

**Sourcing in text comprehension:
A review of interventions targeting sourcing skills**

Eva Wennås Brante & Helge I. Strømsø

Department of Education,
University of Oslo,
Norway

Funding: This study was funded by The Norwegian Research Council (grant number 237981/H20)

Conflict of Interest: The authors declare that they have no conflict of interest.

Corresponding author: Eva Wennås Brante, Department of Education, P.O. Box 1092, Blindern, N-0317 Oslo, Norway.

E-mail: e.w.brante@ped.uio.no

Cell: +46707298429

Phone: +47-22855340

**Sourcing in text comprehension:
A review of interventions targeting sourcing skills.**

Abstract

The process of using information about documents such as the author, genre, and date of publication while evaluating and interpreting those documents' content was labeled "sourcing" in a seminal paper by Wineburg (1991). Studies in various domains have adapted the term sourcing while referring to central reading skills in modern information societies. In this review, we discuss the concept of sourcing grounded in research from social psychology, information sciences, and text comprehension. Based on that, we reviewed 18 intervention studies in educational settings, in order to identify how sourcing was operationalized in the studies, the nature of the interventions, and how successful they were.

The review shows that interventions for younger students emphasized source credibility, whereas interventions among older students also emphasized the role of sourcing in interpretation. None of the studies measured how students search for source features or specifically which features they attend to. Regarding the nature of the studies, the use of multiple partly conflicting documents was common, with that condition positively related to outcome measures. Another characteristic was the use of inquiry tasks. A majority of the studies do not apply findings from persuasion theory and information science indicating that credibility assessment requires effort and motivation. Future interventions should more strongly emphasize the relationship between sourcing and motivation.

Sourcing in text comprehension:

A review of interventions targeting sourcing skills. When the OECD categorized levels of literacy in a 2013 report, the highest level involved “searching for and integrating information across multiple dense texts.” A person reaching that level is able to “evaluate the reliability of evidentiary sources and select key information.” (OECD, 2014, p. 26).

Considering the easily accessible yet overwhelming amount of information on the Internet, the OECD describes vital skills for people living in modern information societies. Whereas various types of publishers have traditionally acted as gatekeepers of the information to which people have been exposed, the responsibility for evaluating information has now increasingly been transferred to readers. Accordingly, whether readers are reflecting on the content of a single document or struggling to make sense of contradictory statements presented in different documents, they need to consider the source of the information being communicated.

The process of using information about documents such as the author, genre, and date of publication while evaluating and interpreting those documents’ content was labeled “sourcing” in a seminal paper by Wineburg (1991). Since that study, increasing attention has been paid both to the importance of readers’ sourcing skills and to their apparent lack thereof. Recent empirical work has shown that readers’ attention to and memory for source information relate to their text-based learning and comprehension (Anmarkrud, Bråten, & Strømsø, 2014; Barzilai & Eshet-Alkalai, 2015; Barzilai, Tzadok, & Eshet-Alkalai, 2015; Bråten, Strømsø, & Britt, 2009; Goldman, Braasch, Wiley, Graesser, & Brodowinska, 2012; Strømsø, Bråten, & Britt, 2010; Wiley et al., 2009). However, prior research also indicates that students often seem either to ignore source information, process it superficially, or fail to take that information into account in discourse processing (Britt & Aglinskias, 2002; Nokes, Dole, & Hacker, 2007; Walraven, Brand-Gruwel, & Boshuizen, 2008; Wiley et al., 2009; Wineburg, 1991). For example, students may only rely on superficial features such as the

name of a website or author, choose less relevant documents based on source credibility, or judge credibility of a source or content based on personal experience or stance (Bråten, Strømsø, & Andreassen, 2016; Brem, Russell & Weems, 2001; Bromme, Scharrer, Stadtler, Hömberg, & Torspecken, 2015; Salmerón, Macedo-Rouet, & Rouet, 2015). Consequently, more effort needs to be invested in educational programs emphasizing the role of sourcing in discourse comprehension.

In the current review, we will first discuss the concept of sourcing based on research from social psychology, information sciences, and text comprehension. Based on this discussion, we will review intervention studies in educational settings, intending to improve students' abilities to use source information while reading. Our goals are to identify how sourcing was operationalized in the studies, the nature of the interventions, and how successful they were.

Sourcing across various disciplines

In this section we will briefly discuss the role of sourcing within three different theoretical frameworks. Two of those (persuasion and information/communication research) have not been oriented towards educational settings in particular, whereas research in the literacy/discourse tradition has closer links to such settings. Hence, we will emphasize the latter tradition somewhat more than the others as our aim is to review interventions directed at student populations.

Persuasion research. There are a number of theoretical approaches to the role of sourcing in peoples' use of different types of information sources. For example, decades of social psychology research on attitudes and persuasion has demonstrated that source credibility may induce attitude changes. Several prominent models suggest two routes to persuasion ("dual-process theories"), with effortful thinking about a message sometimes resulting in attitude changes and reliance on simpler heuristics leading to persuasion at other

times (e.g., Chen & Chaiken, 1999; Petty & Wegener, 1999). A simple heuristic would, for example, imply that a user has confidence in a message only based on the perceived source credibility, whereas people investing more effort in evaluating the message will thoroughly scrutinize both the source information and the site's content (Petty & Briñol, 2012). Reliance on source features only, such as expertise or attractiveness, is considered a heuristic used when the reader is not motivated or able to reflect deeply on the communicated message (Chen & Chaiken, 1999), whereas motivated and able readers may process all available information, including source features, in evaluating the message (Petty & Briñol, 2012). Research has repeatedly demonstrated that source characteristics are often important in people's judgments of persuasive messages (Stiff & Mongeau, 2003). Different source characteristics, such as perceived similarity and physical attractiveness, have been demonstrated to influence persuasion, but the effects of source credibility have most often been highlighted. Source credibility has been defined as a combination of perceived source expertise and source trustworthiness, with expertise indicating knowledge or competence and trustworthiness representing the communicator's intention to communicate valid information (Hovland, Janis, & Kelley, 1953). Perceptions of source credibility may affect both how people process a message and how they judge it, with those processes also depending on people's motivation and ability to reflect on both the source and the message. Thus, an important contribution from social psychology to research on readers' sourcing is the highlighting of how source credibility and readers' motivation may affect their attention to and use of source features in processing and evaluating texts.

Information and communication research. In line with research from social psychology, researchers in information and communication sciences have, in the last decades, paid particular attention to information users' perceptions of source credibility when approaching information in digital media (e.g., Metzger, 2007; Rieh, 2002). An information

seeker's motivation is considered important for the effort invested in evaluating source information, and the motivation may vary according to the task and context. For example, people seem to be more concerned about source credibility when searching for health-related information than for product-related information, indicating that the potential consequences of the information use may impact the credibility evaluation (Rieh & Hilligoss, 2007). However, it has been suggested that Internet users, in general, are unlikely to mobilize the necessary effort needed for the credibility assessment of online information (Metzger, 2007). Given the complexity of source evaluation on the Web, information seekers might apply certain heuristics unless their information need requires a close inspection of the source's credibility. For example, Metzger, Flanagin, and Medders (2010) identified five heuristics applied by Internet users: a source's *reputation*, other people's *endorsement*, *consistency* between information sources, *expectancy* compliance, and perceived *persuasive intent*. The credibility of a message may thus be evaluated according to numerous criteria, with the source features emphasized in a majority of the heuristics used by participants in the above study (Metzger et al., 2010). In summary, research from information and communication sciences on sourcing has been closely related to persuasion research regarding the role of source credibility and information users' motivation. Information and communication researchers particularly pay attention to how people search for and select Web-based information and are concerned about credibility assessment as an information skill.

Literacy and discourse research. Assessing source credibility could certainly be important when deciding which information one should pay attention to and whether to trust persuasive messages, but research in literacy and discourse has also shown that source information could play an important role in the reading and comprehension of informational documents. Several studies have indicated that students' processing, interpretation, evaluation, and comprehension of documents may be affected by their knowledge about the

text, such as the category of texts to which it belongs, who wrote it and when it was published. For example, Zwaan (1994) found that students' genre expectations affected both their reading time and comprehension, whereas Gibbs, Kushner, and Mills (1991) found that knowledge about authors affected readers' judgments of the text's meaningfulness. Other studies have shown how sourcing may affect readers' evaluation of authors' perspectives (Paxton, 2002), reading behaviors (Bazerman, 1985) and expectations regarding document content (Lundeberg, 1987). In Wineburg's (1991) study on expert historians and high school students reading a set of historical texts, he noted that the students regarded the texts as "bearers of information" and that they hardly noticed the source features. In contrast, the experts interpreted the texts in light of the source characteristics, and Wineburg (1994) referred to this approach as the reading of the documents' subtext, implying that the experts believed texts to be both a rhetorical and a human artifact. Thus, in the domain of history, Wineburg (1994) describes sourcing both as a way of constructing meaning from documents and as epistemic beliefs concerning the nature of texts. Therefore, expert historians seemed to be less concerned about source credibility and more focused on how the source characteristics could help them better interpret a document.

Also based on studies of reading in history, the influential documents model framework (DMF) was somewhat later developed by Britt, Perfetti and Rouet (Britt, Perfetti, Sandak, & Rouet, 1999; Perfetti, Rouet, & Britt, 1999) and applied to research on students' reading of multiple documents in domains other than history, such as social sciences, law, and science (e.g., Le Bigot & Rouet, 2007; Strømsø & Bråten, 2002, Wiley et al., 2009). The DMF illustrates how multiple documents about the same issue or situation may be comprehended to create a global understanding across documents. In line with Wineburg (1994), texts are considered to represent more than purely linguistic constructions; they are

also social entities written by authors with specific backgrounds and intentions and published in certain contexts, cultures and points in time (Britt & Rouet, 2012).

The DMF is an extension of Kintsch's construction-integration model (CI) (Kintsch, 1998; Kintsch & Rawson, 2007) in which the mental representations of texts are essentially described at two levels. The text base represents a text's literal meaning, whereas the situation model represents an integration of the text's information and the reader's prior world knowledge. The CI model developed from research on single-text reading, whereas today, people frequently read from *multiple* documents. According to the DMF (Britt et al., 1999; Britt & Rouet, 2012; Perfetti et al., 1999), reading from multiple documents requires the reader to construct two additional text-representing layers, i.e., the intertext model and the integrated mental model of the situations, with the intertext model attracting special interest when the role of source information is considered.

The intertext model develops when readers "tag" content to source features such as who wrote the text, when it was published, and its genre. Thus, the reader constructs a mental representation of the sources of various information units and how those units are interrelated. However, this is a demanding task that the reader will probably only undertake when it is relevant to the reading task and the reader's aims; for example, when writing an argumentative text (Britt & Rouet, 2012) or when exposed to contradictory documents (Strømsø, Bråten, Britt, & Ferguson, 2013). Such tasks may instigate the construction of an intertext model, as attention to source information could help readers in understanding why different documents represent diverse perspectives or positions. Moreover, source features are important for judging whether content should be emphasized or ignored. However, source information also facilitates expert readers' interpretation of documents' meaning and their organization of an integrated mental representation of multiple documents.

Defining sourcing

One challenge when reviewing research on readers' sourcing behavior is that the concept is not used in a consistent manner. Goldman and Scardamalia (2013) note that concepts such as source, sourcing and document are not always clearly distinguishable in research on literacy and discourse. They suggest that sourcing should

“... refer to the processes of identifying and representing metadata, including the author(s) or spokesperson(s) who created the information resource, their purpose in creating it, the intended audience, where and when it appeared initially, as well as subsequent publication.” (p 259)

The term metadata seems to capture what is commonly referred to as source information, source features, or source properties. Metadata represent information about information and, in our context, information about documents. We prefer to use the term “source features” to avoid confusion with the specific use of metadata in fields such as information or computer sciences. Persuasion research typically refers to source features as characteristics of the creator of a message (e.g. gender, expertise), whereas some other forms of metadata are labeled channel variables (Pornpitakpan, 2004). In the present review, source features will refer to the categories suggested by Perfetti et al. (1999), with *author identification* referring to his or her characteristics and motives, *setting* referring to where the document was created, date, and cultural context, whereas *document form* indicates language style and document type. Thus, source features may reveal the author's intention, the intended audience, the author's knowledge of the phenomena described, the setting, the nature of the publication, and other information that may be helpful in understanding the role and context of a document, as well as in judging the quality of the content.

Wineburg (1991) initially referred to sourcing as a discipline-based heuristic that expert historians use in predicting, interpreting and creating an integrated mental representation of a set of documents. Recently, Wineburg and Reisman (2014) claimed that

sourcing goes beyond disciplinary boundaries and is more of “a way of apprehending the world” (p 636). Still, a reader’s degree of discipline knowledge will affect to what extent the reader is able to take advantage of information about documents. For example, Gottlieb and Wineburg (2011) found that both historians and non-historians noted source information but that the two groups used that information differently. The act of paying attention to and using source features may be a general strategy, whereas how source features are used seems to be related to readers’ knowledge about documents’ content and context. Thus, it seems reasonable to define sourcing as both a general and discipline-based approach to documents. For the purpose of identifying how sourcing is operationalized in intervention studies, we will define sourcing according to a set of activities that is involved. Hence, sourcing is here defined as *identifying and representing source features to predict, interpret and evaluate documents’ content and relevance according to a reading task.*

Identifying and attending to source features is obviously a prerequisite for other sourcing activities. However, students apparently often ignore source features, and no sourcing will thus occur (e.g., Bråten et al., 2016; Britt & Aglinskias, 2002; Wineburg, 1991). Representing source features implies that students remember “who said what” in order to use that information when considering the nature, purpose, and reliability of a message. Regarding prediction, it has been demonstrated how source features form the expectations that experts express when selecting or interpreting relevant documents (Bazerman, 1985; Lundeberg, 1987) and that genre expectations may affect students’ text comprehension (Wolfe & Woodwyk, 2010; Zwaan, 1994). The think-aloud studies of Wineburg (1991; Gottlieb & Wineburg, 2012) clearly demonstrate how experts use source information in interpreting documents’ content. Finally, research on source credibility shows that readers could evaluate source features and use them in evaluating content (e.g., Metzger et al., 2010; Stiff & Mongeau, 2003). Thus, our suggested definition of sourcing seems to be in accordance

with results from prior research, and it implies that sourcing may comprise several different activities. To what extent students at different grade levels are able to master those sourcing activities is unclear, although studies in developmental psychology indicate that also students in primary school have the potential to do so (Bråten, Stadler, & Salmerón, in press).

Sourcing at different grade levels.

Studies have shown that even pre-school children are selective in whom they learn from. For example, small children show a tendency to prefer information supplied by someone who has been proven to be a reliable source of information in the past (Harris & Corriveau, 2011). However, being a member of the information society requires a broader repertoire of validation strategies. In a literature review, Walraven et al. (2008) investigated what it takes for learners to become proficient in solving information-based problems among different age groups. Regarding the subskill to ‘judge search results’, Walraven et al. found that the youngest children displayed three strategies as a base for their decisions to access a site, based on the title, the summary describing the results and the rank-order in the hit lists. Successful adults, on the other hand, evaluated the results by looking at the title, the origin of the information and identifiers in the URL. The review also showed that young children rarely corroborate information across sites, and they do not actively search for source information about a site. According to Macedo-Rouet et al. (2013), “children’s ability to evaluate information while reading develops across grades” (p 207). Accordingly, one could expect intervention studies at different levels to target sourcing skills of different complexity.

Rationale of the study

What follows is a comprehensive review of intervention studies intended to improve students’ sourcing approach when reading expository documents in educational settings. Given the limited numbers of such studies, we decided to include intervention studies from primary school to higher education. We will identify how sourcing was operationalized in the

studies, the nature of the interventions, and how successful they were. Thus, our review will be guided by two research questions:

1. Which aspects of sourcing were targeted in interventions and outcome measures?
2. What features of instructional interventions (e.g., type of instruction, duration, and grade level) are associated with improved outcomes?

Search methodology

To detect relevant interventions, we started by conducting a literature search in relevant research databases. The search was conducted in January 2017. We looked for student populations in school/educational settings receiving a treatment with the goal of increasing students' skills related to sourcing. Table 1 displays the five categories of keywords that were used in the search formula.

Table 1 about here

With sourcing not necessarily a concept used in all relevant studies, we also included keywords from research fields that presumptively could emphasize that students need to consider source features when searching for and reading informational texts, for example, studies of information problem-solving (e.g., Walraven, Brand-Gruwel, & Boshuizen, 2013), web search (e.g., Salmerón, Kammerer, & García-Carrión, 2013) or critical literacy (e.g., Behrman, 2006). With respect to the keywords not to be included, we wished to focus on studies in educational settings that addressed reading beyond elementary skills (decoding). Thus, interventions in kindergarten were not relevant, nor were studies that were less related to reading.

The inclusion criteria were set as follows: 1) the intervention should be carried out in a school/educational setting, 2) there should be a control group, 3) part of the intervention should aim to increase students' sourcing skills, 4) only peer-reviewed and published studies would be considered, and 5) English language was required. We limited the search to papers

from 1991 (the year of Wineburg's seminal paper) to January 2017 and conducted searches in four databases related to educational psychology: ProQuest, EBSCOhost, PsychInfo and Web of Science. The first search yielded 9632 hits on ProQuest, 1400 hits on EBSCOhost, 1743 hits on PsychInfo, and 2494 hits on Web of Science. By the use of filters of publication date, English language, and peer-reviewed and by once again excluding studies from, for example, kindergarten or case studies and with more irrelevant features, we narrowed the results to 1744 hits on ProQuest, 701 hits on EBSCOhost, 501 on PsychInfo and 742 hits on Web of Science. All titles and abstracts were scanned, thus we found ten relevant abstracts on ProQuest, 17 abstracts on EBSCOhost, one new one from PSYCHinfo and no new abstracts from Web on Science, as the databases overlapped one another. The first author read the abstracts of the potential studies, using the inclusion criteria to decide which studies should be incorporated into the review. As a result of this process, 15 papers were identified. These papers were then back-searched. Of the articles identified by the back-search, those that failed to provide an explicit definition or were duplicates of those already retrieved were excluded, leaving three remaining manuscripts. In total, we found 18 papers fulfilling the inclusion criteria, and these are included in the review.

Calculation of effect sizes

Effect sizes (ES) were calculated for all studies that provided adequate statistical information using the common standardized mean difference statistic Hedges' g . Positive g -values indicate that the intervention groups have a higher outcome score than the control groups. A majority of the included studies represented quasi-experimental pre-/post-test designs. In those cases, we calculated an "adjusted" ES by subtracting the pre-ES from the post-ES, following the guidelines suggested by Durlak (2009). When more than one intervention group were included in a study, we calculated a separate ES for each intervention group. The ESs show how much difference there is between groups, but there is no direct

relationship between a p -value and the degree of the effect, as the p -value is a function of both the sample size and ES. To indicate p , non-significant (ns) results are marked in the tables in the result section. We did not consider it fruitful to calculate an overall mean effect size for all the studies because comparison groups and dependent measures varied so widely across studies.

Results

The results are reported according to the educational level where the interventions occurred, with three tables displaying the studies for elementary and lower secondary school (7 studies), upper secondary school (7 studies), and higher education (4 studies), respectively. Each table presents the educational level, number of participants, length of the intervention, a brief description of the intervention, dependent measures, and results in terms of effect sizes. Whenever the difference between the control and intervention groups is reported as not significant, this is marked as ns with the corresponding ES in the column. The outcome measures were, with one exception (Reisman, 2012), researcher-developed and, as such, not directly comparable.

Interventions in elementary and lower secondary school

The ESs in Table 2 show that the intervention groups generally do better than the control groups on outcome measures. The ESs in the Walraven et al. (2013) study should be commented on initially. Regarding the students' evaluation of hit lists, the positive ESs are due to a decrease in the control group's results, not to an increase for the intervention group.

Table 2 about here

The contexts of the seven studies varied with respect to several factors. In four of the studies, the teachers were responsible for the implementation over several lessons, ranging from two to sixty hours, with three embedded in the curriculum content (Argelagós & Pifarré, 2012; Kingsley et al., 2015; Walraven et al., 2013; Zhang & Duke, 2011). Argelagós and

Pifarré (2012) implemented their program in math, science, social science, and technology, with teachers also taking part in developing materials, whereas Kingsley et al. (2015) ran the program in language art lessons and Walraven et al. (2013) embedded the intervention in history lessons. The potential challenges of a teacher-led design were demonstrated in Walraven et al. (2013), where field notes showed that the intervention program was only partly executed in the groups. In contrast to the teacher-led interventions, the three researcher-led interventions lasted for only one lesson, and none were connected specifically to the curriculum (Kammerer et al., 2016; Macedo-Rouet et al., 2013; Mason et al., 2014).

Instructions and activities. All but one (Macedo-Rouet et al., 2013) of the studies had participants working on multiple web-based documents to solve a task, with the students' work facilitated by specific guidelines communicated by worksheets (Kammerer et al., 2016; Mason et al., 2014; Walraven et al., 2013; Zhang & Duke, 2011), prompts (Argelagós & Pifarré, 2012), and/or orally by modelling and discussion (Kingsley et al., 2015; Macedo-Rouet et al., 2013; Walraven et al., 2013, Zhang & Duke, 2011). Whereas sourcing was embedded in more comprehensive programs regarding information problem solving in three studies (Argelagós & Pifarré, 2012; Kingsley et al., 2015; Walraven et al., 2013), the other four more specifically focused on sourcing. All seven interventions emphasized attention to/identification of source information as an important step in evaluating a source's credibility. Two studies also explicitly accentuated the importance of linking a source with its content (Kammerer et al., 2016; Macedo-Rouet et al., 2013). None of the studies seemed to explicitly focus on how source information may help students interpret the documents' content, although some of the dependent measures indicate an anticipated learning gain.

Dependent measures. The dependent measures reported reflect that increased skills in the evaluation of source credibility are important in the above studies. Different measures are used, such as credibility justifications and the rank-ordering of documents according to their

credibility. Additionally, two studies included measures on source-content links (Kammerer et al., 2016; Macedo-Rouet et al., 2013), and learning gains in terms of essays, synthesizing information, and exams are included in four studies (Argelagós & Pifarré, 2012; Kingsley et al., 2016; Mason et al., 2014; Walraven et al., 2013).

Outcomes. Considering the results on different types of credibility measures, students in primary school seem to profit from the interventions to a varying degree. In Kingsley et al. (2015), the intervention group did not score significantly higher on the evaluation task than the control group. However, one should note that the task required both evaluating the authors' credibility and verifying the content using other websites. In the Marcedo-Rouet et al. (2013) study, only less skilled readers profited from the intervention regarding identifying source expertise on a multiple-choice test, whereas the intervention students in Zhang and Duke (2011) did not score higher than the control students in evaluating websites' trustworthiness. However, the intervention students did better when asked to give reasons for their evaluations and when rank-ordering a set of websites. In the studies including lower secondary school students, all intervention groups scored higher on the credibility measures than the control groups, though not with statistical significant differences on all measures (Kammerer et al., 2016; Walraven et al., 2013). In particular, one notes the strong ES ($g=1.84$) in the 60-hour study of Argelagós and Pifarré (2012). Regarding measures of source memory and source-content links, the Kammerer et al. (2016) study presents strong results in favor of the intervention group ($g=2.26$; $g=1.20$). Finally, the results on learning as represented in essays or summaries (Argelagós & Pifarré, 2012; Kingsley et al., 2016; Mason et al., 2014; Walraven et al., 2013) indicate that the intervention groups outperform control groups ($g=1.71$ to $g=.23$). In the case of the Argelagós and Pifarré (2012) study, one should note that the essays were scored according to the correctness of factual information. In Kingsley et al. (2015), 5th graders had to synthesize information across three websites, whereas the essays in Mason et al.

(2014) were scored for both factual information and argumentation, and Walraven et al. (2013) scored responses on ten open questions according to their correctness. Thus, only the Mason et al. (2014) study ($g=.54$) seems to take into consideration how source information may affect students' argumentation skills.

Interventions in upper secondary school

Table 3 shows that in three out of the seven studies in upper secondary schools, teachers were responsible for running the interventions over several lessons (10 hours to 50 lessons) in the domain of history (De La Paz & Felton, 2010; Nokes et al., 2007; Reisman, 2012). In these three studies, the interventions were embedded in the schools' curriculum, and in two of the studies, teachers were also involved in planning the interventions (De La Paz & Felton, 2010; Nokes et al., 2007). The four researcher-led interventions lasted for no more than two lessons (60-120 min). Two of those studies also adapted the interventions to the domain of history (Britt & Aglinskas, 2002; Goldberg et al., 2011), whereas the last two included material related to socio-scientific topics (Braasch et al., 2013; Stadler et al., 2016).

Table 3 about here

Instructions and activities. In all seven studies, the reading of multiple documents was a central part of the intervention. Although this was also the case in the majority of the above interventions at the lower levels, the number of documents available to students in upper secondary school is generally higher and thus represent more perspectives on the documents' topic. Whereas improving students' sourcing was the main aim of three studies (Braasch et al., 2013; Britt & Aglinskas, 2002; Stadler et al., 2016), the other four studies have somewhat broader goals related to expertise in history, with sourcing regarded as an integrated part of historical thinking. Thus, activities related to argumentative writing (De La Paz & Felton, 2010; Goldberg et al., 2011) and corroboration and contextualization heuristics (Nokes et al., 2007; Reisman, 2012) were also included in those interventions. The use of

worksheets, modeling and guided practice is common for these studies, whereas only Britt and Aglinskas (2002) used prompts in a computer application. In the study of Braasch et al. (2013), the principle of contrasting cases was used by having participants identify and discuss strategies used in a set of presented student-cases. Stadtler et al. (2016) specifically focused on the pertinence of a source's expertise and invited participants to discuss how that could be evaluated. During the longer interventions in history, the activities switched between cognitive modeling, group and/or whole class discussions, and practice, for example, argumentative writing (De La Paz & Felton, 2010). In all seven studies, attention to the source information was emphasized—not only as a resource for credibility evaluation but also for the interpretation of the content (Braasch et al., 2013; Britt & Aglinskas, 2002, Stadtler et al. 2016) and as an important element in argumentation and historical thinking (De La Paz & Felton, 2010; Goldberg et al., 2011; Nokes et al., 2007; Reisman, 2012).

Dependent measures. In five out of the seven intervention studies in upper secondary schools, the coding of essays was central in the operationalization of the dependent measures. Students' skills in argumentation about historical topics were coded, with such skills regarded central to historical thinking (De La Paz & Felton, 2010; Goldberg et al., 2011). Such thinking was also measured by having students respond to open questions (Nokes et al., 2007) or multiple-choice questions (Reisman, 2012) on issues requiring the use of relevant heuristics. In the Goldberg et al. (2012) study, participants read documents about a controversial historical incident on which they could be expected to have an opinion in accordance with their social identities. Thus, the essays were used to measure students' interpretations of the plot and their stances towards the incident before and after the intervention. In another study, the students' use of core scientific concepts in essays was used as a measure of whether they attended to relevant documents (Braasch et al., 2013), whereas explicit references were coded both according to whether they occurred in the essays and how the students used them in two

of the studies (De La Paz & Felton, 2010; Nokes et al., 2007). The sourcing test used in Britt and Aglinskis (2002) included open questions related to source information and credibility. The measurement of the skills in credibility evaluation was less explicit in the upper secondary sample of studies than in the studies on younger students. However, a rank-order measure was included in Braasch et al. (2013), and participants in the study of Stadtler et al. (2016) had to decide on whom to trust. It is noteworthy that only in Reisman (2012) were standardized tests included, one on world history (California Standards Test) and one reading comprehension test (Gates-MacGinitie Reading Test).

Outcomes. One should note that the control groups in all studies but one (Stadtler et al., 2016) followed regular classroom activities normally anchored in textbooks. Studies measuring the credibility and credibility justifications explicitly show positive results in favor of the intervention groups ($g=.35$ to $g=1.24$) (Braasch et al., 2013; Stadtler et al., 2016). Credibility judgments are also a central part of the sourcing test in Britt and Aglinskis (2002), but that test represents several aspects of sourcing (identification, evaluation, source-content tagging) and, as such, a broader sourcing concept. Though the samples in that series of studies are small, the results clearly indicate positive results from the intervention program ($g=.93$ to $g=1.87$). Three studies demonstrated statistically significant differences in favor of the intervention groups on post-knowledge measures ($g=.29$ to $g=1.47$) (Braasch et al., 2013; Nokes et al., 2007; Reisman, 2012). In Nokes et al. (2007), simply using multiple documents instead of textbooks seems to be the best solution, though the multiple text/heuristics group also outperformed the controls. Of special note is the moderate effect on a standardized knowledge test in Reisman (2012). That study also demonstrated improved reading comprehension for the intervention group on a standardized test ($g=.32$). The longer intervention implemented by teachers that focused on historical thinking, with sourcing as an integrated part, also demonstrated positive results. Nokes et al. (2007) showed a positive

effect for sourcing ($g=1.08$), but only weakly for students' use of documents ($g=.19$), whereas the intervention group in the De La Paz and Felton (2010) study outperformed the controls on document use ($g=.75$) and argumentation, with rebuttals representing different perspectives and sources ($g=.71$). Improved historical thinking ($g=.77$) was also shown in Goldberg et al. (2011), with an increasing "argumentation level" indicating the inclusion of multiple perspectives, use of evidence, etc. From that study, it is also interesting to note that discussions based on multiple sources representing different perspectives affected participants' interpretation of ($g=.77$) and opinion about ($g=.66$) the incident. In summary, in the studies from upper secondary schools, we find fewer results specifically on credibility than in the studies on younger students. The majority of the studies is in the domain of history and seems to see sourcing as an essential part of historical thinking, with sourcing also implying interpretation and argumentation.

Interventions in higher education

Among the four intervention studies reported here, only the Kim and Hannafin (2016) study was implemented as part of a running course and executed by teachers (biology). The three other studies were shorter and researcher-led, with two presenting documents containing medical information (Stadtler & Bromme, 2007; 2008) and one using documents containing science-related information (Wiley et al., 2009). In all four studies, the participants worked on web-based documents.

Table 4 about here

Instructions and activities. In the higher education sourcing interventions, the use of multiple documents is also central. All studies but one had students working with an inquiry task supported by a fixed set of web-documents, whereas the participants in Kim & Hannafin (2016) had unlimited access to information sources. In both the Stadtler and Bromme studies (2007; 2008), the students were provided with a computer tool (*meta.a.ware*) containing

monitoring and evaluation prompts. The monitoring prompts targeted the students' awareness of comprehension issues, whereas the evaluation prompts required students to attend to and evaluate the source credibility and validate the content. In the Kim and Hannafin (2016) study, the students used an annotation tool scaffolding explicit representations of the argumentation structure, source features, and source evaluation. One should note that both the control and intervention groups were provided a checklist for evaluating the information sources they accessed. A similar checklist (template) was provided only for students in the intervention group in Wiley et al. (2009). Additionally, the intervention group in that study had declarative information (text) on determining the source reliability at the start of the session and a rank-order task when they had finished reading the documents. Afterwards, they were provided an example of experts' ranking of the same documents. While the students in the Wiley et al. (2009) and Stadtler and Bromme (2007; 2008) studies worked individually on the inquiry tasks, participants in Kim and Hannafin (2016) worked both individually and in groups over several sessions. They first constructed a document-based argument model in the annotation tool individually and then collaborated to construct a common model.

Dependent measures. Essays were used as dependent measures in all four studies. In both Stadtler and Bromme (2007; 2008) and Wiley et al. (2009), references to sources in students' essays were used. One should note that the participants in the Stadtler and Bromme studies were explicitly instructed to name the sources of the arguments they used. In those studies, open questions were used to measure comprehension, while a multiple-choice test was used to assess factual knowledge. Furthermore, they measured source memory by asking participants to recall source features for the different documents and had them justify the documents' credibility. Wiley et al. (2009) also collected justifications related to the credibility rankings. Additionally, they looked at processing measures, specifically, how often students returned to reliable websites. Based on the essay, a system for scoring students'

conceptual models (of volcanic activity) was developed. Students' conceptual models, as represented in essays, were also measured in Kim and Hannafin (2016), as well as the evidence quality and reasoning to assess how students used sources from the annotation tool in argumentative essays.

Outcomes. In the Stadtler and Bromme (2007; 2008) studies, there is a general tendency towards a positive effect of the interventions, but few group differences are statistical significant. However, in both studies, the intervention groups exposed to evaluation prompts have statistically significant better results than the control groups in the use of sources in essays ($g=.66$; $g=1.04$) and on referring to source features while justifying the credibility ($g=.69$; $g=.66$). The Wiley et al. (2009) also shows statistically significant differences in favor of the intervention group on students' references to source information in essays ($g=.67$), as well as on credibility ranking ($g=.73$) and attention to reliable websites ($g=1.03$). That study also demonstrates that the students in the intervention groups constructed more accurate mental models of the phenomena described in the documents than the participants in the control group ($g=1.00$). In contrast, the most comprehensive intervention (Kim & Hannafin, 2016) did not demonstrate any effect on the dependent measures related to argumentation quality ($g=.06$ to $g=-.19$). The researchers do, however, report that the intervention groups scored statistically higher than the control groups on the two tasks they worked on during the intervention sessions. Hence, the individual work on the annotation tool was then followed up by work in groups, whereas the students worked only individually on the transfer task in the final test.

Discussion

Sourcing in the interventions. Our first research question concerns which aspects of sourcing were targeted in the interventions and the outcome measures. In the present review, we have defined sourcing as identifying and representing source features to predict, interpret

and evaluate documents' content and relevance according to a reading task. The task of identifying source features was more or less explicitly included in the instructional activities in all reviewed studies, whether by the use of worksheets (e.g., Walraven et al. 2013; Zhang & Duke, 2011), prompting (e.g., Britt & Aglinskias, 2002; Stadtler & Bromme, 2007;2008), or modeling (e.g., Kingsley et al, 2016; Reisman, 2012). Not all interventions, however, made participants aware of which source features they should pay attention to. Different information about a source may be more or less relevant and could be a tool for different types of interpretations depending on the nature of the reading task or the domain. In history, for example, the date of publication may facilitate contextualization and interpretation, whereas in a technology course, the date of publication may indicate that information is outdated. Thus, students should not only be instructed to pay attention to the source information, often implying the author, but also to attend to different source features of relevance. Sometimes, specifically on the Internet, they also may need to be guided on how to find different source features. Regarding the outcome measures in the reviewed studies, none seem to specifically measure participants' attention to source features, although a number implicitly do so by measuring memory for or use of sources. Those measures, however, do not show how students go about searching for source features or which features they attend to.

One could assume that the identification of source features also implies that those features will be stored in memory and potentially tagged to relevant content. That is not necessarily the case, and thus several of the studies instructed students to use scaffolds of different types (worksheets, computer applications) to register the source features and connect them to the content (e.g., Britt & Aglinskias, 2002; Kammerer et al. 2016; Kim & Hannafin, 2016). Within the Documents Model Framework (Perfetti et al., 1999), this source-content tagging is an important condition for constructing a mental intertext model, with such a model presumably facilitating the readers' organization of information from multiple documents and

hence text comprehension. The memory for source features is measured in several ways in the reviewed studies, such as the recall of source information (Kammerer et al., 2016), tagging source features to content (Macedo-Rouet et al., 2013; Stadler & Bromme, 2007; 2008; Stadler et al., 2016), and references to sources in essays (De La Paz & Felton, 2010; Nokes et al., 2007; Wiley et al., 2009). Recalling source features is a simpler task than connecting those features to specific content, whereas references in essays do not necessarily show all source features that students remember. In the context of multiple-documents literacy, source-content measures will be the best indicators of the role of source memory in interpretation and evaluation.

Sourcing in terms of using source information to predict content is identified as an important skill among experts (e.g., Lundeberg, 1987; Wineburg, 1991). Regarding the prediction of content based on source features, several of the studies more or less implicitly include such skills when the students are instructed to select documents based on source information. However, credibility is most often highlighted as a criterion for selecting documents (e.g., Argelagós & Pifarré, 2012; Kingsley et al., 2015), whereas relevance is less often emphasized (but see Walraven et al., 2013). Thus, when predicting a document's content based on source information, the students should learn to consider both the credibility and relevance. In the measures used that indicate prediction skills, an evaluation of hit lists –, e.g., a Search Engine Results Page (SERP) – according to the documents' potential credibility was used (e.g., Mason et al., 2014; Wiley et al., 2009), as well as according to less specific criteria like “useful” (Braasch et al., 2013) or “appropriate” (Walraven et al., 2013).

The role of sourcing in the interpretation of documents is most salient in interventions within the domain of history. In those studies, source features are regarded less as indications of a source's reliability and more as indications of a source's perspective (e.g., De La Paz & Felton, 2010; Reisman, 2012). Hence, teachers model how one may reflect on a source's

position, the significance of when and where a document was published, and the author's potential intentions. This has typically been done by the use of modeling, worksheets, group discussions, and essay writing (De La Paz & Felton, 2010; Goldberg et al., 2011; Nokes et al., 2007; Reisman, 2012). However, studies related to the socio-scientific/scientific domains have also emphasized that sourcing should facilitate interpretation (Braasch et al., 2013; Mason et al., 2014; Wiley et al., 2009). In the Braasch et al. study, the instructional design implied individual work, as well as group and whole-class discussion related to the identification of good sourcing strategies in student cases, whereas both Mason et al. and Wiley et al. had students work with declarative information about sourcing and with worksheets. In all the above studies, essays constitute the dependent measures, with conceptual understanding (Braasch et al., 2013; Wiley et al., 2009) or argumentation skills (De La Paz & Felton, 2010; Goldberg et al., 2011; Reisman, 2012) measured, although the outcome measures are not specifically measuring sourcing. In history, however, students' use of source information is regarded a central dimension in argumentation skills.

Finally, evaluation of source credibility was a central part in the majority of the reviewed studies and even more salient in interventions at the elementary and lower secondary than the other grade levels. Modeling, group discussions, and worksheets were common ways of working in those interventions (e.g., Kingsley et al., 2015; Macedo-Rouet et al., 2013; Mason et al., 2014; Zhang & Duke, 2011), and the outcome measures were typically the rank-order of documents and responses on credibility scales for single documents. The distinction between expertise and trustworthiness made by Hovland et al. (1953) is not highlighted in those measures, although some asked students to evaluate trustworthiness (e.g., Zhang & Duke, 2011) and others emphasized the author's expertise (e.g., Stadtler et al., 2016). Several studies also asked for students' justifications of evaluation of source credibility (e.g., Braasch et al., 2013; Mason et al., 2014; Stadtler et al., 2007; 2008; 2016), which could

demonstrate more in detail which criteria students used. Hendriks, Kienhues, and Bromme (2015) suggest an interesting distinction between credibility, as conceptualized in persuasion research, and epistemic trustworthiness, with the latter concept more closely linked to evaluation of expertise in knowledge domains. They developed and tested an inventory including three dimensions; expertise, integrity, and benevolence. Given that educational interventions are more concerned about epistemic development than persuasion, the Hendriks et al. (2015) inventory could be considered a relevant measure for future studies.

In summary, the different aspects of sourcing we have included in our definition are all represented in the sample of intervention studies reviewed above. The use of source features in the interpretation of document content was, however, only represented in some of the studies, primarily at higher grade levels, and was the skill most difficult to identify in the measures. Additionally, we note that far from all studies emphasized the role of task relevance in the instructional activities. A source may be credible but not relevant, and students should be made aware of that.

Improved outcomes. Our second research question concerns what features of instructional interventions are associated with improved outcomes. The use of multiple, partly conflicting documents seems to be a common feature of sourcing interventions. In the study of Nokes et al. (2007), just introducing multiple documents improved students' results on a content test considerably. There are reasonable arguments in favor of exposing students to multiple documents that represent different perspectives during their training on sourcing skills. First, multiple documents representing the same perspective do not indicate that one needs to look for source features that indicate differences between them. Second, one single document representing one perspective does not represent a strong incentive to look for source features unless that perspective strongly opposes the reader's point of view. Third, if students need to complete a task by reading multiple, partly contradicting documents, it seems

more reasonable to reflect on the source features to solve contradictions. However, readers with significant expertise on the topic in question could address this contradiction by judging the extent to which the content of the documents is reasonable. Thus, laypersons faced with contradictions could be more dependent on source features than experts (Bromme & Goldman, 2014). Finally, although different perspectives may also be presented in a single text (Macedo-Rouet et al., 2013), this format represents a poorer pool of source features. Therefore, different perspectives presented in multiple documents seem to be more appropriate for training in sourcing skills than the same perspectives presented in a single document.

It is worth noting that in the majority of the studies (16 out of 18), documents included in the training session/s were pre-selected by researchers in order to represent variation in perspectives, credibility, source features, and relevance. In two studies participants were presented pre-selected documents only in the initial phase, whereas they during the main parts of the intervention searched for useful information sources on the Internet (Kim & Hannafin, 2016; Walraven et al., 2013). One potential challenge with the free search approach is that participants were not necessarily exposed to documents representing the variation included in pre-selected sets of documents. Also, in the Kim and Hannafin (2016) study it implied that common documents were not discussed in the groups. Hence, the absence of improved sourcing skills among participants in the two studies could be partly due to the lack of common pre-selected document sets in the interventions. Future studies should investigate whether interventions including a set of diverse pre-selected documents demonstrate better results than interventions based on participants' search for and use of information sources at the Internet.

One more reason for including pre-selected document sets in sourcing interventions is the possibility to adapt the documents to participants' levels of prior knowledge and reading

skills (e.g., Kammerer et al., 2015; Macedo-Rouet et al., 2013; Mason et al., 2014). Prior studies have showed a relationship between multiple documents comprehension and both prior knowledge (Bråten et al., in press) and reading skills (e.g., Bråten, Ferguson, Anmarkrud, & Strømsø, 2013), and requiring students to apply sourcing strategies on documents beyond their level of domain knowledge and skills may prove to be too demanding.

Another common feature of the majority of the intervention studies in upper secondary and higher education was the use of inquiry tasks. What constitutes inquiry tasks has been debated (e.g., Chinn & Malhotra, 2002; Loyens & Rikers, 2011), but it seems to be agreed that some types of research questions are the point of departure for those tasks. Searching for, selecting and reading textual resources are often part of that process. Thus, evaluating, reflecting on, and understanding multiple texts will often be part of students' work with inquiry tasks. For example, Mason et al. (2014) asked students to consider whether using a cell phone may be harmful to a user's health, whereas Reisman (2012) asked students the following question: "Why did the Homestead strike of 1892 turn violent?" When investigating such questions by approaching textual resources, students are encouraged to also consider the nature and quality of texts based on certain criteria. In other words, what counts as knowledge in this field? It has been argued that students learn the "ins and outs of scientific processes" (Loyens & Rikers, 2011, p 362)—i.e., a subject's epistemology—when working on inquiry tasks, and sourcing is evidently a part of that epistemology. Therefore, such tasks should constitute appropriate tools for improving students' sourcing.

There are indications that readers' beliefs about knowledge and how people come to know in different domains might be related to how they consider and evaluate source information (Barzilai et al., 2015; Bråten et al., in press). Hence, students' credibility evaluations could, for example, rely on what they know about how expertise is acquired in a certain domain. We note that different kinds of credibility measures are dominating the

sample of dependent measures in lower grades interventions, whereas other measures indicating beliefs about how knowledge develop and should be interpreted are added at higher grade levels. Although the assumption about developmental trajectories in students' epistemic cognition is a debated issue (Hofer, 2016), we believe a constrained focus on source credibility in lower grades is wise as other sourcing skills, such as using source features in predicting and interpretation of information, will require more domain knowledge on the part of the reader.

Some of the short researcher-led interventions demonstrated convincing effect sizes by introducing rather simple scaffolds (information and work sheets) on a number of outcome measures (Kammerer et al., 2016; Mason et al., 2014; Wiley et al., 2009). Scaffolds have been successfully included in several of the longer teacher-led implementations (e.g., Argelagós & Pifarré, 2012; De La Paz & Felton, 2010) but also in some not so successful teacher-led studies (Kim & Hannafin, 2016; Walraven et a., 2013). Regarding the Kim and Hannafin study, we note that both the control and intervention students worked on a simple checklist. Given the positive results demonstrated by using simple scaffolds, it may be that a more sophisticated computer application does not represent sufficient added support for intervention students to outperform control students. The field notes collected in Walraven et al. indicate that the intervention was not executed as intended, due to both students' lack of motivation and teachers' administration of the schedule.

Though several of the interventions lasted for some time (e.g. Argelagós & Pifarree, 2012; De La Paz & Felton, 2010; Kim & Hannafin, 2016; Reisman, 2012; Walraven et al., 2013), none of them tested potential long-term effects of the interventions. Admittedly, not all studies have details about the time points for post-tests, but the described procedures indicate that only five studies have a somewhat (1 week) delayed post-test (Britt & Aglinskias, 2002; Mason et al. 2014; Walraven et al. 2013; Wiley et al., 2009; Zhang & Duke, 2011). Hence,

studies including measures of long-term effects of sourcing interventions are needed. A potential indication of such effects is the use of transfer tasks as post-tests in 14 of the 18 reviewed studies.

Regarding the students' motivation, we are slightly surprised that a majority of the studies do not seem to explicitly (however, see Britt & Aglinskias, 2002) take into consideration findings from persuasion theory and information science indicating that credibility assessment requires ability and motivation (Metzger, 2007; Petty & Briñol, 2012; Petty & Cacioppo, 1990). In a recent paper by List and Alexander (2017) this point is elaborated in the context of multiple source use, and they develop a model balancing the importance of both ability and motivation. Specifically they emphasize how affective variables like interest and attitude in concert with ability may affect sourcing. According to Petty and Cacioppo (1990) perceived personal relevance is of great significance if people are to scrutinize all relevant information when considering a message. Hence, they may be able to critically evaluate the message but are not motivated to do so. This was recently demonstrated in an interview study on 9th graders perspectives on sourcing (Paul, Macedo-Rouet, Rouet, & Stadtler, 2017). Participants' demonstrated knowledge about sourcing strategies, but most often refrained to use them in school contexts due to lack of motivation (lack of interest and/or external incentives). One main reason for non-sourcing was lack of encouragement from teachers. This may imply that results from seemingly successful researcher-led interventions are not sustainable when students are back in their regular classrooms (Bråten et al., in press). Although teacher-led interventions may be more challenging (e.g., Walraven et al., 2013), our review show that six out of eight teacher-led interventions demonstrated significant effect sizes. Due to the lack of delayed post-tests, we are not able to tell whether teacher or researcher-led interventions produce more or less sustainable results. However, based on theoretical assumptions regarding the need for motivation and relevance, as well as

results from Paul et al., (2017), we suggest that the more ecological valid interventions increase the possibilities of successful sourcing interventions.

Our review has several limitations. Terminology issues might have restricted our conceptualization of sourcing. In other words, sourcing is not necessarily the term used when researchers study how readers take note of and use source features, and thus, our search for intervention studies may be too restricted. Other keywords could have increased the sample, although we carefully considered relevant keywords. Another limitation is the variety in participants' ages, intervention lengths and measurements. However, we do note that sourcing interventions seem to work for participants from 4th grade to university and across numerous subjects. The sample of studies also represents short- and long-term interventions, with even very short interventions demonstrating positive results. Although the long-term effects are unknown, we believe that the current results point to some elements of sourcing interventions that should be emphasized both in schools and in further studies. Thus, we encourage future interventions to, in line with the OECD's apprehension of advanced literacy, use multiple documents with different perspectives, separate between sources' credibility and relevance and *explicitly* instruct in all functions of sourcing.

References:

- 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. doi: 10.1016/j.lindif.2013.01.007
- Argelagós, E., & Pifarré, M. (2012). Improving information problem solving skills in secondary education through embedded instruction. *Computers in Human Behavior*, 28, 515-526. doi:10.1016/j.chb.2011.10.024

- Barzilai, S., & Eshet-Alkalai, Y. (2015). The role of epistemic perspectives in comprehension of multiple author viewpoints. *Learning and Instruction, 36*, 86-103.
doi:10.1016/j.learninstruc.2014.12.003
- Barzilai, S., Tzadok, E., & Eshet-Alkalai, Y. (2015). Sourcing while reading divergent expert accounts: Pathways from views of knowing to written argumentation. *Instructional Science, 43*, 737-766. doi:10.1007/s11251-015-9359-4
- Bazerman, C. (1985). Physicists reading physics: Schema-laden purposes and purpose-laden schema. *Written Communication, 2*, 3-23. doi:10.1177/0741088385002001001
- Behrman, E. H. (2006). Teaching about Language, Power, and Text: A Review of Classroom Practices that Support Critical Literacy, *Journal of Adolescent & Adult Literacy, 49*, 490-498. doi:10.1598/JAAL.49.6.4
- Braasch, J. L., Bråten, I., Strømsø, H. I., Anmarkrud, Ø., & Ferguson, L. E. (2013). Promoting secondary school students' evaluation of source features of multiple documents. *Contemporary Educational Psychology, 38*(3), 180-195. doi: 10.1016/j.cedpsych.2013.03.003
- Braasch, J.L.G., Rouet, J.F., Vibert, N., & Britt, M.A. (2012). Readers' use of source information in text comprehension. *Memory & Cognition, 40*, 450-465. doi: 10.3758/s13421-011-0160-6
- Bråten, I., Ferguson, L., Anmarkrud, Ø., & Strømsø, H. (2013). Prediction of learning and comprehension when adolescents read multiple texts: The roles of word-level processing, strategic approach, and reading motivation. *Reading and Writing, 26*, 321-348. doi: 10.1007/s11145-012-9371-x
- Bråten, I., Stadtler, M., & Salmerón, L. (in press). The role of sourcing in discourse comprehension. In M. F. Schober, D. N. Rapp & M. A. Britt (Eds.), *Handbook of Discourse Processes* (2nd ed.). New York: Routledge.

- Bråten, I., Strømsø, H. I., & Andreassen, R. (2016). Sourcing in professional education: Do text factors make any difference? *Reading and Writing, 29*, 1599-1628. doi: 10.1007/s11145-015-9611-y
- Bråten, I., Strømsø, H. I., & Britt, M. A. (2009). Trust matters: Examining the role of source evaluation in students' construction of meaning within and across multiple texts. *Reading Research Quarterly, 44*, 6-28. doi: 10.1598/RRQ.41.1.1
- Brem, S. K., Russell, J., & Weems, L. (2001). Science on the Web: Student evaluations of scientific arguments. *Discourse Processes, 32*, 191-213. doi: 10.1080/0163853X.2001.9651598
- Britt, M. A., & Aglinskias, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction, 20*, 485-522. doi: 10.1207/S1532690XCI2004_2
- Britt, M. A., Perfetti, C. A., Sandak, R., & Rouet, J.-F. (1999). Content integration and source separation in learning from multiple texts. In S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 209-233). Mahwah, NJ: Erlbaum.
- Britt, M. A., & Rouet, J.-F. (2012). Learning with Multiple Documents: Component skills and their acquisition. In M. J. Lawson & J. R. Kirby (Eds.), *Enhancing the Quality of Learning* (pp. 276-314). New York: Cambridge University Press.
- Bromme, R., & Goldman, S.R. (2014). The public's bounded understanding of science. *Educational Psychologist, 49*, 59-69. doi:10.1080/00461520.2014.921572
- Bromme, R., Scharrer, L., Stadler, M., Hömberg, J., & Torspecken, R. (2015). Is it believable when it's scientific? How scientific discourse style influences laypeople's resolution of conflicts. *Journal of Research in Science Teaching, 52*, 36-57. doi: 10.1002/tea.21172

- Chen, S., & Chaiken, S. (1999). The Heuristic-Systematic Model in its broader context. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 73-96). New York: Guilford.
- Chinn, C.A., & Malhotra, B.A. (2002). Epistemologically authentic inquiry in Schools: A theoretical framework for evaluating inquiry tasks. *Science Education*, 86, 175-218. doi: 10.1002/sce.1001
- 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, 174-192. doi:10.1016/j.cedpsych.2010.03.001
- Durlak., J. A. (2009). How to select, calculate, and interpret effect sizes. *Journal of Pediatric Psychology*, 34, 917-928. doi:10.1093/jpepsy/jsp004
- Gibbs, R.W., Kushner, J.M., & Mills, W.R. (1991). Authorial intentions and metaphor comprehension. *Journal of Psycholinguistic Research*, 20, 11-29. doi: 10.1007/BF01076917
- Goldberg, T., Schwarz, B. B., & Porat, D. (2011). "Could They Do It Differently?": Narrative and Argumentative Changes in Students' Writing Following Discussion of "Hot" Historical Issues. *Cognition and Instruction*, 29, 185-217. doi: 10.1080/07370008.2011.556832
- Goldman, S.R., Braasch, J.L.G., Wiley, J., Graesser, A.C., & Brodowinska, K. (2012) Comprehending and learning from internet sources: Processing patterns of better and poorer learners. *Reading Research Quarterly*, 47, 356-381. doi: 10.1002/RRQ.027
- Goldman, S.R., & Scardamalia, M. (2013). Managing, understanding, applying, and creating knowledge in the information age: next-generation challenges and opportunities. *Cognition and Instruction*, 31, 255-269. doi: 10.1080/10824669.2013.773217

- Gottlieb, E., & Wineburg, S. (2012). Between Veritas and Communitas: Epistemic switching in the reading of academic and sacred history. *Journal of the Learning Sciences, 21*, 84-129. doi: 10.1080/10508406.2011.582376
- Harris, P. L., & Corriveau, K. H. (2011). Young children's selective trust in informants. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 366*(1567), 1179-1187. doi: 10.1098/rstb.2010.0321
- Hendriks, F., Kienhues, D., and Bromme, R. (2015). Measuring laypeople's trust in experts in a digital age: The Muenster Epistemic Trustworthiness Inventory (METI). *PLoS ONE 10*(10): e0139309. doi:10.1371/journal.pone.0139309
- Hofer, B. K. (2016). Epistemic cognition as a psychological construct. In J. A. Greene, Sandoval, W. A., & Bråten, I. (2016). *Handbook of epistemic cognition* (pp. 19-38). New York: Routledge.
- Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion. Psychological studies of opinion change*. New Haven, CT: Yale University Press.
- 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.
<http://dx.doi.org/10.1016/j.compedu.2016.07.001>
- Kim, S. M., & Hannafin, M. J. (2016). Synergies: effects of source representation and goal instructions on evidence quality, reasoning, and conceptual integration during argumentation-driven inquiry. *Instructional Science, 44*, 441-476. doi: 10.1007/s11251-016-9381-1
- Kingsley, T. L., Cassady, J. C., & Tancock, S. M. (2015). Successfully Promoting 21st Century Online Research Skills: Interventions in 5th-Grade Classrooms. *Reading Horizons (Online), 54*, 0_1,0_2,92-134.

- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.
- Kintsch, W., & Rawson, K. (2007). Comprehension. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 209-226). Malden, MA: Blackwell publishing.
- Le Bigot, L., & Rouet, J.-F. (2007). The impact of presentation format, task assignment, and prior knowledge on students' comprehension of multiple online documents. *Journal of Literacy Research, 39*, 445-470. doi: 10.1080/10862960701675317
- List, A., & Alexander, P. A. (2017). Cognitive affective engagement model of multiple source use. *Educational Psychologist, 52*, 182-199. doi: 10.1080/00461520.2017.1329014
- Loyens, S.M.M., & Rikers, R.M.J.P. (2011). Instruction based on inquiry. In R.E. Mayer, & P.A. Alexander (Eds.). *Handbook of research on learning and instruction* (pp 361-381). New York: Routledge.
- Lundeberg, M.A. (1987). Metacognitive aspects of reading comprehension: Studying understanding in legal case analysis. *Reading Research Quarterly, 22*, 407-432. doi:10.2307/747700
- Macedo-Rouet, M., Braasch, J. L., Britt, M. A., & Rouet, J.-F. (2013). Teaching fourth and fifth graders to evaluate information sources during text comprehension. *Cognition and Instruction, 31*, 204-226. doi:10.1080/07370008.2013.769995
- Mason, L., Junyent, A. A., & Tornatora, M. C. (2014). Epistemic evaluation and comprehension of web-source information on controversial science-related topics: Effects of a short-term instructional intervention. *Computers & Education, 76*, 143-157. doi:10.1016/j.compedu.2014.03.016
- 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, 2078-2091. doi: 10.1002/asi.20672
- Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and Heuristic Approaches to Credibility Evaluation Online. *Journal of Communication*, 60, 413–439.
doi:10.1111/j.1460-2466.2010.01488.x
- Nokes, J.D., Dole, J.A., & Hacker, D.J. (2007). Teaching high school students to use heuristics while reading historical texts. *Journal of Educational Psychology*, 99, 492-504. doi.org/10.1037/0022-0663.99.3.492
- OECD. (2014). *Education at a Glance: OECD indicators. 2014*. Retrieved from <http://www.oecd.org/edu/Education-at-a-Glance-2014.pdf>
- Paul, J., Macedo-Rouet, M., Rouet, J.-F., & Stadler, M. (2017). Why attend to source information when reading online? The perspective on ninth grad students from two different countries, *Computers & Education*, 113, 339-354. doi: 10.1016/j.compedu.2017.05.020
- Paxton, R. J. (2002). The influence of author visibility on high school students solving a historical problem. *Cognition and Instruction*, 20, 197-248.
doi: 10.1207/S1532690XCI2002_3
- Perfetti, C. A., Rouet, J.-F., & Britt, M. A. (1999). Toward a theory of documents representation. In H. V. Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 88-108). Mahwah, N.J.: Erlbaum.
- Petty, R. E., & Briñol, P. (2012). The Elaboration Likelihood Model. In P.A.M. Van Lange, A. Kruglanski, & E.T. Higgins (Eds.), *Handbook of theories of social psychology* (Vol. 1, pp. 224-245). London: Sage.
- Petty, R.E., & Cacioppo, J.T. (1990). Involvement and persuasion: Tradition versus integration. *Psychological Bulletin*, 107, 367-374.

- Petty, R.E., & Wegener, D.T. (1999). The Elaboration Likelihood Model: Current status and controversies. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 41-72). New York: Guilford.
- Reisman, A. (2012). Reading like a historian: A document-based history curriculum intervention in urban high schools. *Cognition and Instruction, 30*, 86-112.
doi: 10.1080/07370008.2011.634081
- Rieh, S. Y. (2002). Judgment of information quality and cognitive authority in the web. *Journal of the American Society for Information Science and Technology, 53*, 145-161.
doi: 10.1002/asi.10017
- Rieh, S. Y., & Hilligoss, B. (2007). College Students' Credibility Judgments in the Information-Seeking Process. In M. J. Metzger & A. J. Flanagin (Eds.), *Digital Media, Youth, and Credibility* (pp. 49–72). Cambridge, MA: The MIT Press.
- Salmerón, L., Kammerer, Y., & García-Carrión, P. (2013). Searching the Web for conflicting topics: Page and user factors. *Computers in Human Behavior, 29*, 2161-2171.
doi: 10.1016/j.chb.2013.04.034
- Salmerón, L., Macedo-Rouet, M., & Rouet, J. F. (2015). Multiple viewpoints increase students' attention to source features in social question and answer forum messages. *Journal of the Association for Information Science and Technology*.
doi:10.1002/asi.23585
- Stadtler, M., & Bromme, R. (2007). Dealing with multiple documents on the WWW: The role of metacognition in the formation of documents models. *International Journal of Computer-Supported Collaborative Learning, 2*, 191-210.
doi:10.1007/s11412-007-9015-3

- Stadtler, M., & Bromme, R. (2008). Effects of the metacognitive computer-tool met. a. ware on the web search of laypersons. *Computers in Human Behavior*, 24(3), 716-737.
<http://dx.doi.org/10.1016/j.chb.2007.01.023>
- Stadtler, M., Scharrer, L., Macedo-Rouet, M., Rouet, J. F., & Bromme, R. (2016). Improving vocational students' consideration of source information when deciding about science controversies. *Reading and Writing*, 29(4), 705-729.
doi:10.1007/s11145-016-9623-2
- Stiff, J. B., & Mongeau, P. A. (2003). *Persuasive communication*. New York: Guilford.
- Strømsø, H. I. & Bråten, I. (2002). Norwegian law students' use of multiple sources while reading expository texts. *Reading Research Quarterly*, 37, 208-227. doi:
10.1598/RRQ.37.2.5
- Strømsø, H. I., Bråten, I., & Britt, M. A. (2010). Reading multiple texts about climate change: The relationship between memory for sources and text comprehension. *Learning and Instruction*, 20, 192-204. doi: 10.1016/j.learninstruc.2009.02.001
- Strømsø, H. I., Bråten, I., Britt, M. A., & Ferguson, L. E. (2013). Spontaneous sourcing among students reading multiple documents. *Cognition and Instruction*, 31, 176-203.
doi: 10.1080/07370008.2013.769994
- Walraven, A., Brand-Gruwel, S., & Boshuizen, H. P. (2008). Information-problem solving: A review of problems students encounter and instructional solutions. *Computers in Human Behavior*, 24, 623-648. doi: 10.1016/j.chb.2007.01.030
- Walraven, A., Brand-Gruwel, S., & Boshuizen, H. P. (2008). Information-problem solving: A review of problems students encounter and instructional solutions. *Computers in Human Behavior*, 24, 623-648. <http://dx.doi.org/10.1016/j.chb.2007.01.030>
- Walraven, A., Brand-Gruwel, S., & Boshuizen, H. P. (2013). Fostering students' evaluation behaviour while searching the internet. *Instructional science*, 41, 125-146.

doi: 10.1007/s11251-012-9221-x

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*, 1060-1106.

doi:10.3102/0002831209333183

Wineburg, S. 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*, 73. doi:10.1037/0022-0663.83.1.73

Wineburg, S. (1994). The cognitive representation of historical texts. In G. Leinhardt, I. L. Beck, & C. Stainton (Eds.), *Teaching and learning in history* (pp. 85-135). Hillsdale, NJ: Erlbaum.

Wineburg, S., & Reisman, A. (2015). Disciplinary literacy in history. A toolkit for digital citizenship. *Journal of Adolescent & Adult Literacy*, *58*, 636-639. doi:10.1002/jaal.410

Wolfe, M.B.W., & Woodwyk, J.M. (2010). Processing and memory of information presented in narrative or expository texts. *British Journal of Educational Psychology*, *80*, 341-362. doi: 10.1348/000709910X485700

Zhang, S., & Duke, N. K. (2011). The impact of instruction in the WWWDOT framework on students' disposition and ability to evaluate web sites as sources of information. *The Elementary School Journal*, *112*, 132-154. doi:10.1086/660687

Zwaan, R.A. (1994). Effects of genre expectations on text comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 920-933.

doi: 10.1037/0278-7393.20.4.920

Table 1

Eligibility criteria

Category	Keywords
Population	children; youth; adolescents; students; pupils
Setting	school; classroom; social science; natural science; information science
Treatment method	intervention; training; experiment; programme; treatment; evaluation
Treatment objective	source evaluation; critical literacy; reading comprehension; search strategies; sourcing;
Variable	Knowledge (ability)

Table 2
Interventions in elementary and lower secondary school

Study	Length of intervention	Activities and instructions related to sourcing in the interventions	Dependent measures	Effect size (Hedges' <i>g</i>)
Argelagós & Pifarré (2012) 7th and 8th graders (N=40)	60 hours	Scaffolds facilitating information problem solving (IPS) were embedded in authentic learning tasks. Students received structured web-based instructions prompting them to define the problem, search for information, <i>identify</i> a site's author and purpose, and present information.	Searching for information Scanning-processing information Selected results evaluation (usability & reliability) Task performance (essay)	.35 ^a ns .83 ^a ns 1.84 ^a 1.71 ^a
Kammerer et al. (2016) 9th graders (N=45)	40 min	Students prompted via worksheet to <i>attend to</i> and <i>tag</i> content to source by classifying websites according to category.	Memory for sources Source-to-content links Credibility official institution websites Credibility journalistic websites Credibility private websites Credibility commercial website	2.26 1.20 .32 ns -.78 -.12 ns -.60
Kingsley et al. (2015) 5th graders (N=418)	10 hours	Students taught via modeling and guided practice to <i>locate</i> source information, <i>question</i> author, <i>checking</i> accuracy of information, and <i>detect</i> bias by triangulating data. In total 13 lessons, with each including teacher modeling, guided practice, and independent inquiry.	Online research skills Sub skills: Locating information Evaluating information Synthesizing information	.27 .34 .10 ns .23
Macedo-Rouet et al. (2013) 4th and 5th graders	30 min	Students guided in discussions to <i>identify</i> the sources present in the texts, to <i>identify</i> who said what and to <i>assess</i>	Source-content questions (less skilled readers) Source-content questions (skilled	.76 -.01 ns

(N=96)		the knowledge level of the source with respect to the topic of the text.	readers) Source evaluation – less skilled Source evaluation –skilled	.60 .03 ns
Mason et al. (2014) 9th graders (N=134)	60 min	Students <i>read</i> declarative knowledge on what to <i>consider</i> when <i>assessing</i> how reliable websites are; answered questions about the author’s knowledge and motivation. Used worksheets during practice task.	Time on least reliable sites Rank-order according to reliability Justification criteria most reliable sites Justification criteria least reliable sites Argumentation in essay	.51 .62 .39 .37 .54
Walraven et al. (2013) 9th graders (N=100)	15 x 50 min	Intervention started with a confrontation lesson, then students read declarative information about information problem solving (IPS). Then students <i>attended to</i> source features via process worksheet with questions as When was the site last updated, What does the author of the information want to achieve?	Instruction domain - Evaluation of hit list - Evaluation of websites Transfer domain - Evaluation of hit list - Evaluation of websites Learning (final exam)	1.12 ^a .64 ^a 1.33 ^a .62 ^a ns .45
Zhang & Duke (2011) 4th and 5th graders (N=242)	4 x 30 min	Students <i>encouraged</i> via worksheets to <i>pay attention to</i> who wrote this, why did they write it and when was it written? The students should <i>check</i> available source information and <i>be aware</i> of publishing purposes.	Website evaluation questionnaire Website judgments (trustworthiness) Website judgments (trust reasons) Website rank-order	.79 ^a -.03 ^a 1.01 ^a .97 ^a

Note: ^a: Adjusted ES; ns: Non significant

Table 3
Interventions in upper secondary school

Study	Length of intervention	Activities and instructions related to sourcing in the interventions	Dependent measures	Effect size (Hedges' <i>g</i>)
Braasch et al. (2013) 13th graders (N=130)	60 min	Contrasting student cases were used to discuss and identify strategies <i>directing students' attention towards</i> source characteristics, and <i>to check and consider</i> the best way to <i>evaluate</i> source characteristics.	Scientific concepts in essays Rank-order discrimination Rank-order justifications: - Author - Credentials - Venue - Type - Date	.34 .37 1.22 .77 .63 1.24 .89
Britt & Aglinskas (2002) Study 1: Undergr. (N=49) Study 2a: 11th gr (N=15) Study 2b: 11th gr (N=29) Study 3: 11th gr (N=23)	Study 1: 35 min Study 2-3: 2 x 40 min	Study 1: Instructed to <i>pay attention to</i> and <i>evaluate</i> source information when reading printed multiple documents. Study 2-3: Students prompted in computer application to <i>attend to, notice</i> and <i>identify</i> source features and then fill in structured note cards with relevant source information.	Study 1: Sourcing test Study 2a: Sourcing test Study 2b: Sourcing test Study 3: Sourcing test Sourcing test included: - Identification of source features - Source evaluation - Memory for central content	.28 1.66 ^a .93 ^a 1.87 ^a
De La Paz & Felton (2010) 11th graders (N=160)	Instructions spread over several weeks	Students instructed via modeling and worksheets to <i>analyze</i> sources, for example <i>consider</i> the author (position, knowledge), the date the document was published, <i>identify</i> the form of document and <i>understand</i> the underlying assumptions in the document.	Document use Claims in essay Rebuttals in essay	.75 ^a .36 ^a ns .71 ^a

Goldberg et al. (2011) 12th graders (N=64)	90-120 min	Students taught via modeling to <i>question</i> the author, the context and the rhetoric in historical sources/documents. Students were instructed to convince each other in small group discussion, and next write an argumentative essay.	Change in plot scheme Change in stand Argument level	.76 .65 .77
Nokes et al. (2007) 11th graders (N=246)	10 hours	The intervention included three heuristics; sourcing, corroboration, and contextualization. The control group had traditional content-focused instruction with the use of a textbook. Regarding the sourcing heuristic, students in the intervention groups were instructed via modeling to <i>look at</i> source information before reading, <i>keep source in mind</i> while reading, and <i>consider</i> source information to <i>make sense</i> of the document. Students then practiced the heuristic in small groups.	Content test: - Control vs Multiple texts/content - Control vs Textbook/heuristics - Control vs Multiple texts/heuristics Sourcing: - Control vs Multiple texts/content - Control vs Textbook/heuristics - Control vs Multiple texts/heuristics Corroboration: - Control vs Multiple texts/content - Control vs Textbook/heuristics - Control vs Multiple texts/heuristics Using documents: - Control vs Multiple texts/content - Control vs Textbook/heuristics - Control vs Multiple texts/heuristics	1.47 ^a .12 ^a ns .56 ^a .34 ^a ns .25 ^a ns 1.08 ^a -.25 ^a ns .31 ^a ns .85 ^a .06 ^a ns -.07 ^a ns .19 ^a ns
Reisman (2012) 11th graders (N=182)	36-50 lessons	Students instructed via modeling and guided practice, and worked on historical questions in small groups and whole	Historical thinking Transfer of Historical thinking Factual knowledge	.20 ^a .55 .29

		class discussions. Reading was guided by questions to <i>identify</i> author's position on historical event, <i>identify and evaluate</i> author's purpose in producing document, <i>predict</i> what author will say BEFORE reading document.	Reading comprehension	.32 ^a
Stadtler et al. (2016) Vocational students (N=56)	90 min	Introductory presentation about the "division of cognitive labor". Next, instructors <i>activated</i> students' prior knowledge about how people acquire expertise. Students <i>practiced</i> individually and in groups to <i>identify</i> the source with expertise that is pertinent to the topic at hand.	Agreement with expert claim - Asymmetrical Judgments supported by references - Symmetrical - Asymmetrical Memory for sources - Symmetrical - Asymmetrical Symmetrical: conflict between two pertinent experts. Asymmetrical: conflict between pertinent and low-pertinent experts.	.93 ^a .35 ^a ns .63 ^a ns .00 ^a ns .28 ^a ns

Note: ^a: Adjusted ES; ns: Non significant

Table 4
Interventions in higher education

Study	Length of intervention	Activities and instructions related to sourcing in the interventions	Dependent measures	Effect size (Hedges' <i>g</i>)
Kim & Hannafin (2016) Undergraduates (N=482)	Four weeks embedded in a large-enrollment course	Both control and intervention groups used a checklist with evaluation criteria. Students in intervention groups also used an online annotation tool to <i>record, examine</i> and <i>reflect</i> on source representations within argumentation structure – both individually and in groups. Group task was either to persuade or to reason balanced.	Argumentation quality <ul style="list-style-type: none"> - Evidence quality - Reasoning - Conceptual integration 	.06 ^a ns -.16 ^a ns -.19 ^a ns
Stadtler & Bromme (2007) Undergraduates (N=79)	40 min	Students were asked to conduct an Internet research on a specific topic and were prompted via a computer tool to <i>indicate</i> source information, <i>rate</i> author's credentials and author's possible bias. (Control condition: plain text window)	Comprehension <ul style="list-style-type: none"> - Control vs Evaluation prompts - Control vs Monitoring prompts - Control vs Evaluation + Monitoring Source memory <ul style="list-style-type: none"> - Control vs Evaluation prompts - Control vs Monitoring prompts - Control vs Evaluation + Monitoring Sourcing (in essay) <ul style="list-style-type: none"> - Control vs Evaluation prompts - Control vs Monitoring prompts - Control vs Evaluation + Monitoring Justification of credibility by source <ul style="list-style-type: none"> - Control vs Evaluation prompts - Control vs Evaluation + 	.16 ns .42 ns .17 ns .59 ns -.31 ns .50 ns .61 ns .12 ns .66 .69 .57 ns

			Monitoring	
Stadtler & Bromme (2008) Undergraduates (N=118)	40 min	Students were asked to conduct an Internet research on a specific topic and were prompted via a computer tool to <i>indicate</i> source information, <i>rate</i> author's credentials and author's possible bias. (Control condition: paper & pencil)	Factual knowledge - Control vs Monitoring prompts - Control vs Evaluation + Monitoring - Control vs Evaluation prompts Comprehension - Control vs Monitoring prompts - Control vs Evaluation + Monitoring - Control vs Evaluation prompts Source memory - Control vs Monitoring prompts - Control vs Evaluation + Monitoring - Control vs Evaluation prompts Justification of credibility by source - Control vs Evaluation prompts - Control vs Evaluation + Monitoring	1.05 .75 ns .70 .66 ns .36 ns .34 ns .08 ns .89 1.04 .66 .54 ns
Wiley et al. (2009) Undergraduates (N=60)	60-120 min	Students presented with declarative knowledge of what to <i>consider</i> when evaluating source features and then completed a worksheet for each site they read. Students asked to <i>identify</i> the author, the author's knowledge, and motivation to write. Feedback on ranking of websites.	Ranking discrimination Rereading reliable websites Conceptual model in essay References to source Justification of source ranking	.73 1.03 1.00 .67 .44 ns

Note: ^a: Adjusted ES; ns: Non significant