Postscript: In pursuit of integration

Alexandra Lista, Patricia A. Alexander

Abstract
In this conclusion to the Special Issue, features of multiple documents and multi-modal documents are systematized. Specifically, documents are defined according to their (a) intended message, (b) author and purpose for being created, (c) symbol system(s) used, and (d) related nature. Indeed, documents are conceptualized as presenting complementary content across symbol systems (i.e., multi-modal documents) or as presenting information on a common topic across multiple texts (i.e., multiple documents). Processes associated with each of these document features (i.e., encoding, comprehension, evaluation, and relation/integration) are specified. Finally, articles included as a part of this Special Issue are conceptualized according to the features and processing of multiple documents and multi-modal documents that they exemplify. The aim of this postscript is to move the field toward a model of multiple, multi-modal document comprehension and integration.

1. Postscript: in pursuit of integration

The last 25 years have seen the development of a literature examining how students engage with, evaluate, and comprehend information coming from multiple, primarily textual, sources (Perfetti, Rouet, & Britt, 1999; Wineburg, 1991). The premise of this literature is that processing and integrating information presented across multiple, disparate sources of information requires behaviors and strategies beyond those necessary when students have only to make sense of information coming from a single source (Britt & Aglinskas, 2002; Wiley & Voss, 1999). Concurrently, a literature has developed around students’ comprehension of multi-modal sources, or documents including information presented via multiple representations or symbol systems (e.g., via text and graphics; Azevedo & Cromley, 2004; Kozma, 2003a,b; Mason, Pluchino, Tornatora, & Ariasi, 2013). This literature has likewise been built around the idea that integrating redundant or complementary information presented across multiple modalities (e.g., text and image), requires particular processing beyond that necessary when making sense of textual information alone.

Independently, the literatures on multiple documents (MD) and multi-modal document (MMD) comprehension have identified an array of processes and competencies critical to living and learning in the 21st century. Yet for too long, the MD and MMD literatures have continued along separate tracks, with few researchers tackling how students may make sense of information from multiple and multi-modal documents. Nonetheless, as Ainsworth (this issue) poignantly illustrated in her commentary for this Special Issue, it is inevitable that today’s learners will repeatedly confront the challenges of dealing with multiple, multi-modal documents. Consequently, what is needed is a model of multiple, multi-modal document processing that adopts an expanded conception of literacy, beyond the comprehension of single, exclusively textual, sources of information (Alexander & The Disciplined Reading and Learning Research Laboratory, 2012; Goldman & Scardamalia, 2013; Leu, Kinzer, Coiro, & Cammack, 2004). With this pressing need in mind, our goal as editors of this Special Issue was to feature articles representative of these two lines of inquiry (i.e., multiple document and MMD processing) that exemplify the aims of this Special Issue to move the field toward a model of multiple, multi-modal document comprehension and integration.

Developing a unified model of comprehension requires conceptualizing what is meant by multiple documents vis-à-vis multi-modal documents and the processes both common and unique to each. In the sections that follow, we first discuss the characteristics of MDs and MMDs. Then, we examine the processes that underlie both multiple document and multi-modal document comprehension. Finally, we consider how the papers included in this Special Issue inform our understanding of the commonalities and differences between MD and MMD processing and move us toward a unified model of comprehension, as well as identifying directions for future research.

Received 27 January 2018; Received in revised form 4 April 2018; Accepted 9 April 2018
Available online 22 April 2018
0959-4752/ © 2018 Elsevier Ltd. All rights reserved.
2. Characteristics of multiple documents and multimedia documents

There are at least four characteristics of multiple documents and multi-modal documents that can be used to characterize these. Examining these four characteristics results in identifying the common and divergent features of MDs vis-à-vis MMDs. It should be noted that focusing specifically on multi-modal documents, to the exclusion of other multi-modal information (e.g., simulations), fails to fully capture all of the literature on multi-modal processing. Nevertheless, we considered this delimitation to be a necessary starting point to conceptualizing the relations between MDs and MMDs. To start, both multiple documents and multi-modal documents are means of conveying information to learners (Gernsbacher, Varner, & Faust, 1990). Moreover, the nature of this information is often complex necessitating the use of multiple documents or multiple representational formats (Ainsworth, 2006; Britt, Rouet, & Braasch, 2013). Second, both MDs and MMDs constitute authored entities (Alexander & Fox, 2004; Alexander & The Disciplined Reading and Learning Research Laboratory, 2012). In other words, these are documents produced by someone for some purpose. The authored nature of these works, and correspondingly, their degree of trustworthiness or lack-there-of, has received much attention in the multiple documents literature, but has been considerably less examined within the context of multi-modal comprehension (Bråten, Strømø, & Britt, 2009; Rouet, Britt, Mason, & Perfetti, 1996). Within the context of multiple documents comprehension, author credentials as well as other information about source origin (i.e., document information, like publisher) have been used as indicators of text authoritativeness, reliability, or potential bias. Such information has been considered to help students determine which information, from across documents, to prioritize or privilege, particularly in the face of conflict (Bråasch & Bråten, 2017; Britt, Perfetti, Sandak, & Rouet, 1999).

Conversely, the multi-modal document comprehension literature has considered issues of document information and trustworthiness evaluation to a much more limited extent. This has been due to multi-modal sources typically taking the form of multiple representations included within a single document (i.e., text with a graphic), reducing the need for students to reconcile conflicts across distinct documents and, therefore, to consider documents’ relative trustworthiness. Moreover, much of the literature on multi-modal document comprehension has focused on the use of multiple representations to teach students scientific content, primarily simulating textbook learning (Brenner et al., 1997; Dufresne, Gerace, & Leonard, 1997; Rozma, 2003). A textbook-based format of information presentations carries with it an implication of document trustworthiness. Whereas, such trustworthiness is often manipulated within studies of multiple document comprehension, by providing students with trustworthy and untrustworthy sources of information (Bråten et al., 2009; Wiley et al., 2009).

As a third defining feature, both MDs and MMDs are encoded via symbolic systems (DeLoache, 2002; Salomon, 1979). The key difference being that while multiple documents typically rely on a single symbol system to relay information (i.e., textual/alphabetic), multimedia documents are defined by their use of multiple symbol systems (e.g., graphic, numeric) to convey meaning. Nevertheless, regardless of the symbol system in use, students are required to understand both the explicit, face value of the symbols introduced, as well as their more latent or implicit meaning. This requires understanding the systematic rules and conventions governing each symbol system as well as its relative strengths and limitations. Within the context of text processing, this requires understanding both the literal definition of terms as well as their connotation or literary use. Within the context of graphic processing, this requires both the recognition of coordinates explicitly plotted on a graph as well as an interpretation of the trends these convey.

Finally, MDs and MMDs can be conceptualized relationally. In other words, multiple document sets are defined by the presence of several documents, related to one another in a variety of ways, all addressing aspects of a common topic. Likewise, multi-modal documents are defined by the presence of multiple representational systems, delivering the same or related information, within a single source. Ainsworth (1999) introduced a taxonomy of the possible relations arising primarily between multiple representations within a single document, but also potentially across documents. Specifically, Ainsworth suggests that multiple representations can have a (a) complementary relationship, either presenting complementary information or facilitating complementary processing, a (b) constraining relationship, whereby one representation scaffolds the interpretation of another, or a (c) constructive relationship, whereby one representation leads to the deeper processing of another. Nevertheless, all of these cross-modal relations can be considered to be complementary in nature, agreeing or buttressing one another in a variety of ways (Ainsworth, 1999). Such relations appear consistent with the multi-modal comprehension literature that typically introduces students to multiple representations of information within a single, and therefore likely internally consistent, document.

While a taxonomy of the possible relations among multiple documents has yet to be developed, a broad range of potential relations can be identified. To start, many of the complementary relations specified by Ainsworth and revisited in her commentary for this Special Issue (1999, this issue) may be evidenced in the relations among multiple documents within a text set. In particular, multiple documents may provide corroborating information, that is redundant across texts, complementary information that is distinct across texts but easy to synthesize, or facilitative information, with content in one document improving or deepening students’ understanding of another (Wiley et al., 2009; Wineburg, 1991). Beyond these relations, much of the multiple documents literature has focused on examining students’ understanding of conflicting information presented across documents (Kobayashi, 2010, 2015; Mason, Ariasi, & Boldrin, 2011; Stadtl & Bromme, 2014; Wolfe & Goldman, 2005). Such conflicting information may be directly oppositional or may simply be discrepant across texts, making reconciliation and integration a challenge.

In sum, while both multiple documents and multi-modal documents can be considered to be authored entities, conveying information to learners through various symbolic systems, these symbolic systems vary in the case of multi-modal comprehension, while nevertheless, relating to one another in a complementary fashion. Conversely, within the context of multiple document comprehension, while the symbol system used to convey information is typically linguistic, the relations among multiple documents are more varied, including possible discrepancy or opposition, necessitating attendance to author information or information about document origin.

3. Processes involved in multiple document and multi-modal document comprehension

The four characteristics of documents previously described (i.e., their intended meaning, author characteristics, symbolic representation, and relational nature) support students’ reasoning about each of these characteristics in a variety of ways. Specifically, each of these characteristics may be (a) encoded, (b) understood, and (c) evaluated by learners, as well as used to (d) facilitate the relating or comparison of information, either across symbol systems or across documents. As an initial step (i.e., encoding), learners must attend to and realize salient features of documents, including who wrote them and what symbol system was used. We know from work on both MD and MMD comprehension that this encoding process is hardly automatic (Bråten et al., 2009). For instance, students ignore document information, like author credentials or publisher, when processing texts (Britt & Aglinskas, 2002; List, Alexander, & Stephens, 2017) and fail to attend to diagrams in
documents, accompanying textual content (Cromley, Snyder-Hogan, & Luciw-Dubas, 2010; Mayer & Anderson, 1992).

Once learners have encoded various document features, prior knowledge is required to comprehend these. Again, the literature on both MD and MMD comprehension indicates the importance of this step. When processing graphic notation, even adults have been found to experience difficulties with knowing the common conventions of this symbol system (e.g., the meaning of an x-axis, van den Bosch, Espin, & Chung, 2017; Okan, Garcia-Retamero, Cokely, & Maldonado, 2017). At the same time, students' knowledge of particular conventions within certain symbol systems (e.g., text including a title to convey the main idea) have been found to facilitate comprehension (Justice & Ezzell, 2002; Meyer, Brandt, & Bluth, 1980).

Comprehension of salient document features may result in their evaluation. The literature on multiple document comprehension has commonly examined how students may evaluate aspects of document information (e.g., author expertise) to determine text trustworthiness (Bråten, Strømso, & Salmerón, 2011). Nevertheless, additional document features likewise need to be evaluated. For instance, a critical question for students to consider may be determining why certain information was presented via numerical or graphical notation, rather than through text. Likewise, students may need to be taught to question why political cartoons versus photographs were used in certain situations. Evaluation based on symbolic information (van Amelsvoort & Schilperoord, this issue; Van Meter & Cameron, this issue), although well-exemplified in this Special Issue, has only been begun to be examined in the literature more broadly.

As a final point, the four features of documents described herein also support relational thinking (Alexander, Dumas, Grossnickle, List, & Faretto, 2016; Grossnickle, Dumas, Alexander, & Baggetta, 2016). These include recognizing the similarities and differences in meaning that arise across symbolic representations within the same document, as well as the relations in ideas conveyed across documents. Indeed, a unified theory of multiple, multi-modal document comprehension requires considering how these aspects of documents may be integrated with one another for coherent understanding. Although fully specifying a model of multiple, multi-modal document integration is beyond the scope of this commentary, we can, at the least, suggest this process will unfold in a logical, albeit recursive, manner. In particular, we may expect students to first integrate a document's specific cross-modal features, with one another. For instance, this would involve synthesizing the information presented via text and an accompanying image, for holistic understanding. Moreover, individual documents would also be related to one-another, in whole or in part, or integrated, in a variety of ways. Such integration may include comparing photographs presented in two different newspapers, viewing an instructional video to better understand a concept in the textbook, or corroborating statistics on a website against census data (Wineburg, 1991). This document comparison and integration may further be expected to entail the recursive enactment of the processes of encoding, comprehension and evaluation.

Integration has been considered to include a range of processes, including those associated with synthesis, to link compatible information, and reconciliation, to recognize and resolve conflicts when these arise (Britt et al., 1999; Perfetti et al., 1999). Each of these processes may be considered to be quite complex onto themselves, as exemplified in Ainsworth's commentary for this Special Issue. Synthesis includes the combining, organization, and prioritizing of complementary information to build unified understanding. The inclusion of multi-modal content may aid cross-textual synthesis in a variety of ways, including by indicating key content to be integrated (e.g., content presented via multiple modalities) or by facilitating students' organizing of information via a variety of symbolic systems (e.g., text arranged in a tabular format; drawing an integrated model of a phenomenon).

Reconciliation includes the comparing of information within and across documents to identify discrepant and potentially conflicting content and the resolution of such discrepancies or conflicts in a variety of ways (Braasch, Rouet, Vibert, & Britt, 2012; Stadler & Bromme, 2014). Although a variety of mechanisms for reconciliation have been introduced (e.g., comparing author expertise; Stadler, Scharrer, Brummernhenrich, & Bromme, 2013), a unique strategy arising in multi-modal contexts may be reconciliation through the consideration of multiple modalities. In other words, including certain multi-modal information may aid students' reconciliation, for instance, by aiding their comprehension (i.e., helping them to decide what is true, as suggested by Stadler & Bromme, 2014) or by providing evidence (e.g., a photograph) that students can more easily evaluate and may consider to be more reliable than textual content. At the same time, the inclusion of such multi-modal information doubtlessly adds another level of complexity to the reconciliation task that students are asked to carry out and may overwhelm even competent readers (Ainsworth, this issue).

4. Articles in this Special Issue

Any model of multiple document, multi-modal comprehension needs to consider both the nature of such documents and the processes involved in their comprehension and integration. The articles included in this Special Issue speak to possible features of such a model in a variety of ways. Given the number of rich commentaries included in this Special Issue, only a partial description of each paper's contributions to such a model are introduced. First, Singer Trakhman, Alexander, and Silverman (this issue) point to the salient elements of documents that students may encode. For instance, students may recognize text, as a symbol system, somewhat differently when this symbol system is introduced via print versus digitally. In part, this may be because the conventions used to process this symbol system (e.g., scrolling) are somewhat different across these two delivery systems. Van Amelsvoort and Schilperoord (this issue) demonstrate both the explicit and implicit aspects of various symbol systems that students have to encode and comprehend. Specifically, van Amelsvoort and Schilperoord demonstrate that when processing boxes with arguments, as part of argument diagrams, students perceive these both literally (i.e., some boxes are larger than others) and figuratively (i.e., larger boxes correspond to arguments that are more prominent or important). Moreover, these authors discuss how such literal and figurative interpretations are the result of prior knowledge and students' experiences comprehending information presented via a variety of modalities.

Three of the papers in this Special Issue specifically demonstrate the complexities involved in multiple, multi-modal document integration. Schüler (this issue) systematically investigates the types of relations that may arise in multiple, multi-modal integration, by varying picture presentation alongside consistent or inconsistent information within a text. In accordance with the multi-modal relational taxonomy introduced by Ainsworth (1999), Schüler (this issue) demonstrates how graphic information can be used to facilitate text comprehension, particularly in the face of conflict. Van Meter and Cameron (this issue) presented students with images (i.e., photographs and political cartoons) alongside two historical documents, demonstrating both that some students may not explicitly attend to or process non-textual information and that images, in particular, may elicit affective processing. Such affective reasoning may support students' evaluation of information presented via various modalities. List (this issue) examined comprehension and integration when students were presented with information via two videos versus two texts, and created a coding scheme demonstrating how cross-document integration may emerge when documents are textual or multi-modal in nature. Taken together, these papers demonstrate each of the processes suggested in the outlines of our model of multiple, multi-modal document comprehension. These include encoding (Singer Trakhman, Alexander, & Silverman, this issue), comprehension (van Amelsvoort & Schilperoord, this issue), evaluation (Van Meter & Cameron, this issue), and relation/integration.
both across modalities (Schüler, this issue) and across documents (List, this issue).

## 5. Coda

As the Internet has become integral to learning in the 21st century, students today are increasingly expected to comprehend and integrate information from multiple sources (i.e., *multiple document processing*) and from sources that convey information not only through text, but also through a variety of media, including images, graphs, diagrams, and videos (i.e., *multi-modal processing*). As such, we contend that it is imperative to begin the process of formulating a model of comprehension that encompasses the cognitive and affective processes that are foundational to meaning making when multiple, multi-modal documents are involved. We hope that this Special Issue serves as a catalyst for the development of such an integrated model.

## Acknowledgment

We would like to thank Jan Vermunt for his stewardship and support in completing this Special Issue. We appreciate his efforts to make the entire process, from proposal to production, as smooth as possible. We would also like to sincerely thank all of the anonymous reviewers who generously provided feedback on articles in this Special Issue.

## References


