Abstract

This quasi-experimental study examined the effects of an intervention designed to teach upper-secondary school students to take source information, for example about author expertise, into consideration when selecting, processing, and using textual resources to complete particular multiple document literacy tasks. The intervention centered on a contrasting cases approach framed by authentic curriculum-based classroom activities and was implemented over six weeks by teachers who had participated in professional development seminars. The findings demonstrated that students who had participated in the sourcing intervention placed more value on source information when selecting texts, invested more time and effort in processing the texts they selected, and more frequently attributed textual ideas to their respective sources compared to students who instead had participated in typical classroom activities. These effects were observed on far transfer tasks where students worked with multiple documents on different topics in different situational contexts for different purposes and were sustained over a period of five and a half weeks. The discussion highlights the uniqueness of the current intervention work and centers on the aspects of the sourcing intervention that likely promoted these broad, sustainable, and transferable sourcing skills in students. Attention is also directed to several possible lines of future research in this area.

Keywords: Sourcing; sourcing intervention; multiple document literacy; contrasting cases approach.

Teaching Sourcing in Upper-Secondary School:

A Comprehensive Sourcing Intervention with Follow-Up Data

Introduction

In the 21st century, there has been a surge of interest in how students select, process, and use multiple information resources (Bråten, Braasch, & Salmerón, in press; Magliano, McCrudden, Rouet, & Sabatini, 2018). This situation is associated with the exponential growth of information resources and the almost unlimited availability and instantaneous accessibility of those resources to any reader at any moment (Kinzer & Leu, 2017; Lawless & Schrader, 2008). At the same time, most of what researchers know about reading is still based on individuals reading single, often narrative, texts (Kintsch, 1998; McNamara & Magliano, 2009). That research is limited in at least two respects: (a) it does not address the integration of information from diverse textual resources, which is required to construct a coherent, meaningful representation of a situation or issue discussed across multiple texts, and (b) it does not address the consideration of source information (e.g., the author or publication), which is required to select, process, and use reliable information instead of wasting time and effort on information of dubious quality (Britt, Rouet, & Durik, 2018).

In this study, we built on theoretical perspectives within multiple document literacy (Britt, Rouet, & Braasch, 2013; Perfetti, Rouet, & Britt, 1999; Rouet & Britt, 2011) and designed an intervention that targeted students' sourcing skills, with the process of *sourcing* defined as attending to, representing, evaluating, and using available or accessible information about the sources of textual content, for example about the author or publisher (Bråten, Stadtler, & Salmerón, 2018). Compared to typical intervention work in the area (Brante & Strømsø, 2018), our intervention was comprehensive in terms of the time frame of the study, the aspects of sourcing that were taught, the ways the effects of the intervention were

assessed, and the use of a delayed posttest to gauge the long-term effects of the intervention.

Why is Sourcing Important?

People often read about complex issues of which they have only limited background knowledge, such as controversial socio-scientific issues regarding the causes of climate change or the safety of nuclear power plants (Bromme & Goldman, 2014; Sinatra, Kienhues, & Hofer, 2014). In such reading contexts, it may be difficult, if not impossible, to determine the accuracy of information and the truth of claims directly, making it highly pertinent to evaluate information in light of the features of the sources (Stadtler & Bromme, 2014). Following Bråten and Braasch (2018), we define *sources* as information about individuals and organizations that create and publish text or document content, including information about when, where, and for what purpose the content is created and published. Accordingly, *source features* involve such aspects of a source as the author (including the author's credentials and affiliation), the text type (e.g., a textbook or a blog), the publication venue, and the date of creation (Britt & Aglinskas, 2002; Rouet, Britt, Mason, & Perfetti, 1996).

Several frameworks highlight the importance of taking source information into consideration when dealing with multiple information resources (e.g., Leu, Kinzer, Coiro, Castek, & Henry, 2013; List & Alexander, 2019; Lucassen, Muilwijk, Noordzij, & Schraagen, 2013). Still, the documents model is arguably the most influential framework for explaining why readers need to note and remember source information when trying to understand an issue discussed across multiple texts (Britt, Perfetti, Sandak, & Rouet, 1999; Perfetti et al., 1999; Rouet, 2006). This framework explains that linking source information and semantic content (i.e., noting and remembering "who says what"), as well as linking different sources of information (e.g., noting and remembering that Author A contradicts Author B), will help readers construct a high quality, integrated understanding of an issue discussed across texts. This is because such source-content and source-source links, respectively, are essential for

prioritizing information from more credible sources and understanding reasons for conflicting views on the issue (e.g., because authors differ with respect to their competencies or motives).

The Multiple-Document Task-based Relevance Assessment and Content Extraction (MD-TRACE) model of Rouet and Britt (2011) complements the documents model by describing a sequence of processes that learners cycle through when using multiple information resources to complete a task (e.g., to write an essay on the safety of nuclear power plants). In a first step, learners set goals based on task instructions and plan procedures that may help them achieve their goals. In a second step, learners assess their information needs to determine whether, and to what extent, external information resources are required to complete the task. In a third step, learners select external information resources and process those resources, including relating information across resources to construct an integrated understanding of the issue. In a fourth step, learners use the information resources to generate their task products (e.g., their essays). Finally, in a fifth step, learners evaluate their task products to determine whether their goals have been achieved or whether they need to recycle through previous processing steps to achieve those goals.

Sourcing plays a pivotal role in the third and fourth steps of the MD-TRACE model (Rouet & Britt, 2011). Thus, when selecting information resources in the third step, learners need to take features of the sources, such as the credentials of the authors, into consideration when evaluating source credibility as a basis for selection. Moreover, when further processing the selected resources, attention to source information can facilitate understanding by helping learners to predict upcoming content information, interpret content information, and evaluate the trustworthiness of content information (Barzilai, Tzadok, & Eshet-Alkalai, 2015; Strømsø, Bråten, Britt, & Ferguson, 2013; Wineburg, 1991). As highlighted by the documents model, building source-content and source-source links in the third step is also likely to help learners construct an integrated understanding of an issue discussed across information resources. In

the fourth step, sourcing comes into play when learners include specific references to sources in their task products, which may serve purposes such as attributing ideas to their respective sources and strengthening (or weakening) arguments by referring to more (or less) credible sources (Salmerón, Gil, & Bråten, 2018; Strømsø et al., 2013). In this intervention, we addressed sourcing involved in the third and fourth processing steps of the MD-TRACE model quite broadly, targeting students' sourcing during selection and processing of documents (Step 3), as well as during the creation of a task product (Step 4).

Finally, the MD-TRACE model (Rouet & Britt, 2011) highlights internal resources that likely help learners navigate the described sequence of processes. These include reading comprehension abilities, prior topic knowledge, and working memory, as well as general knowledge about sources and source features (see also, Rouet, Britt, & Durik, 2017). Given the importance of motivation and engagement in multiple document contexts (List & Alexander, 2017), learners' interest in the topic may also influence how they cycle through these processing steps.

Learning Through Contrasting Cases

Schwartz and Bransford (1998) launched the idea that having students contrast cases in the classroom may improve learning because it helps them develop a more elaborated knowledge base that is applicable across contexts. Essentially, this instructional approach involves that students are simultaneously presented with two cases that illustrate different procedures or conceptual understandings in relation to a problem and asked to compare and contrast the two cases. During this process, students may receive more or less guidance and be presented with more or less detailed case descriptions. For example, whereas Rittle-Johnson and Star (2007), who asked seventh-grade students to compare and contrast two alternative procedures to the same algebraic problem, explicitly directed students' attention to the differences between the procedures, Nagarajan and Hmelo-Silver (2006), who asked

undergraduates to compare and contrast two approaches to formal assessment, did not provide students with any guidance to help them pinpoint key differences between the approaches. Conversely, Nagarajan and Hmelo-Silver (2006) presented the two cases as detailed video recordings of teachers assessing student learning, whereas Rittle-Johnson and Star (2007) presented the two cases as minimal worked examples.

Based on a meta-analysis, Alfieri, Nokes-Malach, and Schunn (2013) suggested that having students compare cases may enhance their understanding of an underlying principle. Moreover, with respect to guidance, Alfieri et al. (2013) suggested that "providing either directive instructions initially that will guide students to the relevant information or the features of at least one case would not detract from the benefits of case comparisons" (p. 110). More recent research targeting conceptual learning seems to allow clearer conclusions regarding the effectiveness of instructional guidance, however.

For example, Roelle and Berthold (2015, 2016) showed that, at least when the contrasting cases are complex, learners may profit from high amounts of instructional guidance, and, conversely, that contrasting cases may be less beneficial when learners are just prompted to compare the cases. In these studies, high guidance conditions involved that students were provided with model answers to the comparison prompts, which highlighted the dimensions in which the two cases differed. Similarly, Sidney, Hattikudur, and Alibali (2015) found that just asking undergraduates to compare contrasting cases did not enhance learning, whereas combining prompts to compare cases with prompts to self-explain (i.e., make sense of the problem) did, suggesting that self-explanation is an important ingredient of contrasting cases. Taken together, these studies suggest that providing more guidance while students compare contrasting cases may facilitate self-explanation and thus help them make sense of the procedural or conceptual problem in question.

In classroom contexts, several activities may be used to implement a contrasting cases

approach (Braasch, Bråten, Strømsø, Anmarkrud, & Ferguson, 2013; Rittle-Johnson & Star, 2007). Thus, direct, teacher-led instruction that highlights differences between cases may be used in combination with independent study of contrasting cases and social-interactive approaches such as discussing the cases with peers or whole-class discussion.

Sourcing Interventions

Students at different educational levels often disregard source information and pay attention only to the content when reading to comprehend and learn from diverse textual resources (Barzilai et al., 2015; Bråten, Strømsø, & Andreassen, 2016; Brem, Russels, & Weems, 2001; Britt & Aglinskas, 2002; Kiili, Laurinen, & Marttunen, 2008; VanSledright & Kelly, 1998; Wineburg, 1991). At the same time, correlational work indicates that sourcing is linked to comprehension and learning outcomes (Anmarkrud, Bråten, & Strømsø, 2014; Barzilai & Eshet-Alkalai, 2015; Bråten, Strømsø, & Britt, 2009; Cho, Woodward, & Li, 2017, 2018; Goldman, Braasch, Wiley, Graesser, & Brodowinska, 2012; Strømsø, Bråten, & Britt, 2010; Wiley et al., 2009). As a consequence, interventions to promote students' sourcing when working with multiple information resources have become an increasingly important agenda for reading research during the last two decades (for recent reviews, see Brante & Strømsø, 2018; Bråten, Stadtler, & Salmerón, 2018).

Intervention studies have been conducted at elementary (e.g., Macedo-Rouet, Braasch, Britt, & Rouet, 2013; Paul, Stadtler, & Bromme, 2019), secondary (e.g., Braasch et al., 2013; Britt & Aglinskas, 2002; De La Paz & Feldon, 2010; Kammerer, Meier, & Stahl, 2016; Mason, Junyent, & Tornatora, 2014; Nokes, Dole, & Hacker, 2007; Pérez et al., 2018; Stadtler, Scharrer, Macedo-Rouet, Rouet, & Bromme, 2016), and postsecondary levels (e.g., Stadtler & Bromme, 2008; Wiley et al., 2009). While there is currently not much evidence that sourcing can be effectively and efficiently promoted at elementary school level, existing evidence indicates that notable improvements in sourcing can result from interventions at

secondary and postsecondary levels. However, most of these interventions have been quite brief (40-120 min), most of the effect sizes have been medium (at best), and next to nothing is known about the sustainability of the improvements in sourcing that have resulted from such interventions (Brante & Strømsø, 2018; Bråten, Stadtler, & Salmerón, 2018).

Regarding sourcing interventions conducted in upper-secondary school, in particular, Brante and Strømsø (2018) identified seven studies that tested the effects on 38 dependent measures altogether. These dependent measures operationalized sourcing in different ways, including the rank ordering of documents differing in reliability, providing justifications for the ranking of documents, memory for source information, and references to sources in post-reading essays. Of the 38 tests, 18 resulted in small effects, while 10 tests resulted in medium effects and 10 tests resulted in large effects (Cohen, 1988). Moreover, five of the upper-secondary school interventions were conducted in history (Britt & Aglinskas, 2002; De La Paz & Feldon, 2010; Goldberg, Schwarz, & Porat, 2011; Nokes et al., 2007; Reisman, 2012), while only two were conducted in science (Braasch et al., 2013; Stadtler et al., 2016).

Exemplifying work in science, Stadtler et al. (2016) conducted a 90-min intervention in which they gave vocational students an introduction to the concept of the division of cognitive labor (Keil, Stein, Webb, Billings, & Rozenblit, 2008), highlighting that people often have to rely on the knowledge of others (i.e., experts) when wanting to learn about complex issues. In the following three training modules, students discussed how people acquire expertise and read pairs of texts that presented conflicting views on a scientific issue, trying to determine, both individually and collectively, which text in each pair was authored by a person with pertinent expertise on the issue and whose view therefore could be trusted. The results showed that the students who had participated in the intervention outperformed those in a control group with respect to how often they agreed with the view of an author with pertinent expertise on the issue versus an author who lacked pertinent expertise on the issue.

Thus far, very few studies have used contrasting cases as an instructional approach in the area of literacy research (Beitzel & Derry, 2009; Braasch et al., 2013; Salmerón & Llorens, in press), and only one prior study building on a contrasting cases approach targeted students' sourcing skills (Braasch et al., 2013). In that study, Braasch et al. (2013) presented upper-secondary students with two hypothetical students' verbal protocols, purportedly resulting from their efforts to try to evaluate the usefulness and trustworthiness of textual resources on a controversial socio-scientific issue. One of these students displayed strategies typically used by secondary school students, focusing solely on the relevance of the content and disregarding information about the sources. The other, in contrast, displayed more sophisticated source-based strategies typical for better college students and domain experts, taking the source features of author, text type, publication venue, and date of publication into consideration. Through a series of classroom activities involving independent seatwork as well as dyadic and whole-class discussions, intervention students were required to compare and contrast the two hypothetical students' verbal protocols to decide which displayed the best strategies when analyzing multiple information resources on the issue and why.

When later working with multiple science texts on a different issue, participants in the Braasch et al. (2013) intervention ranked the usefulness of the texts for understanding the issue in a more expert-like way than did control students and also referred much more to the source features of the texts when justifying their rankings. Further, intervention students included more scientific concepts from reliable texts in essays that they wrote about the new issue. While these results were promising, Braasch et al.'s (2013) intervention was limited by being very brief (i.e., 60 min) and implemented by researchers rather than regular class teachers. Moreover, participants did not practice the strategies learned by comparing and contrasting cases in the context of using multiple information resources to complete authentic academic tasks that really mattered for them, and both the resulting sourcing skills and their

learning from texts were assessed quite narrowly. Finally, no follow-up data were available.

The Present Study

We designed and implemented a six-week intervention targeting sourcing in upper-secondary school. Students were taught to take relevant source features into consideration when selecting textual resources; to predict, interpret, and evaluate content by means of source feature information when reading; and to include specific references to sources in their written task products (Barzilai & Eshet-Alkalai, 2015; Rouet & Britt, 2011; Strømsø et al., 2013). These aspects of sourcing were taught by means of a contrasting cases approach that juxtaposed adaptive sourcing with a lack of sourcing typically displayed by secondary school students (Braasch et al., 2013). This approach provided students with rich case descriptions (Nagarajan & Hmelo-Silver. 2006) and high amounts of instructional guidance that directed their attention to the differences between the cases and the distinguishing features of adaptive sourcing (Roelle & Berthold, 2015. 2016). Further, the intervention was implemented by classroom teachers within the context of regular subject-matter teaching.

After the intervention, all students from intervention and control classrooms were presented with sets of texts on two different socio-scientific topics – climate change and nuclear power – in order to write a letter to the editor on one of those topics. While students in the intervention classrooms had worked with diverse natural and social studies texts as part of the intervention, none of those texts concerned climate change or nuclear power. Students were asked to select the texts they wanted to use when writing their letters and justify their text selections, read the selected texts, and use information from these texts in their letters. First, we investigated the selection process, asking whether intervention and control students differed in terms of the time devoted to text selection and their justifications for text selections. Compared to control students, we expected that intervention students would process information more deeply in this step and therefore devote more time to the selection

process, as well as generate more source-based justifications for their text selections.

Second, we investigated the reading process, asking whether intervention and control students differed in terms of how they processed the selected texts when writing from multiple texts, that is, when working on their letters to the editor. Because the intervention students had been trained to predict, interpret, and evaluate content in light of source-feature information during reading, we expected that they would process the selected texts more intensely and thoroughly in this stage than would the control students.

Third, we investigated the written task products, asking whether intervention and control students differed in terms of their inclusion of specific references to sources, as well as in terms of how well they covered the content of the texts and integrated content across texts. We expected that intervention students would display more evidence of adaptive sourcing compared to control students, thus including more source feature information in their letters to the editor. Also, we expected that the intervention students would display better content coverage and content integration. Presumably, intervention students would outperform controls on these content-related measures because they would be more likely to represent perspectives from different sources and reconcile these perspectives in constructing a more integrated understanding of the issue (Britt & Aglinskas, 2002; Britt et al., 2013).

Lastly, we investigated long-term effects by asking intervention and control students to perform the same tasks five and a half weeks after the intervention. Given the length and comprehensiveness of the intervention, as well as its integration into regular subject-matter teaching, we expected that intervention students would still devote more time to the selection process, generate more source-based justifications for their text selections, process the selected texts more intensely and thoroughly, include more specific source references in their task products, and display better content coverage and content integration compared to control students. As an additional dependent measure, we assessed students' memory for the sources

of the texts after they had completed their letters to the editors, addressing the research question of whether intervention and control students differed in terms of source memory performance. In this regard, we expected that intervention students would display better memory for source feature information compared to controls.

Method

Participants

Participants were 250 students (*M* age = 16.23, *SD* = 0.75, 41% female) attending the first and second years of upper-secondary school at a large public upper-secondary school in southeast Norway. All participants attended college preparatory courses. The majority (74%) were native-born Norwegians who learned Norwegian as their first language, and the rest were bilingual, raised in Norway but with parents from different parts of Europe and Asia. The sample was relatively homogeneous (i.e., middle class) in regard to socioeconomic status.

One-hundred and seventeen participants in six classes (3 first year, 3 second year) constituted the intervention group and 133 participants in six classes (3 first year, 3 second year) constituted the control group. The intervention classes were taught by three female and two male teachers (one male teacher taught two classes), ranging in instructional practice from 1 to 15 years (*M* = 6 years), and the control classes were taught by six male teachers, ranging in instructional practice from 1 to 40 years (*M* = 17 years). Although the teachers of the control classes, on average, had longer instructional practice than those of the intervention classes, there were no differences with respect to formal qualifications. Further, observations and interviews did indicate any differences regarding their engagement in teaching or the quality of their teaching. Collection and handling of data met the requirements of the Personal Data Registers Act and the guidelines of the Norwegian Social Science Data Services.

Pre-Intervention Materials

Demographic survey. A brief survey requested demographic information about age,

gender, native language, area of academic specialization, and previous academic achievement.

Reading comprehension measure. Reading comprehension was assessed using a Norwegian adaptation of the cloze comprehension test developed and validated with Danish adults and young adults by Gellert and Elbro (2013). This measure consisted of five narrative and five expository texts ranging from 40 to 330 words, with a total of 1340 words. The texts contained 41 word gaps in all, with four alternative words provided for each gap. Correct refilling of the gaps required bridging inferences. Participants were instructed to read the texts and refill as many gaps as possible during a period of 10 minutes. The scoring was done by counting the number of correctly refilled gaps (max score = 41). Cronbach's α was .85. Gellert and Elbro (2013) have demonstrated that scores on this measure are highly correlated with scores on standardized question-answering tests of reading comprehension.

Topic knowledge measures. Because texts on climate change and nuclear power were used to assess the effects of the intervention, knowledge about these topics was measured before the intervention. We used a 14-item multiple-choice test to assess knowledge of scientific (e.g., climate gases) and political (e.g., the Kyoto Protocol) aspects of *climate change* (max score = 14). Cronbach's α was .61. This measure has been used and validated in a number of prior studies (e.g., Bråten et al., 2009; Bråten, McCrudden, Stang Lund, Brante, & Strømsø, 2018; McCrudden, Stenseth, Bråten, & Strømsø, 2016). Test-retest reliabilities have ranged from .73 to .77 (Salmerón et al., 2018).

A parallel 14-item multiple-choice test assessed knowledge of scientific (e.g., nuclear fission) and political (e.g., the International Atomic Energy Agency) aspects of *nuclear power* (max score = 14). Cronbach's α was .68. This measure also has been used and validated in a number of previous studies (e.g., Bråten, McCrudden, et al., 2018; McCrudden et al., 2016; Stenseth, Bråten, & Strømsø, 2016). A test-retest reliability of .72 has been reported (McCrudden et al., 2016).

Topic interest measures. Before the intervention, we also assessed interest in the topics used at the posttests. We administered a 12-item measure using a 10-point scale (1 = *not at all true of me*, 10 = *very true of me*) to assess interest in *climate change* (sample item: *I am interested in international climate issues*). The scores on the measure ranged from 1 to 10. Cronbach's α was .93. This measure has been used and validated in a number of previous studies (e.g., Bråten et al., 2009; Bråten, McCrudden, et al., 2018; Strømsø et al., 2010).

Interest in *nuclear power* was assessed with a parallel 12-item measure using an identical 10-point scale (sample item: *I am interested in issues concerning the safety of nuclear power plants*). The scores ranged from 1 to 10. Cronbach's α was .92. This measure also has been used and validated in prior research (e.g., Brandmo & Bråten, 2018; McCrudden et al., 2016; Stenseth et al., 2016).

Working memory measure. We administered a Norwegian adaptation of Swanson and Trahan's (1992) Working Memory Span. Twelve sets of unrelated sentences were read aloud to participants with a 2-s interval between each sentence. The number of sentences in each set was gradually increased from two to five. Participants were required to simultaneously comprehend the sentences so that they could answer a comprehension question about an unknown sentence immediately after the final sentence in a set was read, and remember the final word of each sentence. Thus, immediately after the test administrator asked a question related to one of the sentences, participants first wrote their answer to that question and then the final word of each sentence on the same response sheet. For each of the 12 sets, participants received one point if they answered the comprehension question correctly and one extra point for each of the final words they recalled. If the comprehension question was not answered correctly, they did not get any point for that set regardless of how many final words they remembered. Maximum score was 54. Cronbach's α was .61. Presumably, the relatively low α was due to the high difficulty level of some items. Still, this reliability

estimate may be considered acceptable for research purposes (Hair, Black, Babin, Anderson, & Tatham, 2006; Kerlinger & Lee, 2000).

Source knowledge measure. To assess general knowledge about sources and source features, we administered a Norwegian adaptation of Stadtler, Thomm, Babiel, Hentschke, and Bromme's (2013) German version of the Source Knowledge Inventory, originally developed in French by Rouet, Ros, de Pereyra, Macedo-Rouet, and Salmerón (2013). This measure consisted of 12 tasks. On the first five tasks, participants were introduced to brief excerpts from different sources (e.g., a newspaper) and asked to identify a particular source feature (e.g., author credentials) among several distractors. On the following five tasks, participants were presented with brief texts on different natural and social science topics (e.g., nutrition or demography) and asked to rate the sources of each text (e.g., a nutritionist or a historian) with respect to expertise and potential bias. On the two final tasks, participants were presented with two fictitious search engine results pages (SERPs), one displaying four results on biodiversity and one displaying four results on freshwater on Earth. For each SERP, participants were asked to rate each result with respect to whether they wanted to use information from that website in preparing a presentation on the topic. Following Rouet et al. (2013) and Stadtler et al. (2013), we considered the first five tasks to address identification of source features, the next five tasks to address prompted source feature evaluation, and the last two tasks to address implicit (i.e., unprompted) source feature evaluation. A higher total score indicated more general knowledge about sources and source features. Cronbach's α was .73.

Intervention

Professional development. The teachers of the intervention classes participated in three 3-h professional development seminars organized and led by the authors.[1] In the first, the teachers were informed about the purpose of the project and introduced to "critical reading and learning" as an area of research, highlighting the role of sourcing. Further, the design of

the study and the contrasting cases approach were presented and discussed with the teachers. Finally, the authors shared preliminary versions of the intervention materials.

In the second seminar, a teaching manual for the first two weeks of the implementation period was presented and thoroughly discussed with the teachers. This manual included detailed scripts for three 90-min lessons and all the instructional materials to be presented and distributed to the students during these lessons. The teachers' (mostly minor) suggestions for changes were taken into account before the manual was finalized.

In the third seminar, the authors and the teachers discussed how the principles of adaptive sourcing taught during the scripted lessons could be applied when students worked on authentic assignments. It was decided that one assignment would be an individual writing assignment (writing an essay) and that another assignment would be a group-based oral assignment (preparing and giving a PowerPoint presentation). It was also decided that both assignments should be graded, with grades based on students' sourcing in addition to content.

In addition to these seminars, the authors met with the teachers of the intervention classes two weeks into the implementation period to discuss the implementation so far and make final adjustments regarding the upcoming assignment tasks. A detailed description of the professional development seminars is provided in Appendix A, which is available as supporting information for the online version of this article.

Implementation. The implementation of the intervention started two weeks after the last professional development seminar and one week after collection of the pre-intervention data. Six classes were randomly assigned to the implementation, which was conducted in language-arts class and led by the teachers who had participated in the professional development seminars.[2] The first 90-min scripted lesson was termed "When you select." In this lesson, students read and discussed texts on a controversial topic that differed in terms of source information in order to prepare a presentation on the topic. In this process, they were

shown two hypothetical students' thinking when deciding which text to select, one taking source feature information into consideration and another basing the decision on text interestingness and own opinion. The class studied and discussed these contrasting cases, in particular how the adaptive sourcer had reflected on the source features and their implications as a basis for text selection. In the last part of this lesson, the students were presented with a new set of texts that varied with respect to relevant source features and asked to select the texts they wanted to use when preparing a presentation. In this process, their text selection was scaffolded by the reflections of the adaptive sourcer.

The second 90-min scripted lesson, which was termed "When you read," followed a similar procedure. The first part of the lesson focused on predicting upcoming content by means of source feature information, and the following parts focused on interpreting and evaluating content in light of source feature information. Participants studied and discussed how two hypothetical students (i.e., contrasting cases) had been thinking while predicting, interpreting, and evaluating content, in particular how the adaptive sourcer had reflected on source feature information as a basis for successful prediction, interpretation, and evaluation. Finally, students read new texts and, like the adaptive sourcer, focused on source feature information and its implications for predicting, interpreting, and evaluating text content.

The third 90-min scripted lesson was termed "When you write." Students were presented with two essays (i.e., contrasting cases), one linking different perspectives to their respective sources and the other not including any source references. The students read and discussed the two essays with a focus on differences in sourcing before they studied and discussed how the adaptive sourcer had been thinking about source references when producing an essay.

After the three 90-min scripted lessons, the students worked on the individual writing assignment designed to train the application of the principles of adaptive sourcing. This phase

also consisted of three 90-min lessons distributed over two weeks. In these lessons, students read authentic texts on the topic of the assignment, selected the texts they wanted to use, studied the selected texts, and wrote 500-750 word essays drawing on information discussed in the texts they had selected and read. Source feature considerations were scaffolded by the adaptive sourcer's thinking, as displayed in the scripted lessons.

In the two last weeks of the implementation, the students worked on the group-based oral assignment designed to further train the application of the principles of adaptive sourcing. This phase also consisted of three 90-min lessons, in which groups of 5-6 students selected textual information resources on the assigned issue, studied the selected texts, and created a PowerPoint presentation while taking source features into consideration during selection, reading, and writing. In the last lesson, each group gave a 15-min PowerPoint presentation that drew on the selected resources, with one third of the time devoted to reflection on their sourcing activities while working on the assignment. A detailed description of the nine implementation lessons is provided in Appendix B, which is available as supporting information for the online version of this article.

Observations. Of the 18 scripted lessons (three in each class), 33% were observed to assess the extent to which the teachers followed the teaching manual. These observations used a check-list procedure (Judd, Smith, & Kidder, 1991), with items on the check list equaling the instructional activities described in the teaching manual for a particular lesson. In the observed lessons, 83% of the scripted instructional activities were completed. The activities that were not completed mostly involved that teachers did not distribute materials on paper that they already had shown on PowerPoint slides (and had uploaded on a learning platform).

Observations were also conducted in 10% of the lessons where students worked on the individual writing and the group-based oral assignments. Because these lessons were not scripted, no attempt was made to quantify the extent to which teachers completed specific

instructional activities. However, these observations suggested that the teachers followed the more general guidelines agreed upon. Thus, the teachers consistently reminded students of the sourcing activities taught during the scripted lessons and ensured that they took source features into consideration when working on the assignments. That the teachers' instructional focus on sourcing was maintained through the application lessons was also confirmed in a meeting with all the intervention teachers after the intervention.

The Control Classes

The six control classes received typical classroom instruction in language arts in the lessons that targeted sourcing in the six intervention classes. Still, there were several reasons they could be regarded as an appropriate baseline for the intervention. First, the national curriculum for upper-secondary school emphasizes that students should learn to evaluate and use sources critically when reading and to attribute ideas to the respective sources when writing (Norwegian Directorate for Education and Teaching, 2016a). Second, in evaluating students' written performance, national guidelines for upper-secondary school highlight that appropriate sourcing is required to obtain good grades (Norwegian Directorate for Education and Training, 2016b). Finally, academic writing skills were considered particularly important in this school because the school collaborated with a teacher education department on a long-term project on academic and argumentative writing, which emphasized students' understanding of the appropriate use of sources when writing. Taken together, this gave us reason to expect that all students attending college preparatory courses in this school would have acquired a certain level of sourcing skills during regular classroom instruction.

Still, when we observed two-language arts lessons in the control classes during the implementation period, no teaching mentioning sourcing or sources occurred. Likewise, when meeting with the language-arts teachers of the control classes to learn about the content of their teaching during this period, there were no indications that this had been a prioritized

area. Thus, although the control classes worked with the same curriculum and used the same textbooks as the intervention classes, as well as attended a school with a strong, general emphasis on sourcing in reading and writing, the instructional activities in the control classes seemed to focus much less on promoting adaptive sourcing during the implementation period.

Post-Intervention Materials

Texts and computer application at the immediate posttest. The week after the implementation period ended, students in both conditions were presented with a list of 10 texts about either climate change or nuclear power. Students in each of the 12 classes were randomly assigned to work with texts on one of these topics. In each of the 10 texts, source feature information (author, credentials, affiliation, text type, venue, and date) was displayed on the first two lines, followed by three sentences of content information. The sources ranged from blog postings written by secondary school students to textbooks written by high-school teachers and journal articles written by science professors. The content information was always relevant and consisted of neutral, factual information about the two topics, as well as information regarded as controversial (e.g., concerning the safe storage of CO_2 and radioactive waste, respectively). The source and content information of each text is shown in Appendix E, which is available as supporting information for the online version of this article.

Students accessed the 10 texts on their assigned topic through a web-based application program. First, they selected the texts they wanted to use when writing a letter to the editor on the topic. Then, on a page displaying only the selected texts, they justified in writing why they had selected each of these texts. On a third page, they got access to expanded versions of the selected texts by clicking on them. That is, by clicking on a text, students gained access to an expanded text of approximately 100 words in addition to the source information, and by clicking on another text, that text was expanded and the previous one reduced to three-sentence length again. Students could re-access and reread the expanded texts as many times

as they wanted by clicking on the selected texts, and they could go back and forth between a page where they were writing their letter and the page on which their selected texts were located. After finishing their letters, students submitted their written products to a server.

Measures of selection at the immediate posttest. Regarding selection, we used the *total time spent on the initial selection task* (i.e., when the list of all 10 texts was available) as a dependent measure. This was logged by the application program. As two other measures of the selection process, we used the *source-feature* and *content-based justifications* that students offered for their text selections. Following Braasch et al. (2013), we coded justifications as source-feature based when text selections were based on information about the author, author credentials, author affiliation, text type, venue, and date. Justifications were coded as content-based when text selections were based on the relevance, believability, format/style, comprehensibility, and interestingness of the content information (see Braasch et al., 2013, for further description of the categories of source-feature based and content-based justifications). Two independent raters scored a random selection of 10% of the justifications, resulting in 92% agreement on the type of justification provided for text selections. Statistical analyses were based on weighted justification scores, computed by dividing the number of source-feature and content-based justifications by the number of selected texts.

Measures of reading at the immediate posttest. Regarding reading, we used the *number of times students accessed the expanded texts* as a dependent measure. Also, we used their *reading time for the expanded texts* as a dependent measure. These measures, both based on the log produced by the application program, were used as indicators of how intensely and thoroughly students processed the selected texts. In statistical analyses, these measures also were weighted by taking the number of selected texts into account.

Measures of writing at the immediate posttest. Regarding writing, we coded students' written products in terms of the *number of references to source features* (i.e., author,

author credentials, author affiliation, text type, venue, and date), indicating the extent to which accurate, specific source information was linked to information units from the texts (Salmerón et al., 2018). Two raters independently scored a random selection of 10% of the written products for number of source-feature references, yielding an interrater reliability (Pearson's *r*) of .99. Moreover, we used the *number of textual information units* in the written products as a measure of content coverage (Gil, Bråten, Vidal-Abarca, & Strømsø, 2010). Finally, the *number of switches between information units from different texts* was used to measure the degree of content integration (Britt & Sommer, 2004; Gil et al., 2010). Following Bråten, McCrudden, et al. (2018), when a sentence or part of a sentence in the written product contained information that corresponded to information contained in a particular part of one of the selected texts, it was coded as an information unit coming from that text. Independent scoring of a random selection of 10% of the written products resulted in 92% agreement on which texts the information units came from. Statistical analyses including the writing measures were performed after the number of source-feature references, information units, and switches between information units had been divided by the number of selected texts.

Texts, computer application, and measures at the delayed posttest. Five and a half weeks after the implementation period, students in both conditions were again presented with a list of 10 texts about either climate change or nuclear power, with all students randomly assigned to the climate change topic at the immediate posttest now assigned to the nuclear power topic and vice versa. The same textual materials and the same application program were used at the delayed posttest, with students performing the same selection, justification, reading, and writing tasks with a different topic. Thus, the same dependent measures were also used at the delayed posttest. In addition to these dependent measures, however, we asked all students at the delayed posttest to write down everything they could remember about the author, author credentials, author affiliation, publication venue, and publication date for each

of the selected texts after they had submitted their letters and closed the application. This source memory test was scored by counting the number of correctly recalled source features for the selected texts. Two raters independently scored a random selection of 20% of the source memory tests, reaching an agreement of 82% on the source memory scores. As with all the other dependent measures used at both the immediate and the delayed posttest (except for initial selection time), students' source memory scores were weighted by taking the number of selected texts into consideration before they were analyzed statistically.

Procedure

The authors and trained research assistants administered the pre-intervention measures to the students in all 12 classes during a 90-min class period the week before the implementation period started. First, participants completed the demographic survey, the reading comprehension measure, the working memory measure, and the source knowledge measure in that order. Afterwards, they completed the topic interest measure and the topic knowledge measure in that order for one of the topics (i.e., climate change or nuclear power) and then for the other, with the order of the topics counterbalanced.

The professional development seminars that also preceded the implementation phase consisted of three 3-h meetings distributed over 10 weeks (the first meeting was before the summer break), in which the three authors and the five teachers of the intervention classes thoroughly discussed the instructional activities focusing on the promotion of adaptive sourcing. Also, the authors and the teachers met two weeks into the implementation period to further discuss the assignments designed to train application of the principles of adaptive sourcing. During the six-week implementation period, these teachers implemented the planned sourcing intervention in nine 90-min lessons while the control classes were taught according to the same language-arts curriculum using ordinary practices. The observations of the intervention and control classes during the implementation period, as well as the interview

with the teachers of the control classes, were conducted by the second author.

The data for the immediate posttest were collected the week after the implementation period ended and the data for the delayed posttest were collected five and a half weeks after the implementation. On both occasions, the authors and trained research assistants administered the assessment tasks in all 12 classes during a 90-min class period. Participants used the application program to perform the selection, justification, reading, and writing tasks, with participants in each class randomly assigned to the topic of climate change or the topic of nuclear power at the immediate posttest and then working with texts on the other topic at the delayed posttest. Before logging on with their laptops to access the application, participants received a brief introduction to their respective topic on paper. This introduction provided some factual background information and mentioned a controversy surrounding the issue (viz., concerning the extent to which climate change is caused by human activities and the safety of nuclear power plants). After this introduction to the topic, the task instruction read: *You will be writing a letter to the editor where you discuss human-induced climate change/the safety of nuclear power plants. When you log on, you will see a list referring to 10 web texts. From this list, you are going to select the web texts you want to use when writing the letter to the editor*. Of note is that on the first page of the application, the 10 text were listed in random order for each participant. At the delayed posttest, after they had submitted their letters and closed the application, participants were given a sheet of paper with only the following instruction: *For each of the texts you selected when working on the computer, write down everything you remember about the author's name, the author's credentials, the author's affiliation, where the text was published, and when the text was published.*

Between the immediate and delayed posttests, the teachers of the intervention classes returned to typical classroom activities in their language-arts classes, that is, without teaching sourcing by means of contrasting cases or scaffolding students' use of the learned sourcing

strategies. This was ensured in a meeting with the intervention teachers after the intervention ended. Table 1 gives an overview of the different components of the entire intervention study.

Results

In the first main section of results, we present analyses that establish comparability across the intervention and control classrooms. This is important due to the quasi-experimental nature of the study (Shadish, Cook, & Campbell, 2002). In the second main section, we provide analyses that document differences between the two conditions at the immediate posttest, and in the third main section, we present analyses that document differences between the two conditions at the delayed posttest. Finally, we provide some qualitative information about student performance as a result of the intervention. We established comparability between the two conditions by means of two-tailed tests and used one-tailed, directional tests for our specific hypotheses at the immediate and delayed posttests. We used chi-square tests to analyze categorical data, independent-samples *t*-tests to analyze differences when scores were approximately normally distributed, and nonparametric Mann-Whitney *U* tests to analyze differences when scores deviated substantially from normality.[3] Because the two topics (i.e., climate change and nuclear power) were counterbalanced across the immediate and delayed posttests, we also compared the performance of students assigned to the two topics on all dependent measures at both posttests, with no statistically significant difference between the two topics found for any dependent measure (*p*s > .10).

Establishing Comparability Across Conditions

Chi-square tests were performed to determine if females and males, native and nonnative speakers of Norwegian, and students specializing in science and humanities, respectively, were equally distributed across the two conditions. None of these tests indicated a statistically significant difference, with $\chi^2(1) = 0.32$, $p = .57$, for gender, $\chi^2(1) = 1.25$, $p = .26$, for language background, and $\chi^2(1) = 0.58$, $p = .45$, for area of academic specialization (see

Table 2). With respect to academic achievement, there were also no statistically significant differences between the two conditions, with $t(236) = 0.17$, $p = .87$, for GPA obtained at the end of lower-secondary school, $t(125) = -.0.38$, $p = .71$, for natural science grade at the end of the first year of upper-secondary school, and $t(125) = -1.34$, $p = .18$, for language-arts grade obtained at the end of the first year of upper-secondary school.[4]

Further, the two conditions were similar with respect to reading comprehension, $t(244) = -0.29$, $p = .78$, working memory capacity, $t(243) = -0.93$, $p = .36$, prior knowledge about climate change, $t(230) = -0.88$, $p = .38$, prior knowledge about nuclear power, $t(225) = -1.07$, $p = .28$, topic interest in climate change, $t(240) = -1.57$, $p = .12$, topic interest in nuclear power, $t(239) = -0.95$, $p = .34$, and general knowledge about sources and source features, $t(235) = 0.06$, $p = .96$ (see Table 2 for all means and standard deviations by condition).

Taken together, these analyses excluded a number of pre-existing differences that might account for performance on the measures administered at the posttests. Thus, the intervention students did not, by chance, have a demographic or educational background that might be beneficial to posttest performance compared to the control students. Nor did they outperform control students on a number of individual difference variables (e.g., reading comprehension, working memory, prior knowledge) relevant to performance on the posttest measures. These analyses thus support that any intervention benefits at the immediate and delayed posttests were not indicative of pre-existing differences between the conditions.

Immediate Posttest Results

When selecting. A series of independent-samples *t*-tests were conducted to investigate whether the intervention had the expected effects on the selection process. The results of these analyses are displayed in Table 3. First, the results were not consistent with the prediction that the intervention students ($M = 198.16$, $SD = 91.62$) would devote more time to the initial phase of the selection process than would the control students ($M = 182.83$, $SD = 94.58$), with

t(247) = 1.29, *p* = .098. In other words, selection time did not indicate that intervention students had processed information more deeply in this step. However, as expected, there was a statistically significant difference between the intervention (*M* = 1.80, *SD* = 1.39) and control students (*M* = 0.67, *SD* = 0.86) with respect to source-feature based justifications, with *t*(188.73) = 7.60, *p* = .000, Cohen's *d* = 1.003. Thus, students who had participated in the sourcing intervention generated much more source-feature based justifications for their text selections than did those in the control classrooms. There was no statistically significant difference between the intervention (*M* = 0.59, *SD* = 0.44) and control students (*M* = 0.66, *SD* = 0.43) with respect to content-based justifications, with *t*(248) = -1.30, *p* = .196.

When reading. At the immediate posttest, both our predictions concerning the reading process were confirmed. Thus, the intervention students (*M* = 8.12, *SD* = 6.80) were more likely to revisit the selected texts than were the control students (*M* = 7.11, *SD* = 7.42), with Mann-Whitney *U*, *Z* = -2.33, *p* = .010, *r* = -.15. Moreover, the intervention students (*M* = 99.26, *SD* = 81.20) devoted more time to reading the selected texts than did the control students (*M* = 73.19, *SD* = 64.24), with Mann-Whitney *U*, *Z* = -2.80, *p* = .0025, *r* = -.18. Taken together, these results suggested that students who had participated in the sourcing intervention processed the expanded versions of the selected texts more intensely and thoroughly than did those who had participated in typical classroom activities (see Table 3)

When writing. As expected, there was a statistically significant difference between the intervention (*M* = 0.22, *SD* = 0.36) and control students (*M* = 0.08, *SD* = 0.20) with respect to the number of references to source features included in the written products, with Mann-Whitney *U*, *Z* = -3.45, *p* = .0005, *r* = -.224. However, the results were not consistent with our prediction that the intervention students (*M* = 1.92, *SD* = 1.44) would include more textual information units in their written products than would the control students (*M* = 1.72, *SD* = 1.41), with Mann-Whitney *U*, *Z* = -1.25, *p* = .106. Still, our analyses of the immediate

posttest data confirmed our prediction that the intervention students (M = 0.47, SD = 0.33) would outperform the control students (M = 0.38, SD = 0.35) with respect to the number of switches between information units from different texts that they included in their written products, with $t(235) = 1.97$, $p = .025$, Cohen's $d = 0.26$. Taken together, these results indicated that the students who had participated in the sourcing intervention were more likely to link information units from the texts to specific source information and integrate those information units in their written products, although they did not cover the content of the selected texts any better compared to the control students (see Table 3).

Delayed Posttest Results

When selecting. At the delayed posttest, results were consistent with our prediction that the intervention students (M = 210.22, SD = 134.39) would devote more time to the initial phase of the selection process than would the control students (M = 165.76, SD = 91.60), with $t(198.66) = 3.01$, $p = .0015$, Cohen's $d = 0.39$. Also as expected, the intervention students (M = 1.19, SD = 1.05) generated more source-feature based justifications than did the control students (M = 0.64, SD = 0.75), with $t(208.10) = 4.76$, $p = .000$, Cohen's $d = 0.62$. Finally, there was no statistically significant difference between two conditions with respect to content-based justifications (intervention: M = 0.54, SD = 0.45; control: M = 0.59, SD = 0.45; $t(248) = -0.93$, $p = .35$). Taken together, these results indicated that at the delayed posttest, the students who had participated in the sourcing intervention processed the information more deeply and took source-feature information more into account when selecting texts than did the students who had received typical classroom instruction (see Table 4).

When reading. At the delayed posttest, both our predictions concerning the reading process were once again confirmed. Thus, the intervention students (M = 6.69, SD = 6.41) were more likely to re-access the expanded versions of the selected texts than were the control students (M = 4.90, SD = 5.75), with Mann-Whitney U, $Z = -3.16$, $p = .001$, $r = -.20$.

Likewise, the intervention students (M = 74.39, SD = 69.92) devoted more time to process the expanded versions of the selected texts than did the control students (M = 54.42, SD = 43.07), with Mann-Whitney U, Z = -2.46, p = .007, r = -.16. These results were consistent with the assumption that participants in the sourcing intervention would process the selected texts more intensely and thoroughly than would students in the control classrooms (see Table 4).

When writing. As expected, there was still a statistically significant difference between the intervention (M = 0.20, SD = 0.32) and control students (M = 0.07, SD = 0.15) at the delayed posttest with respect to the number of references to source features that they included in their written products, with Mann-Whitney U, Z = -2.68, p = .0035, r = -.17. However, it was not consistent with our predictions that neither the difference in number of textual information units (intervention: M = 1.53, SD = 1.11; control: M = 1.32, SD = 1.00; $t(242)$ = 1.54, p = .0625), nor the difference in number of switches between information units from different texts (intervention: M = 0.38, SD = 0.36; control: M = 0.36, SD = 0.33; $t(242)$ = 0.47, p = .321), reached statistical significance at the delayed posttest (see Table 4).

Source memory test. Finally, when we asked students to recall the sources of the selected texts after they had submitted their written products and closed the application, the intervention students (M = 1.65, SD = 1.23) clearly outperformed the control students (M = 0.93, SD = 0.77) in terms of source-feature memory performance, with $t(167.70)$ = 5.20. p = .000, Cohen's d = 0.72. Thus, as expected, increased consideration of source-feature information when working on the textual materials seemed to have improved the intervention students' memory for relevant source features relative to controls (see Table 4).

Some Qualitative Information

Of note is that at both posttests, intervention students' source-based justifications varied considerably with regard to quality, with some of them merely mentioning source features without further reasons (e.g., "The author is a professor") while others elaborated on

the significance of particular source features. As an example of the latter, one student wrote: “This is a paragraph from an upper-secondary textbook. This means that the text has been checked a number of times before it was published. It was published in 2012 but I believe it is still relevant.” As can be seen, this student gave a reasonable source-based justification for why the text was considered credible, although the emphasis put on the date of publication was less clear. Another student justified the selection of a seemingly biased text in the following way: “I selected this text because it probably will represent a different perspective on the topic of climate change. The author is an engineer in Statoil, Norway’s largest oil company, and she can be expected to promote Statoil’s view on how to handle human-induced climate change. The article is published in Technology Weekly, which I regard as a trustworthy magazine.” Thus, this student used source features not only in considering credibility but also in predicting content, exemplifying the adoption of a broader conceptualization of sourcing, which also was focused during the intervention.

Likewise, at both posttests, the intervention students varied considerably in terms of how they included references to sources in their written products. Some seemed to take for granted that readers had accessed the same information resources that they had, writing for example: “Professor Per Haugen believes that if radioactive waste is stored safely, there should be no problems.” For readers, it is not obvious that Haugen is a professor in natural science at a well-respected university, however, or that he expressed his views in a letter to the editor. Although other references were more elaborate, for example, “… professor Kari Lunde at the Faculty of Natural Sciences, University of Stavanger, claims that …,” students, in general, did not seem to have internalized a clear standard for when and how they should include references. Still, source references were mostly used in a reasonable way, either to provide evidence for a claim or to justify the introduction of multiple perspectives. Also, source feature information was linked to content by showing the origin of claims, statements,

and explanations (i.e., who said what). Sometimes such source-content links took the form of quotes but mostly students paraphrased statements from the texts they had read.

Discussion

The present study examined the effects of an intervention designed to teach sourcing to upper-secondary school students through a series of classroom activities centered on a contrasting cases approach. Teachers implemented the intervention after having participated in researcher led professional development seminars, with sourcing targeted during the selection, processing, and use of multiple textual information resources to complete particular tasks relevant to the language-arts curriculum. The outcome measures included the time used for text selection and the justifications for text selections; the number of times the selected texts were accessed and the time devoted to reading those texts; and the number of source references, textual information units, and switches between information units from different texts in written task products. The outcome measures were completed at both immediate and delayed posttests when students worked with multiple texts on controversial socio-scientific topics that had not been targeted as part of the sourcing intervention.

With respect to the selection process, our hypotheses were confirmed regarding the justifications that students provided for their text selections. Thus, at both the immediate and delayed posttests, intervention students generated more source-based justifications than did control students, while there was no difference between the two groups with respect to content-based justifications. Our hypotheses were only partially confirmed with respect to the time used for text selection, however. Thus, only at the delayed posttest did the intervention students devote more time to the initial selection process than the control students. Whereas the intervention students actually increased the selection time somewhat from the immediate to the delayed posttest, and thereby were able to retain a relatively strong focus on source information (as indicated by their source-based justifications), the control students reduced

the selection time during the same period and also did not seem to use the remaining time for considering source information in the selection process.

With respect to the reading process, our hypotheses were confirmed at both posttests. The intervention students accessed the selected texts more frequently and read them for a longer period of time compared to the control students. This suggests that the intervention students processed the selected texts more intensely and thoroughly than did the control students, as modeled in the contrasting cases approach and practiced when they predicted, interpreted, and evaluated textual content in light of source feature information.

That the intervention students kept source information in mind when reading the selected texts is also supported by their inclusion of more source feature information in their written products at both posttests, compared to the control students. Thus, as hypothesized, the intervention students to a greater extent linked textual information units to their respective sources in the letters they wrote to the editor about their assigned topic, and also displayed better memory for source feature information at the delayed posttest. Regarding our measures of content coverage and content integration, however, only the difference in number of switches between information units from different texts reached statistical significance at the immediate posttest, and there were no statistically significant differences between the groups in number of textual information units included in the written products or number of switches between information units from different texts at the delayed posttest. One likely reason for this lack of consistent effects with respect to content coverage and content integration is that those aspects of the writing process were not really modeled by the “adaptive sourcer” or explicitly emphasized during the subsequent assignments. Based on the literature (Bråten, Stadtler, & Salmerón, 2018), sourcing was considered to be the main issue for our participants and, accordingly, pivotal to improvement on multiple document literacy tasks. It is therefore an important finding that our intervention was successful in promoting sourcing in the written

task products, although influencing the content of those products may seem to require that processes such as content coverage and content integration are targeted more directly (Mateos, Martin, Cuevas, Villalón, Martinez, & González-Lamas, 2018).

Taken together, the effects of this sourcing intervention are far from trivial. In their recent review of interventions targeting sourcing skills, Brante and Strømsø (2018) found that no prior study had tested potential long-term effects of the intervention (see, however, Pérez et al., 2018). This study is therefore one of the first to demonstrate that not only immediate but also longer term effects on students' sourcing skills can be obtained. Moreover, these effects were demonstrated on a range of assessments distributed across the selection, reading, and writing phases of multiple document use (Rouet & Britt, 2011), with intervention students found to place more value on source information when selecting documents, invest more time and effort in processing those documents, and be more heedful in attributing textual ideas to their respective sources relative to controls. Importantly, these effects were shown on multiple document "far transfer tasks" (Barnett & Ceci, 2002), involving that students applied their sourcing skills when working with different topics in different situational contexts for different purposes. This may indicate that the intervention students had acquired the principles of adaptive sourcing and were able to bring those principles to bear on altogether different tasks not only immediately but also several weeks after the intervention.

Presumably, the broad, sustainable effects that we observed are attributable to several unique aspects of the intervention. First, the intervention was comprehensive in the sense that it targeted sourcing quite broadly, focusing on sourcing not only as a means to select reliable documents, but also as a means to understand (i.e., predict and interpret) and evaluate document content and generate task products that appropriately credit the sources of ideas. Second, the intervention was implemented by classroom teachers and integrated in authentic curriculum-based tasks that mattered for the students in terms of academic performance,

which likely increased their will to engage in sourcing activities (Paul, Macedo-Rouet, Rouet, & Stadtler, 2017). Third, the detailed case descriptions and the strong focus on the distinguishing features of adaptive sourcing that characterized the contrasting cases approach gave the intervention students rich opportunities to learn the principles of adaptive sourcing required for transfer (Nagarajan & Hmelo-Silver, 2006; Roelle & Berthold, 2015, 2016; Sidney et al., 2015). Finally, the intervention lasted longer than most other sourcing interventions implemented to date (Brante & Strømsø, 2018; Bråten, Stadtler, & Salmerón, 2018). Taken together, this unique configuration of instructional features likely explains why this intervention was distinguished by its broad, sustainable, and transferable effects when compared to previous interventions targeting sourcing skills (Brante & Strømsø, 2018; Bråten, Stadtler, & Salmerón, 2018).

Although the complexity of this intervention makes it impossible to attribute its effects to any one component, such as the contrasting cases approach, it might be possible to isolate and test the effects of each of the components in experimental work. However, we would argue that it seems pedagogically more meaningful to integrate them in real classroom contexts. For example, the principles of adaptive sourcing taught via a contrasting cases approach need to be practiced and elaborated both individually and socially, with instructional activities targeting sourcing gaining relevance by being framed by authentic, curriculum-based tasks that matter in students' academic lives. As such, the present intervention can serve as a model for how different design features of sourcing interventions may be combined in a pedagogically meaningful way. That said, the effectivity and efficiency of this intervention should be compared with those of other sets of practices designed to teach sourcing to secondary school students. Preferably, such research should also include a condition to control for the use of multiple information resources per se (Nokes et al. (2007).

Moreover, it seems highly pertinent to test to what extent effects of sourcing

interventions transfer to authentic task contexts beyond the assessment (i.e., posttest) situations, such as when students select, process, and use Internet-based information resources both in and out of school. To better understand the cognitive mechanisms underlying the effects of instructional activities targeting sourcing, it also seems pertinent to supplement the dependent measures used in this study with data based on online methodologies such as eye movements (Leinenger & Rayner, 2017) and verbal protocols (Ericsson & Simon, 1993).

Finally, although we remain enthusiastic about the broad, sustainable, and transferable effects that we obtained, these effects were not stellar compared to Cohen's (1988) effect size benchmarks. However, those benchmarks refer to the entire domain of social science rather than a particular area of research. With respect to previous sourcing interventions in upper-secondary school, most effect sizes have been small to medium (Brante & Strømsø, 2018), which means that the effects of our intervention fared quite well when interpreted in the context of interventions to promote sourcing skills. That said, more research is needed to investigate how sourcing interventions may have even greater impact on students' sourcing activities in multiple document contexts. One obvious possibility is to increase the amount of professional support offered to teachers responsible for the intervention, with continuous support also provided during the implementation period. Another possibility is to increase the duration of the implementation to ensure that the instructional activities targeting sourcing become firmly established and integrated within subject-matter teaching. Last, but not least, it seems important that the sourcing intervention is conducted not only within one subject, as in our case, but across different subjects. This is to ensure that the effects of the intervention are not watered out or even counteracted by teaching going on in other subjects, because source information is not made salient or valued in those subjects, and because students' sourcing is not really taken into account when evaluating their academic performance.

Because sourcing can be considered a foundational competence for the development of

informed citizenship and participation in genuine democratic discourse (Bråten & Braasch, 2017), no student should leave secondary school without understanding the value and importance of adaptive sourcing. For example, the increased use of social media as a source of information about important issues pertaining to people's lives, as well as many people's mistrust of traditional media channels, require that individuals learn to take more responsibility for judging the relevance and reliability of information. Not least does the flow of misinformation about controversial issues, such as climate change and immigration, in most Western countries, highlight the need to emphasize critical reading skills within the curricula (Kahne & Bowyer, 2017). Thus, teaching sourcing is not only a matter of improving students' academic thinking and writing; it is a matter of educating for democracy.

Notes

[1] One teacher who did not participate in the first seminar met with the second author after this seminar to receive an introduction to the project and a summary of the discussions.

[2] In Norwegian upper-secondary school, an essential aim of the language arts curriculum is to teach critical literacy skills, especially skills related to analyzing and producing expository texts, including argumentative texts. As such, the expository natural and social studies texts that students worked with during the three scripted lessons, as well as the expository texts on socio-scientific topics included in the posttests, can be considered highly relevant to language arts class. Also, the texts on the language situation in Norway and Scandinavia that students worked with in the application task lessons following the scripted lessons concerned topics highly relevant to the language arts curriculum in Norwegian upper-secondary school.

[3] Please note that we used one-tailed, directional independent-samples *t*-tests and the nonparametric equivalent to such tests, that is, Mann-Whitney *U* tests, to compare the intervention and control students on the dependent measures because we tested a set of specific, theory-driven hypotheses concerning the effects of the sourcing intervention. In such instances, this approach can be considered appropriate or even optimal (Levin, 1985; Tabachnick & Fidell, 2014).

[4] Grades in natural science and language arts were thus available for only the second year participants in this study.

References

Alfieri, L., Nokes-Malach, T.J., & Schunn, C.D. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist, 48,* 87-113.

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.

Barnett, S.M., & Ceci, S.J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128,* 612-637.

Barzilai, S., & Eseth-Alkalai, Y. (2015). The role of epistemic perspectives in comprehension of multiple author viewpoints. *Learning and Instruction, 36,* 86-103.

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.

Beitzel, B.D., & Derry, S.J. (2009). When the book is better than the movie: How contrasting video cases influence learning. *Journal of Educational Computing Research, 40,* 337-355.

Braasch, J.L.G., 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,* 180-195.

Brandmo, C., & Bråten, I. (2018). Investigating relations between beliefs about justification for knowing, interest, and knowledge across two socio-scientific topics. *Learning and Individual Differences, 62,* 89-97.

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

Bråten, I., & Braasch, J.L.G. (2017). Key issues in research on students' critical reading

and learning in the 21st century information society. In C. Ng & B. Bartlett (Eds.), *Improving reading and reading engagement in the 21st century: International research and innovations*. (pp. 77-98). Singapore: Springer.

Bråten, I., & Braasch, J.L.G. (2018). The role of conflict in multiple source use. In J.L.G Braasch, I. Bråten, & M.T. McCrudden (Eds.), *Handbook of multiple source use* (pp. 184-201). New York: Routledge.

Bråten, I., Braasch, J.L.G., & Salmerón, L. (in press). Reading multiple and non-traditional texts: New opportunities and new challenges. In E.B. Moje, P. Afflerbach, P. Enciso, & N.K. Lesaux (Eds.), *Handbook of reading research (Vol. V)*. New York: Routledge.

Bråten, I., McCrudden, M.T., Stang Lund, E., Brante, E.W., & Strømsø, H.I. (2018). Task-oriented reading with multiple documents: Effects of topic familiarity, author expertise, and content relevance on document selection, processing, and use. *Reading Research Quarterly*, *53*, 345-365.

Bråten, I., Stadtler, M., & Salmerón, L. (2018). The role of sourcing in discourse comprehension. In M.F. Schober, & D.N. Rapp, & M.A. Britt (Eds.), *Handbook of discourse processes* (2nd ed., pp. 141-166). New York: Routledge.

Bråten, I., Strømsø, H.I., & Andreassen, R. (2016). Sourcing in professional education: Do text factors make any difference? *Reading and Writing: An Interdisciplinary Journal*, *29*, 1599-1628.

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.

Brem, S.K., Russell, J., & Weems, L. (2001). Science on the Web: Student evaluations of scientific arguments. *Discourse Processes*, *32*, 191-213.

Britt, M.A., & Aglinskas, C. (2002). Improving students' ability to identify and use source

information. *Cognition and Instruction, 20,* 485-522.

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., & Braasch, J.L.G. (2013). Documents as entities: Extending the situation model theory of comprehension. In M.A. Britt, S.R. Goldman, & J.F. Rouet (Eds.), *Reading: From words to multiple texts* (pp. 160–179). New York: Routledge.

Britt, M.A., Rouet, J.F., & Durik, A. (2018). *Literacy beyond text comprehension: A theory of purposeful reading*. New York: Routledge.

Britt, M.A., & Sommer, J. (2004). Facilitating text integration with macro-structure focusing tasks. *Reading Psychology, 25*, 313-339.

Bromme, R., & Goldman, S.R. (2014). The public's bounded understanding of science. *Educational Psychologist, 49,* 59-69.

Cho, B.-Y., Woodward, L., & Li, D. (2017). Examining adolescents' strategic processing during online reading with a question-generating task. *American Educational Research Journal, 54*, 691-724.

Cho, B.-Y., Woodward, L., & Li, D. (2018). Epistemic processing when adolescents read online: A verbal protocol analysis of more and less successful online readers. *Reading Research Quarterly, 53*, 197-221.

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

De La Paz, S., & Felton, M.F. (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.

Ericsson, K.A., & Simon, H.A. (1993) *Protocol analyses: Verbal reports as data.* Cambridge, MA: The MIT Press.

Gellert, A.S., & Elbro, C. (2013). Cloze tests may be quick but are they dirty? Development and preliminary validation of a cloze test of reading comprehension. *Journal of Psychoeducational Assessment, 31*, 16-28.

Gil, L., Bråten, I., Vidal-Abarca, E., & Strømsø, H.I. (2010). Summary versus argument tasks when working with multiple documents: Which is better for whom? *Contemporary Educational Psychology, 35,* 157-173.

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.

Goldman, S.R., Braasch, J.L.G., Wiley, J., Graesser, A.C., & Brodowinska. K.M. (2012). Comprehending and learning from Internet sources: Processing patterns of better and poorer learners. *Reading Research Quarterly, 47*, 356–381.

Hair, J., Black, W., Babin, B., Anderson, R., & Tatham, R. (2006). *Multivariate data analysis* (6th ed.). Upper Saddle River, NJ: Pearson.

Judd, C.M., Smith, E.R., & Kidder, L.H. (1991). *Research methods in social relations* (6th ed.). Orlando, FL: Holt, Rinehart, & Winston.

Kahne, J., & Bowyer, B. (2017). Educating for democracy in a partisan age: Confronting the challenges of motivated reasoning and misinformation. *American Educational Research Journal, 54*, 3-34.

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.

Keil, F.C., Stein, C., Webb, L., Billings V.D., & Rozenblit, L. (2008). Discerning the division

of cognitive labor: An emerging understanding of how knowledge is clustered in other minds. *Cognitive Science, 32*, 259-300.

Kerlinger, F.N., & Lee, H.B. (2000). *Foundations of behavioral research* (4th ed.). Forth Worth, TX: Harcourt College Publishers.

Kiili, C., Laurinen, L., & Marttunen, M. (2008). Students evaluating Internet sources: From versatile evaluators to uncritical readers. *Journal of Educational Computing Research, 39*, 75-95.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: University Press.

Kinzer, C.K., & Leu, D.J. (2017). New Literacies and new literacies within changing digital environments. In M.A. Peters (Ed.), *Encyclopedia of educational philosophy and theory*. Singapore: Springer.

Lawless, K.A., & Schrader, P.G. (2008). Where do we go now? Understanding research on navigation in complex digital environments. In J. Coiro, M. Knobel, C. Lankshear, & D.J. Leu (Eds.), *Handbook of new literacies* (pp. 267-296). Hillsdale, NJ: Erlbaum.

Leinenger, M., & Rayner, K. (2017). What we know about skilled, beginning, and older readers from monitoring their eye movements. J.A. León & I. Escudero (Eds.), *Reading comprehension in educational settings* (pp. 1-17). Amsterdam: John Benjamins.

Leu, D.J., Kinzer, C.K., Coiro, J., Castek, J., & Henry, L.A. (2013). New literacies: A dual-level theory of the changing nature of literacy, instruction, and assessment. In D.E. Alvermann, N.J. Unrau, & R.B. Ruddell (Eds.), *Theoretical models and processes of reading* (6th ed., pp. 1150-1181). Newark, DE: International Reading Association.

Levin, J.R. (1985). Some methodological and statistical ”bugs” in research on children’s learning. In M. Pressley & C.J. Brainerd (Eds.), *Cognitive learning and memory in children* (pp. 205-233). New York: Springer-Verlag.

List, A., & Alexander, P.A. (2017). Cognitive Affective Engagement Model of multiple source use. *Educational Psychologist, 52*, 182-199.

List, A., & Alexander, P.A. (2019). Toward an integrated framework of multiple text use. *Educational Psychologist, 54*, 20-39.

Lucassen, T., Muilwijk, R., Noordzij, M.L., & Schraagen, J.M. (2013). Topic familiarity and information skills in online credibility evaluation. *Journal of the American Society for Information Science and Technology, 64*, 254-264.

Macedo-Rouet, M., Braasch, J.L.G., 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.

Magliano, J.P., McCrudden, M.T., Rouet, J.F., & Sabatini, J. (2018). The modern reader: Should changes to how we read affect theory and research? In M.F. Schober, D.N. Rapp, & M.A. Britt (Eds.), *Handbook of discourse processes* (2nd ed.; pp. 343-361). New York: Routledge.

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.

Mateos, M., Martin, E., Cuevas, I., Villalón, R., Martinez, I., & González-Lamas, J. (2018). Improving written argumentative synthesis by teaching the integration of conflicting information from multiple sources. *Cognition and Instruction, 36*, 119-138.

McCrudden, M.T., Stenseth, T., Bråten, I., & Strømsø, H.I. (2016). The effects of author expertise and content relevance on document selection: A mixed methods study. *Journal of Educational Psychology, 108*, 147-162.

McNamara, D.S., & Magliano, J. (2009). Toward a comprehensive model of comprehension.

Psychology of Learning and Motivation, 51, 297-384.

Nagarajan, A., & Hmelo-Silver, C. (2006). Scaffolding learning from contrasting video cases. In S.A. Barab, K.E. Hay, N.B. Songer, & D.T. Hickey (Eds.), *Proceedings of the 7th International Conference on the Learning Sciences* (pp. 495-501). Bloomington, IN: International Society of the Learning Sciences.

Nokes, J., Dole, J., Hacker, D.J. (2007). Teaching high school students to be critical and strategic readers of historical texts. *Journal of Educational Psychology, 99*, 492-504.

Norwegian Directorate for Education and Training (2016a). *Curriculum for upper-secondary school*. Oslo: Norwegian Directorate for Education and Training.

Norwegian Directorate for Education and Training (2016b). *Examination guide: Evaluating writing performance in upper-secondary school.* Oslo: Norwegian Directorate for Education and Training.

Paul. J., Macedo-Rouet, M., Rouet, J.-F., & Stadtler, 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.

Paul, J., Stadtler, M., & Bromme, R. (2019). Effects of a sourcing prompt and conflicts in reading materials on elementary students' use of source information. *Discourse Processes, 56*, 155-169.

Pérez, A., Potocki, A., Stadtler, 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.

Perfetti, C.A., Rouet, J.F., & Britt, M.A. (1999). Towards a theory of documents representation. In H. van Oostendorp & S.R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 99–122). Mahwah, NJ: Erlbaum.

Reisman, A. (2012). A document-based history curriculum intervention in urban high schools. *Cognition and Instruction, 30*,86-112.

Rittle-Johnson, B., & Star, J.R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology, 99*, 561-574.

Roelle, J., & Berthold, K. (2015). Effects of comparing contrasting cases on learning from subsequent explanations. *Cognition and Instruction, 33*, 199-225.

Roelle, J., & Berthold, K. (2016). Effects of comparing contrasting cases and inventing on learning from subsequent instructional explanations. *Instructional Science, 44*, 147-176.

Rouet, J.F. (2006). *The skills of document use*. Mahwah, NJ: Erlbaum.

Rouet, J.F., & Britt, M.A. (2011). Relevance processes in multiple document comprehension. In M.T. McCrudden, J.P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 19-52). Greenwich, CT: Information Age.

Rouet, J.F., Britt, M.A., & Durik, A.M. (2017). RESOLV: Readers‘ representation of reading contexts and tasks. *Educational Psychologist, 52*, 200-215.

Rouet, J.F., Britt, M.A., Mason, R.A., & Perfetti, C.A. (1996). Using multiple sources of evidence to reason about history. *Journal of Educational Psychology, 88*, 478–493.

Rouet, J.F., Ros, C., de Pereyra, G., Macedo-Rouet, M., & Salmerón, L. (2013). *Teenagers' developing awareness of source quality*. Paper presented at the annual meeting of the Society for Text and Discourse, Valenica, Spain.

Salmerón, L., Gil. L., & Bråten, I. (2018). Effects of reading real versus print-out versions of multiple documents on students' sourcing and integrated understanding. *Contemporary Educational Psychology, 52*, 25-35.

Salmerón, L., & Llorens, A. (in press). Instruction of digital reading strategies based on eye-

movements modeling examples. *Journal of Educational Computing Research*.

Schwartz, D.L., & Bransford, J.D. (1998). A time for telling. *Cognition and Instruction, 16*, 475-522.

Shadish, W.R., Cook, T.D., & Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. New York: Houghton Miffin Company.

Sidney, P.G., Hattikudur, S., & Alibali, M.W. (2015). How do contrasting cases and self-explanation promote learning? Evidence from fraction division. *Learning and Instruction, 40*, 29-38.

Sinatra, G.M., Kienhues, D., & Hofer, B.K. (2014). Addressing challenges to public understanding of science: Epistemic cognition, motivated reasoning, and conceptual change. *Educational Psychologist, 49*, 123-138.

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*, 716-737.

Stadtler, M., & Bromme, R. (2014). The content-source integration model: A taxonomic description of how readers comprehend conflicting scientific information. In D.N. Rapp & J.L.G. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (pp. 379-402). Cambridge, MA: The MIT Press.

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: An Interdisciplinary Journal, 29*, 705-729.

Stadtler, M., Thomm, E., Babiel, S., Hentschke, J., & Bromme, R. (2013). *Ignorant albeit competent: Examining students' sourcing competencies and spontaneous use of source information while reading conflicting scientific texts*. Paper presented at the annual Workshop on Multiple Document Literacy, Münster, Germany.

Stenseth, T., Bråten, I., & Strømsø, H.I. (2016). Investigating interest and knowledge as predictors of students' attitudes towards socio-scientific issues. *Learning and Individual Differences, 47,* 274-280.

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, 18,* 513-527.

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.

Swanson, H.L., & Trahan, M.F. (1992). Learning disabled readers' comprehension of computer mediated text: The influence of working memory, metacognition, and attribution. *Learning Disabilities Research and Practice, 7,* 74-86.

Tabachnick, B.G., & Fidell, L.S. (2014). *Using multivariate statistics* (6th ed.). New York: Pearson.

VanSledright, B., & Kelly, C. (1998). Reading American history: The influence of multiple sources on six fifth graders. *The Elementary School Journal, 98,* 239-265.

Wiley, J., Goldman, S.R., Graesser, A.C., Sanchez, C.A., Ash, I.K., & Hemmerick, J. (2009). Source evaluation, comprehension, and learning in internet science inquiry tasks. *American Educational Research Association Journal, 46,* 1060-1106.

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,* 73-87.

Table 1

Overview of the different components of the intervention study

	Intervention condition	Control condition
Timeline	Professional development seminars (3 x 3 hours)	Activity
Week 1	Seminar 1: Project information and discussion of design and first versions of instructional materials	Business as usual
Weeks 2 - 9	Summer vacation	Summer vacation
Week 10	Seminar 2: Discussion of teaching manual for scripted lessons and instructional materials Seminar 3: Co-designing application lessons	Business as usual
Week 11	Data collection pre-intervention materials	Data collection
	Implementation (9 x 90 minutes)	
Week 12	Lesson 1 (scripted): «When you select» Lesson 2 (scripted): «When you read»	Typical instruction
Week 13	Lesson 3 (scripted): «When you write»	Typical instruction
Week 14	Lesson 4 (application): Writing assignment (selecting) Lesson 5 (application): Writing assignment (reading)	Typical instruction
Week 15	Lesson 6 (application): Writing assignment (writing)	Typical instruction
Week 16	Lesson 7 (application): Oral assignment (searching and selecting) Lesson 8 (application): Oral assignment (reading and writing)	Typical instruction
Week 17	Lesson 9 (application): Oral assignment (presenting and reflecting)	Typical instruction
Week 18	Immediate post-test	Immediate post-test

Week 19 - 22	Typical instruction	Typical instruction
Week 23	Delayed post-test	Delayed post-test

Table 2
Descriptive data and statistical values for pre-intervention comparisons between students in the intervention and control conditions

Variable	Intervention *M* (*SD*)	Control *M* (*SD*)	Test statistic
Gender (% females)	39	43	$\chi^2 = 0.32$, *ns*
Language background (% Norwegian)	77	71	$\chi^2 = 1.25$, *ns*
Academic specialization (% science)	19	22	$\chi^2 = 0.58$, *ns*
Lower secondary GPA	4.51 (0.42)	4.50 (0.35)	$t = 0.17$, *ns*
First-year science grade	4.37 (1.02)	4.43 (0.94)	$t = -0.38$, *ns*
First-year language-arts grade	4.10 (0.68)	4.27 (0.73)	$t = -1.34$, *ns*
Reading comprehension	24.53 (6.29)	24.75 (5.70)	$t = -0.29$, *ns*
Working memory	31.26 (10.35)	32.48 (10.19)	$t = -0.93$, *ns*
Prior knowledge climate	7.29 (2.50)	7.58 (2.59)	$t = -0.88$, *ns*
Prior knowledge nuclear	6.89 (2.81)	7.31 (2.98)	$t = -1.07$, *ns*
Topic interest climate	4.78 (2.28)	5.21 (1.97)	$t = -1.57$, *ns*
Topic interest nuclear	3.27 (1.72)	3.49 (1.82)	$t = -0.95$, *ns*
Source knowledge	104.49 (15.07)	104.39 (14.27)	$t = 0.06$, *ns*

Table 3

Descriptive data, statistical values, and effect sizes for comparisons between students in the intervention and control conditions at the immediate posttest

Variable	Intervention *M* (*SD*)	Control *M* (*SD*)	Test statistic/effect size
When selecting			
Selection time (sec)	198.16 (91.62)	182.83 (94.58)	$t = 1.29$, *ns*
Source-feature based justifications	1.80 (1.39)	0.67 (0.86)	$t = 7.60$, $p < .001$, $d = 1.003$
Content-based justifications	0.59 (0.44)	0.66 (0.43)	$t = -1.30$, *ns*
When reading			
Number of revisits	8.12 (6.80)	7.11 (7.42)	$Z = -2.33$, $p = .01$, $r = -.15$
Reading time (sec)	99.26 (81.20)	73.19 (64.24)	$Z = -2.80$, $p < .01$, $r = -.18$
When writing			
Source feature references	0.22 (0.36)	0.08 (0.20)	$Z = -3.45$, $p < .001$, $r = -.224$
Number of information units	1.92 (1.44)	1.72 (1.41)	$Z = -1.25$, *ns*
Number of switches	0.47 (0.33)	0.38 (0.35)	$t = 1.97$, $p < .05$, $d = 0.26$

Table 4

Descriptive data, statistical values, and effect sizes for comparisons between students in the intervention and control conditions at the delayed posttest

Variable	Intervention *M* (*SD*)	Control *M* (*SD*)	Test statistic/effect size
When selecting			
Selection time (sec)	210.22 (134.39)	165.76 (91.60)	$t = 3.01$, $p < .01$, $d = 0.39$
Source-feature based justifications	1.19 (1.05)	0.64 (0.75)	$t = 4.76$, $p < .001$, $d = 0.62$
Content-based justifications	0.54 (0.45)	0.59 (0.45)	$t = -0.93$, *ns*
When reading			
Number of revisits	6.69 (6.41)	4.90 (5.75)	$Z = -3.16$, $p = .001$, $r = -.20$
Reading time (sec)	74.39 (69.92)	54.42 (43.07)	$Z = -2.46$, $p < .01$, $r = -.16$
When writing			
Source feature references	0.20 (0.32)	0.07 (0.15)	$Z = -2.68$, $p < .01$, $r = -.17$
Number of information units	1.53 (1.11)	1.32 (1.00)	$t = 1.54$, *ns*
Number of switches	0.38 (0.35)	0.36 (0.33)	$t = 0.47$, *ns*
Source memory	1.65 (1.23)	0.93 (0.77)	$t = 5.20$, $p < .001$, $d = 0.72$.

Supplemental Materials for Online Publication

Appendix A

The Professional Development Seminars

The teachers of the six intervention classes participated in three 3-h professional development seminars organized and led by the three authors before the implementation of the classroom-based sourcing activities. In the first seminar, the teachers were informed about the purpose of the project (i.e., the intervention study) and introduced to "critical reading and learning" as an area of research, with a particular emphasis on the essential role of sourcing in 21st century reading literacy. Further, the design of the study, as well as the contrasting cases approach, was introduced to and discussed with the teachers. In this context, concrete examples were used to highlight the challenge of identifying and evaluating source feature information on many authentic websites. In the last part of the first seminar, the authors shared preliminary versions of the intervention materials with the teachers, especially materials to be used when teaching sourcing during the selection of information resources. The teachers' views and comments on these materials were taken into account when the materials were further revised and adapted for use in classroom-based activities.

In the second seminar, a teaching manual for the first two weeks of the implementation period, created by the authors, was presented and thoroughly discussed with the teachers. This manual included detailed scripts for three 90-min lessons and all the instructional materials to be presented and distributed to the students during these lessons. The first lesson focused on sourcing during selection of information resources, the second lesson focused on sourcing during the processing of information resources (i.e., when predicting, interpreting, and evaluating content information), and the third lesson focused on sourcing during the creation of task products. The instructional materials for these lessons consisted of texts to be read and discussed by the students, verbal protocols from hypothetical students displaying more or less adaptive sourcing when selecting, processing, and using information resources, completed worksheets that highlighted the "adaptive sourcer's" thinking about relevant source features when selecting, processing, and using information resources, and similar worksheets to be completed by the students when selecting, processing, and using information resources. These worksheets consisted of relevant questions to ask oneself about the source features of the texts that they worked with and the implications of those source features when selecting, processing, and using texts to complete particular tasks (see Appendix B in the supplemental materials for further information about how the instructional materials were used). In the process of reviewing the scripts and materials together with the teachers, their (mostly minor) suggestions for changes were taken into consideration before the teaching manual was finalized and this part of the intervention was implemented in their classrooms. The final version of the teaching manual, including all instructional materials, was available on paper as well as digitally, with all materials to be presented to the whole class also available in the form of PowerPoint slides on a USB flash drive provided to each teacher.

In the third seminar, the authors and the teachers discussed how the principles of adaptive sourcing that were taught during the three scripted lessons could be applied when students worked on authentic assignments integrated into curriculum-based teaching in language arts class. Through this discussion, it was decided that one such assignment should be an individual writing assignment where students were tasked to write a 500-750-word essay addressing whether the Norwegian language is threatened due to globalization, and that another such assignment should be a group-based oral assignment where students were tasked to work in groups of 5-6 students to prepare a PowerPoint presentation addressing an issue

concerning the language situation in Norway or Scandinavia. On the first assignment, students should select three texts from a set of seven authentic, printed texts collected by the teachers and use those texts when writing their essays, taking source feature information into consideration when selecting, reading, and using the three texts to complete the writing assignment. On the second assignment, students should search the Internet to select at least three information resources that they would use in their oral PowerPoint presentations, taking source feature information into consideration when selecting, reading, and using those resources in their presentations. The worksheets that were used in the three scripted lessons should also be completed by the students when working with these two assignments. Finally, it was decided in the third seminar that both these assignments should be graded by the teachers, with grades based on students' sourcing in addition to content, and with the criteria for grading made explicit to the students when the assignments were introduced in class. Of note is that these two assignments were designed in close collaboration between the three authors and the five teachers in this seminar. After the seminar, the authors produced a written summary of the discussion and the decisions made about the design, implementation, and scheduling of the two assignment tasks, which were sent all teachers for their approval before the intervention was taken to the classrooms (for further information about the two assignments, see Appendix B in the supplemental materials).

In addition to these seminars, the three authors met with all teachers of the intervention classes two weeks into the implementation period to discuss the implementation so far and make final adjustments regarding the upcoming assignment tasks.

Appendix B

The Implementation Lessons

The implementation of the sourcing intervention started two weeks after the last professional development seminar and one week after the pre-intervention data were collected. Six of the classes were randomly assigned to the classroom-based instructional activities focusing on sourcing, which were integrated into language-arts class and led by the teachers who had participated in the professional development seminars.

The implementation started with the teachers presenting a 113-word text without any source information and initiating a whole-class discussion about this text ("What do you think about this text?").The author of this text described his great efforts to save lives through humanitarian work and his generally unselfish regard for others. After a brief discussion, the students were showed the same text with source information included, revealing that the text was an excerpt from the defense speech of Vidkun Quisling at his trial in 1945, published in connection with his 100th birthday in 1987. Of note is that Quisling led a pro-Nazi puppet government in Norway during World War II and was found guilty of high treason and sentenced to death after the war. In English, his name has become synonymous with "traitor". After a whole-class discussion of the text in light of this source information, the teachers led a whole-class discussion guided by the following questions: What is content information and what is source information? What does source information consist of (exemplified by source features such as author, date, etc.)? Why is source information important when selecting, reading, and using information? After this discussion, the students were told that the class should focus on source information during a number of upcoming lessons.

The activities described above were the start of the first 90-min scripted lesson, which was termed "When you select." It continued with the presentation of two short texts about sun exposure and health, one written by a professor and published in a medical journal and one

written by a journalist and published in the magazine of an outdoor activities organization. The students were tasked to read these texts and discuss with a peer which text they would select in order to prepare a presentation on the topic. The content of both texts was relevant in this regard, such that the main differences between them concerned the source information. After the discussion, students were shown two hypothetical students' thinking when deciding which text to select, one taking source feature information (e.g., author credentials and text type) into consideration and another basing the decision on text interestingness and own opinion (see Appendix C in the supplemental materials). The teacher introduced the contrasting cases by saying, "Now I will show you how two other students selected a text and justified their text selection," and then asked them to study both students' thinking about text selection thoroughly. Afterwards, the class briefly discussed these contrasting cases, with this discussion initiated by the teacher asking, "Who do you think justified the text selection best, and why do you think this student did so?" After this brief discussion, the class studied a completed worksheet that displayed questions the adaptive sourcer had asked about the source features of the two texts (e.g., What kind of text (genre) is this?) and their implications for the trustworthiness of the information (e.g., What significance may the genre have for the trustworthiness of the information?), as well as the adaptive sourcer's responses to those questions (e.g., One is a scholarly article in the journal of a medical association and the other is an article in the membership magazine of a trekking association. I trust what the scholarly article says more than what's in the membership magazine because a scholarly article is written by an expert) (see Appendix D in the supplemental materials). The teachers reviewed this completed worksheet together with the students and highlighted how the adaptive sourcer had reflected on the source features and their implications as a basis for text selection. In doing so, the teachers directed students' attention to each of the questions the adapted sourcer had asked in the process of text selection as well as the response to each question, also informing the class that they should complete similar worksheets when selecting texts in the following lessons. In the last part of this lesson, the students were presented with a new set of four brief texts on the potential health risks of sun exposure that varied with respect to relevant source features (e.g., author credentials, text type, and venue). They were asked to select the two texts they wanted to use when preparing a presentation on the topic. In this process, they completed a worksheet for each of the four texts, individually responding to the questions posed by the adaptive sourcer about source features and their implications as a basis for deciding whether they would select that text. In conclusion, the teacher led a brief whole-class discussion about which texts the students had selected and why.

At the start of the second scripted lesson, which was termed "When you read," the students were told that this lesson concerned how source information can be used actively during reading. The first part of the lesson focused on predicting upcoming content by means of source feature information. The students were presented with a text that consisted of only a heading in addition to the source feature information and tried to predict the content of the text on the basis of this information. After a brief whole-class discussion, the students were shown and discussed two hypothetical students' thinking when trying to predict the content, with one of the cases illustrating how source feature information (in particular author, text type, venue, and date) was successfully used to predict the content of the text. Together, the teachers and students then reviewed a completed worksheet that displayed the questions asked and the responses constructed by the adaptive sourcer when using source feature information in the process of predicting the text's content. Finally, the students were presented with a new text consisting of only source feature information followed by a heading and completed a worksheet on their own by responding to the questions therein when trying to predict content on the basis of the source feature information. A brief, teacher-led whole-class discussion concluded this part of the lesson.

In the following parts of the second lesson, which focused on, respectively, interpreting and evaluating content in light of source feature information, a similar procedure was used. That is, texts including source feature information were first presented and discussed, with students trying to interpret and evaluate content, respectively, by means of source feature information. Then, students were shown how two hypothetical students (i.e., contrasting cases) had been thinking while interpreting and evaluating content, before they reviewed completed worksheets displaying the adaptive sourcer's questions and responses focusing on source feature information as a basis for successful interpretation and evaluation, respectively. Finally, students independently read new texts and completed worksheets responding to questions focusing on source feature information and its implications for interpreting and evaluating the content of the texts. The part of this lesson focusing on interpretation, as well as the part focusing on evaluation, was concluded with a brief, teacher-led whole-class discussion about the interpretations and evaluations of the students and how those were justified.

The third scripted lesson targeted sourcing during the creation of a task product and was termed "When you write." It started by presenting students with two student essays (i.e., contrasting cases) on the topic of doping in sports, both well formulated and organized and with relevant content. However, one of the cases included specific references to sources and thus linked different perspectives on the issue to their respective sources, whereas the other did not include any references to sources. The students were asked to read the two essays carefully and analyze them with respect to differences in sourcing, followed by a brief discussion about which essay displayed adequate sourcing and why. Students then studied a completed worksheet that displayed how the adaptive sourcer had been thinking when producing the essay on doping, asking questions about the accuracy and completeness of the source references and the extent to which different perspectives were attributed to their respective sources, and responding to these questions as a basis for judging the adequacy of the essay in terms of sourcing. This completed worksheet was thoroughly discussed in class and the students were informed that they should complete similar worksheets when working on assignments in the following weeks. Importantly, in this lesson, the teachers also presented their criteria for evaluating students' written products, using the contrasting cases essays as a point of departure. In particular, they explained that they would place much emphasis on students' sourcing when grading their written products, also anchoring this decision in national curricular goals and guidelines for evaluating writing performance in upper-secondary school (Norwegian Directorate for Education and Training, 2016a, 2016b).

In the third week of the implementation period, after the three 90-min scripted lessons described above had been conducted, the students started on the individual writing assignment designed to train application of the principles of adaptive sourcing previously taught. This implementation phase also consisted of three 90-min lessons distributed over two weeks. In the first lesson, the students read seven authentic texts on the topic of the assignment (i.e., whether the Norwegian language is threatened by globalization), which represented different perspectives on the issue and originated from sources differing with respect to source feature information (e.g., author expertise, text type, and venue). For each of the seven texts, they completed the same worksheet that was used in the scripted lesson termed "When you select," with the three texts that they wanted to use for their writing assignment selected on the basis of source feature considerations scaffolded by these worksheets. In the second lesson, students studied the texts they had selected while completing the worksheets concerning interpretation and evaluation that were used in the scripted lesson termed "When you read." (The worksheet concerning prediction was not used in this lesson because students had already read the entire texts during the selection process.) The students also created outlines of their essays during this lesson. In the third lesson, the students worked on their 500-750-

word essays, drawing on information discussed in the texts they had selected and read. While working on their essays, students also completed the same worksheet that was used in the scripted lesson termed “When you write.” When submitting their essays, students also handed in the worksheets they had completed when selecting, reading, and writing, and their assignments were graded by the teachers on the basis of their source feature considerations as evidenced by the completed worksheets as well as the qualities of their final written products. These grading criteria were made explicit at the start of the writing assignment.

In the fifth and sixth weeks of the implementation period, the students worked on the group-based oral assignment designed to further train the application of the principles of adaptive sourcing. This implementation phase also was scheduled for three 90-min lessons, and the assignment concerned a controversial language policy issue (each group within the same class addressed a different question). In these lessons, each group of 5-6 students first searched the Internet and selected three textual information resources on which they wanted to base their oral PowerPoint presentation, collectively completing a “When you select” worksheet for each of the three texts in justifying their selections. Next, the students in each group studied the selected texts and collectively completed the interpretation and evaluation “When you read” worksheets for each text, before they created a PowerPoint presentation while completing a “When you write” worksheet. In the last lesson, each group gave a 15-min PowerPoint presentation on the assigned issue that drew on the three selected resources, with one third of the time devoted to reflection on their sourcing activities while working on the assignment. The basis for grading the oral assignment was both the quality of the presentation (including both content and source information) and the group’s reflection on sourcing. In addition, the competed worksheets were taken into consideration by the teachers when determining the total assignment grade. The grading criteria were made explicit to the students at the start of the oral assignment.

Appendix C

Contrasting Cases Thinking about Text Selection

Student A

One text has the title "Sun counteracts cancer" and the other has the title "Lack of sunlight causes depression." I find the article about depression more exciting than the one about cancer and it is easier to create an interesting presentation to the class about sunlight and depression. I know several who say they are more depressed in the winter than in the summer, and the importance of exercise is known by everyone. The text about cancer seems boring and I don’t believe that sun exposure can prevent cancer in any case. I've actually never heard of anyone else claiming that - on the contrary - many warn against sunbathing without using sunscreen with a high sun protection factor. The text about depression also is written by a journalist, and a journalist is skilled in presenting technical material in an understandable way, while a professor often may be more difficult to understand. I would definitely choose the text about sunlight and depression when preparing my presentation.

Student B

I see that the text called "Sun counteracts cancer" is written by a professor at Oslo University Hospital, while the text called "Lack of sunlight causes depression" is written by a journalist in the Norwegian Trekking Association. It therefore seems that the text about cancer is written by a person who is a specialist in the area, while the journalist probably knows less about what she writes about. Besides, the professor has written a scholarly article in the journal of the medical association, which makes me I rely more on this text than on the article in the magazine of the trekking association. Both authors seemingly want to inform about

sunbathing and health, but I still think that the professor is better qualified than the journalist to provide accurate information on this topic. Both articles were published in 2016 so both seem well updated. The quality of the content seems OK in both texts, but I choose to use "Sun counteracts cancer" because it stems from a more trustworthy source.

Appendix D

The Adaptive Sourcer's Thinking about Text Selection

Consider	**Ask yourself**	**Response**
The quality of source information	What competence does the author have that can be relevant to this topic?	*One of the authors is a professor at Oslo University Hospital, while the other text is written by a journalist in the Norwegian Trekking Association. I think the professor knows more about sun exposure and health.*
	What kind of text (genre) is this?	*One is a scholarly article in the journal of a medical association and the other is an article in the membership magazine of a trekking association.*
	What significance may the genre have for the trustworthiness of the information?	*I trust what the scholarly article says more than what's in the membership magazine because a scholarly article is written by an expert.*
	Who has published this text (e.g., publishing company, website, journal, newspaper)?	*One of the texts is published by the Norwegian Medical Association and the other is published by the Norwegian Trekking Association.*
	What significance may the publisher have for the trustworthiness of the information?	*I think the Norwegian Medical Association is more trustworthy than the Norwegian Trekking Association as far as sun exposure and health is concerned.*
	What is the author's main intention with writing this text?	*Both seems to have the intention to inform about sun exposure and health.*
	What significance may the intention have for the trustworthiness of the information?	*None of these authors seem to have any other intention than to inform.*

	When was this text published or last updated?	*Both articles were published in 2016.*
Conclusion		*I choose to use "Sun counteracts cancer" because the quality of the source information is higher.*

Appendix E

Source and Three-Sentence Content Information Included in the Texts for Each Topic

Climate change

Science teacher Svein Strand, Steinkjer Upper Secondary School
Textbook in natural studies for upper secondary school, Gyldendal Publishers, 2012

When we use oil, CO_2 is generated. This is a greenhouse gas. Many think that an increase in CO_2 causes global warming.

Professor Kari Lunde, Faculty of Natural Sciences, University of Stavanger
Nordic Journal of Astrophysics, no 2, 2013

The sun has a surface temperature of approx. 6000 °C. Without the sun, the Earth would be uninhabitable. Even small changes in the sun's radiation will influence the climate.

Student Kristian Paulsen, Åsane Upper Secondary School
Posting on kristiansenvironmentblog.wordpress.com, June 9, 2016

Some of the gases in the atmosphere are called climate gases. The most important are water vapour, carbon dioxide, and methane. The climate gases affect the Earth's temperature.

Journalist Liv Moe, News Section, Norwegian Broadcasting Corporation
Feature story on nrk.no, May 14, 2015

Heat from the sun penetrates the atmosphere. Most of this heat cannot escape out again. This creates a greenhouse effect.

Journalist Ole Martinsen,
Article in the newspaper VG, January 12, 2015

The temperature at the North Pole is rising faster than anywhere else on the planet. The ice is melting more rapidly than before. This leads to a higher sea level.

Director Øyvind Mathisen, Ministry of Climate and Environment
Report on Government.no/mce/ March, 2016

Coal, oil, and gas have been stored in the ground for millions of years. When they are extracted and burned, the CO_2 level in the atmosphere rises. CO_2 is an important climate gas.

Science teacher Linda Bakke, Sandefjord Upper Secondary School
Article in Nature and the Environment, the membership journal of the Norwegian Society for the Conservation of Nature, no 2, 2011

Can climate change cause more hurricanes? Climate researchers are trying to answer this question. A larger number of strong and devastating hurricanes may possibly occur.

Chief engineer Bente Svendsen, Statoil
Technology Weekly, no 10, 2011

Storage of CO_2 under the seabed began in the 1990s. Such a technology can reduce climate gas emissions. This can counteract global warming.

Professor Anders Kristoffersen, Faculty of Natural Sciences, University of Bergen
Letter to the editor in the newspaper Aftenposten, February 24, 2013

Much carbon is stored in the rainforest. By preserving the rainforest, we can prevent liberation of CO_2. This may effectively prevent global warming.

Student Lene Nygård, Halden Upper Seconday School
Excerpt from project work posted on dario.no, 2014.

The UN Climate Convention is an agreement on climate gas emissions. The goal is to reduce global warming. The agreement places a special responsibility on rich countries.

Nuclear power

Science teacher Jan Karlsen, Røros Upper Secondary School

Textbook in natural studies for upper secondary school, Gyldendal Publishers, 2012

The nuclear power plants of the future will produce less radioactive waste. This is because the waste can be used over and over again as fuel. Thereby, the problem of safe long-term storage can be solved.

Professor Anita Lund, Faculty of Natural Sciences, University of Stavanger
Nordic Journal of Nuclear Physics, no 3, 2013

Earthquakes can damage nuclear power plants. If the earthquake destroys the supply of electricity, overheating can occur. This can cause a meltdown of uranium.

Student Terje Andreassen, Bryne Upper Secondary School
Posting on terjesenvironmentblog.wordpress.com, June 14, 2016

Radioactive waste is a problem. The waste represents a serious health risk if not stored properly. The waste consists of spent uranium fuel.

Journalist Marianne Gundersen, News Section, Norwegian Broadcasting Corporation
Feature story on nrk.no, June, 2015

Sellafield is a recycling facility for radioactive waste. The facility is located on the coast of Great Britain. In Norway, there are concerns that the old tanks in which the waste is stored are insufficiently safe.

Journalist Geir Johnsen
Article in the newspaper VG, March 11, 2015

Several accidents have occurred in nuclear power plants. The worst accident happened in Chernobyl, Ukraine, in 1986. Several hundred people perished.

Director Bjørn Eriksen, Ministry of Climate and Environment
Report on Government.no/mce/ April, 2016

Construction of nuclear power plants must comply with international agreements on safety. The International Atomic Energy Agency is responsible for inspections of nuclear power plants. This is intended to prevent serious accidents.

Science teacher Hilde Dahl, Lillehammer Upper Secondary School
Article in Nature and the Environment, the membership journal of the Norwegian Society for the Conservation of Nature, no 4, 2011

In nuclear power plants, uranium is split. If this splitting gets out of control, an explosion may

occur. Harmful radioactive dust may then spread to the environment.

Chief engineer Hilde Jacobsen, The Halden Reactor
Technology Weekly, no 7, 2011

New technology makes nuclear power plants safer. Now, nuclear power plants that will shut down by themselves no matter what goes wrong are being built. Such power plants are under construction in China and Finland.

Professor Per Haugen, Faculty of Natural Sciences, University of Bergen
Letter to the editor in the newspaper Aftenposten, February 20, 2013

Some radioactive waste can remain dangerous for 100 000 years. Such waste must therefore be stored under safe conditions, for example in mine shafts or in specially constructed tunnels. Good storage sites are hard to find.

Student Kristin Iversen, Kongsberg Upper Secondary School
Excerpt from project work posted on dario.no, 2014.

Many countries want to close down their nuclear power plants. When nuclear power plants are demolished, much radioactive waste must be cleared. This is a difficult and hazardous job.

Author Note

Ivar Bråten (corresponding author) is a professor in the Department of Education at the University of Oslo, Norway; email ivar.braten@iped.uio.no.

Eva W. Brante was a postdoctoral research fellow in the Department of Education at the University of Oslo, Norway, when this study was conducted. She is now a senior lecturer in the Department of Culture, Languages, and Media at Malmö University, Sweden; email eva.wennas-brante@mau.se.

Helge I. Strømsø is a professor in the Department of Education at the University of Oslo, Norway; email h.i.stromso@iped.uio.no.