

Effects of reading real versus print-out versions of multiple documents on students' sourcing and integrated understanding

Ladislao Salmerón^a, Laura Gil^a, Ivar Bråten^{b,*}

^a University of Valencia, Spain

^b University of Oslo, Norway

ARTICLE INFO

Keywords:

Multiple document reading
Sourcing
Comprehension

ABSTRACT

This study investigated the extent to which students' sourcing and comprehension can be supported by the reading of real, as opposed to print-out versions of multiple documents. It was found that the reading of real rather than print-out versions of multiple documents on the issue of climate change increased students' memory for source information and made them include more specific references to document sources in argument essays that they wrote about the issue. In turn, such increased sourcing in essays mediated the positive effect of reading real versus print-out versions of documents on students' construction of coherent representations of the documents' content information. Theoretical and instructional implications of the findings are discussed, and directions for future research are provided.

1. Introduction

Research on multiple document reading has grown immensely during the last decades, building on seminal empirical and theoretical work that was published in the 1990s (e.g., Britt, Perfetti, Sandak, & Rouet, 1999; Perfetti, Britt, & Georgi, 1995; Wineburg, 1991) to become a vital and influential line of literacy research in recent years (Braasch, Bråten, & McCrudden, *in press*; Britt, Rouet, & Durik, 2018; Bråten, Braasch, & Salmerón, *in press*). An important insight gained from this line of research is that when reading multiple documents, paying attention to the sources of content information may be essential (Bråten, Stadtler, & Salmerón, 2018; Britt, Rouet, & Braasch, 2013; Rouet, 2006). In the context of multiple document reading, *sources* can be defined as information about individuals and organizations that create and publish document content, including information about when, where, and for what purpose the content is created and published (Bråten & Braasch, *in press*; Britt et al., 2013; Goldman & Scardamalia, 2013). Accordingly, *sourcing* can be defined as the process of attending to, representing, evaluating, and using available or accessible information about the sources of document content, for example about the author, publisher, or document type (Bråten et al., 2018). In keeping with this definition, Strømsø, Bråten, Britt, and Ferguson (2013), in a think-aloud study, distinguished between sourcing

activities where readers noted and remembered source information, evaluated the trustworthiness of sources, and used source information to predict and interpret document content (see also, Barzilai, Tzadok, & Eshet-Alkalai, 2015).

Although sourcing may help readers to read more critically and construct more integrated, accurate mental representations from multiple documents (Britt et al., 2013; Rouet, 2006), it is a somewhat disheartening fact that students at all educational levels often disregard source information and pay attention only to document content (for a recent review, see Bråten et al., 2018). At the same time, however, research has documented that characteristics of individuals as well as documents may influence the extent to which students source when reading multiple documents, with individual difference factors such as prior knowledge (Bråten, Strømsø, & Salmerón, 2011) and epistemic beliefs (Barzilai & Eshet-Alkalai, 2015) and document factors such as conflicts between documents (Kammerer, Kalbfell, & Gerjets, 2016) and topic familiarity (McCrudden, Stenseth, Bråten, & Strømsø, 2016) seemingly influencing their sourcing behavior.

Our study continues this line of research by addressing a specific document factor that might influence students' sourcing and, in turn, their comprehension of multiple documents: whether they read real or print-out versions of documents. Much previous research on students' sourcing in multiple document contexts has used print-out versions of

* Corresponding author at: Department of Education, University of Oslo, P.O. Box 1092 Blindern, N-0317 Oslo, Norway.
E-mail address: ivar.braten@ped.uio.no (I. Bråten).

real documents, presenting readers with excerpts from books, magazines, and newspapers on separate sheets of paper rather than asking them to study the information in the books, magazines, and newspapers from which those excerpts were taken (e.g., Bråten, Strømso, & Britt, 2009; Strømso, Bråten, & Britt, 2010).¹ For example, Strømso et al. (2010), who had upper secondary school students read multiple documents on climate change, used documents coming from a textbook, popular science magazines, newspapers, and project reports. However, each of these documents consisted of an excerpt from a real document that was printed on a separate sheet of paper. This design strategy has merit because it allows researchers to vary document content and source features such as author, publication, and date of creation, while keeping a range of other document characteristics, such as format, shape, and size, constant. However, our main assumption in the present study is that considering such real document characteristics as experimental noise rather than taking them into account when researching students' sourcing and comprehension may have unforeseen consequences. Crucial to our argument is the possibility that using print-out versions instead of real documents may blur boundaries between documents that are made salient by real document characteristics, making it harder for readers to identify important source features such as publisher or document type. Therefore, previous research may have underestimated students' sourcing abilities in natural, real-document contexts. Before specifying the rationale and the hypotheses for the present study, we discuss relevant theoretical assumptions and prior empirical work.

1.1. Theoretical assumptions

It is a basic idea within text comprehension research that text comprehension involves the construction of a coherent mental representation in which various textual ideas are connected in meaningful ways (McNamara & Magliano, 2009). In the context of multiple document reading, such integrated understanding at the intratext level can be considered a necessary first-step in multiple document comprehension (Rouet & Britt, 2011; van den Broek & Kendeou, 2015). Thus, according to the documents model framework of Britt and colleagues (Britt et al., 1999; Perfetti, Rouet, & Britt, 1999; Rouet, 2006), readers of multiple documents will ideally construct two representational structures in addition to those described in models of single text comprehension (e.g., Kintsch, 1998). First, readers need to construct a *mental model* that represents an integrated understanding of the situations or phenomena described across documents. Second, they need to construct an *intertext model* that represents links between source information and semantic content included in the mental model (i.e., who says what) as well as links between the different sources of information (e.g., Author A opposes Author B). According to the documents model framework, when readers construct links between sources and content as well as between sources (i.e., construct intertext models), this will help them understand conflicts that may be prevalent among multiple documents and reconcile the different perspectives. Of course, constructing intertext models requires that readers note and remember source feature information, referred to as source nodes, in the first place. In this framework, source nodes contain information about relevant source features (e.g., author, publisher, and document type) for each document, while associations between source nodes and semantic content represent source-content links and associations between different source nodes represent source-source links. Taken together, the

source-content links and the source-source links constitute the intertext model (Perfetti et al., 1999; Rouet, 2006).

More recently, Britt et al. (2013) extended the documents model framework by discussing the documents-as-entities assumption, emphasizing that proficient multiple document reading involves considering documents as social artifacts that are written by a particular author, at a particular time, for a particular purpose, and so forth. Further, these authors assumed that moving beyond the semantic content of documents to experience and represent them as such entities, is facilitated when there are distinct boundaries between documents. This is the case, for example, when readers interact with traditional books and magazines that have clearly demarcated boundaries and typically have source information prominently displayed on their covers, making it likely that readers create source nodes that can form the basis for intertext model construction in addition to processing the semantic content. In contrast, readers may interact with web pages with similar visual formatting, which is likely to blur or obscure the boundaries between the documents and, thus, make it more difficult to distinguish content derived from different sources (Britt et al., 2013). Presumably, situations where readers interact with multiple documents in the form of similarly looking printed excerpts also will reduce the distinctiveness of document boundaries compared to real documents and make it harder to identify them as unique instances with specific source characteristics.

The distinctiveness of document boundaries may not only be influenced by visual experiences, however; haptic experiences also come into play. That is, documents have haptic properties such as weight and texture that may define boundaries between them (e.g., the weight of a textbook vs. the weight of a newspaper), with haptic exploratory procedures that underlie representations of haptic information in memory instinctively employed when interacting with documents (Klatzky & Lederman, 2002; Klatzky, Lederman, & Reed, 1987). And, because interaction with real documents such as traditional textbooks, newspapers, and magazines provides readers with richer, more differentiated haptic experiences than interaction with web pages with similar formatting or printed excerpts of the documents (Mangen, 2008), they are likely to increase the distinctiveness of document boundaries, which may help readers tag document content for its source (Britt et al., 2013). Because attending to sources as well as content may help readers understand and reconcile different perspectives on a particular issue (e.g., by realizing that different authors may have different motives or competencies), increased sourcing, in turn, can be assumed to promote integrated understanding when reading multiple documents (Britt et al., 1999; Perfetti et al., 1999; Rouet, 2006).

Of note is that this emphasis on the benefits of haptic experiences when reading real documents is also consistent with the source-monitoring framework of Johnson and colleagues (e.g., Higgins & Johnson, 2012; Mitchell & Johnson, 2000). According to this framework, encoding of effective, that is, detailed and distinct cues, is required to remember origins of mental representations. Presumably, interaction with real documents provides readers with a range of such cues. Specifically, interaction with real documents provides readers with physical (e.g., concerning weight, texture, brightness, and odor) and proprioceptive (e.g., the position of the arms when holding a book vs. a newspaper) cues that may increase the distinctiveness of document boundaries, especially with respect to document type, and help them distinguish content derived from different sources. In contrast, when documents have been used in multiple document research, either in web-based or print-out versions (e.g., Bråten, Strømso, & Britt, 2009; Strømso et al., 2010; Wiley & Voss, 1999), such cues have been absent. For example, print-out versions of documents have typically been standardized in the sense that they are printed on identical sheets of paper with identical visual features, such as font and font style, line spacing, length, page layout, and so forth. Presumably, this way of presenting multiple documents may have removed a number of features that specify particular document types and, thus, obscured document

¹ An analysis of the materials used in 84 multiple document studies described in a recent comprehensive review of the literature (Bråten et al., 2018) showed that only 20 could be considered to have used real documents. Those studies used only digital materials, however, and allowed participants to freely navigate the Web to read the documents. In the other 64 studies, readers were either presented with print-out versions of real documents or with digital documents that were simplified and non-navigable adaptations of real web pages.

boundaries in ways that makes it harder to create source nodes while reading.

1.2. Prior research

No previous study compared the reading of real versus print-out versions of multiple documents, like we did in the current study. However, when reviewing the literature for prior empirical work that can be brought to bear on this issue, we identified two relevant, albeit somewhat different, lines of research. The first compared the reading of the same content information in one single document versus in multiple documents (Bråten & Strømso, 2006; Britt & Aglinskas, 2002; Stadler, Scharer, Brummernhenrich, & Bromme, 2013; Wiley & Voss, 1996, 1999), yielding results consistent with the view that more distinct boundaries between documents may facilitate students' document model construction, including their sourcing. The second compared the reading of documents differing with respect to haptic properties, such as a traditional printed newspaper versus an online version of the newspaper (Eshet-Alkalai & Geri, 2007; Neijens & Voorveld, in press; Tewksbury & Althaus, 2000) or multiple printed documents versus digital versions of the documents (Haber, Nacenta, & Carpendale, 2014; Takano, Shibata, & Omura, 2015), yielding results consistent with the idea that haptic experiences may influence processing and handling of documents. In the next two subsections, we further elaborate and exemplify these lines of research.

1.2.1. Reading one versus multiple documents

A much cited experiment by Wiley and Voss (1999) exemplifies this line of research. These authors had undergraduates read the same content about a historical topic that was presented either in a textbook-like chapter or in separate source documents in a web-like computer environment. Results showed that students who read multiple documents gained deeper, more integrated understanding of the topic and produced more transformed and integrated essays than did students who read the same content information in one single document.

In the same vein, Britt and Aglinskas (2002, Experiment 3) found that when high school students read about the same historical controversy in a single textbook-like document or in seven different documents presented in a computer application program called the Sourcer's Apprentice, those who interacted with multiple documents not only performed better with respect to sourcing (i.e., identification and evaluation of sources) on a transfer test, but also wrote essays on the controversy that were more integrated, cited more sources, and referenced more information from primary and secondary sources than did those in the single-document condition. Of note is that the students in the two conditions were presented with exactly the same information about sources as well as content, with source information embedded (i.e., cited) within the document in the textbook-like version of the materials.

In an example from the domain of science, Stadler et al. (2013) compared two groups of undergraduates who were given the same information about a controversial medical issue either as a single website written by one author or as four websites written by four different authors. This study showed that students who interacted with multiple documents rather than one single document, displayed better memory for conflicting information included in the materials and also wrote more balanced essays on the issue that took conflicting perspectives into consideration.

Finally, a study by Stadler, Scharer, Skodzik, and Bromme (2014) is relevant in this context. In that study, undergraduates read multiple documents from different sources on a controversial medical issue. These documents were presented as nine different websites. In one condition, differences between the documents were clearly signaled through rhetorical means, for example, by starting a document with the following phrase: "In contrast to what some health professionals believe ...". Participants who read documents where differences were explicitly

signaled had better memory for conflicts after reading and also wrote essays on the issue that reported more conflicts in a balanced way and included more references to sources, compared to participants who read the documents without such signaling.

Although none of these studies compared the reading of real versus print-out versions of multiple documents, they suggest that signaling document boundaries may increase students' sourcing and integrated understanding of document content. Presumably, document boundaries are weakened when documents are wrapped in a single textbook-like source, as in the single-document conditions described above (Britt & Aglinskas, 2002; Stadler et al., 2013; Wiley & Voss, 1999), or do not contain any explicit signaling of differences between the documents (Stadler et al., 2014). The results of these studies are therefore consistent with the documents-as-entities assumption of Britt et al. (2013). Of note is, however, that the multiple-document conditions in these studies involved the presentation of similarly looking websites.

1.2.2. The influence of haptic experiences

Thus far, haptic reading experiences have received most attention in the areas of media literacy and digital literacy (Mangen, 2008; Mangen, Walgermo, & Brønnick, 2013; Morineau, Blanche, Tobin, & Gueguen, 2005). In particular, research on undergraduates' reading of newspapers has suggested that real printed newspapers (i.e., not printed excerpts as used in multiple document research) better support critical reading and comprehension than do digital, online versions (Eshet-Alkalai & Geri, 2007; Neijens & Voorveld, in press; Tewksbury & Althaus, 2000). For example, Eshet-Alkalai and Geri (2007), who studied the ability of undergraduates to identify and explain biased, falsified, or manipulative information elements in news as a function of media (i.e., a paper vs an online version of a newspaper), found that those who read the traditional paper version of the newspaper identified and explained more biased information than those who read the digital, online version. Recently, Hou, Rashid, and Lee (2017) studied the influence of haptic experiences when reading comics, finding that undergraduates' feelings of immersion and comprehension were similar when reading comics on paper and in a digital version.

In a study using multiple documents, Takano et al. (2015, Experiment 1) had two groups of adults reading four one-page documents that included texts and graphics. One group read the documents on sheets of paper, while the other read them as PDF files on a desktop computer. In the paper condition, the working space was restricted to 27 in., the same size as the computer monitor. The documents included errors, either a mismatch between the text and the graph in the same document (single document error), or between graphs in different documents (multiple document error). Participants were requested to find inconsistencies between the texts and the graphs. Results showed that participants detected more errors (both single and multiple document errors) in the paper condition than in the desktop computer condition. It was also observed that those in the paper condition handled documents in ways that were not possible in the computer condition. For example, more than 25% of the document actions in the paper condition involved moving multiple documents at the same time, either in the same direction or in opposite directions. In the same vein, Haber et al. (2014) had undergraduate students work with six documents that were presented on either six sheets of paper or six digital tablets. In each condition, groups consisting of three students were tasked to collaboratively produce a journalistic report based on the six documents. Results showed that student behavior changed as a function of the medium, with those in the paper condition holding and using two documents simultaneously, as well as looking at and talking to their colleagues, more often than those in the digital tablet condition. Although these studies did not compare the reading of real and print-out versions of multiple documents, they are relevant in the present context because they suggest that different haptic experiences may afford or facilitate different ways of handling the documents. The potential effects of haptic experiences on individuals' sourcing and comprehension

when reading real versus print-out versions of multiple documents have hitherto not been examined, however.

1.3. The present study

Building on theoretical assumptions and empirical work discussed in the previous sections, the present study uniquely contributes to research on multiple document reading by investigating the extent to which students' sourcing and comprehension can be supported by the reading of real, as opposed to print-out versions of multiple documents. Based on the documents-as-entities assumption (Britt et al., 2013) and the assumption about haptic experiences' influence on reading (Mangen, 2008), as well as relevant research supporting those assumptions (e.g., Britt & Aglinskis, 2002; Eshet-Alkalai & Geri, 2007; Stadler et al., 2013; Tewksbury & Althaus, 2000), we expected that students who read a set of partly conflicting real documents about the issue of climate change would construct more complete document-level representations about the issue than would students who read the same documents printed on separate pages of paper (i.e., in print-out versions). (Please see Section 2.2.3 for more information about the real and print-out versions of the documents.) Specifically, we predicted that when reading multiple real documents, as compared to print-out versions, students would display better memory for the sources of the documents and include more references to sources in argument essays that they wrote about the issue. Consistent with the documents model framework (Britt et al., 1999; Perfetti et al., 1999; Rouet, 2006), we also expected that students who cited more sources when discussing the different perspectives on the issue in their essays would display more integrated understanding of document ideas. This expectation is also consistent with Strømsø et al. (2010), who found that students' ability to link different perspectives to the respective sources uniquely predicted their integrated understanding when reading multiple documents about the issue of climate change. Consequentially, we expected that the students' sourcing as evidenced in their essays would mediate the effects of reading real documents on their integrated understanding of document ideas.

To best isolate variance resulting from our experimental manipulation, we also wanted to control for the individual difference variables of prior knowledge and topic interest. Our choice of these two covariates was based on previous findings indicating that prior knowledge and topic interest may be associated with sourcing as well as integrated understanding when students read multiple documents (e.g., Bråten et al., 2011; Rouet, Britt, Mason, & Perfetti, 1996; Strømsø et al., 2010).

Of note is that this comparison of the reading of real and print-out versions of documents may suggest specific mechanisms that can facilitate or constrain sourcing and multiple document comprehension, such as the distinctiveness of document boundaries affected by haptic experiences. Comparing these conditions with the reading of authentic online documents may also be relevant from a practical perspective. However, authentic online documents involve a higher level of interactivity and include a number of elements not available in printed texts, for example videos, links to additional information, and reader comments and ratings. Thus, restricting our comparison to real and print-out versions of the same documents allowed us to control the level of interactivity and number of textual ideas and more clearly focus on the mechanisms of interest in the present study.

2. Method

2.1. Participants

The sample consisted of 101 undergraduates at a large public university in eastern Spain, who participated in the experiment for extra course credit. Seventy-two students were in the first year of a four-year teacher education program, 15 students were in their second year of the

teacher education program, and seven were in their third year of a four-year psychology degree program. In addition, five participants attended a two-year master program in philology and one participant was in the third year of a four-year speech therapy degree program. None of the study programs included any subject related to the topic discussed in the documents that we used in the current research. The sample included 82 females and 19 males, who ranged in age from 17 to 42 years and had an overall mean age of 21.1 ($SD = 4.6$). All participants were Caucasians and native Spanish speakers. Participants were assigned randomly to conditions.

2.2. Materials

2.2.1. Topic knowledge measure

Prior knowledge about the topic of climate change was assessed with a 17-item multiple-choice measure that has been used and validated in much prior research (e.g., Bråten et al., 2009; Gil, Bråten, Vidal-Abarca, & Strømsø, 2010; McCrudden et al., 2016; Salmerón, Gil, Bråten, & Strømsø, 2010). The measure contains items that are central to the issue of climate change and covers both scientific (e.g., the greenhouse effect) and political (e.g., the Kyoto protocol) aspects of the topic. Participants' scores were the number of correct responses out of 17. The internal consistency reliability (Cronbach's α) for participants' scores on the topic knowledge measure was 0.67. Test-retest reliabilities for the scores on this measure were computed in several independent samples of Spanish and Norwegian undergraduates, resulting in reliability estimates (Pearson's r) ranging from 0.73 to 0.77.

2.2.2. Topic interest measure

Personal interest in issues and activities concerning climate change was assessed with a 12-item multiple-choice measure, which also has been used and validated in much prior research (e.g., Bråten, Gil, Vidal-Abarca, & Strømsø, 2009; Stenseth, Bråten, & Strømsø, 2016). Participants indicated their level of interest by rating each item on a 10-point Likert-type scale ranging from 1 (not at all true of me) to 10 (very true of me). Half of the items assessed interest in the topic without reporting any active involvement (sample item: Global warming is an issue that interests me). The other half focused more on participants' active involvement in the issue, reflecting their willingness to act for the benefit of the Earth's climate (sample item: I try to convince others that we must reduce the discharges of climate gases). The internal consistency reliability (Cronbach's α) for participants' scores on this measure was 0.90.

2.2.3. Documents

Participants read four authentic documents about the topic of climate change. As can be seen in Table 1, the first document was a 526-word excerpt from a textbook that provided basic knowledge about the greenhouse effect and climate change, acknowledged that this is a controversial issue, and explained the function of the Intergovernmental Panel on Climate Change (IPCC). The second document was a 363-word editorial from a major Spanish newspaper that focused on the negative consequences of climate change and criticized the view and actions of the so-called climate sceptics. The third document was a 296-word blog entry written by a university professor who was an expert in climatology that discussed potential benefits of increases in CO_2 . Finally, the fourth document was a 286-word article from a popular science magazine arguing that efforts to prevent global warming, although necessary, will have great economic costs. Of note is that none of the documents contained inaccurate or irrelevant information about the topic of climate change.

As an indication of text difficulty, we computed readability scores for each of the four documents, using the Flesch-Szigriszt Index (Szigriszt, 1992), which is a version of the classic Flesch Index for texts in Spanish. The Flesch-Szigriszt Index is based on the formula, $206.835 - (62.3 \times S/P) - P/F$, where S is the number of syllables, P is the

Table 1
Overview of the four documents.

Document type	Author/Publisher	Content	Embedded sources	Number of words	Readability (Flesch-Szigriszt)	Presentation format
Science textbook	Publishing house (Santillana)	Describes the greenhouse effect and climate change; states that climate change is disputed; explains the function of the IPCC.	Swedish chemist Arrhenius; IPCC.	526	58.95	Textbook vs. print-out
Newspaper editorial	El País	Discusses negative consequences of climate change, such as the melting of icebergs, and criticizes climate sceptics.	Research groups; climate sceptics; IPCC; young voters.	363	55.14	Newspaper vs. print-out
Blog entry	Antón Uriarte, professor	Discusses positive consequences of climate change, such as the fertilization in arid regions, and criticizes ecologists.	Huffington Post; Geophysical Research Letters; Israel Institute of the Negev desert.	296	52.65	Tablet vs. print-out
Popular science magazine article	John Matson, editor of the Scientific American	Discusses economic consequences of change in energy model and the carbon footprint associated with alternative forms of energy.	Researchers at the University of Leiden; recent scientific study.	286	44.64	Popular science magazine vs. print-out

Note. IPCC = The Intergovernmental Panel on Climate Change.

number of words in the text, and F is the number of phrases (defined as the content between two punctuation marks). The mean readability score of the four texts was 52.8, indicating that the reading material was “somewhat difficult” according to the INFLESZ scale (Barrio-Cantalejo et al., 2008). This scale distinguishes five levels of text difficulty ranging from “very difficult” (readability < 40; e.g., science texts and undergraduate textbooks) to “very easy” (readability > 80; e.g., comics and primary school textbooks), with “somewhat difficult” (readability 40–55) characterizing reading materials such as popular science magazines.

In the real documents condition, participants read the textbook excerpt in a printed textbook, the newspaper editorial in a printed newspaper, the blog entry on a tablet, and the popular science article in a printed popular science magazine. In this condition, participants had to adapt their body posture to read each text in the physical context in which it was originally published, for example, changing the position of their arms when switching from a book to a newspaper. In addition to adapting their body posture, participants in this condition perceived differences between the documents with regard to a number of physical properties, such as size, weight, shape, and colors. When presented with the textbook, the newspaper, and the popular science magazine, participants were instructed to look for post-it notes that indicated the start and the end of the texts to be read. For instance, when reading the newspaper editorial, participants had the entire newspaper for that particular day available but were instructed to read only the text that was marked with the post-it notes. With respect to the blog entry, the document appeared directly on the screen when participants turned on their tablets.

In the print-out condition, participants read print-out versions of the same documents, with each text printed in black on a separate A4 sheet of paper, using Times New Roman, font size 12, and line space 1.5. Exactly the same source information that was available in the real document condition (i.e., author, document type, publisher, and date) was presented at the beginning of the texts.

As can be seen in Table 1, all of the documents cited at least two other sources. Following Strømso et al. (2013), we termed other sources cited within the boundaries of a particular document “embedded sources”. Specifically, the textbook excerpt included two embedded sources, while the newspaper editorial included four, the blog entry included three, and the popular science article included two embedded sources. None of the documents referred to each other, but one of the embedded sources occurred in two different documents (i.e., The Intergovernmental Panel on Climate Change was cited in the textbook excerpt as well as in the newspaper editorial).

2.2.4. Writing task

After reading the documents, all participants wrote an argument essay. Specifically, they were given the following written instruction: *The texts that you just read present different perspectives on climate change. Please write a report where you describe and evaluate the different perspectives, taking the arguments and evidence presented in the texts into consideration. Use approximately half a page for your essay.* Although we indicated that they could write approximately a half page of text, they were not restricted with respect to time or space when writing, and many wrote longer essays. Specifically, the average length of the essays was 316.9 words ($SD = 103.9$, $min = 173$, $max = 720$).

Following Magliano, Trabasso, and Graesser (1999), we first segmented every essay into idea units. An idea unit contained a main verb that expressed an event, activity, or state. If an utterance had two verbs and one agent, it was treated as having two separate idea units. Infinitives and complements were included with the main verb (Magliano et al., 1999).

Following segmentation, all essays were coded to indicate students' sourcing. In this analysis, we focused on all the segments that included references to source information from one or more of the documents. Further, we differentiated between three types of references to sources.

In accordance with Rouet et al. (1996), we distinguished between general and specific references to document sources. Segments were coded as containing general references to document sources when they referred to one or more of the documents without mentioning specific features of the sources. Such references could involve general terms such as “the first article”, “the texts” or “the documents” (e.g., “according to the texts that I read”, “one of the articles argued”). Segments were coded as specific references to document sources when they referred to specific, accurate source feature information, such as author, document type, or publication (e.g., “as presented in the newspaper”, “the document from the science textbook explains”, “the article published in the newspaper El País”). In addition to such general and specific references to document sources, we identified specific references to sources cited within the documents (i.e., embedded sources), for example, “at the end of the 19th century Arrhenius prophesied”, “according to the IPCC”).

Regarding our analysis of document content, segments were coded to indicate students’ understanding of the topic. In doing this, we distinguished three types of ideas: single idea paraphrases, intratext inferences, and intertext inferences. Idea units were coded as single idea paraphrases if students used their own words to restate an idea from one of the documents without changing the meaning expressed in the documents (e.g., *The glaciers in the western part of the icy continent are now in a process of fusion, not only inevitably but also irreversibly* [from original newspaper]/*The western glaciers are already suffering from a defrosting process that is irreversible and unstoppable* [from student essay]). Single-idea paraphrases thus contained accurate and relevant information about climate change that was presented in the documents. Still, because they did not reflect a coherent mental representation of document content, single-idea paraphrases can be considered to represent a superficial, piecemeal understanding of the topic of climate change (McNamara & Magliano, 2009; van den Broek & Kendeou, 2015).

Idea units were coded as intratext inferences if they contained some information from the text materials in combination with some information from students’ prior knowledge (e.g., *Both solar energy and wind energy are associated with their own emissions* [from original popular science article]/*I believe that most people rely on this alternative energy without knowing about its damaging side effects* [from student essay]). Idea units were also coded as intratext inferences if they combined two or more pieces of information from one document that were not connected in the text material (e.g., *Both solar energy and wind energy are associated with their own emissions* [from original popular science article] and *According to a study published last year that estimated the natural resources needed to lift new infrastructures, a large solar installation, for example, would take between one and seven years to offset the emissions from a coal-fired power station* [from original science magazine]/*A scientific study shows that renewable energies also pollute, but to a lesser extent* [from student essay]). Of note is that idea units in the essays were not coded as intratext inferences if they included misconceptions of ideas about climate change that were presented in the documents. Following Gil et al. (2010), we regarded misconceptions as false statements in relation to the content of the documents (e.g., *Six years of economic crisis did not prevent continued changes in the atmosphere and the oceans* [from original newspaper, sentence 3] and *The glaciers in the western part of the icy continent are now in a process of fusion, not only inevitably but also irreversibly* [from original newspaper, sentence 4]/*The crisis has much of the blame for the effects of climate change, because the problem with Antarctica increased when the crisis hit* [from student essay]). Because intratext inferences reflect coherent mental representations at the level of single documents, they can be considered to represent an integrated (i.e., within-document) understanding of the topic of climate change (McNamara & Magliano, 2009; van den Broek & Kendeou, 2015).

Finally, idea units were coded as intertext inferences if they combined two or more pieces of information across two or more documents (e.g., *The texts talk about a common issue, climate change, but they present*

it from different perspectives, because in some cases they make positive judgments and in other negative judgments [from student essay]). Again, idea units were not coded as intertext inferences if they included misconceptions of ideas about climate change that were presented in the documents (e.g., *It is certain that an increased concentration of greenhouse gases in the atmosphere contributes to the warming of the surface of the Earth* [from original textbook] and *The extraction of metals such as aluminium and nickel for the construction of solar panels and wind turbines requires huge amount of energy* [from original popular science article]/*We could do many things to prevent the change in the temperature, such as using other energies like solar or wind, but the problem is that some of the materials used to build such energies contaminate* [from student essay]). Because intertext inferences reflect coherent mental representations at the level of multiple documents, they can be considered to represent an integrated, cross-document understanding of the topic of climate change (Rouet & Britt, 2011; van den Broek & Kendeou, 2015).

Two independent judges independently scored a random selection of 10% of the essays. The Cohen’s Kappa agreement was 0.84 for the coding of students’ references to sources and 0.74 for the coding of their understanding of the content. According to Landis and Koch (1977), the agreement on sourcing was almost perfect and the agreement on understanding was substantial. All disagreements were resolved in thorough discussion between the two raters, and one of them scored the remaining essays according to the same coding systems.

2.2.5. Source memory task

Participants were requested to write down all the information they remembered about the documents they had just read, including the author, the document type (e.g., book, magazine), the publisher, and the date. For each document, we coded three source features: the author or publisher, the document type, and the date of publication. The reason author and publisher were not scored separately is that the source information provided for the textbook excerpt and the newspaper editorial did not mention the author of the documents. Because participants received one point for each source feature that they reported correctly, the maximum score for the source memory task was 12 points. Regarding author or publisher, responses were coded as correct if they remembered either the surname of the document’s author or the name of the document’s publisher. Regarding document type, responses were coded as correct if they answered either *book* or *textbook* for the textbook excerpt; *editorial*, *newspaper*, or *journalistic article* for the newspaper editorial; *blog* or *webpage* for the blog entry; and *magazine*, *science magazine* or *science article* for the popular science article. Finally, regarding date, responses were coded as correct if students reported the correct year of publication. No partial credit was given if information was uncertain (e.g., “I don’t remember if it was El País or ABC” [two major Spanish newspapers]). A random sample of 10% of the source memory statements were independently scored by two raters, yielding adequate reliability (Cohen’s Kappa = 0.79). All disagreements were resolved in thorough discussion between the two raters, and one of them scored the remaining source memory statements according to the same coding systems.

2.3. Procedure

Participants were tested individually in sessions lasting approximately 90 min. On arrival, they were sequentially assigned to either the real document condition ($n = 51$) or the print-out condition ($n = 50$). First, they answered a questionnaire on demographics and completed the prior knowledge and interest measures in that order. They completed these measures on a laptop using a spreadsheet application. After having completed these measures, they were presented with the documents and orally given the following instruction: *You are now going to read four texts that present different perspectives on climate change. After reading the texts, you are going to write a report that discusses the different perspectives based on the arguments and evidence presented in the texts.*

Table 2
Descriptive statistics for measured variables.

Variable	M (SD)	Skewness (SE)	Kurtosis (SE)	Minimum	Maximum
Prior topic knowledge	6.15 (2.36)	0.74 (0.24)	0.32 (0.47)	1	13
Topic interest	6.13 (1.37)	-0.45 (0.24)	0.14 (0.48)	2.67	9.42
General document sources	1.21 (1.49)	0.85 (0.24)	-0.81 (0.48)	0	4
Specific document sources	0.87 (1.38)	1.35 (0.24)	0.30 (0.48)	0	4
Embedded sources	1.36 (1.15)	0.74 (0.24)	0.17 (0.48)	0	5
Memory for sources	4.11 (1.84)	0.08 (0.24)	0.41 (0.47)	0	10
Single idea paraphrases	3.18 (2.04)	0.85 (0.24)	0.60 (0.48)	0	9
Intratext inferences	2.02 (1.25)	0.18 (0.24)	-0.49 (0.48)	0	5
Intertext inferences	0.52 (0.70)	0.98 (0.24)	-0.33 (0.48)	0	2

There is no time limit, and you can read and reread the texts in the order you prefer. Please note that you will not have the texts available while writing your report. The presentation order for the four documents was randomized for each participant in the real document condition, and each participant in the print-out condition was presented with the documents in the same random order as the preceding participant in the real document condition. In both conditions, the four documents were piled on a table in random order and all were available to participants during the entire reading task. After participants had finished reading the documents, they completed the writing task and the source memory task in that order on a laptop using a word-processing application.

3. Results

3.1. Descriptive and correlational analyses

Descriptive statistics for all measured variables are displayed in Table 2. As can be seen, all variables were approximately normally distributed. Further, the descriptive data indicated that, on average, participants had relatively low prior knowledge and were only moderately interested in the topic of climate change. On average, they also cited few sources in their essays; approximately one for each category (viz., general document sources, specific document sources, and embedded sources). Their scores on the source memory task indicated that they remembered about one third of the source features on the cued recall task. Finally, student essays mostly contained single idea paraphrases and intratext inferences, whereas very few intertext inferences were included in the essays.

As can be seen in Table 3, zero-order correlations between the measured variables indicated that scores on the source memory task was statistically significantly correlated with scores on all three measures of sourcing in essays. However, while the source memory task was positively related to specific references to document sources and references to embedded sources in the essays, it was negatively related to general references to document sources in the essays. Further, while the

source memory task was not related to number of paraphrases or inferences in the essays, references to both specific document sources and embedded sources in essays were positively correlated with intratext inferences (as was prior knowledge). Condition was only related to scores on the source memory task and number of specific references to document sources in the essays (please see below for an in-depth analysis of these relationships). Thus, as indicated by a lack of correlation between condition and prior knowledge and topic interest, respectively, participants in the two conditions did not differ with respect to prior knowledge or topic interest ($t_s < 1$).

3.2. Effects on sourcing

To test our prediction regarding the effects of condition on sourcing, we computed a set of analyses of covariance (ANCOVAs) with condition (real documents vs. print-out condition) as the independent variable and source memory, specific references to document sources, general references to document sources, and references to embedded sources as dependent variables. Because topic interest was not correlated with any of these dependent variables, only prior knowledge was included as a covariate in these analyses. Results of evaluation of the assumptions for performing ANCOVAs were satisfactory. The reason we computed a set of ANCOVAs instead of a multivariate analysis of covariance (MANCOVA) was that we considered the dependent variables to represent conceptually distinct aspects of sourcing that also might be differentially influenced by the condition (Huberty & Morris, 1989).

Regarding source memory, participants in the real documents condition remembered statistically significantly more source features ($M = 4.55, SD = 0.96$) than did those in the print-out condition ($M = 3.74, SD = 2.31$), $F(1, 98) = 4.97, p = .03, \eta_p^2 = .05$. The effect of the covariate (i.e., prior knowledge) was not statistically significant, with $F(1, 98) = 3.45, p = .07$.

Regarding sourcing in essays, participants in the real documents condition included nearly three times more specific references to document sources in their essays ($M = 1.27, SD = 1.52$) than did those

Table 3
Zero-order correlations between condition and measured variables.

Variable	1	2	3	4	5	6	7	8	9	10
1. Condition	–									
2. Prior topic knowledge	0.03	–								
3. Topic interest	-0.07	0.06	–							
4. General document sources	-0.04	-0.17	0.07	–						
5. Specific document sources	0.30**	0.13	-0.04	-0.29**	–					
6. Embedded sources	0.05	0.16	-0.04	-0.22*	0.18	–				
7. Memory for sources	0.23*	0.22*	0.03	-0.23*	0.43**	0.30**	–			
8. Single idea paraphrases	-0.07	-0.04	0.03	0.13	-0.02	0.32**	-0.06	–		
9. Intratext inferences	0.06	0.19*	-0.12	-0.04	0.37**	0.23*	-0.04	0.25*	–	
10. Intertext inferences	0.09	-0.04	0.12	-0.09	0.00	-0.04	-0.01	-0.28**	-0.13	–

Note.
* $p < .05$.
** $p < .01$.

in the print-out condition ($M = 0.46$, $SD = 1.09$), $F(1, 98) = 9.12$, $p = .003$, $\eta_p^2 = .09$. The effect of the covariate (i.e., prior knowledge) was not statistically significant in this analysis, with $F(1, 98) = 1.63$, $p = .21$. Finally, there were no effects of condition on general references to document sources (real documents: $M = 1.16$, $SD = 1.48$; print-outs: $M = 1.26$, $SD = 1.51$; $F < 1$) or embedded sources in the essays (real documents: $M = 1.41$, $SD = 1.13$; print-outs: $M = 1.30$, $SD = 1.18$; $F < 1$). Thus, consistent with our prediction, participants reading multiple real documents, as compared to print-out versions, displayed better memory for the sources of the documents and included more specific references to document sources in the essays that they wrote about the issue. According to Cohen's (1988) effect size benchmarks, the effect size for source memory was low to medium and the effect size for specific references to document sources in the essays was medium.

3.3. Effect on integrated understanding

Next, we tested our hypothesis that students' sourcing in essays would mediate the effects of reading real documents on their integrated understanding of document ideas. As indicated above, participants in the real documents condition included more specific references to document sources in their essays than did those in the print-out condition. Further, there was a positive relationship between specific document sources and intratext inferences in the essays, with participants including more specific references to document sources also more likely to integrate ideas within documents and across documents and prior knowledge (see Table 3). In turn, this pattern of findings allowed us to test the indirect effect of condition on integrated understanding via sourcing in essays. Of note is that this is possible although the c path (i.e., the effect of condition on intratext inferences) was not statistically significant (see Fig. 1), as would be required by the conventional causal steps approach (Baron & Kenny, 1986). However, this historically popular approach to mediational analysis has been criticized because of reduced power due to the multiple statistical significance tests that are needed to conduct it (Preacher & Selig, 2012), and because it can be considered illogical to assess mediation without directly testing it through the ab path (see Fig. 1) (e.g., Hayes, 2009). Therefore, we opted for the bootstrapping procedure developed by Preacher and Hayes (2008) in the current study, which holds no assumption about the statistical significance of the c path.

Specifically, we tested the effect of condition (real documents vs. print-outs) on the number of intratext inferences included in the essays, using the number of specific references to document sources as a mediator and scores on the prior knowledge measure as a covariate. Topic interest was not included as a covariate because it was not correlated with intratext inferences. In this analysis, condition was contrast coded (real documents = 1, print-out = -1) and the other variables were centered and standardized. The indirect effect was tested using a bootstrap estimation approach with 1000 samples (Preacher & Hayes, 2008). The model accounted for a statistically significant portion of the

variance, $R^2 = 0.17$, $F(3, 97) = 6.49$, $p < .001$. According to Cohen's (1988) effect size benchmarks, an R^2 of 0.17 can be considered a medium effect. Consistent with our hypothesis, the bootstrapped results showed a positive statistically significant indirect effect of condition on intratext inferences via sourcing in essays, yielding an estimate of 0.10 ($CI_{95\%}$: 0.04 to 0.20). As can be seen in Fig. 1, the direct effect of condition on intratext inferences remained statistically non-significant, $b = -0.05$, $SE = 0.10$, $p = .57$, which is consistent with a full mediation. Finally, the covariate of prior topic knowledge ($b = 0.15$, $SE = 0.09$, $p = .11$) was not a statistically significant predictor in this analysis.

4. Discussion

This study uniquely contributes to the area of multiple document reading by showing that the format of the documents may facilitate or constrain essential aspects of readers' document model construction. Specifically, the reading of real rather than print-out versions of multiple documents on the controversial issue of climate change increased readers' memory for source feature information and made them include more specific references to document sources in their essays. In turn, such increased sourcing in essays mediated the positive effect of reading real versus print-out versions of documents on students' construction of coherent representations of the documents' content information. Further, these effects were independent of students' prior knowledge about the topic of the documents. In the following, we discuss how our findings contribute to the literature on sourcing and multiple document comprehension, respectively, before we conclude the article with suggestions for future research and some instructional implications of our findings.

4.1. Effects of reading real documents on sourcing

Our findings regarding students' source memory and sourcing in essays are consistent with the documents-as-entities assumption of Britt and colleagues (Britt et al., 2013) and assumptions that haptic experiences may impact processes and products of reading (Mangen, 2008; Mangen et al., 2013). Presumably, reading real versus print-out versions of the documents offered participants a number of multisensory cues that helped them perceive documents as entities, characterized by different document types or genres, written by particular authors, disseminated by different publishers, and so forth. Examples of such cues are visually salient source information on the cover of the documents, the size, weight, and texture of the documents, and the position of the arms while holding the documents. Such cues may demarcate document boundaries and facilitate the creation of source nodes that can be incorporated into readers' documents model representations (Britt et al., 2013). In the current study, this was reflected in better memory for the sources of the documents and more specific references to document sources in essays for those who read real documents, compared to those who read print-out versions of the same documents, with these findings

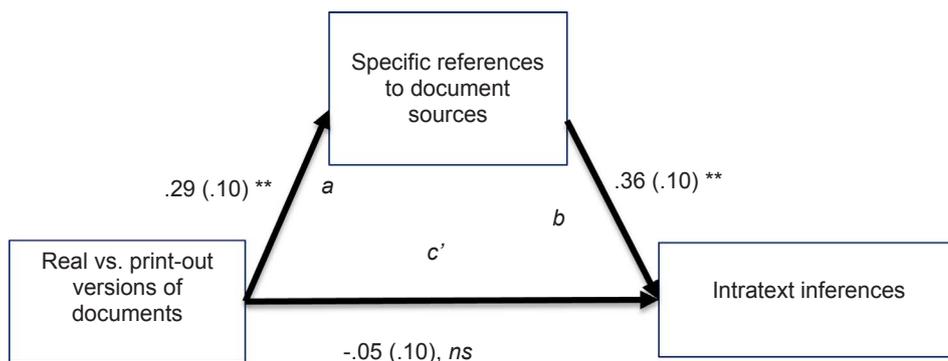


Fig. 1. Mediation model for the effect of condition (contrast coded: 1 = real, -1 = print-outs) on integrated understanding of document ideas with specific references to document sources in essays as a mediator (standardized coefficients).

also consistent with the source-monitoring framework of Johnson and colleagues (Higgins & Johnson, 2012; Mitchell & Johnson, 2000).

Interestingly, the reading of real documents did not influence students' references to embedded sources in their essays. This seems consistent with the documents-as-entities assumption, however. That is, the boundaries between sources cited within documents did not become more clearly demarcated through our experimental manipulation, because in neither condition could the embedded sources be distinguished by different haptic experiences or distinct multisensory cues, as was the case with the specific document sources.

Finally, general references to document sources in the essays were also not influenced by condition in this study. Of note is, however, that inclusion of such references was negatively related to the three other sourcing indices (i.e., memory for sources, specific document sources in essays, and embedded sources in essays). This pattern of findings suggests that such general references to sources do not represent adaptive sourcing in the context of multiple document reading. Rather, it seems like a surrogate that students may use when they do not really care about or remember more specific source information. Similarly, Rouet et al. (1996) found that specific and general references to sources served different purposes in students' essays. In their study, specific references to document sources were more frequently used to support facts and evaluations, whereas general references to document sources were mostly used to support claims, with the latter use of sources resembling situations where writers or speakers make weakly founded or even false assertions by referring to what many people think or say or what research (in general) shows.

4.2. Effects of reading real documents on document model construction

Our finding that sourcing played an important role in the construction of more coherent mental representations of document content is consistent with the documents model framework (Britt et al., 1999; Perfetti et al., 1999; Rouet, 2006) as well as prior research (e.g., Anmarkrud, Bråten, & Strømsø, 2014; Barzilai & Eseth-Alkalai, 2015; Bråten et al., 2009; Goldman, Braasch, Wiley, Graesser, & Brodowinska, 2012; List, Alexander, & Stephens, 2017; Strømsø et al., 2010; Wiley et al., 2009). More important than our finding that students who cited more specific document sources when discussing the different perspectives on the issue in their essays also achieved more integrated understanding of document ideas, however, was the finding that students' sourcing in their essays fully mediated the effect of reading real documents on their integrated understanding of document ideas.

The mediational effect was observed on the integration of ideas within documents and across documents and prior knowledge but not on the inclusion of isolated ideas from single texts or on cross-document integration. As explained above, the reading of real documents seemed to facilitate the creation of source nodes that could be incorporated into readers' document model representations. However, while source information may have scaffolded students' reasoning about and elaboration of the content of each document, they did not seem to have utilized source nodes as organizational elements to construct full document models involving the integration of information across texts. The reason we did not observe any effect on intertext inferences may be that such inferences were very infrequent in student essays, with many participants producing no such inferences at all. Of note is that the document set provided opportunities for several meaningful intertext inferences (e.g., regarding different views on the causes and consequences of climate change and different policies to address it), and that we used an argument writing task previously shown to support students' integration of document information (Naumann, Wechsung, & Krems, 2009; Stadler et al., 2014; Wiley & Voss, 1999; Wiley et al., 2009; see, however, Gil et al., 2010). Future research should therefore investigate whether the reading of real documents may support the cross-text integration of document content, either directly or indirectly via sourcing, among readers with higher levels of domain knowledge or

disciplinary expertise (Rouet, Favart, Britt, & Perfetti, 1997; von der Mühlen, Richter, Schmid, Schmidt, & Berthold, 2016; Wineburg, 1991). That said, constructing integrated understanding within documents can be considered an important process in multiple document comprehension that lays the foundations of full documents models involving cross-document integration (Rouet & Britt, 2011; van den Broek & Kendeou, 2015).

4.3. Implications for multiple document research and instruction

The unique results that we present in this study suggest that findings regarding students' sourcing (or the lack of it) obtained in previous multiple document research should be taken with caution. This is because previous work using versions of multiple documents likely to obscure document boundaries and thereby constrain sourcing may have underestimated students' sourcing activities in multiple document contexts. This certainly does not imply that previous findings in this area should be abandoned; yet, it suggests that the predominant view that students at different educational levels more often than not disregard sources and pay attention only to the content of documents (see, e.g., Bråten et al., 2018) needs to be critically questioned because the materials used in much prior work may have constrained rather than facilitated students' sourcing activities. At the same time, we acknowledge, of course, that the findings of the current study also need to be taken with caution. One reason for this is that the effect sizes that we obtained were not large according to Cohen's (1988) effect size benchmarks. That said, those benchmarks refer to the entire domain of social science rather than a particular area of research. In the area of multiple document literacy, most effect sizes obtained in intervention studies targeting sourcing skills with postsecondary students have been medium (for review, see Brante & Strømsø, *in press*), which means that the effects of our real documents condition actually were substantial when interpreted in the context of interventions to promote sourcing skills. However, another reason for caution is that the generalizability of our findings to younger students at lower educational levels is open to question (see Section 4.4 for further limitations to the current study).

The effect of reading real documents on sourcing that we discovered in this study opens intriguing new avenues for research on multiple document reading. A first issue that should attract the attention of researchers in this area is to what extent the effects found in previous studies still hold when they use multiple real documents, which can be expected to increase readers' attention to and use of specific document sources but not embedded sources. For example, the effect on sourcing of textual factors, such as the level of discrepancy between the views expressed in different documents (Braasch & Bråten, 2017), would not necessarily differ when readers work with real documents. However, other effects such as the critical evaluation of content in light of document type or genre (Bråten et al., 2011) may well increase in real document contexts because certain source features (such as document type) become more salient in such contexts.

Another possible avenue for future research on the effects of reading real documents concerns whether readers might experience the nature of knowledge and the process of knowing differently when reading such documents, as compared to excerpts from the same documents. For example, when reading real documents, readers might more easily grasp the interconnected nature of knowledge and the need to piece together an integrated understanding of a topic than when encountering information detached from a larger textual whole (e.g., a textbook or a magazine). In turn, such implications in terms of epistemic cognition (Sandoval, Greene, & Bråten, 2016) might indirectly affect readers' will and skill to construct integrated understanding from multiple documents.

This study also has instructional implications related to the design of multiple document reading tasks. Based on our findings, instructors may facilitate students' sourcing and integrated understanding of document content by presenting them with real instead of print-out

versions of documents. In this way, our findings seem to support the use of “old-fashioned” text-based documents in instructional contexts because such documents can enhance document boundaries and help readers create document level representations from the reading materials. On a more general note, this suggestion is also consistent with research indicating that digital reading (e.g., on computers or tablets) may have negative effects on comprehension, compared to print reading (e.g., Ackerman & Goldsmith, 2011; Gil, Martínez, & Vidal-Abarca, 2015; Mangen et al., 2013). Of course, this is not to say that only old-fashioned text-based documents should be used in school. This is hardly feasible in the current societal and educational context and may have unwanted economic and practical consequences as well. Still, in the challenging context of multiple document reading, using a combination of real and print-out versions of documents may actually pay off in terms of students’ performance.

Another implication is that when documents are provided online, for example as web pages, instructional designers should attempt to enhance the boundaries of their publications vis-à-vis other publications to facilitate attention to and memory of the particular source. Traditionally, publishers of books and music albums have been careful to create covers and formats that they hope will make their publications stand out from the crowd. In a time where more and more readers and listeners access and download digital materials instead of storing reading materials and music on their bookshelves and in their CD racks, boundaries between works may be obscured with the consequence that they are less frequently experienced and represented as distinct entities from specific sources. It goes without saying that this situation is a formidable challenge for designers and publishers of digital information sources, which in educational contexts also may have implications for learning and comprehension.

4.4. Limitations

Our study does not come without limitations, of course. Because we were interested in the mental representations constructed by students reading real versus print-out versions of multiple documents on a controversial socio-scientific issue, participants were not given access to the documents while writing their essays. It is conceivable that this design feature made the task more difficult and bears some responsibility for low scores on the essay performance measures, especially with respect to intertext inferences. In addition, this task context may not be typical for students reading multiple documents to produce an essay, where they can access and re-access documents while performing the writing task as well (O’Hara, Taylor, Newman, & Sellen, 2002). The extent to which our findings can be generalized to task contexts that differ from the one we created for this study is a question open for future research.

Moreover, our study focused on well-known document types, such as a textbook, a newspaper, and a popular science magazine, in addition to a blog entry. Such document types come in highly idiosyncratic formats, however. For example, newspapers represent a document type that, compared to other types, uses thin paper and is printed on large pages, and that even may come with the smell of fresh ink. Other document types may be more difficult to distinguish on the basis of format, such as a science fiction novel and a popular science book. A similar challenge may occur when people read two real documents of the same type, such as a newspaper, but when those documents have different levels of trustworthiness (e.g., a serious mainstream newspaper versus a tabloid). In such cases, the reading of real documents may not facilitate readers’ sourcing to the same extent and may even interfere with their construction of source-content links. Hopefully, this study will provide an impetus for additional inquiries attempting to clarify such issues.

Acknowledgements

Funding for the research reported in this article was provided by grant EDU2014-59422 from the Spanish Secretaria General de Universidades to Ladislao Salmerón and by grant 237981/H20 from the Research Council of Norway to Ivar Bråten. The authors want to thank Alicia Carrión for help in collecting and coding the data and Kate Ziegelstein for help in reviewing the materials used in prior research on multiple document reading.

References

- Ackerman, R., & Goldsmith, M. (2011). Metacognitive regulation of text learning: On screen versus on paper. *Journal of Experimental Psychology: Applied*, 17, 18–32.
- Anmarkrud, Ø., Bråten, I., & Strømso, H. I. (2014). Multiple-documents literacy: Strategic processing, source awareness, and argumentation when reading multiple conflicting documents. *Learning and Individual Differences*, 30, 64–76.
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182.
- Barrio-Cantalejo, I. M., Simón-Lorda, P., Melguizo, M., Escalona, I., Marijuán, M. I., & Hernando, P. (2008). Validación de la Escala INFLESZ para evaluar la legibilidad de los textos dirigidos a pacientes. *Anales del Sistema Sanitario de Navarra*, 31, 135–152.
- Barzilai, S., & Eshet-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.
- Braasch, J. L. G., & Bråten, I. (2017). The Discrepancy-Induced Source Comprehension (DISC) model: Basic assumptions and preliminary evidence. *Educational Psychologist*, 52, 167–181.
- Braasch, J. L. G., Bråten, I., & McCrudden, M. T. (Eds.) (in press). *Handbook of multiple source use*. New York: Routledge.
- Brante, E. W., & Strømso, H. I. (in press). Sourcing in text comprehension: A review of interventions targeting sourcing skills. *Educational Psychology Review*.
- Bråten, I., & Braasch, J. L. G. (in press). The role of conflict in multiple source use. In J. L. G. Braasch, I. Bråten, & M. T. McCrudden (Eds.), *Handbook of multiple source use*. 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., Gil, L., Strømso, H. I., & Vidal-Abarca, E. (2009). Personal epistemology across cultures: Exploring Norwegian and Spanish university students’ epistemic beliefs about climate change. *Social Psychology of Education*, 12, 529–560.
- Bråten, I., Stadler, 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* (pp. 141–166). (2nd ed.). New York: Routledge.
- Bråten, I., & Strømso, H. I. (2006). Effects of personal epistemology on the understanding of multiple texts. *Reading Psychology*, 27, 457–484.
- Bråten, I., Strømso, 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.
- Bråten, I., Strømso, H. I., & Salmerón, L. (2011). Trust and mistrust when students read multiple information sources about climate change. *Learning and Instruction*, 21, 180–192.
- Britt, M. A., & Aglinskias, 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.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Eshet-Alkalai, Y., & Geri, N. (2007). Does the medium affect the message? The influence of text representation format on critical thinking. *Human Systems Management*, 26, 269–279.
- Gil, L., Bråten, I., Vidal-Abarca, E., & Strømso, H. I. (2010). Summary versus argument tasks when working with multiple documents: Which is better for whom? *Contemporary Educational Psychology*, 35, 157–173.
- Gil, L., Martínez, T., & Vidal-Abarca, E. (2015). Online assessment of strategic reading literacy skills. *Computers & Education*, 82, 50–59.
- 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.
- Goldman, S. R., & Scardamalia, M. (2013). Managing, understanding, applying, and creating knowledge in the information age: Next-generation challenges and

- opportunities. *Cognition and Instruction*, 31, 255–269.
- Haber, J., Nacenta, M. A., & Carpendale, S. (2014). Paper vs. tablets: The effect of document media in co-located collaborative work. In F. Garzotto, P. Paolini, A. de Angeli, G. Jacucci, A. Malizia, M. Matera, & R. Lanzilotti (Eds.). *Proceedings of the 2014 international working conference on advanced visual interfaces* (pp. 89–96). New York: ACM.
- Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, 76, 408–420.
- Higgins, J. A., & Johnson, M. K. (2012). Some thoughts on the interaction between perception and reflection. In J. M. Wolfe, & L. Robertson (Eds.). *From perception to consciousness: Searching with Anne Treisman* (pp. 390–397). New York: Oxford University Press.
- Hou, J., Rashid, J., & Lee, K. M. (2017). Cognitive map or medium materiality? Reading on paper and screen. *Computers in Human Behavior*, 67, 84–94.
- Huberty, C. J., & Morris, J. D. (1989). Multivariate analysis versus multiple univariate analyses. *Psychological Bulletin*, 105, 302–308.
- Kammerer, Y., Kalbfell, E., & Gerjets, P. (2016). Is this information source commercially biased? How contradictions between web pages stimulate the consideration of source information. *Discourse Processes*, 53, 430–456.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.
- Klatzky, R. L., & Lederman, S. J. (2002). Touch. In I. B. Weiner (Series Ed.), A. F. Healy, & R. W. Proctor (Vol. Eds.), *Handbook of psychology: Vol. 4. Experimental psychology* (pp. 147–176). New York: Wiley.
- Klatzky, R. L., Lederman, S. J., & Reed, C. (1987). There's more to touch than meets the eye: The salience of object attributes for haptics with and without vision. *Journal of Experimental Psychology: General*, 116, 356–369.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- List, A., Alexander, P. A., & Stephens, L. A. (2017). Trust but verify: Examining the association between students' sourcing behaviors and ratings of text trustworthiness. *Discourse Processes*, 54, 83–104.
- Magliano, J. P., Trabasso, T., & Graesser, A. C. (1999). Strategic processes during comprehension. *Journal of Educational Psychology*, 91, 615–629.
- Mangen, A. (2008). Hypertext fiction reading: Haptics and immersion. *Journal of Research in Reading*, 31, 404–419.
- Mangen, A., Walgermo, B. R., & Brønneick, K. (2013). Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International Journal of Educational Research*, 58, 61–68.
- McCrudden, M. T., Stenseth, T., Bråten, I., & Strømso, 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. P. (2009). Toward a comprehensive model of comprehension. *Psychology of Learning and Motivation*, 51, 297–384.
- Mitchell, K. J., & Johnson, M. K. (2000). Source monitoring: Attributing mental experiences. In E. Tulving, & F. I. M. Craik (Eds.). *The Oxford handbook of memory* (pp. 179–195). New York: Oxford University Press.
- Morineau, T., Blanche, C., Tobin, L., & Gueguen, N. (2005). The emergence of the contextual role of the e-book in cognitive processes through an ecological and functional analysis. *International Journal of Human-Computer Studies*, 62, 329–348.
- Naumann, A. B., Wechsung, I., & Krems, J. F. (2009). How to support learning from multiple hypertext sources. *Behavior Research Methods*, 41, 639–646.
- Neijens, P. C., & Voorveld, H. (in press). Digital replica editions versus printed newspapers: Different reading styles? Different recall? *New Media & Society*.
- O'Hara, K. P., Taylor, A., Newman, W., & Sellen, A. J. (2002). Understanding the materiality of writing from multiple sources. *International Journal of Human-Computer Studies*, 56, 269–305.
- Perfetti, C. A., Britt, M. A., & Georgi, M. C. (1995). *Text-based learning and reasoning: Studies in history*. Hillsdale, NJ: Erlbaum.
- 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.
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40, 879–891.
- Preacher, K. J., & Selig, J. P. (2012). Advantages of Monte Carlo confidence intervals for indirect effects. *Communication Methods and Measures*, 6, 77–98.
- Rouet, J. F. (2006). *The skills of document use*. Mahwah, NJ: Erlbaum.
- 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., & 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., Favart, M., Britt, M. A., & Perfetti, C. A. (1997). Studying and using multiple sources in history; Effects of discipline expertise. *Cognition and Instruction*, 15, 85–106.
- Salmerón, L., Gil, L., Bråten, I., & Strømso, H. I. (2010). Comprehension effects of signaling relationships between documents in search engines. *Computers in Human Behavior*, 26, 419–426.
- Sandoval, W. A., Greene, J. A., & Bråten, I. (2016). Understanding and promoting thinking about knowledge: Origins, issues, and future directions of research on epistemic cognition. *Review of Research in Education*, 40, 457–496.
- Stadtler, M., Scharrer, L., Brummernhenrich, B., & Bromme, R. (2013). Dealing with uncertainty: Readers' memory for and use of conflicting information from science texts as a function of presentation format and source expertise. *Cognition and Instruction*, 31, 130–150.
- Stadtler, M., Scharrer, L., Skodzik, T., & Bromme, R. (2014). Comprehending multiple documents on scientific controversies: Effects of reading goals and signaling rhetorical relationships. *Discourse Processes*, 51, 93–116.
- Stenseth, T., Bråten, I., & Strømso, H. I. (2016). Investigating interest and knowledge as predictors of students' attitudes towards socio-scientific issues. *Learning and Individual Differences*, 47, 274–280.
- Strømso, H. I., Bråten, I., & Britt, M. A. (2010). Reading multiple texts about climate change: The relationship between memory for sources and text comprehension. *Learning and Instruction*, 20, 192–204.
- Strømso, H. I., Bråten, I., Britt, M. A., & Ferguson, L. E. (2013). Spontaneous sourcing among students reading multiple documents. *Cognition and Instruction*, 31, 176–203.
- Szigriszt, F. (1992). *Sistemas predictivos de legibilidad del mensaje escrito: Fórmula de perspicuidad*. Unpublished doctoral dissertation, Universidad Complutense de Madrid.
- Takano, K., Shibata, H., & Omura, K. (2015). Effects of paper on cross-reference reading for multiple documents: Comparison of reading performances and processes between paper and computer displays. In B. Ploderer, M. Carter, M. Gibbs, W. Smith, & F. Vetere (Eds.). *Proceedings of the annual meeting of the Australian special interest group for computer human interaction* (pp. 497–505). New York: ACM.
- Tewksbury, D., & Althaus, S. L. (2000). Differences in knowledge acquisition among readers of the paper and online versions of a national newspaper. *Journalism & Mass Communication Quarterly*, 77, 457–479.
- van den Broek, P., & Kendeou, P. (2015). Building coherence in web-based and other non-traditional reading environments: Cognitive opportunities and challenges. In R. J. Spiro, M. DeSchryver, M. S. Hagerman, P. M. Morsink, & P. Thompson (Eds.). *Reading at a crossroads? Disjunctures and continuities in current conceptions and practices* (pp. 104–114). New York: Routledge.
- von der Mühlen, S., Richter, T., Schmid, E. M., & Berthold, K. (2016). The use of source-related strategies in evaluating multiple psychology texts: A student–scientist comparison. *Reading and Writing*, 29, 1677–1698.
- 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.
- Wiley, J., & Voss, J. F. (1996). The effects of “playing” historian on learning in history. *Applied Cognitive Psychology*, 10, 63–72.
- Wiley, J., & Voss, J. F. (1999). Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Psychology*, 91, 301–311.
- Wineburg, S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, 83, 73–87.