical sources, this will enhance their performance when working with other sources (Perkins & Salomon, 1992). Thus, the assumption is that when students have learned to evaluate the trustworthiness of sources in history class, they will be able to use this skill in situations outside the classroom as well, for example, when searching on the internet for information about a socio-scientific issue. However, this requires that students realise that they can also use the skill they have learned in the history classroom when reading other sources than historical ones.

Barnett and Ceci (2002) provided a taxonomy in which they distinguished several dimensions of transfer. Looking at the dimension of content, transfer could be considered near, since students must ask questions such as who is behind this information, what is the goal of the maker, what is the evidence, what do other sources say when evaluating both historical and internet sources (e.g. Breakstone et al., 2018; Reisman, 2012). However, regarding other dimensions of the taxonomy, applying the skill on the two kinds of sources can be seen as far transfer. Students must apply the skill in different knowledge domains and in other physical contexts (history classroom vs. other classrooms or at home). Also, students are used to evaluating historical sources when answering textbook or test questions, while evaluating internet sources is usually not done as a classroom exercise but students are expected to do that when, for example, searching for information for an essay or for personal use outside school. Furthermore, there are differences in the modality dimension. When working with historical sources in the classroom, students usually work with printed sources in a fixed format while online sources do not have a fixed format.

A considerable number of studies have been published on how students in secondary education spontaneously (thus, without an intervention) evaluate the trustworthiness of historical sources (e.g. Britt & Aglinskias, 2002; Jacobsen et al., 2018) and of internet sources (e.g. Breakstone et al., 2021; Kiili et al., 2018). These studies, however, only focus on either historical or internet sources. To date, no explorative research has been conducted in which the same group of students evaluates the trustworthiness of both historical and internet sources. In a previous study, we described how Grade 9 students performed when applying the evaluation skill on historical sources (Van der Eem et al., 2023). The same sample of students also performed a task in which they evaluated the trustworthiness of internet sources. The first aim of the present study, therefore, is to establish whether there is a relationship between student evaluation of historical sources and their evaluation of internet sources. Transfer of what students have learned is often difficult (e.g. Hajian, 2019); therefore, it is interesting to investigate whether students spontaneously use similar criteria when evaluating both historical and internet sources. This information can be used when teaching students about the evaluation skill in history education.

The second aim of this study is to gain insight into students' and history teachers' perceptions of the usefulness of the historical evaluation skill outside history classrooms. So far, studies have focused on *how* students apply this historical skill, but not on how useful students consider this skill for purposes other than in history classrooms. Regarding teachers' views, two studies indicated that Dutch history teachers (Sakki & Pirttilä-Backman, 2019) and Dutch prospective history teachers (Wansink et al., 2017) consider teaching students to be critical thinkers to be an important part of their teaching. Learning to evaluate the trustworthiness of sources is seen as an important part of this critical thinking ability. However, these studies did not elaborate on whether teachers thought that this skill could also be used outside their history classrooms.

In this study, we use historical sources and internet sources. Historical sources can, of course, also be found online. It is, therefore, important to define how we distinguish historical and internet sources for the purpose of this study. Within the field of history, the concept of sources has been described in different ways, using terms such as traces, records and accounts to differentiate between different kinds of sources and between primary and secondary sources (Körber, 2016; Seixas, 2016; Van Drie & Van Boxtel, 2008). Studies focusing on internet sources have used the terms printed sources vs. digital sources (McGrew et al., 2019; Scholes et al., 2023) and traditional sources vs. online sources (Gasser et al., 2012). In this study, we use the term 'historical sources' to refer to sources that (1) are concerned with historical events and (2) were not intended to be published online (even though these sources might have been placed on the internet in later times). These are sources that are commonly used in history education to develop students' historical

reasoning competences and contain primary and secondary sources. We define internet sources as sources that are (1) concerned with contemporary issues or events and (2) meant to be published online. These sources are often used by students when searching for information on the internet when working on a presentation or essay, or for personal use not related to school assignments.

THEORETICAL BACKGROUND: EVALUATING THE TRUSTWORTHINESS OF SOURCES

Conceptualisation of the skill related to historical sources

When students evaluate the trustworthiness of a historical source, they must explain whether this source provides evidence to answer a question about the past; a source in itself is not (un)trustworthy but that depends on the question asked (De La Paz et al., 2012; Nokes, 2010; Reisman, 2011; Van Drie & Van Boxtel, 2008). To accomplish this, Wineburg (1991) proposed three heuristics. The heuristics of contextualisation and corroboration can be made concrete by describing what they entail. When students contextualise, they analyse the source in relation to the historical context in which it was made and corroboration means comparing the content of the source with other sources (Hynd et al., 2004; Reisman, 2012; Seixas & Morton, 2013; Wineburg, 1991).

The third heuristic, sourcing, consists of several specific criteria. These criteria require students to attend to the information about the source. In this study, we used the following criteria, based on earlier studies (Britt & Aglinskis, 2002; De La Paz & Felton, 2010; Nokes, 2017; Reisman, 2012; Seixas & Morton, 2013; Sendur et al., 2021): (1) what is the background of the creator; (2) what is the goal of the creator; (3) how did the creator come to know about the event described in the source (evidence); and (4) what was the date the creator made the source (see also: Van der Eem et al., 2023)?

Conceptualisation of the skill related to internet sources

When Wineburg (1991) formulated the three heuristics, the internet was not yet available to the general public. However, since then, the internet has become the primary place to find information and also the place where anyone can share information (Brand-Gruwel et al., 2009). Until recently, most studies on evaluating the trustworthiness of internet sources used the heuristics for historical sources and their corresponding criteria. For example, according to Metzger (2007), the evaluation skill did not have to change after the arrival of the internet; only the mindset of people who use information from the internet had to change. They needed to understand how and when to use this skill.

The idea that the evaluation skill for the discipline of history is related to the evaluation skill for internet sources is reflected in several studies on the trustworthiness of internet sources. In these studies, the authors refer to the three historical heuristics. In the study by Anmarkrud et al. (2014), for example, the heuristic of sourcing is mentioned, as well as the context of the source and the need to evaluate sources in relation to each other. Walraven et al. (2009) also mentioned criteria related to sourcing and corroboration (information is the same or similar across more sites). The criteria for sourcing mentioned in the studies on internet sources are similar to those used in historical studies: who is the author; what are the author's sources; and what are the intentions of the author (Hargittai et al., 2010; Pérez et al., 2018; Potocki et al., 2020; Strømsø et al., 2013)? The exception is the criterion of time, which is quite specific to historical sources.

More recently, it has been argued that evaluating the trustworthiness of internet sources has specific characteristics (Nygren & Guath, 2019) and, thus, requires its own approach (Breakstone et al., 2018). This makes sense, since although there are similarities between internet sources and the historical sources used in the educational setting in classrooms, there are also differences. First, the presentation and layout of the two types of sources are different. When students work with historical sources in the classroom, the sources are usually presented as an isolated text, photo or drawing. An internet source, on the other hand, is more holistic: the text is surrounded by pictures, links to other parts of the website or to other websites, other articles and the hypertext system (Brante, 2019; List et al., 2017; McGrew, 2022).

A second difference is that historical sources used in the classroom are presented to the students with the information they need: above or below the source, there is information about the author, the date the source was produced, where the source was published, etc. When looking for this kind of information on the internet, a student must actively search on the website itself or on other websites (Brante, 2019; McGrew, 2022; McGrew et al., 2019; Metzger, 2007).

Therefore, Breakstone et al. (2018) and McGrew et al. (2018) describe an alternative approach to teaching students how to evaluate the trustworthiness of internet information. Instead of the traditional approach of searching for information on the website itself to answer these questions, students should be taught to look on other websites to find out what these websites have to say about the website under research. Wineburg and McGrew (2017) introduced the term 'lateral reading' for this approach. This approach was applied by all fact checkers in the study of Wineburg and McGrew (2019). In comparison with most of the professional historians in this study, the fact checkers searched for the answer to the question 'who is behind the information' on other websites than the website under study. The fact checkers and the historians used the same criteria, such as who is behind the information, but where they searched for the answers differed.

Evaluating the trustworthiness of sources: Struggling students

Multiple studies have been conducted using historical sources or internet sources, describing how students in secondary education perform when evaluating the trustworthiness of these sources. Over the years, studies have found that students rarely spontaneously apply correct criteria of trustworthiness, such as who made the source and why was the source made, when reading historical sources (Britt & Aglinskas, 2002; Halvorsen et al., 2016; Jacobsen et al., 2018; Wineburg, 1991) or when reading internet sources (Breakstone et al., 2021; Kammerer et al., 2016; Kiili et al., 2018; McGrew et al., 2018; Nygren & Guath, 2019; Paul et al., 2017; Pieschl & Sivyver, 2021; Walraven et al., 2009). The study of Horn and Veermans (2019) appears to be an exception. However, these well performing students in this study were not a representative group, since the participants were the best students in their classes.

Several explanations for low student performance on these kinds of tasks have been suggested. First, students often use incorrect criteria to support their claim on the trustworthiness of a source. McGrew (2021) referred to this as using 'weak heuristics'. For example, students considered a source to be trustworthy because it provided an answer to the question the student had to answer. Both studies on historical sources (Harris et al., 2016; Jacobsen et al., 2018) and internet sources (Breakstone et al., 2021; Kiili et al., 2019; Walraven et al., 2009) found this to be the case. The use of these weak heuristics might be related to the age of the students. It is not an easy skill to apply, which is why, according to Anmarkrud et al. (2022), research on this skill has often been conducted among university

students. It might, thus, be expected that younger students will have even more difficulties when applying the skill.

Coiro et al. (2015) suggested a second explanation for poor student performance. In their study on evaluating the trustworthiness of internet sources, they found that some students demonstrated an understanding of correct criteria but were unable to clarify their reasoning why that criterion made the source more or less trustworthy.

A third explanation is related to motivation to apply the skill. Paul et al. (2017) researched why students refrained from sourcing when working with internet sources. They concluded that (a lack of) motivation played an important part in several ways: sourcing was not expected or needed to receive a grade, most assignments did not challenge students to find the truth themselves, and students were more focused on the content and therefore considered investigating the trustworthiness of the source to be too time-consuming.

The described studies only researched applying the evaluation skill on one of the two kinds of sources we focus on in this article, so it is unknown whether there is a relationship between how students perform when asked to evaluate historical sources and how they perform when evaluating internet sources. It is also unknown if students are aware that the skill learned in history lessons can be applied outside these history classrooms and how teachers perceive this.

RESEARCH QUESTIONS

Previous research has thus shown that students struggle when they evaluate the trustworthiness of both historical and internet sources. However, little is known about the relation between the performances of students on the evaluation of historical and internet sources. Furthermore, little is known about teachers' and students' perceptions of applying the historical skill in other contexts outside history classrooms. Therefore, this research seeks to address the following questions:

1. Is there a relationship between student performance when evaluating the trustworthiness of historical sources and their performance when applying the evaluation skill to internet sources?
2. Is there a relationship between the criteria students use when evaluating the trustworthiness of historical and internet sources?
3. How do students and teachers perceive the usefulness of applying the skill used in history classrooms to other contexts?

METHODS

Participants

In the quantitative part of this research (research questions 1 and 2), Grade 9 students from six different schools in the Netherlands participated ($N=156$). Only students who had parental consent and who were present during both lessons used for data collection were included in the sample, which ultimately consisted of 112 students (mean, M per school 18.67, standard deviation, SD 2.50). The average age of the students was 14.46 years (SD 0.59); 54.5% of the students identified themselves as girls, and 55.5% as boys. All students followed their lessons in the educational track that prepares students for the university of applied sciences. In one of the six participating schools, the students had their own laptop which they used in the lessons. In the other schools, there were a few classrooms with

computers. One of the history teachers used a computer room on a regular basis; the other four teachers did not.

For the qualitative part of this research (research question 3), eight students out of the sample of 112 were selected by their history teachers, based on students' grades (good, average and poor), and they were from three of the six participating schools. Also, eight history teachers (from eight schools), two females and six males, were interviewed. The teachers' ages ranged from 25 to 66 years (M 43.13; SD 15.29; median 43.0), and their teaching experience ranged from 1 to 40 years (M 17.88; SD 14.28; median 17.50). All teachers had experience teaching history in Grade 9. Three of the teachers had also participated with their class in the quantitative part of the research.

Context

In the Dutch national curriculum guidelines for Grades 7–9 (lower secondary education), three of the 58 learning objectives are related to evaluating the trustworthiness of sources. For the subjects Dutch and English, the same objective is formulated: students should be able to 'determine the value of a written or a digital source for themselves and for others'. The third learning objective related to trustworthiness is part of the subject of history: students should be able to 'use historical sources to create an understanding of an era or to find answers to a question'. Thus, there are references to both the internet and historical sources in the educational objectives, but the time spent on teaching students these objectives differs per school and even, within one school, per teacher, since there are no standardised tests at the end of lower secondary education.

Instruments

Evaluation of trustworthiness: Tasks

To determine student performance in evaluating the trustworthiness of historical and internet sources, two tasks were developed: one with historical sources and one with internet sources. Each task contained 15 questions asking students to evaluate the trustworthiness of sources. We used open-ended questions following the format and the phrasing of the Historical Assessments of Thinking (Breakstone, 2014). All questions required the student to read a source and then answer a question about its trustworthiness. In Appendices A and B, a question from each task can be found.

In the historical task, the topics of the (primary and secondary) sources were the First and the Second World Wars. These historical events had been taught in all participating schools. In the internet task, sources about vaccination for childhood diseases¹ were used. Factual information about vaccination is part of the biology curriculum. We also expected that all students would have had personal experience with this topic, since young teenagers in the Netherlands are invited to be vaccinated against the HPV-infection. Table 1 shows the kind of sources that were used in the tasks and the phrasing of the questions, which was similar in both tasks.

Both tasks had three sourcing criteria in common: (1) the background of the creator; (2) the goal of the creator; and (3) the evidence used to create the source. Each of these criteria could be used as a correct criterion several times. These criteria are suited for students in lower secondary education, who can be considered novices when it comes to the evaluation skill and are in line with the criteria mentioned in students' history textbooks.

Since not all schools could provide access to computers, both sets of tasks were printed on paper. In the historical task, students first read information about the source

TABLE 1 Overview of the tasks.

Historical sources	Internet sources	Number of questions	Questions
Eyewitness of the Battle of Stalingrad (1942–1943) in an interview in 2003	The website from a Dutch women's magazine about a couple whose baby died after vaccination	1	This source helps us to understand (...). I agree/I do not agree/I both agree and disagree. Explain your answer
Fragment from the book <i>All Quiet on the Western Front</i> (1929) about the trenches in the First World War	Three parents telling their story on a website called vaccinefreefoundation.nl	3	<i>The students are given five additional facts about the source.</i> (a) Fact ... might cause me to question the trustworthiness of the source. (b) Fact ... might cause me to question the trustworthiness of the source. (c) Fact ... might strengthen the trustworthiness of the source. Explain your answers
Fragment from a documentary (1995) about the battle in the Java Sea (1942)	Fragment from the website infoNow.nl , about the possible relationship between vaccination and autism	4	(a+b) Give two reasons why this source might be useful as evidence for your research on (...). (c+d) Give two reasons why this source might be less useful as evidence for your research on (...)
Fragment from a speech from Hitler in 1934 about the Night of the Long Knives followed by descriptions of three other sources about the same event (a diary from 1934, a newspaper from 1934, and a book written by a historian in 2008)	Fragment about vaccination from the website from a national Dutch news channel followed by three websites about vaccination (one from a website with conspiracy theories, one from the Dutch governmental health service, and one from the Dutch critical vaccination association)	4	(a) explain why the first source probably does not provide evidence for your research on (...). (b + c + d) Answer for each of the following three source descriptions/websites: this source provides me trustworthy information/does not provide me trustworthy information. Explain your answers
A picture of Rotterdam made by a German soldier during the Second World War	A graphic from the website Children's Health Defence	3	(a) Explain why the background information about the source might lead you to question the source's trustworthiness. (b) Which feature in the source might lead you to question the source's trustworthiness? (c) What is one thing about the source or the maker of the source you would like to know to better determine the trustworthiness and explain why that would help you

and then read the source itself. In the internet task, the URL of the website was shown first, and then a screenshot of the website was shown. That way, students had the opportunity to look for other clues about the trustworthiness of the website other than just the information that was given about vaccination. For example, in one of the screenshots, if students paid attention to the whole screenshot, they could see that there was also the possibility of reporting seeing a UFO on this website. Finally, information about the author and website was presented to the students, and if necessary, text from the website was provided.

Students' and teachers' perceptions about the usefulness of the historical skill: Interviews

Another focus of this study was whether students spontaneously referred to the transfer of the historical skill to other settings. After completing the historical task, eight students (from three of the schools) were asked to read statements about the importance and usefulness of learning the historical evaluation skill aloud and explain their thoughts. These statements are part of the Task Value Questionnaire (Pintrich et al., 1991). Two statements were related to importance: (1) it is important for me to learn this historical skill; and (2) it is important for me to understand this historical skill. Two other statements were concerned with usefulness: (1) I can use this historical skill in other subjects; and (2) I think it is useful for me to learn this historical skill.

To gain further insight into teachers' thoughts about the usefulness of the historical skill in other settings, we asked the eight teachers two questions: (1) do you think that this historical skill is also useful outside the setting of the history classroom; and if yes, (2) how do you show that to your students?

Procedure

Each participating class was visited twice. During the first visit, the students performed the historical task in a regular history lesson supervised by their own history teacher. Two weeks later, the students performed the internet task. This was done during a weekly scheduled lesson in which students learn academic and social skills from one of their teachers. In all classrooms, the students were given the same instructions, read aloud by the teacher. The students were told by that they were given a task with historical/internet sources. The students were asked to explain their answers as well as they could and wrote their answers on prepared worksheets. Both times, the students had a maximum of 40 min to accomplish the task; all students were able to complete the task within this time period.

We used three different sequences of each task, so that students' answers would not be influenced by the formulation of the different questions. In each class, there was an even distribution of the three versions. Kruskal–Wallis tests indicated that the order in which the tasks were presented did not significantly influence students' performance on either the historical task $H(2)=0.544$, $p=0.762$ or the internet task $H(2)=3.032$, $p=0.220$.

To answer the third research question, eight students were recorded during an individual think aloud-session, in the presence of the first author or a research assistant. Their answers were transcribed. Another research assistant interviewed the eight teachers via Microsoft Teams. Their answers were also recorded and transcribed.

Data analysis

Student performance

Students' answers on the tasks were scored using a rubric (see Appendices A and B). We used the same scoring principles as in our previous study, in which we analysed students' answers on only the historical task, to find out whether procedural knowledge and task value were predictors of student performance (Van der Eem et al., 2023). Following this rubric, each answer received a score between zero and two. Zero points were given when the answer was wrong or when the student did not answer at all. One point was given to answers that were partly correct. For example, a score of one would be given when the student had written down a correct criterion but failed to give a (good) explanation. A maximum of two points was awarded when a student (1) used the correct criterion and (2) elaborated his or her reason correctly, explaining why that specific criterion made the source trustworthy or not. On the historical task, for instance, when students wrote that the fact that the book *All Quiet on the Western Front* was written in 1928 might cause them to question the trustworthiness of the source because 'it was written a long time after the war', one point was awarded. These students used a correct criterion (when was the source made) but failed to explain why writing a source about an event years later might make the source less trustworthy. The student who wrote 'it was written a while after the war; therefore, he might have forgotten small but important details' received two points because the second part of the answer elaborated on the influence of the criterion of time.

The focus was on whether students were able to use a correct criterion, such as the time in which the source was made, and whether they were able to explain why that specific criterion made the source (un)trustworthy. The participating students were in Grade 9 and can be described as novices when it comes to applying the evaluation skill. Experts, such as historians and media scholars, not only will carefully consider one criterion but will also include other criteria such as the background of the maker of the source, and then decide about the trustworthiness of a source. In the case of the source from the Remarque novel, the students were told that the writer had been in the trenches as a soldier. The students had to reason that Remarque was an eye-witness writing from memory 11 years later. A source made (many) years after the events described is not necessarily untrustworthy. In one of the other history questions, the students read information about a secondary source, written in 2008 about an event in 1934. The information about the source told the students that this was written by a historian who had used multiple primary sources. So, in this case, the criterion of time had to be evaluated in another way than in the Remarque question.

Each task consisted of 15 questions, resulting in a maximum of 30 points per task set. We calculated a total score for each student. To determine the interrater reliability, the first author and a research assistant independently scored 450 answers from 30 students (five per school; 26.79% of the total number of answers per task). For the scores on the history task, Cohen's κ was 0.813 (89.6% agreement); for the scores on the internet task, the κ was 0.791 (88.4% agreement). These kappas can be considered good values. Therefore, after resolving the differences, the first author scored the remaining answers on the history task, and the research assistant scored the remaining answers on the internet task.

Finally, Cronbach's α was calculated based upon the answers of each task to determine the internal consistency of the task constructs. Cronbach's α on the historical task was 0.733, and on the internet task it was 0.752; this means that we can consider both tasks sufficiently internally consistent to indicate student performance on evaluating the trustworthiness of sources.

Use of criteria

For each answer, we analysed whether the student had used a correct criterion of trustworthiness and, if so, which criterion. It is possible that a student used a correct criterion incorrectly. For example, one student wrote that a source about the Second World War was made in 2008 (thus recognising that time is one of the criteria of trustworthiness) but did not notice that the source was made by a historian using primary sources. According to this student, the source was not trustworthy because it was published too long after the war. In these cases, this was counted as a correct criterion, since the student showed that he or she did know the criterion but did not know how to apply it well. Again, the first author and the research assistant independently analysed the criteria used in 450 students' answers per task. Cohen's κ on the history task was 0.839 (86.0% agreement), and on the internet task, it was 0.843 (86.9% agreement). After resolving the differences, the remaining answers on the historical task were categorised by the first author and those on the internet task by the research assistant.

Students' and teachers' perceptions of usefulness

Based on the transcriptions of the interviews with the students, we investigated whether they spontaneously referred to using the skill outside their history classrooms. In the transcriptions of the interviews with the teachers, we looked for indications that the teachers thought that this skill was also useful in other contexts and, if they did, whether and how they explained or showed this to their students.

Statistical analysis

First, descriptive data were generated for all variables of the quantitative part of the research. Paired *t*-tests were used to compare students' performance on both tasks and their use of correct criteria. Two-tailed Pearson's correlation analysis was run to determine the relationship between student performance on the tasks and the use of criteria in students' answers.

RESULTS

Student performance

The first research question concerned the relationship between student performance when evaluating the trustworthiness of historical sources and their performance when applying that skill to internet sources. [Table 2](#) presents the descriptive statistics on the task scores in the two tasks.

TABLE 2 Descriptive statistics: task scores (maximum 30 points).

	Historical task	Internet task
Mean	8.55 (SD 4.856)	8.38 (SD 3.902)
Mode	5.00	5.00
Median	8.00	8.00
Range	0–24	1–18

Table 2 reveals that the students did not perform well in both tasks. Except for the range, student performances on both tasks are similar. A paired *t*-test confirmed that the average scores of the students on the historical task and on the internet task were not significantly different: $t = 0.384$, $d.f. = 111$, $p = 0.702$. The correlation between students' scores on the historical task and the internet task is 0.386 ($p = 0.000$), which can be considered a moderate correlation (Cohen, 1988).

To see whether the distributions of the scores are also similar, the number of times a student's answer was awarded a score of 0, 1 or 2 points was counted and compared through a paired *t*-test. The results are shown in Table 3.

Where there is no significant difference in the average scores of the two tasks, the distribution of the scores differs significantly when the awarded scores per answer are compared. The maximum score of 2 points and the minimum score of 0 points were awarded more frequently to the answers on the historical task than on the internet task, while on the internet task a score of 1 point was more frequently awarded than on the history task.

Use of correct criteria

The second research question asked whether there is a relationship between the criteria students use when evaluating the trustworthiness of historical and internet sources. Table 4 presents the descriptive statistics of students' use of correct criteria in their answers.

Table 4 shows that students used a correct criterion of trustworthiness in just over half of the questions in each task. All students had at least one answer in which a good criterion of trustworthiness was used. A paired *t*-test revealed that the average numbers of correct criteria in students' answers on the historical task and on the internet task were not significantly different: $t = -1.317$, $d.f. = 111$, $p = 0.190$. There is a strong correlation between the numbers of students' answers in which a correct criterion was used in each task ($r = 0.544$, $p < 0.001$).

Three criteria for trustworthiness could be applied in both tasks: (1) the background of the creator; (2) the goal of the creator; and (3) how the creator came to know about the event described in the source (the evidence). In both tasks, these criteria could be used to support students' answers several times. Table 5 shows how many students had used these criteria at least once, and if so, in which of the tasks the criterion was used.

TABLE 3 Average number of answers (per student) awarded 2–1–0 points (paired *t*-test).

	Historical task		Internet task	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
2 points	2.44***	2.20	1.16	1.46
1 point	3.68***	1.86	6.05	2.57
0 points	8.88***	2.94	7.79	2.96

*** $p < 0.001$.

TABLE 4 Descriptive statistics: number of answers in which a correct criterion was used (maximum 15).

	Historical task	Internet task
Mean	7.62 (SD 3.401)	7.98 (SD 3.119)
Mode	6.00	9.00
Median	8.00	8.00
Range	1–15	2–14

TABLE 5 Overview of the use of overlapping criteria (in percentages).

	Used in both tasks	Used only in the historical task	Used only in the internet task	Not used
Background creator	86.6	1.8	10.7	0.9
Goal creator	57.1	8.9	25.0	8.9
Evidence used	86.6	3.6	9.8	0

The results in Table 5 show that almost all students used these criteria at least once in their answers. The use of the criteria about the background of the creator and the evidence used by the creator were applied similarly by the students. The criterion of the goal of the creator shows a different pattern. Not only was it used less by the students in comparison with the other two criteria, but also only a small majority of the students used this criterion in both tasks.

Students' perceptions of the usefulness and importance of the historical skill

When students learn how to evaluate the trustworthiness of sources in history class, it is hoped that they will also use this skill in other contexts, for example, when searching for information on the internet. To gain more insight into how students perceive the usefulness and importance of the historical evaluation skill, eight students responded to statements regarding the usefulness and importance of learning to evaluate the trustworthiness of historical sources. There are no significant differences between these eight students and the other students in their average performances on the history task ($t = 1.40$, $d.f. = 110$, $p = 0.165$) and on the internet task ($t = -0.751$, $d.f. = 110$, $p = 0.454$).

Two statements that the students responded to were concerned with usefulness: (1) I can use this historical skill in other subjects; and (2) I think it is useful for me to learn this historical skill. None of the students spontaneously referred to using the skill when searching for information on the internet. On the first statement, four students referred to other school subjects in which they could use the skill, for example in the languages. On the second statement, five of them only referred to it being useful in school, e.g. for tests and final exams. One student had no idea where he could use this skill. The answers of the other two students were phrased in a more general way, for example, saying that it is useful to learn to decide whether something is real or fake.

The students were also asked to respond to two statements about the importance of (1) learning and (2) understanding the historical skill of evaluating the trustworthiness of sources. Again, none of the students referred to information on the internet in their answers. Some students showed an understanding that they could use this skill outside (or after) school, saying, 'It is important to understand because if you apply it incorrectly, you will apply it incorrectly over your whole life'. Other students only referred to the school context, for example, by saying, 'We have to take a lot of history tests, about the World War, and there are sources included in the tests'.

Teachers' perceptions of the usefulness of the historical skill

When asked whether they thought that the historical skill could also be used outside their own classroom, seven of the eight teachers wholeheartedly said yes. Six of them explicitly referred to the internet, where students would have to use this skill as well. Two of them also

referred to citizenship education. One teacher hesitated. He first said that he considered it to be a specific skill to history. However, after some thought, he said that it might also be useful for students to be aware of trustworthiness when searching online.

Although the teachers agreed that this skill is useful outside their classrooms, most of them were more hesitant when asked how they showed the usefulness of the skill to their students. Only three teachers described concrete situations in their lessons, with two of them referring to fake news. These two teachers showed their students examples of fake news that they might recognise, for example, related to Donald Trump. The third teacher used a (historical) documentary to show his students that an opinion is not automatically the truth.

Two teachers were hesitant, saying that they tried to pay attention to explaining why this skill is important, but did not refer to concrete examples of how they tried to do that. One of them said, 'I should make the connection between past and present more often. I sometimes try to do that. I have noticed that the question "how would this be of use in your life" helps'.

The other three teachers did not discuss with their students why this historical skill might also be useful outside their history classrooms. One of these teachers said that he lacked the time owing to a full history curriculum. The second teacher said that this skill was also used in other subjects, but mostly in higher secondary school, not in Grade 9. Finally, the last teacher simply said, 'I have never done that'.

DISCUSSION

This study examined the relationship between the way students evaluated the trustworthiness of historical sources and of internet sources, and how students and teachers perceive the usefulness of the historical skill outside the history classroom.

Students' average score on both tasks was similar (history task, 8.55; internet task, 8.38; maximum score, 30) and we found a moderate correlation between student performance on both tasks. However, the distribution of the awarded scores differed significantly. On the historical task, more answers received the maximum score of two points, whereas more answers on the internet task received one point. A possible explanation for this finding might be that students have learned how to evaluate the trustworthiness of sources in history class, and have received feedback on their answers, for example on tests. Their ability to evaluate the trustworthiness of internet sources is probably not tested in summative tests since this application of the skill is usually not part of the formal curriculum (Meneses, 2021). Findings of Coiro et al. (2015) support this; they found that students were aware of correct criteria for internet sources but were unable to phrase their answers in a good way. This could explain that relatively more answers on the internet task received the score of 1.

Students in this study scored low on both historical and internet sources, which is in line with studies that focus on only one of the two types of sources (e.g. Breakstone et al., 2021; Britt & Aglinskis, 2002; Jacobsen et al., 2018; Kiili et al., 2018). Research among high school students has shown this skill is not easy to apply (e.g. Jacobsen et al., 2018). So, it could be expected that for students in lower secondary education this would be even more difficult. For this reason we used tasks that only focused on trustworthiness instead of more complicated tasks in which evaluating the trustworthiness is part of a larger assignment, such as document-based questions (Harris et al., 2016). However, even with explicit phrasing, students still had trouble with applying the skill in the correct way. For further research, we would recommend analysing which information students focus on when evaluating sources. More insight into students' reading behaviour would help teachers better guide their students in applying criteria of trustworthiness.

Based on the data, we noticed a discrepancy in the results of this study. On the one hand, the results indicate that students implicitly seem to realise the similarities. On the other hand, it seems that students are less explicitly aware that there is a relationship between applying the skill to history and internet sources. With respect to the former, we found that the three correct criteria of trustworthiness that could be applied in both the historical task and the internet task were indeed used by most of the students in both task sets. This indicates that a student who is aware of correct criteria of trustworthiness in one of the source types also uses those criteria when working with the other source type. It is unknown, however, if students have learned these criteria in history class and then transferred them onto internet sources, or the other way around. In future research, students could be asked if and where they have learned how to evaluate the trustworthiness of sources.

That students are less explicitly aware of the relationship between the two kinds of sources was shown in the think-aloud sessions. None of the students mentioned the word 'internet' or any related term when asked to explain whether they considered learning to evaluate historical sources important or useful. This lack of awareness can be related to the outcomes in the teacher interviews. Most of the interviewed teachers did use the terms internet or social media in relation to the usefulness of the history skill. However, only three of the eight teachers described concrete lesson situations in which they showed their students that evaluating the trustworthiness of information is also useful in other contexts, such as on contemporary examples of fake news. When explaining why no attention was given to using the historical skill in other contexts, one of the teachers referred to the fullness of the history curriculum. Since only eight teachers were interviewed, it is not clear how representative this argument is. However, the history teacher in McGrew's (2022) study mentioned the same argument. It is therefore interesting to further research teacher motives for refraining from giving attention to the use of the evaluation skill outside their classrooms.

To promote transfer of the evaluation skill from historical sources to internet sources, evaluating online historical sources might be a useful intermediate step that could be further investigated. In this way, the lesson is still focused on a historical topic, but the students are confronted with another type of source than is usually found in the textbooks. A possible trajectory could first focus on the historical evaluation skill using sources with the 'traditional' printed layout in textbooks. These sources are specially selected for students and modified so that the information about the source is presented directly above or beneath the source (McGrew, 2022). As the present study and other studies have shown that students often use 'weak heuristics' (e.g. Jacobsen et al., 2018; McGrew, 2021), the focus should be on correct criteria for evaluating sources and how to formulate well elaborated answers, for example through modelling and explicit teaching. Next, historical sources from the internet can be introduced. First, teachers can select these sources and provide some information about the internet source in the same way as when reading the 'traditional' historical sources, for example about the background of the maker of the website on which the source can be found. Gradually, students can be taught to how to find the information about an online source themselves, introducing the term 'lateral reading' (Wineburg & McGrew, 2017).

When moving from printed historical sources to online historical sources, transfer of the skill can gradually move from near to far. The content dimension of the skill remains similar, since the criteria of trustworthiness are the same, such as who is behind the information and what is the evidence. Also, the dimensions knowledge domain (history) and physical context (history classrooms) remain the same. For the other dimensions of the transfer taxonomy (Barnett & Ceci, 2002), there can be a gradual shift from near to far transfer: from printed to online sources, from answering textbook questions to evaluating online sources for writing a historical essay, from history to other subjects. It is recommended that, within the school, teachers from various school subjects discuss which criteria are more general and which

are more subject specific and develop a common learning trajectory for evaluating sources in different situations (Breakstone et al., 2018).

Limitations

Several limitations must be noted regarding the present study. First, we had to present the internet sources printed on paper instead of using computers and directing students to the real sources on the internet. Therefore, the sources in the internet task were not presented in an authentic way, whereas the historical sources were presented to the students in the same way as they would be exposed during history lessons. Students were therefore unable to apply the approach to internet sources described by Breakstone et al. (2018) and use the fact-checking method (Wineburg & McGrew, 2017), in which they would have to look at other websites to find information about the website under study. If we had presented the internet sources in an actual online setting, students might have used different strategies to determine trustworthiness, and they might have given different answers. However, this is less likely since most research has shown that students seldom leave the site under research to find information about that specific website (Wineburg & McGrew, 2017). Since we used screenshots of the websites, students did not get the 'holistic' view, so the internet sources were visually different from the historical sources. We believe, therefore, that the answers on the internet task reflect students' abilities. Nevertheless, for future research, we recommend that not only historical sources but also internet sources be presented in an authentic way.

Second, nowadays, students mostly use the internet for all their research, so they will also look for historical sources online. The sources they find online will look different from the historical sources we used in this research: they were edited in the usual educational style. Combined with the first limitation, this means that we offered the students both the historical and the internet sources in the 'easiest' way. This might lead to the conclusion that the results found in this study are a good reflection of the capabilities of students of this age regarding applying correct criteria of trustworthiness. Scores might have been even lower if we had used the internet sources in an online context.

A third limitation is that we used the same task order in the six participating schools: first the historical task, then the internet task. If we had reversed the task order in three of the schools, our design would have been stronger. Then, we would have been able to conclude whether conducting the historical task influenced student performance on the internet task. Since there was a gap of 2–3 weeks between the tasks, we assume that there was a minimal influence of the first task on the second. However, we cannot be certain.

Fourth, we did not collect data about students' characteristics or dispositions that might influence their performance, such as academic achievement, online experience or their reading skills. A review from Anmarkrud et al. (2022) showed mixed results on the influence of students' characteristics and dispositions on sourcing. However, as all participating students were placed in the same educational track based upon their academic achievement, it might be expected that they are a relative homogeneous group of students. Furthermore, their reading skills should have influenced their answers in both tasks in a similar way. In future research, students' experiences with internet sources could be taken into account, as this might influence their performance.

A final limitation is the small sample size used for the interviews: eight students and eight teachers. Although their answers provide us with some information about their views on the usefulness of the historical skill in other settings, these views might not be representative for Grade 9 students and their teachers. Additionally, these were structured interviews, and it is possible that we would have had more insight into students' and teachers' views in a less structured interview setting.

Conclusion and implications

This study adds to our understanding of the relationship between different student applications of history skill to evaluate the trustworthiness of historical sources and internet sources. Given the fact that there is a significant, moderate correlation between student performance on both tasks, and that students often used overlapping criteria in both tasks, we might conclude that a basic level of transfer is 'spontaneously' happening, although we do not know in which direction this transfer is occurring. Since learning how to evaluate the trustworthiness of sources is part of the history curriculum, we suggest that history classrooms can be an appropriate place to teach students not only how to evaluate the trustworthiness of historical sources, but also how to apply that evaluation skill to information they encounter on the internet. However, students do not explicitly make a connection between the evaluation skill used in history class and applications of that skill when searching for internet sources. History teachers can play an important role in raising students' awareness of this issue. When teachers evaluate the trustworthiness of historical sources with their students, they can demonstrate the differences and similarities between historical sources and internet sources, such as the criteria to use and how to find the information needed to apply these criteria.

Most history textbooks only include historical sources. We recommend using a broader range of sources, including sources from the internet. Since events and topics from the internet era are now included in history textbooks, it is possible to include online sources in the history curriculum. Additionally, internet sources about how historical events are perceived in the present time could be used in history lessons. In this way, the focus of the lesson is still historical, which is important for many history teachers, who are also expected to meet the demands of the curriculum and do not want to or cannot spend too much time on non-historical topics in their lessons.

Since the internet is the primary source for students when they search for information, for school or in their private life, it is important that they are aware that not everything they find online is true. Students using artificial intelligence such as ChatGPT for their history essays cannot automatically expect that these essays are solely based upon trustworthy information. History teachers can play an important role in raising students' critical attitude towards information they find online and helping them with the transfer of the historical skill of evaluating the trustworthiness of sources to internet sources as well.

ACKNOWLEDGEMENTS

The authors would like to thank research assistant Marlieke op de Weegh-Smelt for helping with constructing and scoring the internet task. We also want to thank research assistant Nina van Gastelen for her assistance with the teacher interviews.

FUNDING INFORMATION

This work was supported by Dudoc Alfa (grant number DA2-2018-13). Dudoc Alfa had no involvement in the research and in the article.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available owing to privacy or ethical restrictions.

ETHICS STATEMENT

This research was approved by the Ethical Board of the University of Amsterdam (2018-CDE-9715).

ENDNOTES

¹The students performed these tasks before the COVID19 pandemic, and were therefore not influenced by the discussions about the vaccinations against COVID19.

²Translation: At the top of the website there is a button 'UFO Reporting Center' (*UFO meldpunt*). Underneath the dolphin it says 'Since 2003, NIBURU has been publishing revealing and awareness-raising news, focusing on the disclosure of extraterrestrial life and the transformation of humanity and its planet. Forces are being united to manifest a new, loving, and healthy world' (*Sinds 2003 publiceert NIBURU onthullend en bewustmakend nieuws, waarbij de bekendmaking van buitenaards leven en de transformatie van de mensen en haar planeet centraal staan. Krachten worden gebundeld om een nieuwe, liefdevolle en gezonde wereld te manifesteren*). On the left there are two sponsors of the website: 'Orjana, naturally good' and Lyme's disease.

REFERENCES

- Anmarkrud, Ø., Bråten, I., Florit, E., & Mason, L. (2022). The role of individual differences in sourcing: A systematic review. *Educational Psychology Review*, 34, 749–792. <https://doi.org/10.1007/s10648-021-09640-7>
- Anmarkrud, Ø., Bråten, I., & Strømsø, H. I. (2014). Multiple-documents literacy: Strategic processing, source awareness, and argumentation when reading multiple conflicting documents. *Learning and Individual Differences*, 30, 64–76. <https://doi.org/10.1016/j.lindif.2013.01.007>
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological Bulletin*, 128(4), 612.
- Brand-Gruwel, S., Wopereis, I., & Walraven, A. (2009). A descriptive model of information problem solving while using internet. *Computers & Education*, 53(4), 1207–1217. <https://doi.org/10.1016/j.compedu.2009.06.004>
- Brante, E. W. (2019). A multiple-case study on students' sourcing activities in a group task. *Cogent Education*, 6(1), 1651441. <https://doi.org/10.1080/2331186X.2019.1651441>
- Breakstone, J. (2014). Try, try, try again: The process of designing new history assessments. *Theory & Research in Social Education*, 42(4), 453–485. <https://doi.org/10.1080/00933104.2014.965860>
- Breakstone, J., McGrew, S., Smith, M., Ortega, T., & Wineburg, S. (2018). Why we need a new approach to teaching digital literacy. *Phi Delta Kappan*, 99(6), 27–32. <https://doi.org/10.1177/0031721718762419>
- Breakstone, J., Smith, M., Wineburg, S., Rapaport, A., Carle, J., Garland, M., & Saavedra, A. (2021). Students' civic online reasoning: A national portrait. *Educational Researcher*, 50(8), 505–515. <https://doi.org/10.3102/0013189X211017495>
- Britt, M. A., & Aglinskias, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction*, 20(4), 485–522. https://doi.org/10.1207/S1532690XCI2004_2
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (second ed.). Routledge.
- Coiro, J., Coscarelli, C., Maykel, C., & Forzani, E. (2015). Investigating criteria that seventh graders use to evaluate the quality of online information. *Journal of Adolescent & Adult Literacy*, 59(3), 287–297. <https://doi.org/10.1002/jaal.448>
- De La Paz, S., & Felton, M. K. (2010). Reading and writing from multiple source documents in history: Effects of strategy instruction with low to average high school writers. *Contemporary Educational Psychology*, 35(3), 174–192. <https://doi.org/10.1016/j.cedpsych.2010.03.001>
- De La Paz, S., Ferretti, R., Wissinger, D., Yee, L., & MacArthur, C. (2012). Adolescents' disciplinary use of evidence, argumentative strategies, and organizational structure in writing about historical controversies. *Written Communication*, 29(4), 412–454. <https://doi.org/10.1177/0741088312461591>
- Gasser, U., Cortesi, S., Malik, M. M., & Lee, A. (2012). Youth and digital media: From credibility to information quality. *Berkman Center Research Publication*, 2012, 1. <https://doi.org/10.2139/ssrn.2005272>
- Hajian, S. (2019). Transfer of learning and teaching: A review of transfer theories and effective instructional practices. *IAFOR Journal of Education*, 7(1), 93–111. [10.22492/ije.7.1.06](https://doi.org/10.22492/ije.7.1.06)
- Halvorsen, A.-L., Harris, L. M., Aponte Martinez, G., & Frasier, A. S. (2016). Does students' heritage matter in their performance on and perceptions of historical reasoning tasks? *Journal of Curriculum Studies*, 48(4), 457–478. <https://doi.org/10.1080/00220272.2015.1092585>
- Hargittai, E., Fullerton, L., Menchen-Trevino, E., & Thomas, K. Y. (2010). Trust online: Young adults' evaluation of web content. *International Journal of Communication*, 4, 468–494.
- Harris, L. M., Halvorsen, A.-L., & Aponte-Martínez, G. J. (2016). "[My] family has gone through that": How high school students determine the trustworthiness of historical documents. *The Journal of Social Studies Research*, 40(2), 109–121. <https://doi.org/10.1016/j.jssr.2015.06.007>
- Horn, S., & Veermans, K. (2019). Critical thinking efficacy and transfer skills defend against 'fake news' at an international school in Finland. *Journal of Research in International Education*, 18(1), 23–41. <https://doi.org/10.1177/1475240919830003>

- Hynd, C., Holschuh, J. P., & Hubbard, B. P. (2004). Thinking like a historian: College students' reading of multiple historical documents. *Journal of Literacy Research*, 36(2), 141–176. https://doi.org/10.1207/s15548430jr3602_2
- Jacobsen, R., Halvorsen, A.-L., Frasier, A. S., Schmitt, A., Crocco, M., & Segall, A. (2018). Thinking deeply, thinking emotionally: How high school students make sense of evidence. *Theory & Research in Social Education*, 46(2), 232–276. <https://doi.org/10.1080/00933104.2018.1425170>
- Kammerer, Y., Meier, N., & Stahl, E. (2016). Fostering secondary-school students' intertext model formation when reading a set of websites: The effectiveness of source prompts. *Computers & Education*, 102, 52–64. <https://doi.org/10.1016/j.compedu.2016.07.001>
- Kiili, C., Coiro, J., & Raikkonen, E. (2019). Students' evaluation of information during online inquiry: Working individually or in pairs. *The Australian Journal of Language and Literacy*, 42(3), 167–183. <https://doi.org/10.1007/BF03652036>
- Kiili, C., Leu, D. J., Marttunen, M., Hautala, J., & Leppänen, P. H. (2018). Exploring early adolescents' evaluation of academic and commercial online resources related to health. *Reading and Writing*, 31(3), 533–557. <https://doi.org/10.1007/s11145-017-9797-2>
- Körber, A. (2016). Translation and its discontents II: A German perspective. *Journal of Curriculum Studies*, 48(4), 440–456. <https://doi.org/10.1080/00220272.2016.1171401>
- List, A., Alexander, P. A., & Stephens, L. A. (2017). Trust but verify: Examining the association between students' sourcing behaviors and ratings of text trustworthiness. *Discourse Processes*, 54(2), 83–104. <https://doi.org/10.1080/0163853X.2016.1174654>
- McGrew, S. (2020). Learning to evaluate: An intervention in civic online reasoning. *Computers & Education*, 145, 103711. <https://doi.org/10.1016/j.compedu.2019.103711>
- McGrew, S. (2021). Challenging approaches: Sharing and responding to weak digital heuristics in class discussions. *Teaching and Teacher Education*, 108, 103512. <https://doi.org/10.1016/j.tate.2021.103512>
- McGrew, S. (2022). Bridge or byway? Teaching historical reading and civic online reasoning in a US history class. *Theory & Research in Social Education*, 2, 196–225. <https://doi.org/10.1080/00933104.2021.1997844>
- McGrew, S., Breakstone, J., Ortega, T., Smith, M., & Wineburg, S. (2018). Can students evaluate online sources? Learning from assessments of civic online reasoning. *Theory & Research in Social Education*, 46(2), 165–193. <https://doi.org/10.1080/00933104.2017.1416320>
- McGrew, S., Smith, M., Breakstone, J., Ortega, T., & Wineburg, S. (2019). Improving university students' web savvy: An intervention study. *British Journal of Educational Psychology*, 89, 485–500. <https://doi.org/10.1111/bjep.12279>
- Meneses, L. F. S. (2021). Thinking critically through controversial issues on digital media: Dispositions and key criteria for content evaluation. *Thinking Skills and Creativity*, 42, 100927. <https://doi.org/10.1016/j.tsc.2021.100927>
- Metzger, M. J. (2007). Making sense of credibility on the web: Models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology*, 58(13), 2078–2091. <https://doi.org/10.1002/asi.20672>
- Nokes, J. D. (2010). The evolving concept instructional strategy: Students reflecting on their processing of multiple, conflicting, historical sources. *National Social Science Journal*, 35(1), 104–117.
- Nokes, J. D. (2017). Exploring patterns of historical thinking through eighth-grade students' argumentative writing. *Journal of Writing Research*, 8(3), 437–467. [10.17239/jowr-2017.08.03.02](https://doi.org/10.17239/jowr-2017.08.03.02)
- Nygren, T., & Guath, M. (2019). Swedish teenagers' difficulties and abilities to determine digital news credibility. *Nordicom Review*, 40(1), 23–42. <https://doi.org/10.2478/nor-2019-0002>
- Paul, J., Macedo-Rouet, M., Rouet, J.-F., & Stadler, M. (2017). Why attend to source information when reading online? The perspective of ninth grade students from two different countries. *Computers & Education*, 113, 339–354. <https://doi.org/10.1016/j.compedu.2017.05.020>
- Pérez, A., Potocki, A., Stadler, M., Macedo-Rouet, M., Paul, J., Salmerón, L., & Rouet, J.-F. (2018). Fostering teenagers' assessment of information reliability: Effects of a classroom intervention focused on critical source dimensions. *Learning and Instruction*, 58, 53–64. <https://doi.org/10.1016/j.learninstruc.2018.04.006>
- Perkins, D. N., & Salomon, G. (1992). Transfer of learning. *International Encyclopedia of Education*, 2, 6452–6457.
- Pieschl, S., & Sivyer, D. (2021). Secondary students' epistemic thinking and year as predictors of critical source evaluation of internet blogs. *Computers & Education*, 160, 104038. <https://doi.org/10.1016/j.compedu.2020.104038>
- Pintrich, P. R., Smith, D. A., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)*. National Center for Research to Improve Postsecondary Teaching and Learning.
- Potocki, A., de Pereyra, G., Ros, C., Macedo-Rouet, M., Stadler, M., Salmerón, L., & Rouet, J.-F. (2020). The development of source evaluation skills during adolescence: Exploring different levels of source processing

- and their relationships. *Journal for the Study of Education and Development*, 43(1), 19–59. <https://doi.org/10.1080/02103702.2019.1690848>
- Reisman, A. (2011). The 'document-based lesson': Bringing disciplinary inquiry into high school history classrooms with adolescent struggling readers. *Journal of Curriculum Studies*, 44(2), 233–264. <https://doi.org/10.1080/00220272.2011.591436>
- Reisman, A. (2012). Reading like a historian: A document-based history curriculum intervention in urban high schools. *Cognition and Instruction*, 30(1), 86–112. <https://doi.org/10.1080/07370008.2011.634081>
- Sakki, I., & Pirttilä-Backman, A.-M. (2019). Aims in teaching history and their epistemic correlates: A study of history teachers in ten countries. *Pedagogy, Culture & Society*, 27(1), 65–85. <https://doi.org/10.1080/14681366.2019.1566166>
- Scholes, L., McDonald, S., Stahl, G., & Comber, B. (2023). Many truths, many knowledges, many forms of reason: Understanding middle-school student approaches to sources of information on the internet. *British Educational Research Journal*, 50, 53–72. <https://doi.org/10.1002/berj.3909>
- Seixas, P. (2016). Translation and its discontents: Key concepts in English and German history education. *Journal of Curriculum Studies*, 48(4), 427–439. <https://doi.org/10.1080/00220272.2015.1101618>
- Seixas, P., & Morton, T. (2013). *Six big historical thinking concepts*. Nelson Education Ltd.
- Sendur, K. A., van Boxtel, C., & van Drie, J. (2021). Undergraduate L2 students' performance when evaluating historical sources for reliability. *English for Specific Purposes*, 61, 17–31. <https://doi.org/10.1016/j.esp.2020.08.004>
- Strømsø, H. I., Bråten, I., Britt, M. A., & Ferguson, L. E. (2013). Spontaneous sourcing among students reading multiple documents. *Cognition and Instruction*, 31(2), 176–203. <https://doi.org/10.1080/07370008.2013.769994>
- Van der Eem, M., van Drie, J., Brand-Gruwel, S., & van Boxtel, C. (2023). Students' evaluation of the trustworthiness of historical sources: Procedural knowledge and task value as predictors of student performance. *The Journal of Social Studies Research*, 47(1), 64–76. <https://doi.org/10.1016/j.jssr.2022.05.003>
- Van Drie, J., & Van Boxtel, C. (2008). Historical reasoning: Towards a framework for analyzing students' reasoning about the past. *Educational Psychology Review*, 20(2), 87–110. <https://doi.org/10.1007/s10648-007-9056-1>
- Walraven, A., Brand-Gruwel, S., & Boshuizen, H. P. (2009). How students evaluate information and sources when searching the World Wide Web for information. *Computers & Education*, 52(1), 234–246. <https://doi.org/10.1016/j.compedu.2008.08.003>
- Wansink, B. G., Akkerman, S. F., Vermunt, J. D., Haenen, J. P., & Wubbels, T. (2017). Epistemological tensions in prospective Dutch history teachers' beliefs about the objectives of secondary education. *The Journal of Social Studies Research*, 41(1), 11–24. <https://doi.org/10.1016/j.jssr.2015.10.003>
- Wineburg, S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, 83(1), 73–87. <https://doi.org/10.1037/0022-0663.83.1.73>
- Wineburg, S., & McGrew, S. (2017). *Lateral reading: Reading less and learning more when evaluating digital information*. Stanford University.
- 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), 1–40.

How to cite this article: van der Eem, M., van Drie, J., Brand-Gruwel, S. & van Boxtel, C. (2025). Students evaluating the trustworthiness of historical sources and internet sources: A comparison. *British Educational Research Journal*, 51, 665–686. <https://doi.org/10.1002/berj.4095>

APPENDIX A

Example of a question and rubric of the historical task

Imagine you are researching life in the trenches. You find this excerpt from the book *All Quiet on the Western Front* by the German author Erich Remarque. In this book, he describes life in the trenches in France.

An excerpt from the book *All Quiet on the Western Front*:

Haje Westhus is being dragged away with his back torn open; with every breath, you can see his lung beating through the wound. I can still shake his hand. 'It's over for me', he groans and bites his arm in pain. We see people still alive, even though their skulls have been blown away; we see soldiers walking with both feet shot off; they stagger on their splintered leg stumps towards the nearest cover; a corporal crawls two kilometers on his hands, dragging his shattered knees behind him ...; we see people without mouths, without faces ... The sun sets; night returns; the shells whistle; the world comes to an end.

You find the following five facts as background information for this book:

1. Erich Remarque was a soldier in the First World War and served in the trenches.
2. Remarque wanted to show the horrors and pointlessness of war with his book.
3. Erich Remarque wrote the book in 1929.
4. Within 18 months of its publication, 2.5 million copies had been sold, and a film had already been made.
5. When Hitler came to power in Germany, the book was banned and publicly burned.

Two of the five facts in the background information might make you question the trustworthiness of this source for your research.

Fact ... makes me question the trustworthiness of the source, because (give explanation): ...

Fact ... makes me question the trustworthiness of the source, because (give explanation): ...

The rubric.

Score	Description	Examples of students' answers
0	The student selects the wrong fact or selects the correct fact with an incorrect explanation	<ul style="list-style-type: none"> • Fact 1: if he was at the front, why isn't he dead? • Fact 2: people didn't believe that war was pointless. • Fact 5: because if it was banned, how could this fragment still exist?
1	The student selects the correct fact/criterion but provides no clear elaborated explanation	<ul style="list-style-type: none"> • Fact 2: he thought the war was purposeless. • Fact 3: because it was written so late
2	The student selects a correct fact/criterion and provides a clear explanation	<ul style="list-style-type: none"> • Fact 2: because he wanted to show the horrors and pointlessness of war, maybe he made the book more intense or exaggerated some parts. • Fact 3: because it was written a long time later, so he might have forgotten some small but important details

APPENDIX B

Example of a question and rubric of the internet task² [Colour figure can be viewed at wileyonlinelibrary.com]

http://niburu.co/index.php?option=com_content&view=article&id=13621:nog-eeen-aantal-dwingende-reden-om-vaccinaties-te-vermijden&catid=17:gezondheid&Itemid=30.



On this website, an unknown author describes under the heading 'A few more compelling reasons to avoid vaccinations' that companies always want to make a profit, including the pharmaceutical industry that makes vaccines. To make a profit, they need to sell as much as possible for as little money as possible. The government is their accomplice and ensures that people get vaccinated. 'That's not rocket science, that's a fact'. Furthermore, the author states that vaccinations usually do not work or even have the opposite effect.

This site provides me with trustworthy/not trustworthy/somewhat trustworthy information of the possible dangers of vaccination (circle your choice).

Because (provide explanation): ...

The rubric.

Score	Description	Examples of students' answers
0	The student considers the source (somewhat) trustworthy or gives a wrong argument explaining why the source is not trustworthy	<ul style="list-style-type: none"> Somewhat, it looks childish. Somewhat, it's for money, and there is evidence against the information I read
1	The student considers the source to be untrustworthy mentioning a criterion but the students does not provide a clear elaborated explanation	<ul style="list-style-type: none"> Somewhat, the author is anonymous Not, the site itself already doesn't look good; in addition, most of the arguments are based on opinion No sources
2	The student considers the source to be untrustworthy using a correct criterion and provides an elaboration	<ul style="list-style-type: none"> Not, you don't know who wrote this information or if it's trustworthy. Before I can consider a source trustworthy, I would like to know who wrote it and their intentions. In this source, you can only see their intentions, but not the author, so I don't know if the source is trustworthy Not, you don't know who is writing it and where they got their information from Not, this site is not trustworthy; at the top left it says, 'UFO Reporting Center', showing they don't take journalism very seriously

Note: If you want to see all the tasks, please contact the first author.