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

A R T I C L E   I N F O

A B S T R A C T

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 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 & Eseth-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

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, Strømsø, & Britt, 2009; Strømsø, Bråten, & Britt, 2010).[1] For example, Strømsø 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, Strømsø, & Britt, 2009; Strømsø 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.