The design of technology-rich learning environments as metacognitive tools in history education

Eric Poitras · Susanne Lajoie · Yuan-Jin Hong

Abstract   Research has shown that learners do not always engage in appropriate meta-cognitive and self-regulatory processes while learning complex historical topics. However, little research exists to guide the design of technology-rich learning environments as metacognitive tools in history education. In order to address this issue, we designed a metacognitive tool using a bottom-up approach. Thirty-two undergraduate students read an historical narrative text either with or without the benefit of the metacognitive tool. Results from process and product data suggest that learners using the metacognitive tool had better recall and that the tool helped them (a) notice that particular events in an historical narrative text are unexplained, and (b) generate hypothetical causes to explain the occurrence of such events. We discuss the implications of these findings for the development of the MetaHistoReasoning Tool, a technology-rich learning environment that assist learners in terms of regulating their learning while they accomplish authentic tasks of historical inquiry.

Keywords   Technology-rich learning environments · Metacognitive tools · Top-down and bottom-up approach · History education · MetaHistoReasoning tool

Introduction

The study of history involves the examination of unique events, individuals, groups, movements, institutions, nations, and eras (Leinhardt and Stainton 1994). Critical thinkers explain historical events by weighing various points of view and evaluating a variety of historical evidence. Students generally learn about history from historical narratives. Learning history is a complex task in which learners must seek the causes of historical events from diverse historical sources, evaluate and interpret such information based on

E. Poitras (✉) · S. Lajoie · Y.-J. Hong
Department of Educational and Counselling Psychology, Advanced Technologies for Learning
in Authentic Settings Research Group, McGill University, 3700 McTavish Street, Montreal,
QC H3A 1Y2, Canada
e-mail: eric.poitras@mail.mcgill.ca
URL: http://www.mcgill.ca/atlas-lab/

their own prior knowledge regarding the events, and construct their own historical knowledge (Seixas 1993; Spoehr and Spoehr 1994). Student confusion arises when they read about an unexplained historical event, questioning its occurrence, and seeking its causes (van Drie and van Boxtel 2008).

While learning about complex historical topics students often engage in dysregulated behaviors (Azevedo and Feyzi-Behnagh 2010) during which time they fail to modify their behaviors, thereby resulting in minimal learning. Students often fail to engage in key self-regulatory processes such as setting learning goals, monitoring their progress, and using learning strategies to achieve their goals (Greene et al. 2010). We argue that teaching students to regulate their comprehension of historical narratives can be helpful. For this reason, we describe the literature on metacognition that can lead to the development of metacognitive tools that enhance historical comprehension.

Our goal is to develop, implement, and evaluate a technology-rich learning environment that can help students become more aware of their thought processes during their learning about complex historical topics. Technology-rich learning environments (TREs), as described by Lajoie and Azevedo (2006), are learning environments that are designed for an instructional purpose that use technology to support the learner in achieving the goals of instruction. TREs are designed based on a theory or model of learning and instruction.

The guiding metaphor for the design of computer-based learning environments in history education is that of using computers as cognitive tools (O'Neill and Weiler 2006) rather than metacognitive tools (Greene et al. 2010). Consequently, the existing TREs are not exclusively designed to assist students in being aware of the thought processes that are involved in learning about history. As such, we outline a generic framework that distinguishes between the top-down and bottom-up approaches to the design of metacognitive tools. The framework can be used to derive principles and choose elements of design that will foster new areas of research in the design of metacognitive tools in history education. In this paper we pay particular attention to evaluating the effectiveness of the bottom-up approach to designing a metacognitive tool for history education.

Technology-rich learning environments in history education

Literature reviews on the research and development of TREs in social studies education show progress in aligning the design and implementation of technology with empirical evidence (Ehman and Glenn 1991; Berson 1996; Martorella 1997; Whitworth and Berson 2003; Bolick et al. 2003, 2006, 2007; Bolick 2004; Berson and Balyta 2004; Lee and Hicks 2006; Friedman and Hicks 2006; Thieman 2008; Swan and Hofer 2008; Lee and Friedman 2009). Although a substantial body of empirical research has investigated the design of TREs as cognitive tools in history education (for a review of this literature, see O'Neill and Weiler 2006), research has yet to address the pressing need to design TREs as metacognitive tools to enhance the learning of complex historical topics.

TREs can be designed as cognitive tools to enhance students' learning (Lajoie and Derry 1993; Lajoie 2000, 2005, 2007, 2009; Lajoie and Azevedo 2006; Jonassen 1996, 2003; Jonassen and Reeves 1996; Kommers et al. 1992; Pea 1985; Perkins 1985; Salomon et al. 1991; Salomon and Perkins 2005). Cognitive processes are involved in the transformation, reduction, elaboration, storage, recovery, and use of sensory input (Neisser 1967). Lajoie and Derry (1993) defined cognitive tools as any TRE that assists learners in accomplishing cognitive tasks by (a) supporting cognitive processes such as memory and metacognition, (b) sharing the cognitive load through supporting lower-order thought

processes so that learners may focus on higher-order thought processes, (c) allowing them to engage in thought processes that would otherwise be out of their reach, and (d) enabling learners to generate and test hypotheses in the context of problem solving.

There are several examples of TREs that support historical reasoning and embody the cognitive tools principles. For example, Decision Point (Brush and Saye 2000, 2001, 2002, 2004; Saye and Brush 1999, 2002, 2004, 2005, 2006, 2007, 2009; Wolfe et al. 2003), SCIM Historical Inquiry Tutorial (Doolittle and Hicks 2003; Hicks and Doolittle 2008, 2009; Hicks et al. 2004), and the Sourcer's Apprentice (Britt et al. 1996, 2000, 2004; Rouet et al. 1996; Britt and Aglinskas 2002), among others, provide learners with a variety of different types of cognitive tools which offer assistance as they learn about complex historical topics. Decision Point (Saye and Brush 2002) is a multimedia environment that facilitates learners' cognitive processes by assisting them in collecting, analyzing, and evaluating historical evidence as well as presenting conclusions. Decision Point provides learners with various types of embedded scaffolds that help them handle the cognitive demands of a historical reasoning task. These scaffolding mechanisms embed expert guidance within the learning environment to assist students in better developing their critical reasoning in the context of problem-based learning.

Increasingly, researchers are interested in using technology to foster students' meta-cognitive and self-regulatory processes that are critical to enhancing learning (Azevedo 2005a, b). Metacognitive processes involve the evaluation of ongoing cognitive processes (i.e., metacognitive monitoring processes) as well as their resulting regulation or alteration (i.e., metacognitive control processes) (Flavell 1979; Nelson and Narens 1990, 1994; Nelson 1996). These theoretical constructs have been extended through self-regulated learning theories (Pintrich 2000, 2004; Winne and Hadwin 1998; Winne and Perry 2000; Schunk 2005; Zimmerman 2000, 2008) that account for the monitoring and control of affective, behavioral, contextual, and cognitive processes (see Dinsmore et al. 2008 for a literature review). Technology can be designed with metacognitive tools that detect, trace, model, and prompt students' metacognitive and self-regulatory processes in order to enhance learning (Azevedo et al. 2009b).

For example, TREs such as iStart (Graesser et al. 2005; McNamara et al. 2007a, b; Boonthum et al. 2007), Inquiry Island (White and Frederiksen 2005; Eslinger et al. 2001, 2008; Shimoda et al. 1999, 2002; White et al. 1999), and MetaTutor (Azevedo et al. 2009a, b), among others, support learners in their attempt to regulate their learning about complex science topics. Azevedo et al. (2009b) developed MetaTutor, a multi-agent intelligent system that detects, models, traces, and fosters students' self-regulatory processes while learning about complex and challenging biological systems in a hypermedia-based environment. The design of MetaTutor is based on identifying the self-regulatory processes that human tutors teach their students in order to facilitate learning (see Azevedo 2008, for a literature review). MetaTutor trains students in using these processes through video clips of human agents as well as discrimination and detection tasks in order to enhance science learning.

However, little is known about the potential benefits of designing metacognitive tools to enhance learning about history. We address this issue by creating a framework for metacognitive tools that distinguishes between a bottom-up and a top-down approach to their design (Poitras 2010). We exemplify the framework for the creation of metacognitive tools to assist students in regulating their understanding while learning about complex historical topics. The terminology and assumptions underlying the top-down and bottom-up approaches to designing metacognitive tools are outlined in the following section (see Fig. 1).

Logical Foundations for Deriving Elements and Principles
of Design for Metacognitive Tools

	Theory	**Design**
Top-Down Approach	Effective Metacognitive Processes	Deducing Principles of Design
Bottom-Up Approach	Ineffective Cognitive Processes	Inferring Principles of Design

Fig. 1 The main assumptions underlying the top-down and bottom-up approach

Top-down and bottom-up approaches to designing metacognitive tools

The top-down and bottom-up approaches constitute a generic framework that describes the logic or rationale underlying the design of metacognitive tools. This logical framework defines a set of guidelines for choosing elements and deriving principles of design for metacognitive tools on the basis of theory and empirical evidence in order to enhance learning across disciplines. As such, this framework guides researchers in designing and developing metacognitive tools across different domains, theoretical models, and evidentiary base. In this section, we describe the framework in the context of developing metacognitive tools that enhance learning about complex historical topics.

Our definition of a top-down approach to the design of metacognitive tools is based on two assumptions. The first assumption is that researchers should capture the deployment of metacognitive and self-regulatory processes as they fluctuate over time. These processes, which occur at the meta-level or "top" layer (Nelson and Narens 1990, 1994; Nelson 1996), can then be examined in regards to their impact on learning. Using this process-oriented model, researchers are able to identify the metacognitive and self-regulatory processes that are effective in terms of enhancing learning. The second assumption is that elements and principles of design can be established, based on an instructional model or empirical evidence, in order to foster the effective metacognitive and self-regulatory processes. This logic or rationale underlying the development of metacognitive tools is referred to as the top-down approach.

As an example, Greene et al. (2010) used what we call a top-down approach to develop metacognitive tools for history education. Greene et al. (2010) asked high-school students to perform a concurrent think-aloud protocol while learning complex historical topics embedded in a hypermedia-based environment. Almost all of the students gained both declarative knowledge about the historical topic (i.e., the Regulator Movement) and historical thinking skills, although these gains were modest taking into account the measurement scale. Examining the deployment of macro-level self-regulatory processes and their impact on learning outcomes, these researchers concluded that engaging in planning was related to the amount of historical knowledge measured at post-test. Planning involves stating or restating goals or sub-goals, as well as their time requirements, immediately before taking action. They deduced, based on their model, that a metacognitive tool should be designed to support learners in engaging appropriately in planning self-regulatory processes in order to enhance learning.

Furthermore, the top-down approach is also used to guide the design of metacognitive tools in science education. MetaTutor (see Azevedo et al. 2010a) is a multi-agent intelligent system that aims to foster self-regulatory processes that are critical to learning about the circulatory system. The design of this metacognitive tool is based on previous research that examined how human tutors assist students to effectively regulate their learning. Azevedo and his colleagues deduced, based on their evidence, that MetaTutor should be designed to assist learners in engaging in key self-regulatory processes in order to enhance learning.

Our definition of a bottom-up approach to the design of metacognitive tools is based on two different assumptions. The first assumption is that researchers should capture the deployment of cognitive processes as they fluctuate over time. These processes occur at the object-level or "bottom" layer (Nelson and Narens 1990, 1994; Nelson 1996) and can be examined in regards to their impact on learning. Researchers use this process-oriented model to identify the cognitive processes that are ineffective and detract from learning. The second assumption is that metacognitive tools can be designed to foster metacognitive and self-regulatory processes that are expected to resolve cognitive processing difficulties. This rationale underlies the development of metacognitive tools using a bottom-up approach. In the following sub-sections we describe how the bottom-up approach can be used to design metacognitive tools for enhancing learning about complex historical topics.

First assumption underlying the bottom-up approach

The first assumption underlying the bottom-up approach to designing metacognitive tools is that a process-oriented model should be used to pinpoint the cognitive processes that are ineffective in a specific learning context. In this section, we argue that historical narrative texts with causal coherence gaps are detrimental to learning because they interfere with learners' inferential processing. Moreover, a principled method of discourse analysis (Trabasso et al. 1989) should be used to derive a process-oriented model of learners' inferential processing in order to pinpoint difficulties that are likely to arise due to causal coherence breaks.

Historical narrative texts are written accounts of past events in which actions, events, and their consequences are interpreted and organized. Text coherence is a central factor that influences reader comprehension and is defined as the level of the narrative's organization and focus on a central theme. Historians believe that text coherence is the most important factor in judging the quality of a narrative (Carretero and Voss 1994; Voss and Wiley 2006). Causal coherence breaks occur when historical narrative texts fail to demonstrate that historical events lead to particular consequences (Britt et al. 1994; Linderholm et al. 2000; Vidal-Abarca et al. 2000, 2006; Gilabert et al. 2005; Espin et al. 2007). For example, an historical narrative text may mention the occurrence of the Second World War, but fail to suggest potential causes of the war (i.e., Hitler's lust for power, the political situation in Germany, and impoverished middle class workers; Carretero et al. 1997).

Causal coherence breaks interfere with learners' inferential processes, and subsequently with their construction of a coherent mental representation of the text (van den Broek et al. 1995, 1996, 1999, 2005; Linderholm et al. 2004; Tzeng et al. 2005). Researchers have developed principled methods to resolve coherence breaks in historical narrative texts by making connections amongst ideas in the text more explicit or adding information to facilitate learners' inferential processing (Vidal-Abarca et al. 2006).

Using the discourse analysis method of Trabasso et al. (1989), researchers have made connections amongst ideas in the text more explicit by (a) arranging text events in temporal order, (b) making implicit character goals explicit, and (c) revising inadequate explanations, multiple causalities or distant causal relationships (Linderholm et al. 2000). Both less and more skilled readers who had the benefit of the revised text recalled more information and answered comprehension questions more accurately in regards to events causally related to one another in the more difficult text. Researchers have also added information to texts with causal coherence breaks to facilitate learners' inferential processing (i.e., information intended to trigger learners' causal antecedent and superordinate goal inferences; see Vidal-Abarca et al. 2000, 2006). Readers who had the benefit of the revised text generated more correct answers to inferential questions and recalled more information, regardless of prior knowledge differences (Gilabert et al. 2005; Vidal-Abarca et al. 2000).

In summary, the first assumption of the bottom-up approach states that researchers should model cognitive processing difficulties. Researchers have demonstrated that learner comprehension of historical texts decreases when texts are made less coherent. In the following section, we argue that a metacognitive tool should be designed to foster the effective metacomprehension processes needed to help learners' resolve their inferential processing difficulties.

Second assumption underlying the bottom-up approach

Designing metacognitive tools using a bottom-up approach involves deriving principles of design that support metacognitive processes by helping learners become aware of various types of cognitive processing difficulties (i.e., metacognitive monitoring), and to then resolve them (i.e., metacognitive control) to enhance learning. In this section, we argue that a metacognitive tool should be designed to assist learners in monitoring comprehension failures caused by causal coherence breaks and generating explanatory inferences to reinstate coherence in understanding.

In order to assist students in engaging in metacognitive monitoring and control processes to resolve comprehension failures, we outline the theory that influences the design of the metacognitive tool by integrating constructs taken from research on metacognition (Nelson and Narens 1990, 1994; Nelson 1996; Hacker 1998) and reading comprehension (van den Broek et al. 1996, 1999, 2005). Each construct included in the model is operationally defined in the form of a coding scheme (see Table 1). According to this model, reading an unexplained event at the origin of a coherence gap should ideally trigger students' monitoring of a comprehension failure. However, individual differences in metacomprehension knowledge mediate whether students are able to become aware of inferential processing difficulties. The ability to monitor a comprehension failure is a first step in deploying a strategy to re-establish coherence in understanding. In the event that learners cannot connect or reinstate information from the text that was previously read, they should generate explanatory inferences for unexplained events in order to reinstate coherence in their understanding. Specifically, learners should relate information retrieved from their own prior knowledge to the event that is at the origin of the causal coherence break in an attempt to formulate a tentative historical explanation.

Our metacognitive tool has been designed to foster such metacomprehension processes through the construction of elaborative interrogation prompts. In doing so, the element of design for the metacognitive tool triggers explanations based on learners' prior knowledge (for review, see McCrudden and Schraw 2007). Elaborative interrogations are "why" questions that prompt learners to explain the occurrence of events at the origin of causal

Table 1 Coding scheme of metacomprehension monitoring and control processes

Macro-level processes	Micro-level processes	Definition	Example
Comprehension control			
Associations		Activated prior knowledge unrelated to text coherence	"there is a kid in my class whose name is Bourassa"
Paraphrase/ text repetition		Reread or paraphrased the current sentence	"this is leading to the development of the canadian armed forces"
Backward inferences		Information is retrieved from one of three sources to explain the current sentence	
	Connecting inference	An explanation where information is retrieved from the sentence immediately preceding the current sentence	"so because they were more efficient they went overseas"
	Reinstatement inference	An explanation where information is retrieved from prior sentences that preceded the current sentence and the one prior to it	"Oh so they didn't want to go in south africa"
	Explanatory inference	An explanation where information is retrieved from background knowledge not presented in the text	"because they didn't want to be involved with Great Britain"
Comprehension monitoring			
	Monitoring comprehension successes	Aware that he or she understands	"ok so the liberals are probably lauriers administration and the conservative is bourassa"
	Monitoring comprehension failures	Aware that he or she does not understand	"Why did he do that? I don't know"
Other		Responses that do not fall in these categories	"That's pretty cool"

Coding scheme adapted from Hacker (1998) and Kendeou and van den Broek (2005, 2007)

coherence breaks. For example, a causal coherence break may occur because a text fails to mention the causes of the Second World War. Therefore, the tool prompts the learner by asking the question "Why did the Second World War occur?" thereby leading them to seek potential causes. The process of answering why-questions through generating explanations or inferring potential causes has been referred to as explanation-centered learning (Graesser and Franklin 1990; Golding et al. 1990; Graesser et al. 1991, 1994, 1996, 2005; Long et al. 1992).

Elaborative interrogations enable learners to construct a more complete mental representation of the text through maintaining information in short term memory while at the same time linking it to information retrieved from long term memory (McDaniel and Donnelly 1996; Ozgungor and Guthrie 2004). The elaborative interrogation effect has been extensively studied in a wide range of contexts such as learning factual statements shown in sentences (Wood et al. 1990; Martin and Pressley 1991; Woloshyn et al. 1992, 1994), factual statements embedded within paragraphs (Seifert 1993, 1994; Willoughby et al. 1994), studying technically difficult texts (McDaniel and Donnelly 1996; Ozgungor and

Guthrie 2004; Callender and McDaniel 2007), and studying lengthy texts in computer-based learning environments (Dornisch and Sperling 2004, 2006). However, querying the occurrence of events at the origin of causal coherence breaks through elaborative interrogation has yet to be investigated. This issue is addressed as we empirically evaluate the bottom-up approach to designing metacognitive tools.

Evaluating the bottom-up approach to the design of metacognitive tools

The objective of this study was to empirically evaluate the effectiveness of the bottom-up approach used to develop metacognitive tools for supporting the comprehension of historical narratives. We used a three-pronged approach (for review, see Magliano and Graesser 1991) to evaluate the elements and principles of design for our metacognitive tool. The three prongs were (1) theoretical predictions in regards to the effect of the metacognitive tool, (2) data collected during the task (i.e., online), and (3) data collected after the task was completed (i.e., offline). In doing so, our theoretical predictions were tested through converging evidence obtained from both online verbal protocol measures (i.e., concurrent think-aloud protocol) and offline behavioral measures (i.e., free recall and inferential question assessment).

The offline behavioral measures were used to assess the effects of the metacognitive tools in terms of facilitating the construction of a coherent mental representation (memory representation of text information and understanding of underlying causal structure; see Linderholm et al. 2000). Moreover, concurrent think-aloud protocols (Ericsson and Simon 1993; Pressley and Afflerbach 1995) were used to capture the deployment of metacomprehension processes during reading. Based on the recommendations of Azevedo et al. (2010b), think-aloud protocols were converged with an unobtrusive measure (i.e., written responses to the elaborative interrogations that were stored as log file trace data) to compensate for missing data due to a lack of verbalizations during writing. A silent control condition was also used, as students read silently to act as a control for any interference (positive or negative) that the think-alouds may have caused during the treatment condition.

The research question addressed in this study was "Can the metacognitive tool enhance learners' memory and comprehension of text information by assisting them in monitoring comprehension failures and generating explanatory inferences?" To answer this question, our experiment followed a 2×2 between subjects design with the treatment (i.e., learners either had or did not have the benefit of the metacognitive tool) and measurement conditions (i.e., thinking aloud or learning silently) as the independent variables. Prior knowledge and reading comprehension ability were assessed at pre-test in order to statistically control for their mediating effects. The amount of material recalled and the accuracy of answers to inferential questions were measured at post-test in order to compare learning performance across conditions.

As such, the research hypotheses were as follows:

1. If the use of the metacognitive tool facilitates recall and understanding of the historical narrative text, then learners who have the benefit of the metacognitive tool are expected to show a greater amount of information recalled and more accurate answers to comprehension questions at post-test while statistically controlling for the mediating effects of individual differences in reading comprehension skills.
2. If the use of the metacognitive tool fosters the key metacomprehension monitoring and control processes, then the analysis of concurrent think-aloud protocols converged

with the answers to the elaborative interrogations will show that learners who have the benefit of the metacognitive tool monitor comprehension failures and generate explanatory inferences more frequently.

3. If engaging in the key metacomprehension monitoring and control processes result in facilitating recall and understanding of the historical narrative text, then we expect a positive bivariate correlation between the dependent variables (i.e., amount of information correctly recalled and the accuracy of answers) with the independent variables (i.e., the frequency with which learners engage in monitoring comprehension failures and generated explanatory inferences).

4. If the concurrent think-aloud methodology does not alter metacomprehension processes in a way that affects learning performance, then we expect that learning performance (i.e., amount of information correctly recalled and the accuracy of answers) does not differ across measurement conditions (i.e., thinking aloud and learning silently).

Methods

Participants

A convenience sample was used (i.e., students responded to an online ad posted on the website of the university) and consisted of 32 undergraduate students from a Canadian university (6 men and 26 women). The 32 students were randomly assigned to each of the four conditions (i.e., eight each to the treatment and think-aloud, treatment and silent reading, control and think-aloud, and control and silent reading groups). The average age of the participants was 23 years old (SD = 5.1). A compensation of $10 was given to each participant. Participants were proficient in both reading and writing English and were not studying in the domain of history. The participants' grade point average was 3.4 (SD = .4). A pre-test free recall measure of amount of background knowledge related to the subject of the text showed that they were unfamiliar with the topic. Some learners were able to recognize that Henri Bourassa and Wilfred Laurier were politicians but most of them were unfamiliar with the details surrounding the events described in the text.

Materials

Historical narrative text

The participants read a short excerpt (approximately 500 words) taken from an historical narrative text (see Levitt 1970). The topic of the excerpt pertained to the roles played by Henri Bourassa and Wilfred Laurier in the events surrounding the Imperialist Movement, the Boer War and the foundation of the Canadian Navy. The text was used as part of an undergraduate-level introduction to a history course offered at a Canadian university and was assumed to be at an appropriate level of difficulty for the participants.

Discourse analysis method

The discourse analysis method of Trabasso et al. (1989) was used in order to represent the historical narrative text as a causal network of events. This discourse analysis method had

been previously used with historical narrative texts (Britt et al. 1994; Linderholm et al. 2000; Vidal-Abarca et al. 2000, 2006; Gilabert et al. 2005; Espin et al. 2007). Following the discourse analysis method of Trabasso et al. (1989), the text was first segmented according to main clauses. A main clause is defined as a part of a sentence that contains a verb that meaningfully relates two concepts.

Each main clause was categorized by judges as either referring to a goal, attempt, outcome, reaction or setting. Goals are desired or undesired states, activities or objects. Attempts are actions carried out by the protagonist in order to achieve success with respect to goals. Outcomes are changes in state that index either goal attainment or failure. Reactions are internal states, state changes or internal reactions of protagonists, and include emotion or cognition. Settings introduce characters in time and space, providing background conditions that enable the states and actions of the episodes to occur. The discourse analysis yielded a total of 7 goals, 11 attempts, 21 outcomes, 4 reactions and 7 settings.

Two judges engaged in counterfactual reasoning to identify causal relationships between main clauses. Counterfactual reasoning refers to verifying whether the criterion of necessity in the circumstances applies to a pair of main clauses. The criterion states that event A caused event B if it is true that had event A not happened in the circumstances described in the text, then event B would not have happened. Also, a causal relationship can be established only if event A is temporally prior to event B and event A is active when event B occurs. Consequently, indirect and distal causes can be inferred by transitivity: if event A caused event B, and B caused C, then events A, B and C are causally connected. As a result, the text was represented as a causal network of events (see Fig. 2).

One judge derived the causal network of events while a second judge independently derived the causal network for 20% of the total amount of main clauses. Interrater agreement for categorizing and interrelating main clauses was obtained on 90% of the occasions. Differences between judges were resolved through discussion and revisions were made for the remaining main clauses when necessary.

Elements of design for the metacognitive tool

Using the causal network of the events representation, causal coherence breaks were either identified or created in the circumstances described in the historical narrative text. Causal coherence breaks are operationalized as events categorized as goals or reactions that are implied in the circumstances described in the historical narrative text. In other words, the historical narrative text fails to demonstrate that a cause (i.e., event categorized as a goal or reaction) led to an effect (i.e., event categorized as an action or outcome). Using the causal network representation of events, one causal coherence break was identified. Furthermore, seven additional coherence breaks were introduced for the purposes of evaluating the metacognitive tool by not mentioning main clauses categorized as reactions and goals in the circumstances described in the historical narrative text. No additional information was added and necessary text alterations were kept to a minimum. For example, the text failed to mention the cause, "the imperial defense planners wished the Dominions to contribute to a system of imperial defense," which led to the consequent event, "the colonies were contributing large amounts of money."

Eight elaborative interrogations, each one corresponding to a causal coherence break, were created (see Table 2). Each elaborative interrogation asked learners to explain the occurrence of the consequent event or effect at the origin of a causal coherence break. For example, one of the elaborative interrogations asked learners to respond to the question, "Why would the colonies contribute large amounts of money?" Each elaborative
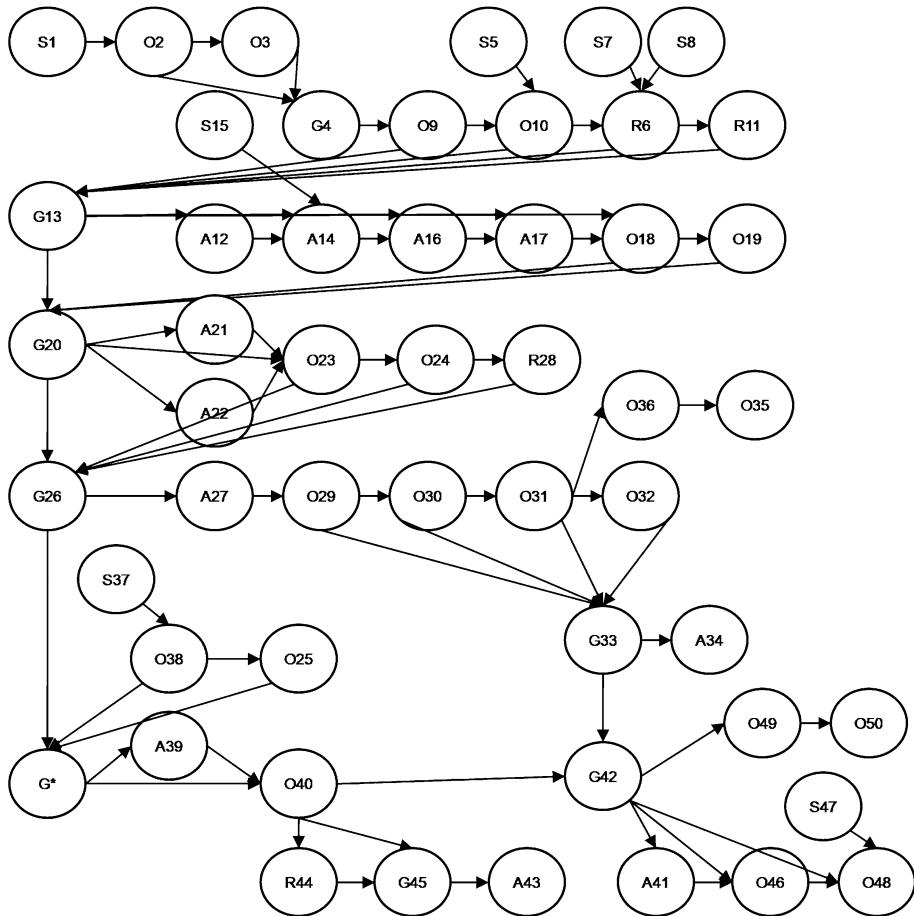
Fig. 2 Causal network representation of the text derived using the discourse analysis method of Trabasso et al. (1989). For a list of each clause, see Poitras (2010)

interrogation had a reference to the paragraph that contained the event at the origin of a causal coherence break. Participants wrote their answers in a textbox. They chose which elaborative interrogation to answer by clicking on the "previous" or "next" button (see Fig. 3).

Process measures

The process data consisted of the participants' concurrent think-aloud protocols augmented with their answers to the elaborative interrogations stored in the log files. Although the validity of concurrent think-aloud protocols in regards to capturing, analyzing, and classifying metacognitive and self-regulatory processes has been established (Bannert and Mengelkamp 2008), researchers should converge and temporally align online trace methodologies so as to account for missing verbal traces (Azevedo et al. 2010b). In this study, we converged the think-aloud protocols with the log file data according to the time at which each utterance or written answer was made. This means that the participants'

Table 2 Summary of elaborative interrogation prompts

| Number	Event	Category	Elaborative interrogation
9	Colonies were contributing large amounts of money	Outcome	Why would the colonies contribute large amounts of money?
12	Laurier rejected any proposals for defense centralization	Attempt	Why did Laurier rejected any proposals for defence centralization?
21	New training schemes for officers were begun	Attempt	Why were new training schemes for officers begun and military institutions in the country made more efficient?
22	Military institutions in the country were made more efficient	Attempt	
27	He stated publicly that soldiers could not be legally sent to South Africa	Attempt	Why did Laurier stated publicly that soldiers could not be legally sent to South Africa?
34	Resigned from his seat in the Commons	Attempt	Why did Bourassa resigned from his seat in the Commons?
39	Laurier proposed forming a small navy	Attempt	Why did Laurier proposed forming a small navy?
41	Bourassa and his supporters opposed the Naval Bill	Attempt	Why did Bourassa and his supporters opposed the Naval Bill?
43	Many Anglo-Canadians, however, objected to Laurier's proposal	Attempt	Why did many Anglo-Canadians objected to Laurier's proposal?

Number, Event and Category refers to the clauses derived from the discourse analysis method of Trabasso et al. (1989). These clauses were at the origin of causal coherence breaks in the circumstances described in a historical narrative text. Elaborative interrogation prompts were created to ask learners to explain the occurrence of these clauses

utterances overlapped according to the order of their occurrence, whether they originated from writing in the textbox or were verbalized while reading.

The result of the transcription was 30 single-spaced pages ($M = 1.88$ pages per participant), with a total of 12,266 words ($M = 766.63$ words per participant). The data were then analyzed according to the following coding scheme: (a) associations, (b) backward inferences including connecting, reinstatement and explanatory inferences, (c) paraphrases/text repetitions, (d) comprehension monitoring including either successes and failures, and (e) other (see Table 1. Interrater agreement obtained for a randomly selected sample consisting of 10% of the total amount of utterances was substantial (Cohen's Kappa $= .64$; percentage of agreement $= 72.2\%$). Disagreements were resolved through discussion and necessary revisions were made to the rest of the transcripts.

Product measures

We controlled for differences in reading comprehension ability by pre-testing participants on the paper-based Form A adult version of the Wechsler Fundamentals: Canadian (The Psychological Corporation 2007). It is an ecologically valid measure of academic skills (i.e., reading comprehension in relation to complex texts) scored according to the Canadian norms, as described in the administration manual. Learners are instructed to read four texts that differ from the task and topic used to evaluate the metacognitive tool. In order to assess their ability to comprehend texts, they answered five multiple-choice inferential questions for the first two texts and then eight questions for the last two texts (e.g., "What first clued scientists into the fact that the ground around Yellowstone was becoming deformed?").
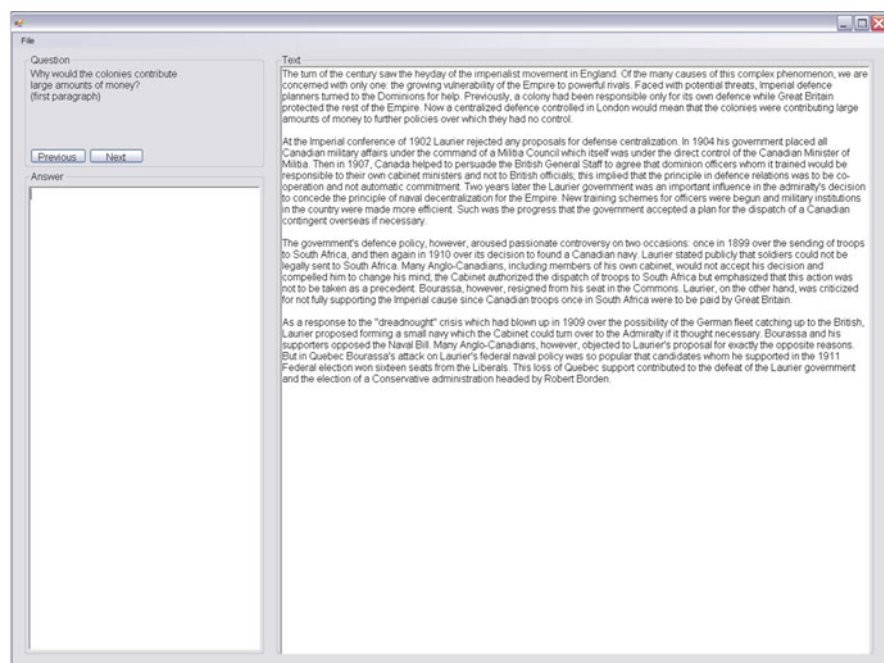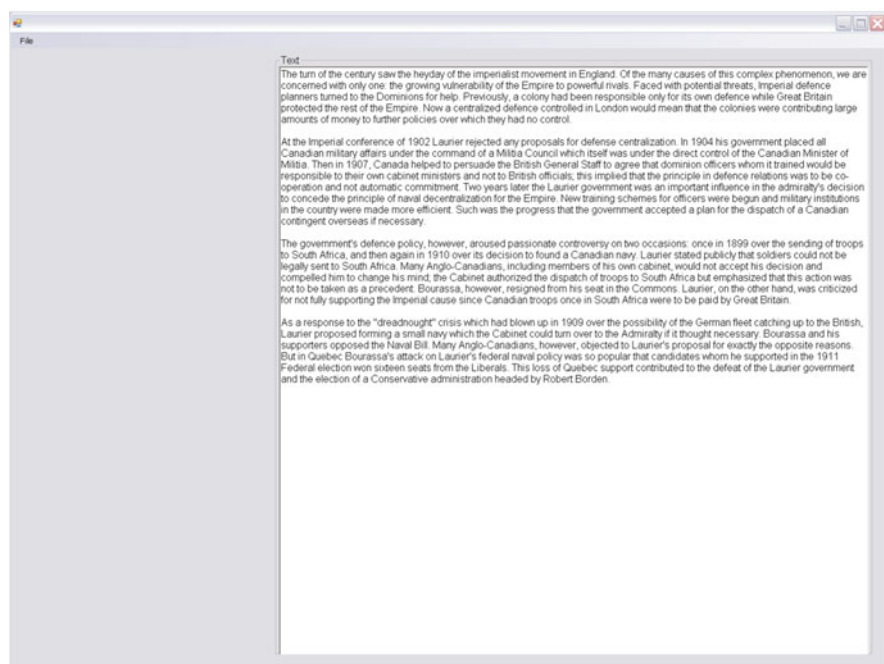
Fig. 3 The screen captures of the software used by participants undergoing the treatment and control condition

Participants' reading comprehension ability levels were determined at pre-test to statistically control for this extraneous factor in the experimental design.

In order to protect against threats to construct validity, we aligned the theoretical foundations of assessment with the procedures used to measure the students' ability to remember (i.e., post-test recall task) and understand (i.e., post-test inferential questions) the text (see Leslie and Caldwell 2009). The psychological validity of the causal network representation of the text derived from the principled discourse analysis method is well established (e.g., Trabasso et al. 1984, 1988; Trabasso and Sperry 1985; Trabasso and van den Broek 1985; van den Broek and Trabasso 1986). As such, we segmented and scored the recall and inferential question assessments using clauses consisting of a main verb as a unit of analysis (e.g., Kendeou and van den Broek 2005; Magliano et al. 1999) as well as using the causal network representation as a basis for comparison (same procedure as in Linderholm et al. 2000).

The post-test free recall task assessed learners' memory of the information mentioned in the text. Students were instructed to recall and write as accurately as possible everything they could remember from the text they had read (e.g., "Please type everything you remember from the text you just read"). Students wrote what they remembered in a textbox on the computer screen. The recall protocols were scored using the causal network representation of the text whereby a point is awarded for each event mentioned that matches an event mentioned in the text. Interrater agreement obtained for a randomly selected sample consisting of 10% of the total amount of utterances was substantial (Cohen's Kappa = .63; percentage of agreement = 84.4%).

The post-test comprehension questions assessed understanding of the causal chain of events. The inferential questions assessed students' understanding of the causal chain of events by prompting them to write the cause or consequence of a particular event that was mentioned in the text (e.g., "What caused the Imperial defense planners to turn to the Dominions for help?"). The answers were scored using the causal network representation, whereby a point is awarded for each answer that matches with either the correct cause or effect for an event mentioned in the text. Interrater agreement obtained for a randomly selected sample consisting of 10% of the total amount of utterances was fair (Cohen's Kappa = .38; percentage of agreement = 80%). For both the post-test free recall task and comprehension questions, disagreements amongst judges were resolved through discussion and necessary revisions were made to the rest of the transcripts.

Procedure

Participants first completed a consent form and demographic questionnaire. Pre-test measures of prior knowledge and reading comprehension ability were then administered in order to determine whether they altered the outcome of the study by influencing the effects of the metacognitive tool towards memory and comprehension. First, the pre-test free recall task was administered. Participants were asked to write everything they knew about Wilfred Laurier and Henri Bourassa and the events surrounding the Imperialist Movement, the Boer War and the foundation of the Canadian Navy. Second, the participants completed the reading comprehension test. They were asked to read the texts once and then answer to the questions, provided on a separate sheet, without looking back at the text.

All of the participants read the historical narrative text with causal coherence gaps. However, they were randomly assigned to one of four experimental conditions. The first condition was reading the historical narrative text while having the benefit of the metacognitive tool and thinking aloud (i.e., metacognitive tool and think-aloud). The second

condition was reading silently while using the metacognitive tool (i.e., metacognitive tool and silent). The third condition was reading the historical narrative text while not having the benefit of the metacognitive tool and thinking aloud (i.e., no metacognitive tool and think-aloud). The fourth condition was reading silently while not using the metacognitive tool (i.e., no metacognitive tool and silent).

In regards to manipulating the availability of the metacognitive tool, all of the participants were instructed to "read the text at a rate that allowed them to remember and understand the text" since they would later be tested on both their recall and understanding. The participants in the treatment condition had the benefit of the metacognitive tool displayed in the right margin of the TRE; those assigned to the control condition did not have the tool since it was not displayed (see Fig. 3). Participants in the experimental group were taught how to use the tool by writing answers in the textbox and using the buttons to change elaborative interrogations.

The participants in the concurrent think-aloud protocol conditions were trained according to the general guidelines outlined in Ericsson and Simon (1993). The training consisted of a researcher reading aloud an abstract from an historical narrative text as well as showing how to verbalize each of his thoughts without any explanation or interpretation (e.g., the experimenter said, "This reminds me of a history class I had taken years ago."). Then, the participants were instructed to practice thinking aloud by reading the following sentences of the abstract. At the end, the participants were asked if they had any questions. In doing so, the researcher did not model any of the processes that were meant to be fostered by the metacognitive tool since the example consists of generating an association (i.e., monitoring comprehension failures, generating explanatory inferences). Also, the abstract used for training is unrelated to the text that was read during the experiment. The instruction given to the participants before undergoing the experiment was, "Please remember that it is very important to say everything that you are thinking while you are reading the text without explaining or interpreting it." The experimenter reminded participants to keep verbalizing when they had been silent for more than 3 s (e.g., "Remember to say what you are thinking as you read through the text."). These reminders were given only three times. The participants assigned to the silent condition did not receive such training.

Subsequently, the post-test measures of free recall and inferential questions were administered to assess learners' memory representation as well as their understanding of the underlying causal structure of text information. First, participants were administered the post-test free recall task. They were asked to write everything they could remember from the text. Second, they completed the post-test inferential questions. They were instructed to write an answer to each open-ended question. The experiment lasted on average about 1 h and was followed by a debriefing session where the purpose of the experiment was explained in more detail.

Results

Does the metacognitive tool enhance learning performance?

A 2×2 between-subjects multivariate analysis of covariance was performed on the product data, specifically the amount of information correctly recalled and the accuracy of answers to open-ended inferential questions (see Table 3). Participants were randomly

Table 3 Does the Metacognitive tool enhance learning performance?

Variables	Recall		Comprehension	
	M	SD	*M*	SD
Experimental & think aloud	10.75	6.11	2.25	.89
Control & think aloud	7.38	1.30	1.13	.84
Experimental & learn silently	16.63	3.82	2.63	.92
Control & learn silently	10.38	7.56	2.50	1.85

The means and standard deviations are presented above for each dependent variable (i.e., recall and comprehension) as a function of the experimental condition

allocated to one of two experimental conditions (treatment or control) and one of two measurement conditions (silent reading or think-aloud). The learning outcomes were adjusted for one covariate, individual differences in reading comprehension skills.

The multivariate analysis of covariance showed that the scores obtained on the two dependent variables differed as a function of the measurement condition, $F(2,26) = 3.92$, $p < .05$. The univariate analysis of covariance showed that the scores obtained on the free recall task differed as a function of group, $F(1,27) = 5.37$, MSE $= 26.72$, $p < .05$. Participants who had the benefit of the metacognitive tool scored higher on the amount of information recalled ($M = 13.44$, SE $= 1.30$) compared to participants who did not have the tool ($M = 9.12$, SE $= 1.30$). However, the effect of the metacognitive tool towards enhancing recall was small ($\eta^2 = .17$).

A similar result was obtained on the univariate analyses of covariance in regards to the post-test free recall task and the comprehension questions measure, $F(1,27) = 7.03$, MSE $= 26.72$, $p < .05$, and $F(1,27) = 5.79$, MSE $= 1.35$, $p < .05$, respectively. Participants who performed a concurrent think aloud scored lower on amount of information recalled ($M = 8.81$, SE $= 1.30$, vs. $M = 13.75$, SE $= 1.30$) and accuracy of answers to comprehension questions ($M = 1.62$, SE $= .30$, vs. $M = 2.63$, SE $= .30$) compared to the participants who read silently. The effect size was small in regards to free recall and comprehension questions ($\eta^2 = .21$, and $\eta^2 = .18$, respectively).

Does the metacognitive tool foster the key metacomprehension processes?

Think-alouds were used to determine the type of metacognitive processes used by participants, and Mann–Whitney U tests were conducted to examine differences between conditions (see Table 4). Frequencies of each learner's use of metacomprehension processes were converted to ranks due to the non-normal distribution of the process data. This analysis allowed us to determine whether participants assigned to the treatment condition engaged in monitoring comprehension failures and generating explanatory inferences more frequently than participants in the control condition.[1]

---

[1] Since our research question required us to compare both the control and treatment conditions in terms of the frequency of each metacognitive variable, we chose a non-parametric test called the Mann–Whitney U (i.e., the non-parametric equivalent to the independent samples *t*-test). We chose this test due to the non-normal distribution of our data, which is addressed since each score is transformed as a rank. The Mann–Whitney U test compares the mean rank across the control and treatment conditions. This allows us to infer whether the frequencies of each metacognitive variable differed across the groups.

Table 4 Does the Metacognitive tool foster the key Metacomprehension processes?

Variable	Control		Treatment		U
	M rank	n	M rank	N	
Association	11.12	8	5.88	8	11.00*
Connecting inference	7.88	8	9.12	8	37.00
Reinstatement inference	4.50	8	12.50	8	64.00*
Explanatory inference	5.25	8	11.75	8	58.00*
Paraphrase/text repetition	7.62	8	9.38	8	39.00
Monitoring comp. failures	5.12	8	11.88	8	59.00*
Monitoring comp. successes	6.88	8	10.12	8	45.00
Other	7.81	8	9.19	8	37.50

The mean ranks are reported above for both treatment and control conditions. We used a Mann–Whitney U test to determine differences across conditions in terms of each process variable

* Mann–Whitney U test is significant at the .05 alpha level

The results obtained from the Mann–Whitney U tests and the examination of the mean ranks across conditions showed a significant difference in terms of monitoring comprehension failures, $U = 59.00$, $p < .05$, with the mean rank equal to 11.88 for treatment and 5.12 for control groups. In a similar manner, there was a significant difference between the treatment (M rank $= 11.75$) and control (M rank $= 5.25$) groups in terms of generating explanatory inferences, $U = 58.00$, $p < .05$. Furthermore, learners assigned to the treatment condition engaged more often in generating reinstatement inferences compared to those in the control condition (M rank $= 12.50$ vs. M rank $= 4.50$), $U = 64.00$, $p < .05$. This finding stands in contrast to learners in the control condition who engaged more often in generating associations compared to those in the treatment condition (M rank $= 11.12$ vs. M rank $= 5.88$), $U = 11.00$, $p < .05$. There were no significant differences found between conditions in terms of engaging in generating connecting inferences, paraphrasing, monitoring comprehension successes, as well as engaging in other types of processes (see Table 4).

Do the key metacomprehension processes facilitate learning performance?

Correlational analyses were performed between the ranked counts of metacomprehension strategy use and the learning outcome measures (i.e., amount of information recalled and accuracy of answers to open-ended inferential questions; see Table 5). In order to address the issue of the non-normal distribution of the process data, we used Spearman's correlation coefficient to convert the data in ranks. The correlation analyses were conducted to determine whether an increasing amount of specific strategy use was related to an increase or decrease in learning performance.

The results showed a statistically significant relationship between the ranked recall and explanatory inference, negative comprehension monitoring, and positive comprehension monitoring scores, $r_s(14) = .68$, $p < .01$, $r_s(14) = .55$, $p < .05$, and $r_s(14) = .60$, $p < .05$, respectively. As such, engaging in comprehension monitoring, by detecting both failures and successes in understanding, as well as generating explanatory inferences, was associated with correctly recalling more text information (i.e., stable memory representation). Furthermore, there was a statistically significant relationship between the ranked

Table 5 Does the Metacognitive tool facilitate learning performance?

Variables	1	2	3	4	5	6	7	8	9	10
1. Association	–									
2. Connecting inference	.05	–								
3. Reinstatement inference	−.65**	.08	–							
4. Explanatory inference	−.43	.20	.68**	–						
5. Paraphrase/text repetition	.00	.29	.19	.58*	–					
6. Monitoring comp. failures	−.40	−.03	.70**	.85**	.60*	–				
7. Monitoring comp. successes	−.11	.44	.35	.58*	.53*	.57*	–			
8. Other	−.10	.61*	.36	.37	.07	.27	.44	–		
9. Recall	−.24	.17	.29	.68**	.43	.55*	.60*	.24	–	
10. Comprehension	−.21	−.07	.46	.63**	.55*	.67**	.28	−.11	.46	–

* Spearman correlation is significant at the .05 level (2-tailed)

** Spearman correlation is significant at the .01 level (2-tailed)

comprehension and explanatory inference, paraphrase/text repetition, and negative comprehension monitoring scores, $r_s(14) = .63$, $p < .01$, $r_s(14) = .55$, $p < .05$, and $r_s(14) = .67$, $p < .01$, respectively. Therefore, detecting failures in understanding as well as generating explanatory inferences and paraphrasing or repeating text information was associated to accurately answering open-ended inferential questions (i.e., understanding the underlying causal structure of text information).

Discussion

The objective of this study was to empirically evaluate the effectiveness of the elements and principles underlying the design of the metacognitive tool that was developed following a bottom-up approach. The rationale was that students should be assisted in monitoring comprehension failures caused by causal coherence breaks as well as generating explanatory inferences in order to reinstate coherence in understanding. We expected that using the metacognitive tool would facilitate monitoring comprehension failures and generating explanatory inferences, thereby enhancing learning (i.e., memory representation and understanding of underlying causal structure of text information). The evaluation of the metacognitive tool was done while accounting for the mediating effects of learners' prior knowledge and reading comprehension ability. Moreover, a silent control condition was used to act as a control for any interference (positive or negative) that the think-alouds may have caused during the treatment condition.

By testing these hypotheses, we determine the advantages and disadvantages in regards to the logic underlying the bottom-up approach as well as the methodology used to empirically evaluate the effectiveness of the design guidelines. In addressing these issues, we suggest future directions for research in using both a top-down and bottom-up approach to designing metacognitive tools to enhance the learning of history.

Evaluating the strengths and limitations of the bottom-up approach

The impact of the metacognitive tool on learning performance

Our first hypothesis stated that learners who had the benefit of the metacognitive tool would recall text information more correctly and answer inferential questions more accurately while statistically controlling for the mediating effects of individual differences in reading comprehension abilities. This research hypothesis is partially supported, since the use of the metacognitive tool facilitated recall but not comprehension. The practical significance of the intervention, however, was limited due to the small effect sizes of the differences that were obtained.

One potential explanation for the lack of effect towards comprehension is that learners' prior knowledge may have impacted their ability to generate explanatory inferences. Since learners' in this study had low levels of prior knowledge at pre-test, they may have benefited more from reading more coherent text in combination with elaborative interrogation prompts. For instance, Ainsworth and Burcham (2007) investigated the impact of self-explanation training on learning performance while reading an incoherent text. Self-explanation training involved generating self-explanations or additional knowledge that went beyond the information mentioned in the text. Text coherence was manipulated by varying the degree of explicitness of relationships between ideas in regards to noun-pronoun references, descriptive elaborations, connectives, topic headers and macropropositions that linked the text to the topic. Learners who read the less coherent text did not learn as much as those who read the more coherent text, despite generating more self-explanations. Furthermore, the benefits of revising texts in order to make them more coherent is apparent for less and more skilled readers (Linderholm et al. 2000) as well as learners with low and high prior knowledge (Gilabert et al. 2005). Future research should investigate this further by replicating the elaborative interrogation effect in regards to learners with low and high prior knowledge, and in regards to both less and more coherent text segments.

Alternatively, fostering learners' inferential processes as a means of facilitating sophisticated levels of understanding in learners with low prior knowledge may require a different instructional model. For instance, the design features of the metacognitive tool could be augmented by providing learners with an inquiry-based learning activity, thereby situating learning in an authentic and meaningful context (for a review of the literature, see Aulls and Shore 2008). Historical inquiry is an activity in which students analyze, evaluate, and synthesize information gathered from a set of historical sources in order to write an historical account (Leinhardt and Young 1996; Voss and Wiley 2006). Furthermore, scaffolding mechanisms could be embedded as part of the TRE in order to enhance learning through historical inquiry (for an excellent debate on the question of scaffolding and inquiry-based learning, see Kirschner et al. 2006 and Hmelo-Silver et al. 2007). On the one hand, the discourse analysis method (Trabasso et al. 1989) could be used to expose learners to an unexplained historical event as a means of prompting them to ask an historical question and to generate tentative explanations in the context of learning through inquiry. On the other hand, providing learners with multiple documents as part of the inquiry task may be critical in facilitating causal inferential processes in learners with low prior knowledge.

The elements and principles of design

We accurately predicted that learners who had the benefit of the metacognitive tool would monitor comprehension failures and generate explanatory inferences more frequently than

would the control group. The metacognitive tool was designed based on a process-oriented model that enabled us to pinpoint inferential processing difficulties that were likely to occur during reading due to causal coherence breaks. The metacognitive tool consisted of elaborative interrogation prompts or why-questions that queried the occurrence of unexplained events during reading, and was effective as a means of fostering key metacomprehension processes that led to better recall.

The findings obtained in this study suggest that manipulating the design of the prompts based on certain text properties (i.e., text coherence) may mediate the effectiveness of the elaborative interrogation effect (for a review of the literature, see McCrudden and Schraw 2007). The elaborative interrogation effect has been attributed to the maintenance of information in short term memory as a means of linking it to information retrieved from the previously read text as well as the learners' background knowledge (McDaniel and Donnelly 1996; Ozgungor and Guthrie 2004). Although the examination of the content of answers to elaborative interrogation prompts suggest that learners may refer more frequently to information found in the text (see McDaniel and Donnelly 1996), learners are nevertheless able to answer more accurately inferential questions which assess their understanding of the text. The elaborative interrogation effect, however, is less robust when reading from lengthy texts in the context of computer-based learning environments (Dornisch and Sperling 2004, 2006).

The mediating effect induced by the design of the prompts itself may have important implications for the types of processes that are fostered as well as the impact towards learning gains when embedded in technology-rich learning environments. In this study, we found that the elaborative interrogation effect was limited to facilitating only recall, rather than the accuracy of answers to inferential questions, if the prompts targeted unexplained events described in the circumstances of an historical narrative text. Future research should examine this effect by assessing and manipulating the degree of text coherence while attempting to replicate the elaborative interrogation effect.

Fostering key metacomprehension processes in order to enhance learning

We expected to obtain a positive relationship between monitoring comprehension failures and generating explanatory inferences with learners' memory and comprehension. We found that engaging in both of these metacomprehension processes was predictive of the amount of information correctly recalled and the accuracy of answers to inferential questions. However, in light of the small effect that was obtained in regards to enhancing learning performance, we postulate that the quality of these metacomprehension processes, rather than their quantity, must be responsible for the learners' failure to develop sophisticated levels of understanding. Our synthesis and operationalization of theoretical constructs from the literature on both metacognition (Nelson and Narens 1990, 1994; Nelson 1996; Hacker 1998) and reading comprehension (van den Broek et al. 1996, 1999; 2005) should be augmented by considering both of these factors, not simply focusing on the frequency, but also the correctness, accuracy, and level of sophistication of learners' metacomprehension processes.

Our findings suggest that the theoretical framework that guides the design of the metacognitive tool as well as its evaluation should be revised. The metacomprehension monitoring construct could be more precise by measuring the time of onset relative to reading an unexplained event. In doing so, the definition of this construct, in addition to the positive and negative valence, should distinguish between whether a learner makes a judgment of learning (JOL) or a feeling of knowing (FOK) (for a review of the literature,

see Azevedo and Witherspoon 2009). Furthermore, examining probabilities in shifting from metacomprehension processes to control processes such as asking questions and generating explanations is critical in order to better understand how learners generate tentative explanations. Distinguishing between the types of questions asked (e.g., explanatory vs. evaluative questions; Logtenberg et al. 2011) as well as the types of explanations that were generated (e.g., intentionalist vs. structuralist; Carretero et al. 1997) represents a shift in the conceptualization of metacognition in terms of its domain-specific properties rather than its domain-general properties.

Reactivity to the online verbal protocol

We expected that the amount of information correctly recalled and the accuracy of answers would not differ across the think-aloud and silent conditions, thereby suggesting that the instruction to think-aloud did not interfere with learners' metacomprehension processes during reading. However, the participants who were instructed to think aloud recalled less information correctly and answered less accurately to inferential questions than those who were learning silently. Since the instruction to think aloud had a negative impact on learning performance it might have in fact hindered the learners' ability to engage in key metacomprehension processes. This phenomenon is referred to as reactivity (see Schraw 2010) since the online measurement used as part of the evaluation of the metacognitive tool altered the learners' performance.

The mechanisms underlying the reactivity effect as well as the factors that lead to increase or decrease its likelihood are still unknown. In their review of the validity of measurement and assessment in regards to metacognition and self-regulated learning constructs, Schraw (2010) identified two unresolved issues in empirical studies evaluating the effectiveness of technology-rich learning environments. First, self-awareness of verbalizing one's own thought processes during reading may interfere with cognitive processing. Second, little is known about how the complexity of technology-rich learning environments affects non-experts' limited cognitive resources and thereby their self-reporting of cognitive processing. We have yet to test competing hypotheses about the mechanisms underlying reactivity effects in an effort to prevent their occurrence.

On the one hand, the findings obtained as part of this study contradict the notion that the validity of concurrent think-aloud protocols has been established beyond any doubt in regards to measuring metacognitive and self-regulatory processes.[2] On the other hand, Bannert and Mengelkamp (2008) also investigated this issue by determining whether learning performance differed depending on the methods of measurement. In accordance with the model of Ericsson and Simon (1993), they showed that learning performance (i.e., post-test scores obtained on a recall, knowledge, and transfer task) did not differ across participants who were either instructed to think aloud or learn silently, which suggests that it did not affect metacognitive processes. These contradictory findings are problematic in that the think-aloud instruction used as part of this study was similar to the one used in Bannert and Mengelkamp (2008) (i.e., level 2 verbalizations according to the model of Ericsson and Simon (1993)).

---

[2] One of the reviewers of this manuscript suggests that this finding may be due to the reminders given to participants to verbalize their thought processes. These reminders were given after the participant fell silent for 3 s, which is more likely to elicit automatic as opposed to strategic processes that are consciously invoked.

Nevertheless, the issue of reactivity in regards to empirically evaluating the effectiveness of metacognitive tools using concurrent think-aloud protocols may affect the validity of the conclusions drawn in this study. Furthermore, the size of our sample limits the generalizability of our findings. We would recommend that the findings obtained as part of this study be replicated in order to better ascertain the strengths and limitations of using a bottom-up approach. In doing so, we raise awareness of this phenomenon and encourage further research to consider measurement reactivity when designing empirical studies. The presence and the magnitude of reactivity effects can be determined through the use of a control condition in which participants undergo the treatment while learning silently rather than performing the think-aloud protocol. Furthermore, corroborating evidence in regards to the deployment of metacognitive processes can be gathered in the control condition through unobtrusive measures such as time on task, video and audio time-stamp data (Azevedo et al. 2010b), and log file data (Aleven et al. 2010).

Conclusion

In summary, this study introduced the notion of a top-down and a bottom-up approach to the design of metacognitive tools. We demonstrated that using a bottom-up approach to the design of metacognitive tools can assist students in (a) noticing that particular events in an historical narrative text are unexplained, and (b) generating hypothetical causes to explain the occurrence of such events. However, students failed to achieve sophisticated levels of understanding, presumably due to their lack of prior knowledge in regards to the topic of the text. The broader implications of these findings include (1) outlining a generic framework that describes the logic or rationale for deriving principles and choosing elements of design for metacognitive tools, and (2) verifying reactivity in regards to the use of obtrusive online measurements. In the following section, we raise several recommendations for future research.

Future research directions

We along with others (Greene et al. 2010) have begun to extend computers as metacognitive tools to enhance learning about history. Learning about complex historical topics requires learners to engage in metacognitive and self-regulatory processes; however, students have difficulties doing so and consequently learn relatively little. In order to assist students in overcoming these challenges, computer-based learning environments can be designed as metacognitive tools using either a top-down or a bottom-up approach.

In regards to using a top-down approach, recent advances in the design of metacognitive tools to enhance the learning of complex science topics could be adapted to research in the domain of history education. For instance, MetaTutor (Azevedo et al. 2008, 2009a, b, 2010a) is a multi-agent intelligent system designed to detect, trace, model, and foster students' self-regulatory processes as they learn about human biological systems. Since theoretical frameworks of self-regulated learning can be used to account for learning from science as well as history (see Greene et al. 2010), MetaTutor could serve as a platform to further identify the key self-regulatory processes that are critical in regards to enhancing learning. Moreover, the scaffolding mechanisms embedded in MetaTutor could be used as a means of fostering these processes while learning about complex and challenging historical topics.

In regards to using a bottom-up approach, our future research will focus on evaluating the effectiveness of a metacognitive tool called the MetaHistoReasoning Tool. The MetaHistoReasoning Tool is a computer-based learning environment that assists students in engaging in domain-specific metacognitive processes while learning through historical inquiry. This line of research builds on and extends the current study by (1) broadening our theoretical framework in terms of the domain-specific properties of metacognition, (2) situating the learning task as an inquiry with a set of historical sources, and (3) augmenting the elements of design of the metacognitive tool in accordance with the new principles of design that were derived. The rationale that underlies the design of the metacognitive tool is to implement each recommendation that was derived in the current study in terms of evaluating the bottom-up approach and scaffolding students with low prior knowledge.

The environment thus serves as a platform to investigate the role of metacognitive processes in overcoming the limitations of low prior knowledge in learning about complex and challenging historical topics. The MetaHistoReasoning Tool assists students in terms of not only noticing unexplained events and formulating historical explanations, but also arguing in favor of these explanations in accordance with disciplinary-based practices. The environment enables them to do so through embedded scaffolding mechanisms (i.e., a training and inquiry module which features a pedagogical agent, an annotation tool, and multimedia instructional videos) and a set of historical sources (i.e., documents, paintings, and websites). Future research will investigate whether using the MetaHistoReasoning Tool results in fostering domain-specific metacognitive processes in regards to conducting historical inquiry, enhancing learning and thereby circumventing the limitations of low prior knowledge.

References

Ainsworth, S., & Burcham, S. (2007). The impact of text coherence on learning by self-explanation. *Learning and Instruction, 17*, 286–303.

Aleven, V., Roll, I., McLaren, B. M., & Koedinger, K. R. (2010). Automated, unobtrusive, action-by-action assessment of self-regulation during learning with an intelligent tutoring system. *Educational Psychologist, 45*(4), 224–233.

Aulls, M. W., & Shore, B. M. (2008). *Inquiry in education (vol. 1): The conceptual foundations for research as a curricular imperative*. New York: Lawrence Erlbaum Associates.

Azevedo, R. (2005a). Computer environments as metacognitive tools for enhancing learning. *Educational Psychologist, 40*, 193–198.

Azevedo, R. (2005b). Using hypermedia as a metacognitive tool for enhancing student learning? The role of self-regulated learning. *Educational Psychologist, 40*(4), 199–209.

Azevedo, R. (2008). The role of self-regulation in learning about science with hypermedia. In D. Robinson & G. Schraw (Eds.), *Recent innovations in educational technology that facilitate student learning* (pp. 127–156). Charlotte, NC: Information Age Publishing.

Azevedo, R., & Feyzi-Behnagh, R. (2010). *Dysregulated learning with advanced learning technologies.* Paper presented at the Fall Symposium of the Association for the Advancement of Artificial Intelligence (AAAI), Arlington, Virginia.

Azevedo, R., Johnson, A., Chauncey, A., & Burkett, C. (2010a). Self-regulated learning with MetaTutor: Advancing the science of learning with MetaCognitive tools. In M. Khine & I. Saleh (Eds.), *New science of learning: Computers, cognition, and collaboration in education*. Amsterdam: Springer.

Azevedo, R., Moos, D. C., Johnson, A. M., & Chauncey, A. D. (2010b). Measuring cognitive and meta-cognitive regulatory processes during hypermedia learning: Issues and challenges. *Educational Psychologist, 45*(4), 210–223.

Azevedo, R., & Witherspoon, A. M. (2009). Self-regulated learning with hypermedia. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 319–339). Mahwah, NJ: Routledge.

Azevedo, R. Witherspoon, A., Chauncey, A., Burkett, C., & Fike, A. (2009, November). *MetaTutor: A MetaCognitive tool for enhancing self-regulated learning*. Paper presented at the annual meeting of the American Association for Artificial Intelligence, symposium on metacognitive and cognitive educational systems, Washington, DC.

Azevedo, R., Witherspoon, A., Graesser, A., McNamara, D., Chauncey, A., Siler, E., et al. (2009). Meta-Tutor: Analyzing self-regulated learning in a tutoring system for biology. In V. Dimitrova, R. Mizoguchi, B. du Boulay, & A. Graesser (Eds.), *Building learning systems that care: From knowledge representation to affective modeling* (pp. 635–637). Amsterdam: IOS Press.

Azevedo, R., Witherspoon, A., Graesser, A., McNamara, D., Rus, V., Cai, Z., & Lintean, M. (2008). *MetaTutor: An adaptive hypermedia system for training and fostering self-regulated learning about complex science topics*. Paper presented at annual meeting of the Society for Computers in Psychology, Chicago, IL.

Bannert, M., & Mengelkamp, C. (2008). Assessment of metacognitive skills by means of instruction to think-aloud and reflect when prompted. Does the verbalisation method affect learning? *Metacognition and Learning, 3*, 39–58.

Berson, M. J. (1996). Effectiveness of computer technology in social studies: A review of the literature. *Journal of Research on Computing in Education, 28*(4), 486–499.

Berson, M. J., & Balyta, P. (2004). Technological thinking and practice in the social studies: Transcending the tumultuous adolescence of reform. *Journal of Computing in Teacher Education, 20*(4), 141–150.

Bolick, C. M. (2004). The giant is waking! *Journal of Computing in Teacher Education, 20*(4), 130–132.

Bolick, C. M., Berson, M. J., Coutts, C., & Heinecke, W. (2003). Technology applications in social studies teacher education: A survey of social studies methods faculty. *Contemporary Issues in Technology and Teacher Education, 3*(3), 300–309.

Bolick, C. M., Berson, M. J., Friedman, A. M., & Porfeli, E. J. (2007). Diffusion of technology innovation in the preservice social studies experience: Results of a national survey. *Theory and Research in Social Education, 35*(2), 174–195.

Bolick, C. M., McGlinn, M. M., & Siko, K. L. (2006). Twenty years of technology: A retrospective view of Social Education's technology themed issues. *Social Education, 69*(3), 155–161.

Boonthum, C., Levinstein, I. B., & McNamara, D. S. (2007). Evaluating self-explanations in iSTART: Word matching, latent semantic analysis, and topic models. In A. Kao & S. Poteet (Eds.), *Text mining and natural language processing*. London: Springer.

Britt, M. A., & Aglinskas, C. (2002). Improving student's ability to use source information. *Cognition & Instruction, 20*, 485–522.

Britt, M. A., Perfetti, C. A., van Dyke, J., & Gabrys, G. (2000). The sourcer's apprentice: A tool for document-supported history instruction. In P. Stearns, P. Seixas, & S. Wineburg (Eds.), *Knowing, teaching and learning history: National and international perspectives* (pp. 437–470). New York: New York University Press.

Britt, M. A., Rouet, J.-F., Georgi, M. C., & Perfetti, C. A. (1994). Learning from history texts: From causal analysis to argument models. In G. Leinhardt, I. L. Beck, & C. Stainton (Eds.), *Teaching and learning in history* (pp. 47–84). Hillsdale, NJ: Erlbaum.

Britt, M. A., Rouet, J. F., & Perfetti, C. A. (1996). Using hypertext to study and reason about historical evidence. In J. F. Rouet, J. Levonen, A. Dillon, & R. Spiro (Eds.), *Hypertext and cognition* (pp. 43–72). Mahwah, NJ: LEA.

Britt, M. A., Wiemer-Hastings, P., Larson, A. A., & Perfetti, C. A. (2004). Using intelligent feedback to improve sourcing and integration in students' essays. *International Journal of Artificial Intelligence in Education, 14*, 359–374.

Brush, T. A., & Saye, J. W. (2000). Implementation and evaluation of a student-centered learning unit: A case study. *Educational Technology Research and Development, 38*(3), 79–100.

Brush, T. A., & Saye, J. W. (2001). The use of embedded scaffolds with hypermedia-supported student-centered learning. *Journal of Educational Multimedia and Hypermedia, 10*(4), 333–356.

Brush, T. A., & Saye, J. W. (2002). A summary of research exploring hard and soft scaffolding for teachers and students using a multimedia supported learning environment. *Journal of Interactive Online Learning, 1*(2), 1–12.

Brush, T. A., & Saye, J. W. (2004). Supporting learners in technology-enhanced student-centred learning environments. *International Journal of Learning Technology, 1*(2), 191–202.

Callender, A. A., & McDaniel, M. A. (2007). The benefits of embedded question adjuncts for low and high structure builders. *Journal of Educational Psychology, 99*, 339–348.

Carretero, M., López-Manjón, A., & Jacott, L. (1997). Explaining historical events. *International Journal of Educational Research, 27*(3), 245–253.

Carretero, M., & Voss, J. F. (1994). *Cognitive and instructional processes in history and the social sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Dinsmore, D., Alexander, P., & Loughlin, S. (2008). Focusing the conceptual lens on metacognition, self-regulation, and self-regulated learning. *Educational Psychology Review, 20*(4), 391–409.

Doolittle, P., & Hicks, D. (2003). Constructivism as a theoretical foundation for the use of technology in Social Studies. *Theory and Research in Social Education, 31*(1), 72–104.

Dornisch, M. M., & Sperling, R. A. (2004). Elaborative questions in web-based text materials. *International Journal of Instructional Media, 31*(1), 59–69.

Dornisch, M. M., & Sperling, R. A. (2006). Facilitating learning from technology-enhanced text: Effects of prompted elaborative interrogation. *Journal of Educational Research, 99*(3), 156–166.

Ehman, L., & Glenn, A. (1991). Interactive technology in the social studies. In J. P. Shaver (Ed.), *Handbook of research on social studies teaching and learning* (pp. 513–522). New York, NY: Macmillan Publishing.

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

Eslinger, E., White, B., & Frederiksen, J. (2001). A modifiable multi-agent system for supporting inquiry learning. In J. D. Moore, C. L. Redfield, & W. L. Johnson (Eds.), *Artificial intelligence in education.* Amsterdam: IOS Press.

Eslinger, E., White, B., Frederiksen, J., & Brobst, J. (2008). Supporting inquiry processes with an interactive learning environment: Inquiry Island. *Journal of Science Education and Technology, 17*, 610–617.

Espin, C. A., Cevasco, J., van den Broek, P., Baker, S., & Gersten, R. (2007). History as narrative: The nature and quality of historical understanding for students with learning disabilities. *Journal of Learning Disabilities, 40*, 174–182.

Flavell, J. H. (1979). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231–236). Hillsdale, NJ: Erlbaum.

Friedman, A. M., & Hicks, D. (2006). The state of the field: Technology, social studies, and teacher education. *Contemporary Issues in Technology and Teacher Education, 6*(2), 246–258.

Gilabert, R., Martínez, G., & Vidal-Abarca, E. (2005). Some good texts are always better: Text revision to foster inferences of readers with high and low prior background knowledge. *Learning and Instruction, 15*(1), 45–68.

Golding, J. M., Graesser, A. C., & Millis, K. K. (1990). What makes a good answer to a question?: Testing a psychological model of question answering. *Discourse Processes, 13*, 305–325.

Graesser, A. C., Baggett, W., & Williams, K. (1996). Question-driven explanatory reasoning. *Applied Cognitive Psychology, 10*, 17–32.

Graesser, A. C., & Franklin, S. P. (1990). QUEST: A cognitive model of question answering. *Discourse Processes, 13*(2), 279–303.

Graesser, A. C., Lang, K. L., & Roberts, R. M. (1991). Question answering in the context of stories. *Journal of Experimental Psychology: General, 120*, 254–277.

Graesser, A. C., McNamara, D. S., & VanLehn, K. (2005). Scaffolding deep comprehension strategies through Point&Query, AutoTutor, and iSTART. *Educational Psychologist, 40*, 225–234.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*, 371–395.

Greene, J. A., Bolick, C. M., & Robertson, J. (2010). Fostering historical knowledge and thinking skills using hypermedia learning environments: The role of self-regulated learning. *Computers & Education, 54*, 230–243.

Hacker, D. J. (1998). Self-regulated comprehension during normal reading. In D. J. Hacker, J. Dunlosky, & A. C. Graeser (Eds.), *Metacognition in educational theory and practice* (pp. 165–191). Mahwah, NJ: Erlbaum.

Hicks, D., & Doolittle, P. E. (2008). Fostering analysis in historical inquiry through multimedia embedded scaffolding. *Theory and Research in Social Education, 36*(3), 206–232.

Hicks, D., & Doolittle, P. E. (2009). Multimedia-based historical inquiry strategy instruction. Do size and form really matter? In J. Lee & A. Friedman (Eds.), *Research on technology in social studies education.* Greenwich, CT: Information Age Publishing.

Hicks, D., Doolittle, P. E., & Ewing, T. (2004). The SCIM-C strategy: Expert historians, historical inquiry, and multimedia. *Social Education, 68*(3), 221–225.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist, 42*, 99–107.

Jonassen, D. H. (1996). *Computers in the classroom: Mindtools for critical thinking*. Englewood Cliffs, NJ: Prentice Hall.

Jonassen, D. H. (2003). Using cognitive tools to represent problems. *Journal of Research in Technology in Education, 35*(3), 362–381.

Jonassen, D. H., & Reeves, T. (1996). Learning with technology: Using computers as cognitive tools. In D. Jonassen (Ed.), *Handbook of research on educational communication and technology* (pp. 693–719). New York: Scholastic.

Kendeou, P., & van den Broek, P. (2005). The effects of readers' misconceptions on comprehension of scientific text. *Journal of Educational Psychology, 97*, 235–245.

Kendeou, P., & van den Broek, P. (2007). Interactions between prior knowledge and text structure during comprehension of scientific texts. *Memory & Cognition, 35*, 1567–1577.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*, 75–86.

Kommers, P., Jonassen, D. H., & Mayes, T. (Eds.). (1992). *Cognitive tools for learning*. Heidelberg, FRG: Springer.

Lajoie, S. P. (Ed.). (2000). *Computers as cognitive tools (Vol. 2): No more walls*. Mahwah, NJ: Erlbaum.

Lajoie, S. P. (2005). Cognitive tools for the mind: The promises of technology: Cognitive amplifiers or bionic prosthetics? In R. J. Sternberg & D. Preiss (Eds.), *Intelligence and technology: Impact of tools on the nature and development of human skills* (pp. 87–102). Mahwah, NJ: Erlbaum.

Lajoie, S. P. (2007). Developing computer based learning environments based on complex performance models. In B. Shuart, W. Spaulding, & J. Poland (Eds.), *Nebraska symposium on motivation. Modeling complex systems* (Vol. 52, pp. 123–144). Lincoln: University of Nebraska Press.

Lajoie, S. P. (2009). Developing professional expertise with a cognitive apprenticeship model: Examples from avionics and medicine. In K. A. Ericsson (Ed.), *Development of professional expertise: Toward measurement of expert performance and design of optimal learning environments* (pp. 61–83). Cambridge, UK: Cambridge University Press.

Lajoie, S. P., & Azevedo, R. (2006). Teaching and learning in technology-rich learning environments. In P. Alexander & P. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 803–821). Mahwah, NJ: Erlbaum.

Lajoie, S. P., & Derry, S. J. (Eds.). (1993). *Computers as cognitive tools*. Hillsdale, NJ: Erlbaum.

Lee, J. K., & Friedman, A. M. (2009). More to follow: The untapped research agenda in social studies and technology. In J. Lee & A. M. Friedman (Eds.), *Research on technology in social studies education*. Charlotte, NC: Information Age Publishing.

Lee, J. K., & Hicks, D. (2006). Editorial: Discourse on technology in social education. *Contemporary Issues in Technology and Teacher Education, 6*(4), 414–417.

Leinhardt, G., & Stainton, C. (1994). A sense of history. *Educational Psychologist, 29*(2), 79–89.

Leinhardt, G., & Young, K. M. (1996). Two texts, three readers: Distance and expertise in reading history. *Cognition and Instruction, 14*, 441–486.

Leslie, L., & Caldwell, J. A. (2009). Formal and Informal measures of reading comprehension. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 403–427). New York: Routledge.

Levitt, J. (1970). *Henri Bourassa on imperialism and biculturalism, 1900–1918*. Toronto, CA: Copp Clark Pub Co.

Linderholm, T., Everson, M., van den Broek, P., Mischinski, M., Crittenden, A., & Samuels, J. (2000). Effects of causal text revisions on more- and less-skilled readers' comprehension of easy and difficult texts. *Cognition and Instruction, 18*, 525–556.

Linderholm, T., Virtue, S., Tzeng, Y., & van den Broek, P. (2004). Fluctuations in the availability of information during reading: Capturing cognitive processes using the Landscape Model. *Discourse Processes, 37*, 165–186.

Logtenberg, A., van Boxtel, C., & van Hout-Wolters, B. (2011). Stimulating situational interest and student questioning through three types of historical introductory texts. *European Journal of Psychology of Education, 26*(2), 179–198.

Long, D. L., Golding, J. M., & Graesser, A. C. (1992). The generation of goal related inferences during narrative comprehension. *Journal of Memory and Language, 31*, 634–647.

Magliano, J. P., & Graesser, A. C. (1991). A three-pronged method for studying inference generation in literary text. *Poetics, 20*, 193–232.

Magliano, J. P., Trabasso, T., & Graesser, A. C. (1999). Strategic processing during comprehension. *Journal of Educational Psychology, 91*(4), 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.

Martorella, P. H. (Ed.). (1997). *Interactive technologies and the social studies: Emerging issues and applications*. Albany, NY: State University of New York Press.

McCrudden, M. T., & Schraw, G. (2007). Relevance and goal-focusing in text processing. *Educational Psychology Review, 19*(2), 113–139.

McDaniel, M., & Donnelly, C. (1996). Learning with analogy and elaborative interrogation. *Journal of Educational Psychology, 88*, 508–519.

McNamara, D. S., Boonthum, C., Levinstein, I. B., & Millis, K. (2007a). Evaluating self-explanations in iSTART: Comparing word-based and LSA algorithms. In T. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of LSA* (pp. 227–241). Mahwah, NJ: Erlbaum.

McNamara, D. S., O'Reilly, T., Rowe, M., Boonthum, C., & Levinstein, I. B. (2007b). iSTART: A web-based tutor that teaches self-explanation and metacognitive reading strategies. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 397–421). Mahwah, NJ: Erlbaum.

Neisser, U. (1967). *Cognitive psychology*. New York: Appleton.

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist, 51*, 102–116.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125–141). New York: Academic Press.

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: MIT Press.

O'Neill, D. K., & Weiler, M. J. (2006). Cognitive tools for understanding history: What more do we need? *Journal of educational computing research, 35*(2), 181–197.

Ozgungor, S., & Guthrie, J. T. (2004). Interactions among elaborative interrogation, knowledge, and interest in the process of constructing knowledge from text. *Journal of Educational Psychology, 96*, 437–443.

Pea, R. D. (1985). Beyond amplification: Using the computer to reorganize mental functioning. *Educational Psychologist, 20*(4), 167–182.

Perkins, D. N. (1985). The fingertip effect: How information processing technology shapes thinking. *Educational Researcher, 14*, 11–17.

Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientations in learning and achievement. *Journal of Educational Psychology, 92*, 544–555.

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16*(4), 385–407.

Poitras, E. (2010). *A metacognitive tool to support reading comprehension of historical narratives*. Unpublished manuscript, Department of Educational and Counseling Psychology, McGill University, Montreal, Canada.

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

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

Salomon, G., & Perkins, D. (2005). Do technologies make us smarter? Intellectual amplification with and through technology. In R. J. Sternberg & D. D. Preiss (Eds.), *Intelligence and technology: The impact of tools on the nature and development of human abilities*. NJ: Lawrence Erlbaum Associates, Inc.

Salomon, G., Perkins, D., & Globerson, T. (1991). Partners in cognition: Extending human intelligence with intelligent technologies. *Educational Researcher, 20*, 10–16.

Saye, J. W., & Brush, T. (1999). Student engagement with social issues in a multimedia-supported learning environment. *Theory and Research in Social Education, 27*(4), 472–504.

Saye, J. W., & Brush, T. (2002). Scaffolding critical reasoning about history and social issues in multimedia-supported learning environments. *Educational Technology Research and Development, 50*(3), 77–96.

Saye, J. W., & Brush, T. (2004). Promoting civic competence through problem-based history learning experiments. In G. E. Hamot, J. J. Patrick, & R. S. Leming (Eds.), *Civic learning in teacher education* (Vol. 3, pp. 123–145). Bloomington, IN: The Social Studies Development Center.

Saye, J. W., & Brush, T. (2005). The persistent issues in history network: Developing civic competence through technology-supported historical inquiry. *Social Education, 69*(4), 168–171.

Saye, J. W., & Brush, T. (2006). Comparing teachers' strategies for supporting student inquiry in a problem-based multimedia-enhanced history unit. *Theory and Research in Social Education, 34*(2), 183–212.

Saye, J. W., & Brush, T. (2007). Using technology-enhanced learning environments to support problem-based historical inquiry in secondary school classrooms. *Theory and Research in Social Education, 35*(2), 196–230.

Saye, J. W., & Brush, T. (2009). Using the affordances of technology to develop teacher expertise in historical inquiry. In J. Lee & A. Friedman (Eds.), *Research on technology in social studies education*. Greenwich, CT: Information Age Publishing.

Schraw, G. (2010). Measuring self-regulation in computer-based learning environments. *Educational Psychologist, 45*(4), 258–266.

Schunk, D. H. (2005). Self-regulated learning: The educational legacy of Paul R. Pintrich. *Educational Psychologist, 40*, 85–94.

Seifert, T. L. (1993). Effects of elaborative interrogation with prose passages. *Journal of Educational Psychology, 85*, 642–651.

Seifert, T. L. (1994). Enhancing memory for main ideas using elaborative interrogation. *Contemporary Educational Psychology, 19*(3), 360–366.

Seixas, P. (1993). The community of inquiry as a basis for knowledge and learning: The case of history. *American Educational Research Journal, 30*(2), 305–324.

Shimoda, T. A., White, B. Y., & Frederiksen, J. R. (1999). *Acquiring and transferring intellectual skills with modifiable software agents in a virtual inquiry support environment*. Proceedings of the 32nd annual Hawai international conference on system sciences. Los Alamitos, CA: IEEE Computer Society.

Shimoda, T. A., White, B. Y., & Frederiksen, J. R. (2002). Student goal orientation in learning inquiry skills with modifiable software advisors. *Science Education, 86*, 244–263.

Spoehr, K. T., & Spoehr, L. W. (1994). Learning to think historically. *Educational Psychologist, 29*, 71–77.

Swan, K. O., & Hofer, M. (2008). Technology and social studies. In L. Levstik & C. A. Tyson (Eds.), *Handbook of research in social studies education* (pp. 307–326). New York: Sage.

The Psychological Corporation. (2007). *Weschler fundamentals: Academic skills-Canadian*. Toronto, ON: Pearson Canada Assessment Inc.

Thieman, G. Y. (2008). Using technology as a tool for learning and developing 21st century citizenship skills: An examination of the NETS and technology use by preservice teachers with their K-12 students. *Contemporary Issues in Technology and Teacher Education, 8*(4), 342–366.

Trabasso, T., Secco, T., & van den Broek, P. W. (1984). Causal cohesion and story coherence. In H. Mandl, N. L. Stein, & T. Trabasso (Eds.), *Learning and comprehension of text* (pp. 83–111). Hillsdale, NJ: Erlbaum.

Trabasso, T., & Sperry, L. L. (1985). Causal relatedness and importance of story events. *Journal of Memory and Language, 24*, 595–611.

Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language, 24*, 612–630.

Trabasso, T., van den Broek, P., & Lui, L. (1988). A model for generating questions that assess and promote comprehension. *Questioning Exchange, 2*, 25–38.

Trabasso, T., van den Broek, P., & Suh, S. Y. (1989). Logical necessity and transitivity of causal relations in stories. *Discourse Processes, 12*, 1–25.

Tzeng, Y., van den Broek, P., Kendeou, P., & Lee, C. (2005). The computational implementation of the Landscape Model: Modeling inferential processes and memory representations of text comprehension. *Behavioral Research Methods, Instruments & Computers, 37*, 277–286.

van den Broek, P., Rapp, D. N., & Kendeou, P. (2005). Integrating memory-based and constructionist approaches in accounts of reading comprehension. *Discourse Processes, 39*, 299–316.

van den Broek, P., Risden, K., Fletcher, C. R., & Thurlow, R. (1996). A "landscape" view of reading: Fluctuating patterns of activation and the construction of a stable memory representation. In B. K. Britton & A. C. Graesser (Eds.), *Models of understanding text* (pp. 165–187). Mahwah, NJ: Erlbaum.

van den Broek, P. W., Risden, K., & Husebye-Hartmann, E. (1995). Comprehension of narrative events: Maintaining sufficient explanations. In R. F. Lorch Jr. & E. O'Brien (Eds.), *Sources of coherence in text comprehension* (pp. 353–373). Hillsdale, NJ: Lawrence Erlbaum Associates.

van den Broek, P., & Trabasso, T. (1986). Causal networks vs. goal hierarchies in summarizing text. *Discourse Processes, 9*, 1–15.

van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 71–98). Mahwah, NJ: Erlbaum.

van Drie, J., & van Boxtel, C. (2008). Historical reasoning: Towards a framework for analyzing students' reasoning about the past. *Educational Psychology Review, 20*, 87–110.

Vidal-Abarca, E., Gilabert, R., Gil, L., & Martínez, T. (2006). How to make good texts for learning: Reviewing text revision research. In A. V. Mitel (Ed.), *Focus on educational psychology*. New York: Nova Science Publishers.

Vidal-Abarca, E., Martínez, T., & Gilabert, R. (2000). Two procedures to improve instructional text: Effects on memory and learning. *Journal of Educational Psychology, 92*, 1–10.

Voss, J. F., & Wiley, J. (2006). Expertise in history. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 569–584). Cambridge: Cambridge University Press.

White, B., & Frederiksen, J. (2005). A theoretical framework and approach for fostering metacognitive development. *Educational Psychologist, 40*, 211–223.

White, B., Shimoda, T., & Frederiksen, J. (1999). Enabling students to construct theories of collaborative inquiry and reflective learning: Computer support for metacognitive development. *International Journal of Artificial Intelligence in Education, 10*(2), 151–182.

Whitworth, S. A., & Berson, M. J. (2003). Computer technology in the social studies: An examination of the effectiveness literature (1996–2001). *Contemporary Issues in Technology and Teacher Education, 2*(4), 472–509.

Willoughby, T., Wood, E., & Kahn, M. (1994). Isolating variables that impact or detract from the effectiveness of elaboration strategies. *Journal of Educational Psychology, 86*, 279–289.

Winne, P. H., & Hadwin, A. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Hillsdale, NJ: Erlbaum.

Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 531–566). San Diego, CA: Academic Press.

Wolfe, S., Brush, T., & Saye, J. (2003). Using an information problem-solving model as a metacognitive scaffold for multimedia-supported information-based problems. *Journal of Research on Technology in Education, 35*(3), 321–341.

Woloshyn, V. E., Paivio, A., & Pressley, M. (1994). Use of elaborative interrogation to help students acquire information consistent with prior knowledge and information inconsistent with prior knowledge. *Journal of Educational Psychology, 86*, 79–89.

Woloshyn, V., Pressley, M., & Schneider, W. (1992). Elaborative interrogation and prior knowledge effects on learning of facts. *Journal of Educational Psychology, 84*, 115–124.

Wood, E., Pressley, M., & Winne, P. H. (1990). Elaborative interrogation effects on children's learning of factual content. *Journal of Educational Psychology, 82*, 741–748.

Zimmerman, B. J. (2000). Attainment of self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). San Diego, CA: Academic Press.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*(1), 166–183.