Measuring strategic processing when students read multiple texts

Ivar Bråten · Helge I. Strømsø

Abstract This study explored the dimensionality of multiple-text comprehension strategies in a sample of 216 Norwegian education undergraduates who read seven separate texts on a science topic and immediately afterwards responded to a self-report inventory focusing on strategic multiple-text processing in that specific task context. Two dimensions were identified through factor analysis: one concerning the accumulation of pieces of information from the different texts and one concerning cross-text elaboration. In a subsample of 71 students who were also administered measures of intratextual and intertextual comprehension after responding to the strategy inventory, hierarchical multiple regression analysis indicated that self-reported accumulation of information and cross-text elaboration explained variance in intertextual comprehension even after variance associated with prior knowledge had been removed.

The main purpose of this study was to design and validate a task-specific self-report measure of strategic multiple-text processing. Building upon the strategy conceptualizations of Weinstein and Mayer (1986) and Alexander et al. (1998), Bråten and Samuelstuen (2004) defined text comprehension strategies as forms of procedural knowledge that readers voluntary use for acquiring, organizing, or transforming text information, as well as for reflecting on and guiding their own text comprehension, in order to reduce a perceived discrepancy between a desired outcome and their current state of understanding. There is much prior research documenting that the use of deeper-level strategies falling in the broad categories of organization, elaboration, and monitoring is particularly important for successful comprehension performance when students read single texts (McNamara 2007; National Reading Panel 2000; Trabasso and Bouchard 2002). Especially, analyses of think-aloud protocols (Ericsson and Simon 1980) have shown that good readers are strategically

I. Bråten (✉) · H. I. Strømsø
Institute for Educational Research, University of Oslo, P.O. Box 1092 Blindern, N-0317 Oslo, Norway
e-mail: ivar.braten@ped.uio.no

active as they read, predicting upcoming text content, drawing inferences, asking and answering questions, creating images, reflecting on main points, and constructing personal interpretations (Pressley and Afflerbach 1995). In addition, good readers often jump back and forth in text and distribute their attention unequally, that is, paying more attention to some parts of the text than to others (Pressley and Harris 2006). In contrast, a hallmark of poorer readers seems to be that they read less actively, failing to produce the strategies when reading that good readers use to comprehend text. Moreover, poorer readers are much more likely to read word by word in a linear way (Duke et al. 2004; Pressley and Harris 2006). However, as shown by Andreassen and Bråten (2010, in press), the importance of deeper-level strategies may vary with the amount of working memory resources and the level of inferential processing demanded by the comprehension task.

Multiple-text comprehension

While there is a long and productive line of research on individuals trying to comprehend single texts, more systematic efforts to research how people try to understand a topic or issue by reading and making connections across multiple, diverse sources of information only dates back to the early 1990s (Spiro et al. 1991; Wineburg 1991). However, because it seems that the single-text paradigm still dominates research on text-based-learning and comprehension, it can be argued that research in this area is somewhat out of step with the intertextual reality encountering most readers in present-day society (Goldman 2004), where readers both in and out of school typically learn about different topics from multiple sources. For example, the role of textbooks is often downplayed in schools as primary and secondary source materials become more prominent, not least because of the rapid access to innumerable sources provided by complex computerized information systems.

While the availability and accessibility of abundant information sources undoubtedly afford unique learning opportunities, they also represent unique challenges for most readers regardless of age (Rouet 2006). This is because high-quality learning in this context requires more than locating and comprehending separate textual resources. In addition, in order to capitalize on the whole range of materials, learners need to integrate information across sources to construct a more complete representation of the topic or issue than any single resource can afford (Van Meter and Firetto 2008). Still, how students try to meet the challenge of comprehending and integrating information from multiple information sources is not well understood. It seems plausible that when content in one text largely overlaps with content in another text, integration across texts may proceed more or less automatically (Kurby et al. 2005). However, in most instances, it can be assumed that the construction of coherent mental representations from the reading of multiple texts requires considerable intentional strategic effort. In general, the comprehension of multiple texts puts considerable demands on readers' working memory resources as well as on their ability to create bridging inferences. Indeed, mental representations constructed from working with multiple texts can be regarded as authored by the readers themselves, more than by the authors of the individual texts. Therefore, it stands to reason that when students try to build an integrated understanding by reading multiple texts on a particular topic, more strategic effort is required than when they try to understand a single text on the same topic (for further discussion, see Bråten, Britt et al. 2011).

Research within the single-text paradigm indicates that good readers strategically combine ideas to construct a coherent representation, whereas poor readers often perform a "piecemeal processing" with minimal integration of information (Garner 1987). In the next

section, we discuss prior research on multiple-text comprehension indicating that corresponding strategy differences seem to exist between readers working with multiple expository texts.

Prior research on strategic processing of multiple texts

In a pioneer think-aloud study, where expert historians and high-school students read multiple texts on a historical event, Wineburg (1991) found that historians typically sought out and considered the source of each text to determine its evidentiary value, also using this source information in their interpretation of the text's content. For example, this "source heuristic" involved attention to the author, text type, and place and date of text creation. Wineburg also noted two other strategies heavily relied on by the historians, first, a systematic comparison of content across texts to examine potential contradictions or discrepancies among them (called a "corroboration heuristic" by Wineburg), and second, using prior knowledge to situate text information in a broad spatial-temporal context (called a "contextualization heuristic" by Wineburg). Through a systematic approach composed of these three heuristics, historians tried to piece together a coherent interpretation of the event described in the texts, at the same time paying close attention to the different sources on which this interpretation was based. In contrast, the high-school students participating in Wineburg's study seldom used these heuristics when reading the same texts, for example, often ignoring source information and regarding the textbook as more trustworthy than texts written by persons directly involved in the event (primary sources) and texts written by persons commenting on the event (secondary sources). Moreover, the students had difficulty resolving and even noticing discrepancies among sources.

Some years later, Stahl et al. (1996) also found that high-school students reading multiple history texts seldom used the three heuristics that Wineburg observed among historians. In Stahl et al.'s study, students were allowed to take notes during reading, and very few of the 20 students who took notes included any comments that could be considered as sourcing, corroboration, or contextualization. For example, only four students included a total of five comments that could be classified as corroboration in their notes. Please note that the heuristics discussed by Wineburg (1991) and Stahl et al. (1996) can be considered strategic processes used by historians when working with multiple texts (see Bråten, Britt et al. 2011).

In another think-aloud study, Strømsø and colleagues (Bråten and Strømsø 2003; Strømsø et al. 2003) had law students read multiple, self-selected texts in preparation for a high-stakes exam, finding that elaboration and monitoring strategies were more involved in the construction of an integrated situation model during multiple-text reading than were other types of strategies. In corroboration of this, Wolfe and Goldman (2005), who studied adolescents reading two contradictory texts explaining the Fall of Rome and thinking aloud after each sentence, demonstrated the value of elaborative processing. In that study, elaborative processing in the form of both within-text causal self-explanations and cross-text causal and comparative self-explanations indicated strategic attempts to develop a coherent account of the historical event based on the similarities and differences across all the information. However, VanSledright and Kelly (1998) noted that adolescents reading multiple texts in history may approach the task as accumulating as much factual information from the different texts as possible, without much attempt to elaborate upon the information or integrate information across the texts. Presumably, this concentration on accumulating historical facts may be reinforced by teacher beliefs in knowledge and

knowing in history as the memorization and reproduction of factual information (cf., VanSledright 2002). At the same time, it should be noted that all elaborative processing is not created equal, and that elaboration may not always be facilitative and sometimes even counterproductive, as when it involves the incorporation of idiosyncratic, irrelevant background knowledge into text representations (Martin and Pressley 1991; Samuelstuen and Bråten 2005; Williams 1993).

Recently, Afflerbach and Cho (2009) reviewed existing research on strategic reading of multiple texts and, drawing heavily on empirical studies cited above, identified three categories of "constructively responsive reading comprehension strategies" as involved in the reading of multiple texts: identifying and learning important information, monitoring, and evaluating. Among the reading behaviors included in the first category were comparing and contrasting the content of the text being read with the content of related texts to develop a coherent account of cross-textual contents, generating causal inferences by searching for relationships between texts and connecting information from current text with previous text contents, and organizing related information across texts by using related strategies (e.g., concept mapping, outlining, summarizing). The second category, monitoring, included reading behaviors such as detecting a comprehension problem with a particular text and trying to solve the detected problem by searching for clarifying information in other available texts, monitoring comprehension strategies and meaning construction with current text in relation to constructed meanings of other relevant texts, and perceiving that multiple texts related to the same topic can provide diverse views about the topic, complementary information about the topic, or both. Finally, the category of evaluating included reading behaviors such as using information about the source of each text to evaluate and interpret text contents, perceiving and distinguishing the characteristics of different texts (e.g., text types, age, author, prose styles) and evaluating texts' trustworthiness based on these features, and evaluating one text in relation to another, using specific information in each text (e.g., comparing claim and evidence in two or more texts).

It should be noted that the constructively responsive comprehension strategies discussed by Afflerbach and Cho (2009) are regarded as typical of expert and accomplished reading of multiple texts. Thus, when Maggioni and Fox (2009) analyzed the think-aloud protocols of American high-school students reading multiple history texts in light of the strategy categories identified by Afflerbach and Cho, not much was found. That is, on average, only 8% of participants' reading behaviors fell in one of the categories of multiple-text comprehension strategies identified by Afflerbach and Cho, with students often treating the different texts as if they were the paragraphs of a single text and extracting pieces of information from each text. Analyzing notes that students took during the reading of multiple science texts, Hagen et al. (2009) found that poorer comprehenders took notes paraphrasing pieces of factual information from single texts, whereas good comprehenders took notes that reflected elaborative integration of contents both within and across texts.

Measurement issues

Apart from a couple of studies analyzing students' note taking, most of our knowledge about the strategic processing of multiple texts is thus far based on think-aloud protocols. Think-aloud protocols contain concurrent verbal reports of how students set about particular learning tasks. This methodology has been regarded as an effective tool for gaining access to on-line strategic processing during reading (e.g., Ericsson and Simon 1980; Pressley 2000; Pressley and Afflerbach 1995), with its validity demonstrated through

convergence with reading time data (Magliano et al. 1999) as well as with data from retrospective debriefing (Taylor and Dionne 2000). From a practical perspective, however, think-alouds are a very time- and labour-intensive methodology in terms of instruction to students, administration, and scoring of protocols, which makes it less suitable for larger samples (Pintrich et al. 2000). In comparison, self-report inventories can be very easily and efficiently administered, completed, and scored, making them more relevant and useful than think-alouds when the purpose is to conduct larger-scale investigations and perform statistical analyses of student scores. However, as recently demonstrated by Samuelstuen and Bråten (2007) in the context of single-text reading, to obtain valid measurements when using self-report strategy inventories, it seems necessary to tailor those inventories to particular reading tasks or contexts. Samuelstuen and Bråten (2007) found that students' scores on an inventory where items referred to a recently completed reading task accounted for their comprehension performance, whereas scores on a general self-report inventory, asking students to report on their strategic processing across tasks and situations, did not relate to comprehension performance, nor to the strategic processing reported on in a specific task context. Please note that the very time- and labour-intensive nature of the think-aloud methodology can also be considered an issue with analyses of students' notes during multiple-text reading, as in the Stahl et al. (1996) and Hagen et al. (2009) studies.

Because learning about different topics from multiple texts containing supporting, complementary, or opposing information is increasingly required in classrooms, it seems pertinent to try to develop assessment tools that can be easily group administered and used to inform teachers as well as students themselves about strengths and weaknesses with respect to multiple-text comprehension strategies. Moreover, such assessments can form the basis of strategy instruction focusing on multiple-text comprehension and integration, also providing data that can be used to evaluate the (hopefully) beneficial effects of such instruction. Of course, an important purpose of such assessment tools would also be to provide quantitative data allowing researchers to investigate relationships between students' multiple-text comprehension strategies and potential antecedents (e.g., topic interest) and consequences (e.g., comprehension performance) of strategy use. So far, however, no self-report inventory focusing specifically on strategic processing in the context of multiple-text reading exists, at least not to our knowledge. In the current research, we therefore tried to fill this gap by designing and validating a task-specific strategy inventory targeting superficial as well as deeper-level strategic processing in the context of multiple-text reading. In doing this, we built on prior work that mainly used think-aloud methodology to describe students' strategy use when reading multiple texts, attempting to create an inventory that would capture a superficial multiple-text strategic approach involving the gathering of pieces of factual information from different texts as well as a deeper-level multiple-text strategic approach involving cross-text elaborative processing. Associating the former multiple-text strategic approach focusing on memorization with superficial processing and the latter focusing on elaboration with deeper-level processing also seems consistent with the terminology used by authors describing general learning strategies (Weinstein and Mayer 1986), single-text comprehension strategies (Alexander and Jetton 2000), and the processing of multiple texts (Goldman 2004). At the same time, we wanted to follow Bråten and Samuelstuen's (2007) four guidelines for constructing task-specific strategy inventories: First, a specific task (i.e., a set of texts) must be administered, to which the items on the inventory are referring. Second, the task must be accompanied by an instruction. The instruction should include information about task purpose (i.e., reading purpose). Additionally, readers should be directed to monitor their strategies during subsequent task performance and informed that they will be asked some questions

afterwards about how they proceeded. Third, to minimize the retention interval, the strategy inventory must be administered immediately after task completion. Fourth, in referring to recent episodes of strategic processing, the wordings of task-specific items must be different from more general statements. This means that general item stems such as "when I study" or "when I read" must be omitted in task-specific items. Moreover, to make it clear that the items refer back to the recently completed task, the verb should be in the past tense (e.g., "I tried to find ideas that recurred in several texts") instead of in the present tense (e.g., "I try to find ideas that recur in several texts"). It should be noted that Bråten and colleagues (Anmarkrud and Bråten 2009; Bråten et al. 2010; Bråten and Samuelstuen 2004, 2007; Samuelstuen and Bråten 2005, 2007) have demonstrated validity for self-reports on such a task-specific strategy inventory in several investigations of single-text reading.

The present study

Given this theoretical background analysis, we set out to design and validate a task-specific self-report measure that could target superficial as well as deeper-level multiple-text comprehension strategies. Because prior work on multiple-text comprehension strategies had not developed and used self-report inventories, such an inventory might complement other methodologies and provide an assessment tool for efficiently measuring how readers try to handle present-day intertextual challenges.

Specifically, we set out to address the following two research questions in our investigation: First, what is the dimensionality of multiple-text comprehension strategies as revealed by a factor analysis of scores on a task-specific self-report strategy inventory? Regarding this question, we expected that dimensions concerning both a superficial facts-gathering approach and a deeper-level cross-text elaborative approach would be identified. Second, do task-specific self-reports of multiple-text comprehension strategies uniquely predict students' comprehension performance when their prior knowledge about the topic of the texts is controlled for? In regard to this, we expected that self-reported use of superficial strategies would negatively predict multiple-text comprehension whereas self-reported use of deeper-level strategies would positively predict multiple-text comprehension. Acknowledging that a superficial facts-gathering approach may help the reader remember loosely organized declarative statements from the texts, we considered such an approach counterproductive in this study because our multiple-text comprehension measure assessed the ability to draw inferences across texts (see below). The expectancy that a superficial facts-gathering approach would negatively predict multiple-text comprehension is also consistent with research using other methodologies (Hagen et al. 2009; Maggioni and Fox 2009; Stahl et al. 1996; VanSledright and Kelly 1998). With respect to deeper-level cross-text elaboration, our expectation that such an approach would positively predict multiple-text comprehension is also consistent with other research on strategic processing of multiple texts (Hagen et al. 2009; Strømsø et al. 2003; Wineburg 1991; Wolfe and Goldman 2005). By including a measure of inferential single-text comprehension as a dependent variable, we hoped to demonstrate that self-reported multiple-text comprehension strategies would be better predictors of multiple-text comprehension than of single-text comprehension. The reason we decided to control for prior knowledge is that several previous studies (e.g., Bråten et al. 2009; Le Bigot and Rouet 2007; Moos and Azevedo 2008; Pieschl et al. 2008; Strømsø et al. 2010; Wiley et al. 2009) have shown relationships between prior knowledge and multiple-text comprehension and integration, with prior

knowledge presumably allowing readers to engage in bridging inferential processing to build links and coherence across texts.

Method

Participants

A total of 216 undergraduates at a large state university in southeast Norway, 177 females and 39 males, with an overall mean age of 22.8 ($SD=5.5$), read seven separate texts about different aspects of climate change and were administered a self-report inventory assessing their multiple-text comprehension strategies immediately after reading (see below). In the current study, these students' scores on the strategy inventory were used to explore the dimensionality of multiple-text comprehension strategies. In a subsample, consisting of 71 students who were also administered a prior knowledge measure before reading and single- and multiple-text comprehension measures after reporting on their strategy use, we examined the unique predictability of the resulting strategy dimensions for comprehension performance after controlling for prior knowledge. The subsample consisted of 59 females and 12 males, with an overall mean age of 23.0 ($SD=4.6$).

With very few exceptions, all participants were white native-born Norwegians who had Norwegian as their first language and had completed their secondary education in a Norwegian school. All participants were following a mandatory foundation course in educational science for students enrolled in bachelor programs in education and special education. This course spanned two semesters and covered 40 ECTS (European Credit Transfer and Accumulation System) credits. ECTS is based on the principle that 60 credits measure the workload for a full-time student during one academic year.

Materials

Prior knowledge measure A 15-item multiple-choice test was used to assess prior knowledge. The items referred to concepts and information central to the issue of climate change that were discussed in the seven texts (e.g., the greenhouse effect, climate gases, and the Kyoto Protocol). Sample items from the prior knowledge measure are displayed in the Appendix. Diverse aspects of the topic were covered by the prior knowledge measure, with items referring to both scientific (e.g., the greenhouse effect) and political (e.g., the Kyoto Protocol) aspects of the topic. Participants' prior knowledge score was the number of correct responses out of the 15 items. In an independent sample of first-year education undergraduates ($n=56$), we computed the test-retest reliability of the scores on this prior knowledge measure, with two weeks between the test and the retest. This yielded a reliability estimate (Pearson's r) of .77.

Texts The participants read seven separate texts about different aspects of climate change. One 362-word text about global warming was obtained from a textbook in nature studies for upper secondary education. This text explained the natural greenhouse effect and the manmade greenhouse effect in relatively neutral, academic terms. The second text was a 251-word popular science text published by the Center for International Climate and Environmental Research at the University of Oslo. This text focused on the causes of the manmade greenhouse effect, that is, on the manmade discharges of climate gases into the atmosphere and their contribution to observed climate changes. The third text was a 277-

word popular science article taken from a research magazine, where a professor of theoretical astrophysics argued that climate changes to a large extent are steered by astronomical conditions and therefore due to natural causes rather than mankind's activities. The fourth text was a 302-word newspaper article written by a journalist in a Norwegian liberal daily, describing the negative consequences of global warming in terms of a potential weakening of ocean currents in the North Atlantic and a melting of ice around the poles. The fifth text was a 231-word newspaper article written by a journalist in a Norwegian conservative daily, describing the positive consequences of a warmer climate in northerly regions in terms of an ice free sea route through the Northwest Passage and the access to natural resources now concealed under the Arctic ice. The sixth text was a 276-word public information text published by the Norwegian Pollution Control Authority, discussing international cooperation within the framework of the UN as a way to reduce the discharges of climate gases. Finally, the seventh text was a 303-word project presentation published by a large Norwegian oil company, describing new technology that could reduce the discharges of carbon dioxide into the atmosphere. Thus, apart from the more neutral textbook excerpt, the six other texts contained partly conflicting information, with two texts presenting different views on the causes of global warming (manmade versus natural), two texts presenting different views on the consequences of global warming (negative versus positive), and two texts presenting different views on the solutions to global warming (international cooperation versus new technology). Table 1 provides an overview of the seven texts that were used in this study.

As an indication of text difficulty, we computed readability scores for each of the texts. To be able to compare the difficulty of these texts with the difficulty of other kinds of Norwegian reading material, we used the formula proposed by Björnsson (1968), which is based on sentence length and word length. When using this formula, readability scores normally range from about 20 to about 60. For example, Vinje (1982) reported that public information texts from the information office of the Norwegian government had a readability score of approximately 45, while texts in the Norwegian code of laws had readability scores ranging from 47 to 63. As can be seen in Table 1, the readability scores of the seven texts ranged from 41 to 61, with a mean of 47.6, suggesting that the reading material represented a sufficient challenge for the participants and was well suited for eliciting effortful, strategic processing.

Each text was printed on one separate sheet of paper, and in addition to a title indicating the topic of the text, the source, author's name and credentials, and date of publication were presented at the beginning of each text. The seven texts were presented to the participants in random order, and they could read them in any order they preferred.

Multiple-text strategy measure In conjunction with the reading task, participants were directed to monitor the strategies that they used while studying the texts and told that they would be asked some questions about what they did during reading. We designed the Multiple-Text Strategy Inventory (MTSI) for this purpose. This 15-item questionnaire was based on prior empirical research on students' reading of multiple texts reviewed above. Thus, we wrote items to assess two hypothesized dimensions of multiple-text comprehension strategies, one concerning accumulation of pieces of information from multiple texts and one concerning cross-text elaboration. The five items concerning accumulation of pieces of information focused on the degree to which students tried to memorize as much information as possible from the seven texts (sample item: I tried to remember as much as possible from all the texts that I read). The 10 items concerning cross-text elaboration focused on the degree to which students tried to compare, contrast, and integrate contents

Table 1 Overview of the seven texts

No.	Type of text	Publisher	Author	Content	Number of words	Readability score
1.	Textbook in nature studies for upper secondary education	Publishing house	Teachers in upper secondary	Explanation of the natural greenhouse effect and the manmade greenhouse effect in relatively neutral, academic terms.	362	43
2.	Popular science text	Center for International Climate and Environmental Research	Not given	Focus on manmade discharges of climate gases into the atmosphere and their contribution to observed climate changes.	251	61
3.	Popular science article	University of Oslo	Professor of theoretical astrophysics	Argues that climate changes to a large extent are steered by astronomical conditions and therefore due to natural causes.	277	49
4.	Newspaper article	Norwegian liberal daily	Journalist	Description of the negative consequences of global warming in terms of a potential weakening of ocean currents in the North Atlantic and a melting of ice around the poles.	302	43
5.	Newspaper article	Norwegian conservative daily	Journalist	Description of the positive consequences of a warmer climate in northerly regions in terms of an ice free sea route through the Northwest Passage and the access to natural resources now concealed under the Arctic ice.	231	48
6.	Public information text	Norwegian Pollution Control Authority	Not given	Discussion of international cooperation within the framework of the UN as a way to reduce the discharges of climate gases.	276	48
7.	Project presentation	Norwegian oil company	Not given	Description of new technology that could reduce the discharges of carbon dioxide into the atmosphere.	303	41

across texts (sample item: I tried to note disagreements between the texts). All items of the MTSI are displayed in Table 2. Each item of the inventory referred to the recently completed reading task as a frame of reference (cf., Bråten and Samuelstuen 2007).

Please note that our intention was to capture two types of *multiple-text* comprehension strategies with the MTSI—one superficial focusing on the accumulation of pieces of factual information from multiple texts and one deeper-level focusing on comparing, contrasting, and integrating contents from multiple texts—it was not to capture one single-text opposed to one multiple-text strategic approach.

Each item was accompanied by a 10-point Likert-type scale, on which the participants rated to what extent they had performed the described activity while reading the texts (1=*not at all*, 10=*to a very large extent*).

Comprehension measures To measure multiple-text comprehension, we used an *intertextual inference verification task* (InterVT) focusing on participants' ability to draw inferences across texts. The task consisted of 20 statements, 8 of which could be inferred by combining information from at least two of the texts (i.e., valid inferences), and 12 of which could not be inferred by combining information from at least two of the texts (i.e., invalid inferences). The participants were instructed to mark the valid inferences *yes* and the invalid inferences *no*. The statements included in the InterVT were not just combinations of statements presented in the texts themselves but rather statements that required bridging inferences between information given in two or more texts. While 12 items combined information from two texts, eight items combined information from three or more texts. Participants' score on the InterVT was the number of correct responses out of the 20 items. The reliability (Cronbach's α) for the 71 participants' scores on this task was .61.

Finally, as mentioned in the introduction, we also assessed inferential single-text comprehension in this study. Specifically, to assess participants' deeper comprehension of each single text we used an *intratextual inference verification task* (IntraVT), based on the procedure developed and validated by Royer et al. (1996). The IntraVT consisted of 29 items that were constructed by combining information from different sentences within one of the texts to form either a valid or an invalid inference. There were 19 valid and 10 invalid inferences, and participants were instructed to mark those sentences *yes* that could be inferred from material presented in one of the texts and those sentences *no* that could not be inferred from material presented in one of the texts. An item did not just combine different statements presented in a text but was constructed to represent a connective inference that went beyond what was explicitly stated in the text. Each text was represented by at least three items. The items from the same text were not grouped together, and it was not signalled in any way on which text an item was based. The items were written so that they were verifiable by only one of the texts, and in a few instances, items were verifiable by one text but contradicted by another. In those instances, a response was judged to be correct if the participant marked the inference as verifiable because the task was to decide whether a statement could be inferred from material presented in one of the texts (se above). Participants' score on this task was the number of correct responses out of the 29 items. The reliability (Cronbach's α) for the 71 participants' scores on this task was also .61. While the reliability estimate for the scores on both comprehension measures were somewhat lower than desirable, reliability estimates in the .60s can still be considered acceptable with measures developed and used for research purposes (Nunnally 1978).

Further information about the design and validity of the InterVT, as well as a number of sample items for both types of verification tasks used in the present study, are contained in prior work (Bråten et al. 2009; Bråten and Strømsø 2010; Bråten et al. 2008; Strømsø et al. 2008).

Procedure

The reading materials and the multiple-text strategy measure were group administered to the participants in the entire sample (n=216) at the beginning of the autumn term. Before they started to read, all participants were given a written instruction asking them to imagine that they after reading should write a brief report to other students about climate change, based on information included in the seven texts. In addition, they were told: "When you study the texts, we want you to notice how you proceed. Afterwards, you will receive some questions about what you did while you studied the texts." Immediately after the participants had finished studying the texts, they were asked to complete the multiple-text strategy inventory to indicate to what extent they had used the procedures described there as they had studied the texts. Please note that the participants did not actually write any reports; they were just told to "imagine" that they should do so while reading. The participants were also informed that they could read the texts in any order they preferred.

Seventy-one participants were randomly assigned to a task condition were they in addition to reading the seven texts and responding to the multiple-text strategy measure were administered the prior knowledge measure before reading and the intertextual inference verification task (InterVT) and the intratextual inference verification task (IntraVT) after reading. Thus, the order of the experimental tasks was as follows for these participants: (a) prior knowledge measure, (b) text reading, (c) multiple-text strategy measure, (d) InterVT, and (e) IntraVT, with the InterVT and the IntraVT administered in counterbalanced order. There was a short written instruction at the beginning of each task. The instruction for text reading focused on the construction of arguments for this subsample, with the written instruction given before reading asking them to imagine that they should write a brief report to other students where they express and justify their personal opinion about climate change, basing their report on information included in the seven texts. The reason we wanted to examine relationships between self-reports of multiple-text strategy use and comprehension performance among students reading to construct arguments in this study, is that an argument task has been considered particularly suitable for eliciting strategies that promote the comprehension of multiple texts, especially strategies in the form of deeper-level elaborative and integrative cross-text processing (Gil et al. 2010; Le Bigot and Rouet 2007; Naumann et al. 2009; Wiley and Voss 1999; Wiley et al. 2009). Based on prior work that also considered other task instructions, we thus decided to focus exclusively on strategic processing of multiple texts in an argument task condition, assuming that such a task would highlight the need for effortful and complex processing of multiple texts (cf., Bråten, Gil and Strømsø 2011). Before starting on the two inference verification tasks, the participants in the subsample were also informed that they should not look back to the texts while answering these tasks. Please note that participants not given the argument task instruction worked on other tasks that were administered for other research purposes.

Results

Our first research question, concerning the dimensionality of multiple-text comprehension strategies, was addressed through factor analysis of 216 students' scores on the Multiple-Text Strategy Inventory (MTSI). Our second research question, concerning the unique contribution of multiple-text comprehension strategies to comprehension performance, was addressed with the subsample including 71 students. With this sample, we examined bivariate relationships between self-reports of strategy use and comprehension performance as well as the unique

predictability of strategy-scale scores for comprehension after variance associated with prior knowledge had been removed through forced-order hierarchical regression.

Factor analysis

In the sample of 216 undergraduates, maximum likelihood exploratory factor analysis was performed on the 15 items of the MTSI. Because we expected the factors to be correlated, we chose to conduct oblique rotation in this analysis. Prior to factor analyzing the data, we examined the item scores descriptively, with this analysis indicating that the score distributions for all items were approximately normal.[1] Initial factor analysis yielded three factors with eigenvalues greater than 1 that explained 59.1% of the total sample variation. However, because two of the four items falling on the third factor also loaded significantly and equally on another factor, because the eigenvalue of the third factor was only 1.05, and because the scree plot appeared consonant with a two-factor solution, we decided to explore this solution further. When we forced a two-factor solution on the 15 items, we identified two factors with high loadings (>.44) and no overlap for any item. This two-factor solution included all 15 items, and the two factors had eigenvalues of 5.08 and 2.74, respectively, explaining 52.1% of the total sample variation. Because all 10 items written to assess cross-text elaboration loaded on the first factor and all five items written to assess accumulation of pieces of information from multiple texts loaded on the second factor, we labelled the two factors Cross-Text Elaboration and Accumulation, respectively. The items assigned to each of the factors, as well as item-to-factor loadings for each factor, are shown in Table 2.

We also examined the reliabilities for the two factors emerging from the MTSI-data in the sample of 216 students. Cronbach's alpha for items loading on Cross-Text Elaboration and Accumulation were .88 and .82, respectively.

Finally, we performed a confirmatory factor analysis with Mplus 6 (Muthén and Muthén 2010) to examine how well the two-factor model suggested by the exploratory factor analysis represented students' multiple-text comprehension strategies. Taken together, the fit indices indicated an acceptable but not spectacular fit of the model, with $\chi^2(89, n=216)=244.47$, $p<.001$, CFI=.88, RMSEA=.089, and SRMR=.065.

Correlational and multiple-regression analyses

Descriptive statistics (means, standard deviations, coefficients of skewness, and coefficients of kurtosis) for all measured variables in the subsample of 71 students are shown in Table 3. These data indicated that all score distributions were approximately normal and, thus, appropriate for use in parametric statistical analyses. In the subsample, the reliabilities (Cronbach's alpha) for scores on the strategy measures based on the factor analysis were .76 for the accumulation measure and .86 for the cross-text elaboration measure.

Intercorrelations between the variables are also shown in Table 3. It can be seen that self-reported accumulation and cross-text elaboration strategies were uncorrelated ($r=.03$).

---

[1] Because we collapsed the group of participants given the argument task instruction with the other participants in the factor analysis, we also insured that no item score distributions were markedly skewed in any of the groups. Moreover, the standard errors of the means were comparable in the two groups, and, although the participants given the argument task instruction tended to obtain slightly higher scores on most of the items, none of these differences were statistically significant when a series of independent-samples t-tests with Bonferroni adjustment were performed. Only on two items were the effect sizes (Cohen's d) greater than .20 (.33 and .28, respectively).

Table 2 Factor analysis of the Multiple-Texts Strategy Inventory (*n*=216)

Variables	Factor loadings	
	Cross-text elaboration	Accumulation
I tried to understand the issues concerning climate by comparing the content of the different texts	.81	
I tried to note disagreements among the texts	.72	
I tried to find out how the content of the different texts hang together	.71	
I tried to find ideas that recurred in several texts	.69	
I considered whether the texts represented contradictory views on global warming	.68	
I considered whether different explanations of climate change can be reconciled	.67	
I tried to compare different causal explanations of climate change	.60	
I tried to summarize content from several texts	.60	
I tried to consider the connection between climate changes and environmental initiatives	.51	
I tried to form a complete picture of the consequences that may be caused by climate change	.49	
Cross-text elaboration Eigenvalue= 5.08		
I concentrated on remembering as much factual information as possible from all the texts		.83
I emphasized picking up as much information as possible from all the texts		.76
While I read the different texts, I gathered as much factual information as possible		.72
I tried to remember as much as possible from all the texts that I read		.68
I tried to remember as accurate as possible what I thought was important in the different texts		.44
Accumulation Eigenvalue= 2.74		

Moreover, albeit not statistically significantly, topic knowledge was negatively correlated with accumulation ($r=-.11$, $p=.184$) and positively correlated with cross-text elaboration ($r=.18$, $p=.066$). With respect to intertextual comprehension, self-reported accumulation was negatively correlated ($r=-.28$, $p=.011$) and self-reported cross-text elaboration was positively correlated ($r=.25$, $p=.021$) with performance. In comparison, the correlations of accumulation ($r=-.19$, $p=.057$) and cross-text elaboration ($r=.16$, $p=.086$) with intra-textual comprehension were lower and in both instances statistically non-significant. Please also note that prior knowledge was more strongly correlated with intratextual ($r=40$, $p=.000$) than with intertextual ($r=.25$, $p=.020$) comprehension.

To further examine the predictability of self-reported multiple-text comprehension strategies for performance, we ran two hierarchical multiple regression analyses, with the IntraVT and the InterVT as the dependent variables. In each analysis, prior knowledge about the topic of the texts was entered into the equation in Step 1. In Step 2, we included participants' self-reported use of accumulation and cross-text elaboration strategies.

The results of the hierarchical regression analysis for variables predicting intratextual comprehension are shown in Table 4. In this analysis, prior knowledge entered into the

Table 3 Descriptive statistics and zero-order correlations for measured variables ($n=71$)

Variable	1	2	3	4	5
1. Topic knowledge	-				
2. Accumulation	−.11	-			
3. Cross-text elaboration	.18	.03	-		
4. Intratextual comprehension	.40***	−.19	.16	-	
5. Intertextual comprehension	.25*	−.28*	.25*	.51***	-
M	6.87	6.13	6.69	19.82	13.37
SD	2.23	1.76	1.66	3.73	2.97
Skewness	.09	−.25	−.63	−.10	−.52
Kurtosis	−.46	−.76	.22	−.85	.17

***$p<.001$, *$p<.05$

equation in Step 1, explained a statistically significant amount of the variance in intratextual comprehension, $F(1, 69)=13.32$, $p=.001$, $R^2=.16$. The addition of accumulation and cross-text elaboration scores in Step 2 did not result in a statistically significant increment in the variance accounted for, $F_{change}(2, 67)=.97$, $p=.385$, $R^2=.18$. In this step, only prior knowledge uniquely predicted scores on the IntraVT, $\beta=.36$, $p=.002$, indicating that students with high prior knowledge were more likely to display good intratextual comprehension than were students with low prior knowledge.

The results of the second hierarchical regression analysis, with the intertextual comprehension measure as the dependent measure, are shown in Table 5. As can be seen, prior knowledge explained a statistically significant amount of the variance in intertextual comprehension in Step 1, $F(1, 66)=4.36$, $p=.041$, $R^2=.06$. The addition of scores on the two strategy measures in Step 2 of the equation resulted in a statistically significant 11% increment in the explained variance, $F_{change}(2, 64)=4.30$, $p=.018$, $R^2=.17$. In this step, self-reported use of accumulation negatively predicted intertextual comprehension, $\beta=−.26$, $p=.024$, whereas self-reported use of cross-text elaboration was a positive predictor, $\beta=.22$, $p=.058$. It should be noted, however, that the unique predictability of cross-text elaboration did not quite reach a conventional level of statistical significance with this sample size. Still, as we expected, the results of this analysis indicate that the contribution of multiple-text comprehension strategies may override the contribution of prior knowledge to readers' comprehension and integration of contents across texts.

Table 4 Results of hierarchical regression analysis for variables predicting intratextual comprehension ($n=71$)

Variable	B	SE B	β	t	p
Step 1					
Topic knowledge	.65	.18	.40	3.65	.001
Step 2					
Topic knowledge	.59	.19	.36	3.19	.002
Accumulation	−.27	.24	−.13	−1.12	.269
Cross-text elaboration	.22	.26	.10	.85	.398

$R^2=.16$ for step 1 ($p=.001$), $\Delta R^2=.02$ for step 2, ns

Table 5 Results of hierarchical regression analysis for variables predicting intertextual comprehension ($n=71$)

Variable	B	$SE\ B$	β	t	p
Step 1					
Topic knowledge	.33	.16	.25	2.09	.041
Step 2					
Topic knowledge	.24	.16	.18	1.53	.130
Accumulation	−.45	.19	−.26	−2.31	.024
Cross-text elaboration	.40	.21	.22	1.93	.058

$R^2=.06$ for step 1 ($p=.041$), $\Delta\ R^2=.11$ for step 2 ($p=.018$)

Discussion

The importance of building integrated representations of controversial topics and events through the reading of multiple information sources is greater than ever, both in and out of school (Rouet 2006). However, our knowledge of the strategic competence required to do so is thus far rather limited and mainly based upon very labor-intensive think-aloud methodologies. In this paper, we presented preliminary results from a first attempt to gauge students' strategic processing of multiple texts by means of a task-specific self-report inventory that is easily administered and scored, suitable for individual assessment as well as for large-scale investigations. These results uniquely contribute to the existing literature on strategic processing of multiple texts by indicating that such an inventory can capture both superficial and deeper-level strategies that differentially predict deeper-level intertextual comprehension, hopefully providing an impetus for more research in the area involving quantitative assessments and more advanced statistical analyses (e.g., structural equation modeling) that allow for testing complex (direct and indirect) relationships between multiple-text comprehension strategies and other constructs collectively.

First, in accordance with our hypothesis, the factor analysis of the MTSI-scores that we conducted indicated that superficial and deeper-level multiple-text comprehension strategies were differentiated by our strategy inventory. Ideas for the items of the MTSI were derived from reviewing prior studies using other methodologies, in particular think-alouds and note-taking, and the two resulting factors seemed to capture essential aspects of superficial and deeper-level multiple-text processing as described in that research (e.g., Hagen et al. 2009; Maggioni and Fox 2009; Stahl et al. 1996; Strømsø et al. 2003; VanSledright and Kelly 1998; Wineburg 1991; Wolfe and Goldman 2005), with one factor concerning the accumulation of pieces of information from different texts and one factor focusing on elaborative processing of contents across texts. The reliability estimates that we obtained for the factors also indicated that psychometrically sound measures could be constructed on the basis of the factors.

Second, the results of our regression of intertextual comprehension on prior knowledge and multiple-text comprehension strategies were also in accordance with our hypothesis. Thus, when the comprehension task involved the building of understanding by synthesizing information across texts, readers concentrating on accumulating as many pieces of information as possible from the different texts seemed to be disadvantaged. At the same time, readers who reported that they elaborated on the information by trying to compare, contrast, and integrate contents across texts were more likely to display good intertextual comprehension. These findings are also consistent with prior research describing novice versus expert processing of multiple texts (Maggioni and Fox 2009; Strømsø et al. 2003;

VanSledright and Kelly 1998; Wineburg 1991). In that research, think-alouds have revealed that whereas novices tend to adopt a piecemeal, facts-gathering approach (e.g., Maggioni and Fox 2009), experts opt for a strategy where they construct meaning by comparing, contrasting, and integrating contents across multiple texts (e.g., Strømsø et al. 2003; Wineburg 1991). Moreover, our finding that self-reports of strategies may contribute to intertextual comprehension over and above prior knowledge seems consistent with prior research. For example, in Wineburgs's (1991) study, differences in use of sourcing, corroboration, and contextualization were not simply due to differences in prior knowledge about the topic, with historians who did not know very much about the topic still using the same heuristics in their think-alouds. With respect to gaining an adequate, inferential understanding of each single text, however, prior knowledge seemed to play a more crucial role, presumably in combination with the use of other strategies not captured by the MTSI.

Our contribution comes with several limitations, of course. First, the two dimensions of multiple-text comprehension strategies that we identified were limited by the content of the inventory and, thus, we were probably not assessing all relevant dimensions of multiple-text strategy use. For example, items focusing on sourcing and contextualization, which may be relevant not only when studying multiple historical documents (Bråten, Strømsø et al. 2011), were not included in the inventory. In other words, our use of the MTSI restricted participants' reports of strategies to those prelisted on the inventory even though they may have used other strategies during reading. In particular, we consider it pertinent to add new items to the MTSI that focus on the strategic attention to and use of source information while reading multiple texts (cf., Afflerbach and Cho 2009; Bråten et al. 2009), with the dimensionality and predictability of such an expanded version examined in future work.

Second, this study included only Norwegian education undergraduates reading texts about climate change in order to construct arguments, and future work should examine the generalizablity of our findings by studying other populations reading multiple texts on other topics for other purposes.

Third, the fact that our data were correlational and gathered at a single point in time seriously restricts inferences that can be reliably made about causal relations. Firmer conclusions concerning causality may thus have to await longitudinal studies or experimental work demonstrating that instructing students to deal deeply and effectively with multiple texts may increase their scores on the cross-text elaboration subscale of the MTSI as well as their multiple-text comprehension. At the same time, such instruction could be expected to decrease students' scores on the accumulation subscale of the MTSI.

Despite existing limitations, our preliminary validation data seem promising and speak for continued work on this instrument. Approaches to further validation and refinement of the MTSI might also include comparing students' scores on this inventory to traces of studying recorded by software (Winne 2004, 2006) or eye movement data (Hyönä et al. 2003). While such methodologies are efficient ways of collecting data on online strategic processing, they also create huge amounts of data in need of considerable interpretation, in addition to requiring particular technical equipment. Other possibilities are to conduct small-scale studies comparing scores on the inventory to think-aloud data or presenting individuals with particular items and interviewing them about the interpretation of those items (Willis 2005). In any case, it seems pertinent to compare MTSI scores with online data because there is no denying the fact that even task-specific self-reports focus on students' perceptions or memories of strategy use. Therefore, they

may suffer from biased processing or reconstructive memory processes, with, especially, the social desirability of response alternatives threatening the validity of students' reports of strategies. Finally, while we considered the MTSI to capture two types of multiple-text comprehension strategies—one superficial and one deeper-level—it is clearly a relevant validity issue to what extent the inventory really measures multiple-text rather than single-text comprehension strategies. In future research, scores on the accumulation and cross-text elaboration scales could therefore be compared with scores on measures focusing on single-text superficial strategies (i.e., memorization) and single-text elaboration, respectively, with measures of both intratextual and intertextual comprehension performance also included in that research. While moderate positive relationships could be expected between single- and multiple-text superficial strategies and between single- and multiple-text deeper-level strategies, it could also be expected that single-text comprehension strategies would be better predictors of intratextual comprehension and multiple-text comprehension strategies better predictors of intertextual comprehension. Hopefully, further validation work along the lines suggested above could result in an instrument that is useful for practitioners and researchers alike.

Appendix

Sample Items for the Prior Knowledge Measure

The Kyoto Protocol deals with

a) trade agreements between rich and poor countries
b) reduction in the discharge of climate gases*
c) the pollution of the Pacific Ocean
d) protection of the ozone layer
e) limitations on international whaling

The greenhouse effect is due to

a) holes in the ozone layer
b) increased use of nuclear energy
c) increased occurrence of acidic precipitation
d) streams of heat that do not get out of the atmosphere*
e) the pollution of the oceans

Climate gases

a) do not occur naturally in the atmosphere
b) are necessary for much of the life on the earth*
c) did not exist in pre-industrial times
d) are exclusively synthetic combinations
e) can cause legionnaires' disease

References

Afflerbach, P., & Cho, B. (2009). Identifying and describing constructively responsive comprehension strategies in new and traditional forms of reading. In S. Israel & G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 69–90). New York: Routledge.

Alexander, P. A., Graham, S., & Harris, K. R. (1998). A perspective on strategy research: progress and prospects. *Educational Psychology Review, 10*, 129–154.

Alexander, P. A., & Jetton, T. L. (2000). Learning from text: A multidimensional and developmental perspective. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 285–310). Mahwah: Erlbaum.

Andreassen, R., & Bråten, I. (2010). Examining the prediction of reading comprehension on different multiple-choice tests. *Journal of Research in Reading, 33*, 263–283.

Andreassen, R., & Bråten, I. (in press). Implementation and effects of explicit reading comprehension instruction in fifth-grade classrooms. *Learning and Instruction.*

Anmarkrud, Ø., & Bråten, I. (2009). Motivation for reading comprehension. *Learning and Individual Differences, 19*, 252–256.

Björnsson, C. H. (1968). *Läsbarhet [Readability]*. Stockholm: Liber.

Bråten, I., & Strømsø, H. I. (2003). A longitudinal think-aloud study of spontaneous strategic processing during the reading of multiple expository texts. *Reading and Writing: An Interdisciplinary Journal, 16*, 195–218.

Bråten, I., & Samuelstuen, M. S. (2004). Does the influence of reading purpose on reports of strategic text processing depend on students' topic knowledge? *Journal of Educational Psychology, 96*, 324–336.

Bråten, I., & Samuelstuen, M. S. (2007). Measuring strategic processing: comparing task-specific self-reports to traces. *Metacognition and Learning, 2*, 1–20.

Bråten, I., & Strømsø, H. I. (2010). Effects of task instruction and personal epistemology on the understanding of multiple texts about climate change. *Discourse Processes, 47*, 1–31.

Bråten, I., Strømsø, H. I., & Samuelstuen, M. S. (2008). Are sophisticated students always better? The role of topic-specific personal epistemology in the understanding of multiple expository texts. *Contemporary Educational Psychology, 33*, 814–840.

Bråten, I., Strømsø, H. I., & Britt, M. A. (2009). Trust matters: examining the role of source evaluation in students' construction of meaning within and across multiple texts. *Reading Research Quarterly, 44*, 6–28.

Bråten, I., Amundsen, A., & Samuelstuen, M. S. (2010). Poor readers—good learners: a study of dyslexic readers learning with and without text. *Reading and Writing Quarterly: Overcoming Learning Difficulties, 26*, 166–187.

Bråten, I., Gil, L., & Strømsø, H. I. (2011). The role of different task instructions and reader characteristics when learning from multiple expository texts. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Relevance instructions and goal-focusing in text learning*. Greenwich: Information Age Publishing.

Bråten, I., Britt, M. A., Strømsø, H. I., & Rouet, J. F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: towards an integrated model. *Educational Psychologist, 46*, 48–70.

Bråten, I., Strømsø, H. I., & Salmerón, L. (2011). Trust and mistrust when students read multiple information sources about climate change. *Learning and Instruction, 21*, 180–192.

Duke, N. K., Pressley, M., & Hilden, K. (2004). Difficulties with reading comprehension. In C. A. Stone, E. R. Silliman, B. J. Ehren, & K. Apel (Eds.), *Handbook of language and literacy: Development and disorders* (pp. 501–520). New York: The Guilford Press.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87*, 215–251.

Garner, R. (1987). Strategies for reading and studying expository text. *Educational Psychologist, 22*, 299–312.

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

Goldman, S. R. (2004). Cognitive aspects of constructing meaning through and across multiple texts. In N. Shuart-Faris & D. Bloome (Eds.), *Uses of intertextuality in classroom and educational research* (pp. 317–351). Greenwich: Information Age Publishing.

Hagen, Å., Strømsø, H.I., & Bråten, I. (2009, August). *Epistemic beliefs and external strategy use when learning from multiple documents.* Paper presented at the biennial conference of the European Association for Research on Learning and Instruction, Amsterdam, The Netherlands

Hyönä, J., Radach, R., & Deubel, H. (Eds.). (2003). *The mind's eye: Cognitive and applied aspects of eye movement research*. Amsterdam: Elsevier.

Kurby, C. A., Britt, M. A., & Magliano, J. P. (2005). The role of top-down and bottom-up processes in between-text integration. *Reading Psychology, 26*, 335–362.

Le Bigot, L., & Rouet, J. F. (2007). The impact of presentation format, task assignment, and prior knowledge on students' comprehension of multiple online documents. *Journal of Literacy Research, 39*, 445–470.

Maggioni, L., & Fox, E. (2009, April). *Adolescents' reading of multiple history texts: An interdisciplinary investigation of historical thinking, intertextual reading, and domain-specific epistemic beliefs.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Magliano, J. P., Trabasso, T., & Graesser, A. C. (1999). Strategic processing during comprehension. *Journal of Educational Psychology, 91*, 615–629.

Martin, V. L., & Pressley, M. (1991). Elaborative-interrogation effects depend on the nature of the question. *Journal of Educational Psychology, 83*, 113–119.

McNamara, D. S. (Ed.). (2007). *Reading comprehension strategies: Theories, interventions and technologies*. New York: Erlbaum.

Moos, D. C., & Azevedo, R. (2008). Self-regulated learning with hypermedia: the role of prior domain knowledge. *Contemporary Educational Psychology, 33*, 270–298.

Muthén, L. K., & Muthén, B. O. (2010). *Mplus: Statistical analysis with latent variables. User's guide (Version 6)*. Los Angeles: Author.

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institute for Child Health and Human Development.

Naumann, A. B., Wechsung, I., & Krems, J. F. (2009). How to support learning from multiple hypertext sources. *Behavior Research Methods, 41*, 639–646.

Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw Hill.

Pieschl, S., Stahl, E., & Bromme, R. (2008). Epistemological beliefs and self-regulated learning with hypertext. *Metacognition and Learning, 3*, 17–37.

Pintrich, P. R., Wolters, C., & Baxter, G. (2000). Assessing metacognition and self-regulated learning. In G. Schraw & J. C. Impara (Eds.), *Issues in the measurement of metacognition* (pp. 43–97). Lincoln: Buros Institute of Mental Measurements.

Pressley, M. (2000). Development of grounded theories of complex cognitive processing: Exhaustive within- and between-study analyses of think aloud data. In G. Schraw & J. C. Impara (Eds.), *Issues in the measurement of metacognition* (pp. 261–296). Lincoln: Buros Institute of Mental Measurements.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale: Erlbaum.

Pressley, M., & Harris, K. R. (2006). Cognitive strategies instruction: From basic research to classroom instruction. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 265–286). Mahwah: Erlbaum.

Rouet, J. F. (2006). *The skills of document use: From text comprehension to Web-based learning*. Mahwah: Erlbaum.

Royer, J. M., Carlo, M. S., Dufresne, R., & Mestre, J. (1996). The assessment of levels of domain expertise while reading. *Cognition and Instruction, 14*, 373–408.

Samuelstuen, M. S., & Bråten, I. (2005). Decoding, knowledge, and strategies in comprehension of expository text. *Scandinavian Journal of Psychology, 46*, 107–117.

Samuelstuen, M. S., & Bråten, I. (2007). Examining the validity of self-reports on scales measuring students' strategic processing. *The British Journal of Educational Psychology, 77*, 351–378.

Spiro, R. J., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L. (1991). Cognitive flexibility, constructivism, and hypertext: random access instruction for advanced knowledge acquisition in ill-structured domains. *Educational Technology, 31*(5), 24–33.

Stahl, S. A., Hynd, C. R., Britton, B. K., McNish, M. M., & Bosquet, D. (1996). What happens when students read multiple source documents in history? *Reading Research Quarterly, 31*, 430–456.

Strømsø, H. I., Bråten, I., & Samuelstuen, M. S. (2003). Students' strategic use of multiple sources during expository text reading. *Cognition and Instruction, 21*, 113–147.

Strømsø, H. I., Bråten, I., & Samuelstuen, M. S. (2008). Dimensions of topic-specific epistemological beliefs as predictors of multiple text understanding. *Learning and Instruction, 18*, 513–527.

Strømsø, H. I., Bråten, I., & Britt, M. A. (2010). Reading multiple texts about climate change: the relationship between memory for sources and text comprehension. *Learning and Instruction, 20*, 192–204.

Taylor, K. L., & Dionne, J. P. (2000). Accessing problem-solving strategy knowledge: the complementary use of concurrent verbal protocols and retrospective debriefing. *Journal of Educational Psychology, 92*, 413–425.

Trabasso, T., & Bouchard, E. (2002). Teaching readers how to comprehend text strategically. In C. C. Block & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 176–200). New York: Guilford Press.

Van Meter, P. N., & Firetto, C. (2008). Intertextuality and the study of new literacies: Research critique and recommendations. In J. Coiro, M. Knobel, C. Lankshear, & D. J. Leu (Eds.), *Handbook of research on new literacies* (pp. 1079–1092). New York: Erlbaum.

VanSledright, B. (2002). *In search of America's past: Learning to read history in elementary school*. New York: Teachers College Press.

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

Vinje, F. E. (1982). *Journalistspråket [The journalist language]*. Fredrikstad: Institute for Journalism.

Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 315–327). New York: Macmillan.

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.

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

Williams, J. P. (1993). Comprehension of students with and without learning disabilities: identification of narrative themes and idiosyncratic text representations. *Journal of Educational Psychology, 85*, 631–641.

Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks: Sage.

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.

Winne, P. H. (2004). Students' calibration of knowledge and learning processes: implications for designing powerful software learning environments. *International Journal of Educational Research, 41*, 468–488.

Winne, P. H. (2006). How software technologies can improve research on learning and bolster school reform. *Educational Psychologist, 41*, 5–17.

Wolfe, M. B. W., & Goldman, S. R. (2005). Relations between adolescents' text processing and reasoning. *Cognition and Instruction, 23*, 467–502.