

Comprehending and learning from Internet sources: A conceptual replication study of Goldman, Braasch, Wiley, Greasser and Brodowinska (2012)

Susanne Rott and Bianca Gavin

Abstract

This investigation reports on a conceptual replication of Goldman et al. (2012) which sought to determine processing patterns of better and poorer learners as they attempted to comprehend and learn from seven textual sources on the Internet in their native language English. The aim of the current investigation was to explore how reading multiple textual sources in the L2 interacts with sense making, monitoring and evaluation processes during reading and researching on the Internet. Nineteen intermediate and advanced learners of German were asked to outline the position of the German government regarding immigration. L2 readers received a selection of five Internet sources which varied in reliability and appropriateness for the research assignment. Students' reading behavior was recorded through think-aloud protocols. The findings only partially replicated Goldman et al. Better learners used significantly more varied reading strategies overall than poorer learners. Their reading behavior was marked by greater use of self-explanation and monitoring. Better learners also skipped more unfamiliar words. Nevertheless, neither group of L2 readers employed reading and research skills that lead to a coherent intertext model fulfilling the goal of the research task.

KEYWORDS: COMPREHENSION STRATEGIES; INTEGRATION OF KNOWLEDGE; READING MULTIPLE TEXTS

Affiliation

University of Illinois at Chicago, School of Literatures, Cultural Studies, and Linguistics, Chicago, IL, USA.
email: srott@uic.edu (corresponding author)

Introduction

Increasingly, students in foreign and second language education conduct research on the Internet. Students are often asked to complete short research assignments and present their findings in oral presentations or create reading materials for the entire class that are disseminated through wiki software (e.g., Arnold, Ducate, & Kost, 2009; Aydin & Yildiz, 2014; Ducate, Anderson, & Moreno, 2011; Kessler, Bikowski & Boggs, 2012; Kost, 2011; Rott & Weber, 2013). Some instructors have even integrated student' research summaries posted in wikis into the content of their courses (e.g., Arnold, Ducate, & Kost, 2009; Rott & Weber, 2013). Because websites are constantly updated they provide language learners with the most current information and perspectives on socio-cultural topics discussed in culture studies courses, such as debates concerning minimum wage and immigration. Likewise, websites generally convey information in the form of concise outlines and summaries of what the author wants to explain to their readers (e.g., newspapers, websites of the government, political parties, museums, etc.). Consequently, easy access – anytime and anywhere where there is an Internet connection – to such relatively short web-based materials has catalyzed student research in the L2 and thereby can present advantages over book-based library research.

Conducting any research – Internet or book-based – requires learners to read multiple sources written by different authors and interest groups for different audiences. Yet, relying on Internet sources can be problematic considering that 'anybody can post about anything to a website' (Goldman *et al.*, 2012: 356). Therefore, critical reading skills are essential to evaluate the reliability of the information and the interests promoted by the source. Accordingly, conducting research entails the integration of information from multiple sources, the acquisition of a knowledge base, and subsequently the selection of information deemed essential to present the research findings in a coherently organized framework.

Motivated by the assumption that college students turn to the Internet to conduct research 'regardless of whether the assignments are intended as Internet research projects or not' (Wiley *et al.*, 2009: 1060), recent studies with college students reading in their first language have sought to better understand how students engage with multiple sources that might or might not be reliable (e.g., Britt & Aglinskas, 2002; Goldman *et al.*, 2012; Lawless *et al.*, 2012; Wiley *et al.*, 2009). These studies identified reading behaviors that lead to the successful integration of reliable and accurate information taken from multiple websites.

Second language (L2) research on reading texts, whether online or in paper format, has tended to focus only on the reading and comprehension of individual texts. This research generally has not, however, examined the role of

evaluative processes in selecting and making sense of information acquired from multiple Internet sources. Therefore, this investigation is a conceptual replication of Goldman *et al.* (2012). The goal of the current study was to determine whether intermediate and advanced second language readers of German exhibit the same reading behaviors when reading in the target language as first language readers reported in the study by Goldman *et al.* (2012). More specifically, the study sought to determine whether and how second language reading skills interact with reading behaviors and comprehension processes with regard to learning from multiple Internet sources. With these objectives in mind, we begin by reviewing the existing literature on second language reading strategies in an online environment and review the original study (Goldman *et al.*, 2012) in the context of multiple source comprehension research. Next we outline the methodology used in this study, present and interpret the findings in light of the original study, and finally summarize the limitations.

Review of literature

Second language reading strategies in an online environment

Studies on second language (L2) reading in an online environment have without exception looked at the reading processes and text comprehension of only one hypertext. For the most part, studies which compared L2 readers' use of and preferences for different hypermedia enhancements and the way in which such enhancements contribute to text comprehension have been inconclusive (e.g., Ariew & Erçetin, 2004; Chun, 2001; Sakar & Erçetin, 2005). Many of the studies, in fact, found a negative relationship between the use of hypermedia enhancements and text comprehension. In general, it seems that the availability of lexical resources continuously provides the L2 reader with the option to leave the text processing task, thereby interrupting the creation of a coherent textual model. Without any further research it is questionable whether it is worthwhile for instructors to enhance authentic online text.

Another line of research concerns which reading strategies L2 students use to make meaning of a text in the hypertext reading situation. Three studies to date have used think-aloud protocols to elicit L2 readers' reading behaviors (Akyel & Erçetin, 2009; Chun, 2001; Konishi, 2003). In the studies by Akyel and Erçetin (2009) and Chun (2001) the online texts were enhanced through hypermedia. By contrast, Konishi's (2003) L2 readers first engaged in a search task and then read authentic and unenhanced websites.

Chun (2001) tracked the strategic hyperlink use of 23 readers of German who read two enhanced texts. For the first text readers had access to text-internal glosses presenting translations, an external link to a bilingual dictionary, and an audio narration of the text. For the second text they had access only to the external bilingual dictionary. In order to determine text comprehension,

learners were asked to recall the content of each text. Quantitative and qualitative results revealed that the L2 readers preferred by a significant margin the internal word translation gloss over the external dictionary. One reason was ease of access. In fact, learners reported during the think-aloud and the post-reading interview that they did not try to infer word meaning from context and that they would have preferred more internal glosses. At the same time, learners lamented that the second text contained many unknown words. Consequently, ‘they lost their train of thought when they had to look up “every other word”’ (p. 390). Comparing the reading behavior of high and low ability learners, Chun (2001) found that high ability learners looked up fewer words than low ability learners. Nevertheless, both groups comprehended the texts equally well. Chun (2001) concluded that low ability learners benefited more from the availability of the online lexical support.

Konishi (2003) and Akyel and Erçetin (2009) compared L2 readers’ strategy use when reading text online with strategies identified in studies where L2 learners read print text. Konishi (2003) used the strategies identified by Carrell (1989; 1998) as a framework for analyzing online reading strategies, whereas Akyel and Erçetin (2009) used Anderson’s (1991). Both taxonomies (Carrell, 1989; and Anderson, 1991) make a distinction between cognitive and metacognitive reading strategies. Konishi (2003) further differentiated between local and global strategies in the cognitive category. He gave the identification of word meaning, grammar, and pronunciation as examples of local strategies and the search for coherence and consistency of the text, the identification of text structure, and the use of inferences and background knowledge as global strategies. Akyel and Erçetin (2009) further listed raising questions, paraphrasing in the L1 and L2, rereading, confirming inferences, and skimming the text for general comprehension in the cognitive category. As metacognitive strategies Konishi (2003) included the use of goal setting to skim the text for specific information, monitoring comprehension, and monitoring the use of strategies. Akyel and Erçetin (2009) further mentioned acknowledging the lack of background information and recognizing the loss of concentration.

Based on the think-aloud protocols of six ESL readers, Konishi (2003) provided a description of the students’ use of reading strategies. Students engaged in either an open reading task, which allowed them to freely browse the webpages of one newspaper and choose one particular article to read. Or the second task, a closed reading task, required learners to find specific information by browsing the Internet about Aboriginal languages. Students’ comprehension of the websites read was not assessed. Only one reader in the closed task compared and integrated information from several websites. The study described that in the online reading environment students used all the strategies outlined in the taxonomy for reading texts in paper format. Additionally,

L2 readers used several strategies to navigate through webpages. As readers were searching for an appropriate website to read, they used strategies such as scrolling up and down, moving between pages, using multiple windows, and changing from one active window to another.

Akyel and Erçetin (2009) enhanced the online text with translation glosses for lexical support as well as topic information annotations to provide background information. A recall protocol was used to assess text comprehension. No inferential statistics were conducted. Instead the results were reported in the form of raw scores and percentages. Additionally, a distinction was made between high and low prior knowledge readers. Like Konishi (2003), Akyel and Erçetin (2009) concluded that while all of their L2 readers used strategies that had been identified for reading printed texts, the high knowledge readers used more of the strategies. Overall, the group of readers with high prior knowledge were 'cognitively more involved' (p. 146) than the low knowledge readers by using more cognitive and metacognitive strategies. Nevertheless, both groups recalled the content of the text equally well. It seems that while high level knowledge readers relied heavily on their background knowledge, and therefore consulted the online annotations less often, low knowledge readers were able to compensate for their limited background knowledge by making use of the annotations. Both groups of learners also used strategies specific to online reading. They monitored their navigation behavior between the individual pages and explicitly stated their decisions on how to proceed. Finally, the researchers emphasized that the most widely used metacognitive strategy was the commentary on the relevance and reliability of both the text and the annotations.

To summarize: The few studies that explored online reading behavior through think-aloud procedures exclusively provided insights into L2 readers' strategy use when reading a single text. These studies reported that beyond reading strategies used when reading texts in print format, L2 readers used navigation strategies. And yet, because of the descriptive nature of the investigations it remains unclear precisely which strategies lead to more or less successful text comprehension in an online environment. Moreover, with the exception of the closed reading task in the study by Konishi (2003), the goal of reading the online texts in the studies reviewed was detailed text comprehension. Neither in the study by Chun (2001) nor in the study by Akyel and Erçetin (2009) were L2 readers engaged in an authentic reading task, namely, to read a text for specific information or to research a topic.

Multiple source comprehension

Acknowledging that college students increasingly conduct research on the Internet, researchers have sought to identify the reading skills required to

integrate information from multiple text sources successfully. The mental representation of information from multiple individual texts is called an intertext model. The aim of this line of research is 'to develop an in-depth theory of inquiry-based, multiple source comprehension' (Lawless *et al.*, 2012: 3). Britt and Rouet (2012) have pointed out that a single text cannot provide comprehensive information about a particular topic. In fact, students need to understand the social dimension of a text in the sense that any author invariably selects a particular perspective or a certain subset of information. Therefore, quality learning goes beyond memorizing facts from a single text. Rather, quality learning from texts involves the understanding of the 'complex relationships between documents and content they present' (p. 281). In fact, research has shown that learning from multiple documents can lead to a deeper and more complete understanding of a topic (Wiley *et al.*, 2009).

Based on their study of how experts read in their fields of study to update their understanding of a specific topic or phenomenon, Lawless *et al.* (2012) conclude that '[t]o achieve deep comprehension successful readers connect ideas within a text to each other and with relevant prior knowledge; explain ideas and connections; and actively engage with the text to construct coherent representations' (p. 4). Reviewing the documents model (e.g., Perfetti, Rouet, & Britt, 1999) and multiple-documents trace theory (Rouet & Britt, 2011) Goldman *et al.* (2012) identify two additional processes to single text comprehension models. They explain that an integrated model of multiple texts captures the connections between the content of text (e.g., overlapping, complementary, or conflicting information) and the information about each text (e.g., author, reason for writing, publication venue). Additionally, metacognitive skills regulate reading behavior, such as navigation between texts, evaluating sources, and monitoring comprehension.

To date a team of researchers has studied various aspects of undergraduate students' reading behavior and intertext construction as the students were completing a research task (Goldman *et al.*, 2012; Wiley *et al.*, 2009). In these studies students were given the task of explaining, in writing, what caused the eruption of Mt St Helens. In each study students read seven websites in the form of a Google mockup interface to ensure that all participating students read the same texts. Additionally, the researchers were able to control the informational content (e.g., overlap and uniqueness) and the reliability of the information presented on the seven chosen sites. Three of the seven web-sites were considered reliable in that they contained accurate and partially overlapping information. Three other sources were deemed unreliable because they were 'hosted on .com URLs by sources with potentially questionable credibility' (Goldman *et al.*, 2012: 361). This second group of sources also contained incomplete and partially erroneous information. The seventh site discussed

other types of volcanoes and was not considered in the analyses. Learning was assessed with respect to two tasks: first, a true-false test assessing comprehension of concepts and relations; second, an essay task which required students to explain or describe the causes of the eruption of Mt St Helens. In order to assess undergraduate students' ability to determine the reliability of each of the seven websites, students received a printout of the Google search page and were asked to rank the reliability on a scale of 1–5.

The first study (Wiley *et al.*, 2009) sought to determine the effect of the writing task on developing an accurate intertext model (that is, a textual model of multiple sources) of the information gathered from the seven websites. Students were asked either to describe or explain the reasons (see below) for the eruption of Mt St Helens. Overall, students who engaged in the research task significantly gained knowledge. Likewise, the explanatory writing task (argument) led students to produce more core causes than students who were asked simply to describe (description) them. Nevertheless, students who had been asked to give reasons also mentioned more erroneous causes than students who simply described causes. This means that, in general, students did not attempt to discriminate between reliable and unreliable Internet sources. In fact, 'the level of critical analysis of information across sites was relatively impoverished' (Wiley *et al.*, 2009: 1085). Nevertheless, better learning was associated with better evaluation behaviors. Students who showed better comprehension also spent more time on reliable sources than students with lower comprehension.

In a follow-up study, Goldman *et al.* (2012) explored which reading processes and behaviors led to the greatest learning from the pretest to the posttest. Using the same reading materials and research tasks as in the Wiley *et al.* (2009) study, they asked 34 students to think aloud while reading the seven texts. Additionally, they made a distinction between better and poorer learners based on the posttest knowledge gain scores. Of the two hour research session students spent the first 50 minutes reading the texts, 40 minutes on the writing task, and about 15 minutes on the forced choice concept posttest. The pretest had been administered several weeks earlier. Think-aloud protocols were coded in idea units corresponding to clause-length events. The researchers identified the following four groups of processing events: (a) Comments about context: repetition/paraphrase, surface connection, self-explanation, irrelevant association, prediction; (b) assessment of one's own understanding through monitoring; (c) information/source evaluation and navigation to reflect movement among pages and sites; and (d) the event of intertext connection when students mentioned what they had earlier read in another text.

In general, Goldman *et al.*'s (2012) study replicated the findings from the Wiley *et al.* (2009) study. The study expanded on the earlier finding that better

learners generally produced essays that integrated concepts from multiple texts. Only one of the better learners as compared to five of the poorer learners integrated an erroneous cause in the essay. Analyzing the use of the different text processing strategies, the study revealed that better and poorer learners used the same kind of strategies. And yet, better learners used more self-explanation and monitoring strategies on reliable sites and tended to differentiate more between reliable and unreliable sites. They also reflect more on goal-directedness, namely to complete the research task, than poorer learners.

To summarize: Research with undergraduate students engaging in multiple source reading to conduct research in their native language strongly suggests that students' ability to select relevant information from reliable and unreliable sites depends on their constant awareness of the task goal (here to outline the causes for the eruption of Mt St Helens). Additionally, the selection of relevant information rests on students' continuous monitoring of what they already know and what they still need to learn from the textual input to accomplish the task successfully. Consequently, better learners are more strategic in their navigation among sources and their selection of what to read.

Statement of the problem

From reading research we know that L2 readers experience difficulties with the linguistic aspects when reading a target language text. Additionally, if readers are lacking the background knowledge for the topic they are researching, they can have difficulty with the text's content as well. Consequently, we can anticipate that L2 readers will encounter these same issues when reading multiple texts from different authors written for different audiences. Moreover, when engaging in online research, L2 readers will need additional skills to develop an intertext model of multiple sources. Studies in which college students engaged in a research task in their native language have determined that it is not a given that students will successfully evaluate the reliability of a source, select the relevant information from each textual source, and integrate the selected information in a coherent intertext model. Consequently, the educationally relevant question is whether L2 readers at the intermediate and advanced level have the necessary skills to engage in online research tasks and report their findings by posting them in a wiki or in a traditional research paper.

The aim of the current investigation, therefore, was to conduct a conceptual replication study, as called for by Chun (2012). The study sought to develop a better understanding of how research skills to establish an intertext model interact with L2 reading abilities. This study adapted the research methodologies of Goldman *et al.* (2012). A key difference, however, was that the participants in the present study read and wrote in their second language, whereas

those in Goldman *et al.*'s (2012) study used their first language. This focus on second language abilities required numerous adjustments to the original study. These modifications are described in detail in the method section below and are summarized in Table 1.

Research questions

1. Do better learners demonstrate more learning from multiple sources than poorer learners?
2. Can better learners differentiate better between reliable and unreliable sources than poorer learners?
3. What kind of reading strategies and evaluation processes do better and poorer learners engage in while reading multiple sources on the Internet?

Method

Participants

The participants in both the original and the current investigation were adult college-level learners from a Midwestern public university. The original study enrolled 21 native English speakers from an Introductory Psychology Subject Pool. In the current investigation, participants were intermediate and advanced second language learners of German who were enrolled in 2nd–4th year language and culture courses. Nineteen native speakers of English completed all parts of the current investigation. The original study divided their participants into better and poorer learners based on pre- and posttest results of the concept knowledge assessment task. In the current study we had to use an alternative method to discriminate between better and poorer learners because we were not able to administer a pre-test due to a lack of time. Therefore, we adopted the method used by Wiley *et al.* (2009) who had used students' correct recall of concepts on the writing task. The researchers had included students who mentioned four or more out of 13 concepts in their essays in the successful group. The current study operationally defined successful learners as students who comprehended the main ideas in multiple texts, selected the ideas appropriate for the task goal (outline the government's position), and presented the ideas comprehensibly in essay format in German. Since this was the first study which investigated foreign language learners' research skills (reading multiple sources, selecting and integrating information) we did not set a specific number of governmental policies on immigration we expected from better learners at the outset of the study. In fact, we counted 18 possible policies (see materials below) which students could have mentioned. But it seemed unreasonable that students would be able to recall all or be able to

outline all 18 in one essay. Therefore, we decided to set a figure after exploring the essays. Participants in the current study mentioned between five and seven governmental policies. Consequently, better learners were those who had outlined four or more governmental policies (10 participants) and poorer learners those who outlined fewer than four. One characteristic of successful learners was that they mentioned at least one idea from the two governmental sites as well as the site that outlined reliable and unreliable ideas. That is, these learners integrated information from multiple websites.

While it would certainly be interesting to see whether more successful learning was related to a certain level on the ACTFL Proficiency scale or the Common European Framework of Reference, these tests were not available. Student categorization into better and poorer learners did not strictly align with their enrollment in intermediate (second year) or advanced (third or fourth year) language and culture courses. Participation was voluntary. Students were able to receive extra credit points for participating in the research study.

Procedure

As in the original study, students participated individually in a research session that lasted approximately two hours. The experiment consisted of three phases: the reading and thinking-aloud phase, the writing phase, and the final assessment phase. A video camera and voice recorder were used to record the entire session. The voice recorder was placed in front of the student, while the camera was set up behind the student to the side to record possible hand gestures and the computer screen. The computer screen was recorded to monitor which text the student was reading.

After students were made comfortable with the set-up they were told that they would engage in a research task reading five texts on the Internet. They were informed that their goal was to learn about the German government's position and laws on immigration and that after reading the texts they would be asked to outline the government's position in an essay written in German.

Next, participants were informed that they would engage in a think-aloud procedure. The instructions for the think-aloud were adapted from Sanz *et al.* (2009 cited in Bowles, 2010) asking learners to say everything aloud while they were making sense of the texts (Appendix A). Following the original study one of the researchers modeled thinking aloud using a short paragraph written in German. Each participant received the same model (see Appendix B). Next, participants practiced thinking aloud with the second paragraph of the same text. In addition, they were informed that they could verbalize their thoughts in English or in German, or switch between the two languages. Students were also made aware that they would be prompted by the researcher with the question 'What are you thinking' if they stopped verbalizing their thoughts. It is

certainly possible that thinking aloud while reading affects second language students' reading behavior. Nevertheless, Bowles (2010) argued that thinking aloud does not significantly influence the comprehension process if learners limit their verbalizations to the content of the reading (non-metacognitive), which was the case in the current investigation.

After the practice session students engaged in the actual research task using five websites on a Google mock interface (see below). They had 50 minutes for the reading and researching task and were allowed to take notes on a piece of paper.

Immediately after the reading and researching phase students engaged in the writing assignment (see below). As in the original study students had 40 minutes to finish the assignment. They did not have access to the five texts on the Internet in order to determine what they had learned while researching the topic and to prevent copying from the source texts. Students were given the option either to use pen and paper or type their essay on the computer. They were also allowed to use the notes they had taken while reading.

After the writing phase participants were asked to rate the reliability of the individual websites and explain their rating. At the end of the session participants completed the 11 comprehension questions. This final assessment phase took about 10 minutes.

Materials

Internet reading materials

In the original study undergraduate students received seven websites about the causes of the eruption of Mt St Helens. Students in the current investigation read about the different positions on immigration in Germany. The topic was chosen because it fit the cultural studies curriculum the participants were pursuing. For this reason it was expected that students would find the reading task meaningful. In addition, the content of the texts concerned living situations, work, travel, and cultural differences, i.e., semantic fields intermediate and advanced students are generally familiar with.

Moreover, in contrast to the original study, students received links to only five texts. Reading seven texts was not feasible within a 50 minute time frame. The researchers were concerned that students would get tired if they had to read for a longer period of time. The five chosen texts were slightly shortened. This was done by deleting paragraphs rather than paragraph internal content. That way the texts remained authentic. In addition, ten words across the five texts were annotated by providing the translation in parentheses next to the German word. These ten words were deemed essential to the overall understanding of the text. The texts were chosen by two native German speakers who were experienced teachers in the program.

The goal of the research task was to identify and outline the position of the German government towards immigration. Texts that presented stances of the German government were considered reliable sources, texts that did not present the position of the government were considered unreliable. Two of the reliable sources had been posted by the German government on the Internet. The text ‘Studie zeigt: Deutschland ist ein weltoffenes Land – Integration gelingt’ was 106 words in length and ‘Das Zuwanderungsrecht’ was 127 words in length. A third article ‘EU untersucht Armutszuwanderung’ was posted by the news outlet Deutsche Welle. It consisted of 199 words and presented reliable (positions of the German government) and unreliable (non-governmental positions) information. The two websites considered unreliable were hosted by political parties. The 197 word long text ‘Für Respekt voreinander – und eine Zukunft miteinander!’ was posted by the social democratic party (SPD). The text ‘Qualifizierte Einwanderung bleibt Phantasieprodukt’ was 149 words long and was hosted by the national democratic party (NPD).

Like in the original study, a mock-up Google search result page was the gateway to the five individual websites (Figure 1). This mock-up Google site and the individual websites were stored on a local server to provide the students with a sheltered web environment. This way the researchers controlled which websites students read as it was important that all participants read the same texts. The mock Google interface preserved the naturalistic feel of an Internet search as explained in the original study. The Google interface showed the titles of the individual web pages, the first sentence of the text on the webpage, and the original URL. As in the original study two versions of the mock interface were created to present the texts in different orders. Finally, students received the following instructions ‘We did a search on Google using the phrase “Regierung und Immigration.” You are presented with the top five hits from this search.’

Instrumentation

Comprehension questions

The original study implemented a pretest-posttest design to assess learning outcome. This assessment included 20 true-false questions to tap the learning of concepts and causal relations. Likewise, for the current investigation 11 true-false questions were created assessing the learning about the German government’s perspective towards immigration. The two researchers read each text individually and coded it for governmental and non-governmental positions. Based on the coding both researchers created comprehension questions individually and then discussed the suitability and relevance of each question to arrive at the final 11 questions.

In contrast to the original study it was not possible to administer a pretest before the research session. Therefore, during the posttest, participants were

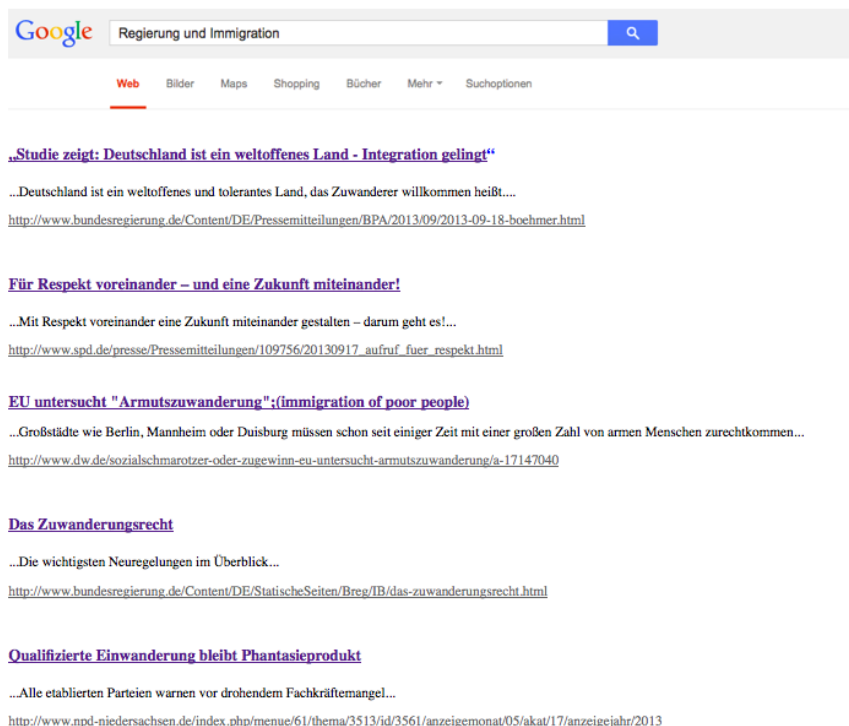


Figure 1: Google search results page that served as the interface for the research activity

asked to indicate whether they had learned the facts by reading the texts or whether they had been familiar with the information before the reading session. Finally, in order to avoid students' random guesses the study did not adopt a forced-choice format. Students had the opportunity to choose a 'Don't know' option. None of the participants made use of this option.

Essay task

Participants were asked to outline the German government's position and laws on immigration (Beschreiben Sie die Perspektive (Auffassung) und Gesetze der deutschen Regierung zum Thema Immigration). The prompt was provided in German and in English. This task required students to integrate the information they had gathered from the two reliable texts (and potentially the third text that presented governmental and non-governmental stances) and create a coherent intertext model. Students were required to write the essay in German.

Reliability ranking task

This task followed the original study in that students were given a paper print-out of the Google interface page and the individual websites. They were then asked to rank the five sites on a scale of 1–5 based on how reliable the student thought the information was. Five corresponding to most ‘reliable’. Reliable here refers to the representation of the German government’s position.

Table 1 compares the procedures and materials used in the current investigation and the original study (Goldman *et al.*, 2012).

Table 1: Participants, Procedures, and Materials in Original and the Current Investigation

	Goldman <i>et al.</i> (2012)	Current Investigation
Participants	21 total; 10 better and 11 poorer learners	19 total; 10 better and nine poorer learners
Division into two groups	Based on comparison of the pretest-posttest results.	Based on number of governmental positions outlined in writing assignment.
Native language (L1)	English	English
L2 level	N/A	Intermediate (second year) and advanced (third and fourth year) learners of German
Language of think-aloud protocol	English (L1)	English (L1) or German (L2)
Materials:	7 texts in English (L1); about the eruption of Mt St Helens 3 reliable, 3 unreliable, 1 text about a different volcano presenting unreliable information	5 texts in German (L2); about immigration in Germany; 2 reliable, 2 unreliable, 1 text presenting reliable and unreliable information
Design:		
Preparation phase	1. Knowledge pretest with 20 forced yes-no questions on the topic of the texts; 2. Instructions, model, and practice of think-aloud procedure	1. Instructions, model, and practice of think-aloud procedure
Experimental phase	3. Reading with think-aloud protocol, videotaped: 50 minutes 4. Essay writing: 40 minutes: argumentative or descriptive essay	2. Reading with think-aloud protocol, videotaped: 50 minutes 3. Essay writing: 40 minutes: descriptive essay
Assessment phase	5. Reliability ranking of the source texts/sites 6. Post-test: same as pretest; 20 forced yes-no questions on the topic of the texts (volcano eruption)	4. Reliability ranking of the source texts/sites 5. Post-test: 11 true-false items assessing comprehension and previous topical knowledge on topic

Scoring

Comprehension questions

The true-false comprehension questions were scored as correct or incorrect. Correct responses received one point, incorrect responses received zero points.

Writing task

Participants' essays were scored for the presence or absence of the German government's policies and laws about immigration. For this analysis the individual policies and ideas on the reliable as well as unreliable websites had to be determined. Two raters independently identified policies and ideas in each text. Cohen's Kappa between the two raters was 0.97. Any discrepancies were resolved through discussion. The raters identified 18 ideas representing the government's policies and 19 ideas from unreliable texts that did present perspectives of other groups, such as political parties. Ideas from the text that presented reliable and unreliable information ('EU untersucht Armutzuwanderung') were included in the tally. For each correctly identified governmental policy participants received one point. For example, one of the governmental policies identified was that immigrants have to take courses to foster integration (Integrationskurse um Integration zu unterstützen). A student who wrote that an immigrant can take a course and demonstrate knowledge of the German language and culture before becoming a citizen received one point (*Der Immigrant muss eine Kenntnisse den deutschen Sprache, Kultur, Rechte und Gesellschaft zeigen bevor einen Bürger sein konnte. Orientierungskurse ...*). Partial credit (0.5 points) was given if parts of the idea were incorrect or missing. For example when students wrote that immigrants receive 190 Euros for unemployment instead of 190 Euros for child support. Grammatically correct language use was not counted into the analysis, only comprehensibility of the governmental policies outlined was essential.

Even though the goal of the research task was to only outline the government's policies many participants mentioned several stances from other groups in their essays. As a post hoc analysis we tallied the non-governmental ideas if they were identified as non-governmental. Students received one point under the category 'Other Perspectives'.

Each learner's essay was independently coded by the two researchers. Cohen's Kappa between the two raters was 0.86. All discrepancies were resolved through discussion.

Reliability ranking task

Students' rankings for the two reliable and two unreliable sites were tallied. The site presenting reliable and unreliable information was not used in this analysis. The reliability ranking was scored as correct when learners ranked

the reliable sites as either five or four and the unreliable sites as a two or one. For each correct ranking students received one point, for an incorrect ranking they received zero points.

Processing measures derived from think-aloud protocols

The eight processing behaviors identified by Goldman *et al.* (2012) were used for the current study to determine students' text processing strategies and research behaviors. Several categories were added to the current study: Since students read the websites in their L2 they reread and repeated words and sentences in German (Repetition in German), they commented about lacking linguistic knowledge (Language monitoring), they made statements about lacking word knowledge (Skipping words), they verbalized the ability to use linguistic knowledge to make sense of a word or phrase, and they attempted to infer the meaning of a word based on its phonemic and graphemic aspects (Misinterpretation of words). Table 2 presents an overview of the 13 processing behaviors the L2 readers engaged in. The think-aloud recordings were transcribed. This included students' reading text aloud and everything they verbalized while they were making sense of the texts.

Table 2: Description of the Think-aloud Processing Categories

Category	Description
Repetition in German*	Repetition of a word or sentence of the text without adding information.
Paraphrasing in English	Restatement of the gist of a segment in the text without adding additional information.
Self-explanation	Elaboration, interpretation, or reasoning with or about a focal segment; brings new information to the focal segment, including relating it to prior knowledge or information in other information segments.
Surface connection	Vague references to previously read information but without any other type of processing (e.g., 'I just read that'; 'I heard that somewhere before').
Irrelevant association	Associations to the content of the focal segment that bear little or no obvious relevance to the task: These often relate personal experiences that do not contribute to understanding the information in the context in the verbal text.
Prediction	Statements about what the learner expects to find out next or what the next segment of text is likely to be about.
Monitoring	Statements that confirm comprehension or indicate lack of comprehension, or awareness of prior knowledge (e.g., 'I didn't know that').
Information/source evaluation	Judgments about some aspects of the sources, including relevance of the content, consistency with other information, author credentials, credibility, style, or appearance.

Navigation	Descriptions of movement within or across pages, including readers' intentions about where to go next, why they wanted to go there or what they were looking for (goals), and reasons for leaving pages that they were in the process of reading.
Language monitoring*	Statements that indicate the use or the lack of linguistic information to comprehend the segment.
Skipping words*	Statements about non-comprehension of a word without attempting to infer its meaning.
Linguistic knowledge*	Statements about the use of linguistic knowledge to infer the meaning of a word or phrase.
Misinterpretation of words*	Wrong inference of a word's meaning based on its phonemic or graphemic aspects.

Note: Adapted from Goldman *et al.* (2012); * categories were added in the current investigation and were not used in the original study.

As in the original study transcripts were 'parsed into comments, defined as speech bursts following the reading of a sentence or group of sentences' (Goldman *et al.*, 2012: 362). Each speech burst was further divided into events which captured a specific strategy, because any speech burst could contain several processing behaviors. Examples of speech bursts and their coding is provided in Table 3.

Table 3: Examples of Coding of Think-aloud Comments

Focal Segment being Read*	Think-Aloud Comments Parsed Into Events	Coding Category (Annotation)
'Wir wollen, das Vielfalt selbstverständlich ist. Kulturelle, weltanschauliche und religiöse Vielfalt muss als selbstverständlicher Teil der Gesellschaft in Politik, Wirtschaft, Kunst und Kultur mitgedacht und gefördert werden.' (SPD)	(1)'Ahm. Ha. I think they are trying to go for I guess more cultural exposure or not, I guess. (2) I mean ahm we did talk in some of my German classes about how even though there is like a lot of cultures in Germany their is still under representation to some extent, (3) but I'm not sure if that is even related to what I'm reading.'	Event 1: Self-explanation (Explains own understanding of what this sentence is about) Event 2: Irrelevant association (Draws on prior knowledge and connects it to the content of the sentence) Event 3: Monitoring (Monitors own thoughts in light of the text)
'Deutschland ist ein weltoffenes und tolerantes Land, das Zuwanderer willkommen heißt. Mit dieser Einstellung ist unsere Gesellschaft für die Herausforderung des demografischen Wandels und des Fachkräftemangels gut gerüstet.' (Studie)	(1)'Ahm. Are you saying that Germany is welcoming to working immigrants? Ahm. (2) Well it's kind of half and half. (3) The other article doesn't seem that immigrants would get a lot of jobs when they get into Germany.'	Event 1: Paraphrase in English (Translates the first part of the text) Event 2: Self-explanation (Draws on prior knowledge to comment on the information) Event 3: Surface connection (Compares this information to something just read in a different text)

<p>'Laut Studie sind 54 Prozent der Personen, die gefragt wurde, mit der Integrationspolitik der Bundesregierung zufrieden.' (Studie)</p>	<p>(1) 'So. 54 Percent of people that ask [/] that were asked about the politics of integration for those dates. (2) I don't know what <i>zufriend</i> is."</p>	<p>Event 1: Paraphrase in English Event 2: Language Monitoring (shows lack of knowledge of word meaning)</p>
<p>'Laut Studie sind 54 Prozent der Personen, die gefragt wurde, mit der Integrationspolitik der Bundesregierung zufrieden.' Die Studie zeigt deutlich: Das Umsteuern und der 2005 eingeleitete Kurs bei der Integration haben einen positiven Effekt. Der Zusammenhalt in unserer Gesellschaft ist gestärkt. Die Mehrheit versteht Zuwanderung als Chance.' (Studie)</p>	<p>(1) 'Mh. People say what or say ahm or claim what interesting. Ahm. <i>Bundesregierung zufrieden</i>. (2) I probably have to come back to that one. Ahm. (3) The <i>der</i>, people have, ahm Germanic people have ah 5 integration, say ahm, the course of integration has a positive effect. Mh. (4) Cause the thing, is that <i>zusammen</i> is together and <i>der Zusammenhalt</i> would probably be, I'm guessing maybe togetherness, maybe it would be assimilation perhaps. Ahm. (6) <i>Is gestärkt. Die Mehrheit versteht Zuwanderung als Chance</i>. Mh. (6) Maybe the next paragraph will shed a little light on that.'</p>	<p>Event 1: Paraphrase in English Event 2: Monitoring (Monitors understanding or lack of understanding and returning to this segment after reading other texts) Event 3: Language: Monitoring (gathers words understood, but does not understand the meaning of the sentence correctly) Event 4: Language: Linguistic knowledge (uses knowledge of a word to guess the meaning of a similar word) Event 5: Repetition in German (Repeats reading part of the sentence in German) Event 6: Prediction (anticipates what the next paragraph might be about)</p>
<p>'EU-Kommissar Laszlo Andor sagt: Die große Mehrheit der Rumänen und Bulgaren arbeitet und trägt stark zum Wachstum Deutschlands bei denn sie zahlt Steuern und Sozialversicherungsbeiträge und gibt in Deutschland Geld aus.' (DW)</p>	<p>(1) 'Ok. Where does he get his information from?'</p>	<p>Event 1: Information/source evaluation (Evaluates the credibility of the person being quoted)</p>
<p>"Die Städte sorgen sich vor allem um die Zukunft. Denn ab dem kommenden Jahr haben Bulgaren und Rumänen die gleichen Rechte wie alle anderen EU-Bürger: Das bedeutet, dass sie während der ersten drei Monate in Deutschland Arbeitslosengeld bekommen können. Dadurch, so die Befürchtung, könnte die Armutzuwanderung dramatisch zunehmen.' (DW)</p>	<p>(1) 'Ah let's see. Ahm. Cities trying to make room for all of them, I guess. Ahm. So then during the first three month in Germany, ahm. (2) So I guess they are trying to make room for them, ah, in the future. They got a lot of ah people coming from, ah, Romania and Bulgaria. (3) Let's try another site.'</p>	<p>Event 1: Paraphrase in English Event 2: Monitoring (Shows understanding of the text) Event 3: Navigation (Discontinues with this site/text)</p>

*The segments in column 1 are taken from the following texts (in order): SPD-Für *Respect voreinander und eine Zukunft miteinander*, *Bundesregierung* (segment 2-4) *Studie zeigt: Deutschland ist ein weltoffenes Land-Intgration gelingt*, DW (segment 5-6) *EU untersucht 'Armutzuwanderung'*.

Two raters independently coded seven of the transcripts. Cohen's Kappa was 0.90. Any discrepancies were discussed and resolved. Each of the raters coded an additional six protocols independently. Each strategy use was scored with one point and tallied in a spreadsheet.

Results

Research question one sought to determine whether better learners of German were able to demonstrate more knowledge about the German government's position on immigration from reading multiple input passages. Table 4 presents means, standard deviations and results of statistical tests for differences of learning outcome. To address the first research question multiple *t*-tests were conducted. Both groups of learners had very little content knowledge before they engaged in the reading of the websites ($M = 0.68$; $SD = 0.88$). Regarding learning outcomes a paired-samples *t*-test indicated that both groups of learners (better and poorer) knew significantly more about the government policies on immigration after reading the five texts ($M = 5.32$; $SD = 1.63$). The effect size was medium. Nevertheless, an independent samples *t*-test showed that better learners did not gain significantly more knowledge than poorer learners as indicated on the content knowledge post-test.

Students' performance on the essay task functioned as a second measure to determine learning outcome from the readings. The essay task asked students to outline the government's policies and laws on immigration. The independent samples *t*-test showed that better learners outlined significantly more governmental policies than poorer learners. The effect size of the analysis was large. Even though it was not part of the writing assignment students outlined ideas on immigration from other sources they read. There was no significant difference in the number of other perspectives better and poorer learners mentioned in their essays. Yet, when reliable (government policies) and unreliable (other perspectives) were combined better learners outlined significantly more ideas than poorer learners. These findings are partially consistent with the Goldman *et al.* (2012) study. Like in the current investigation, better students in their study demonstrated significantly more learning from multiple texts. However, their students demonstrated higher content knowledge gain on the post-test but not on the essay assignment.

Table 4: Means, Standard Deviations, Statistical Tests in Differences of Learning Outcomes in Better and Poorer Learners

Measure	Better Learners <i>M</i> (SD)	Poorer Learners <i>M</i> (SD)	Statistical test
Content knowledge gain	(pretest-posttest)		$t(18) = -13.47$ $p = 0.00^{***}$ $d = 0.36$
Content knowledge before reading	0.60 (0.84)	0.78 (0.97)	$t(17) = 0.43$ $p = 0.67$
Content knowledge posttest	5.00 (1.49)	5.66 (1.80)	$t(17) = 0.88$ $p = 0.39$

Essays: government policies	5.05 (1.21)	2.39 (1.05)	$t(17) = -5.08$ $p = 0.00^{***}$ $d = 1.55$
Essays: other perspectives	3.70 (1.99)	3.17 (1.82)	$t(17) = -0.61$ $p = 0.55$
Essay Total	8.75 (2.47)	5.55 (2.23)	$t(17) = -2.94$ $p = 0.01^{**}$ $d = 1.36$

Note. Better Learners $n = 10$; Poorer Learners $n = 9$; content knowledge maximum score = 11; essay government maximum score = 18; essay other perspective maximum score = 19; d = effect size Cohen's d ; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Research question 2 sought to develop further insights into whether better and poorer L2 readers differentiated between more and less reliable text sources. Means, standard deviations, and statistical tests are reported in Table 5. Again, multiple independent sample t -tests were used for the analyses. Goldman *et al.* (2012) found that better learners were more successful in identifying which texts were more reliable than poorer students. The current study did not replicate this finding. Based on the rating task better learners did not categorize reliable and unreliable texts any better than poorer learners. In fact, both groups of learners were quite successful in their ratings attaining $M = 3.80$ (better) and $M = 3.44$ (poorer) out of a maximum score of 4.

Yet, better learners ($M = 20.68$; $SD = 4.59$) spent significantly more time on reliable texts as compared to poorer learners ($M = 16.22$; $SD = 4.06$), as had been the case for learners in the Goldman *et al.* (2012) study. No differences were found on the overall reading time or the time spent on unreliable texts.

Table 5: Means, Standard Deviations, Statistical Tests in Differences on Reliability Rankings and Reading Times in Better and Poorer Learners

Measure	Better Learners M (SD)	Poorer Learners M (SD)	Statistical test
<i>Difference in Ranking</i>			
Reliable – unreliable	3.80 (1.40)	3.44 (1.51)	$t(17) = -0.53$ $p = 0.60$
<i>Reading Times</i>			
Time: total (minutes)	36.07 (7.50)	30.06 (11.03)	$t(17) = -0.57$ $p = 0.58$
Time: reliable texts	20.68 (4.59)	16.22 (4.06)	$t(17) = -2.23$ $p = 0.04^*$ $d = 1.03$
Time: unreliable texts	15.39 (3.49)	14.38 (5.19)	$t(17) = -1.50$ $p = 0.15$

Note. Better Learners $n = 10$; Poorer Learners $n = 9$; maximum score for reliable-unreliable ranking = 4; time was recorded in minutes; d = effect size Cohen's d ; * $p < 0.05$.

Research question 3 investigated the use of text processing and research strategies of better and poorer learners on reliable and unreliable websites. In order to determine whether more and less experienced learners used different text processing strategies data were submitted to a 2 (groups: better and poorer learners) \times 2 (processing of reliable and unreliable texts) mixed ANOVA. There was a significant effect for learner group, $F(1, 17) = 195.17$, $p = 0.00$, $\eta^2 = 0.92$, indicating that better and poorer learners approached the online reading and researching task differently. Better learners ($M = 138.75$; $SD = 33.82$) used overall more processing strategies than poorer learners ($M = 88.06$; $SD = 36.77$). However, there was no significant effect for approaching reliable and unreliable texts, $F(1, 17) = 2.39$, $p = 0.14$. That is, L2 readers in the current investigation behaved differently than college students who read in their native language. Goldman *et al.* (2012) had found that better students used significantly more processing strategies on reliable sites than poorer learners. On unreliable sites college students' processing strategies did not differ. Consequently, in the current investigation processing strategies for reliable and unreliable texts were combined for the follow-up analyses. Means, standard deviations, and statistical measures for strategies used by both groups (better and poorer learners) are reported in Table 6.

Before further exploring differences in processing strategies of better and poorer learners all data were submitted to a test of normality to determine whether the scores were normally distributed. The K-S test of normality determined that several processing categories significantly diverted from normal. Scores for making Surface Connections within a text, $D(18) = 0.29$, $p = 0.01$, making Irrelevant Associations, $D(18) = 0.26$, $p = 0.04$, Prediction, $D(18) = 0.41$, $p = 0.00$, and Navigation between different texts, $D(18) = 0.34$, $p = 0.00$, were significantly non-normal. Since in all of these categories only few instances were recorded they were combined in one score denoted as Other in Table 6. Also as combined scores, the scores were significantly non-normal, $D(18) = 0.29$, $p = 0.05$. Since all remaining categories did not deviate significantly from normal, independent sample *t*-tests were conducted to compare the use of processing strategies between better and poorer learners of German. Results partially replicated findings from Goldman *et al.* (2012) whose better college students outperformed poorer learners on the use of self-explanation, monitoring, evaluation and navigation strategies. Like in Goldman *et al.* better learners in the current investigation used the self-explanation and monitoring strategies significantly more than poorer learners of German. Additionally, better learners also skipped more unfamiliar words than poorer learners.

Table 6: Means, Standard Deviations and Statistical Measures of the Distribution of Comments Coded from Think-Aloud Statements for Upper and Lower Level Learners

Processing Category	Better Learners M (SD)	Poorer Learners M (SD)	Statistical test
<i>Processing Strategies</i>			
Reliable texts	62.20 (23.04)	44.11 (15.64)	
Unreliable texts	76.75 (19.46)	43.83 (22.07)	
Repetition in German	10.80 (11.89)	4.77 (4.35)	$t(17) = -1.43$ $p = 0.17$
Paraphrasing in English	35.15 (11.72)	27.61 (12.60)	$t(17) = -1.35$ $p = 0.19$
Self-Explanation	12.35 (11.41)	10.33 (7.04)	$t(17) = -2.49$ $p = 0.02^*$ $d = 1.10$
Monitoring	36.10 (13.40)	21.89 (13.38)	$t(17) = -2.31$ $p = 0.03^*$ $d = 1.06$
Information Source Evaluation	2.60 (1.37)	1.00 (1.23)	$t(17) = -1.19$ $p = 0.24$
Other ^a	7.10 (9.43)	4.67 (4.69)	$t(17) = -0.70$ $p = 0.49$
Language Monitoring	15.40 (11.41)	7.33 (5.41)	$t(17) = -1.93$ $p = 0.07$
Skipping Words	8.60 (7.32)	2.44 (3.32)	$t(17) = -2.31$ $p = 0.03^*$ $d = 1.08$
Linguistic Knowledge	5.10 (1.28)	3.56 (4.42)	$t(17) = -0.87$ $p = 0.40$
Graphemic Misinterpretation	2.50 (1.18)	1.56 (2.55)	$t(17) = -1.01$ $p = 0.33$

Note. Better Learners $n = 10$; Poorer Learners $n = 9$; each strategy use was awarded one point; d = effect size Cohen's d ; $*p < 0.05$; ^aOther includes Surface Connection, Irrelevant Association, Prediction, and Navigation, each of which occurred infrequently. None of the four categories were normally distributed and were analyzed with non-parametric statistics.

Qualitative analysis of individual cases

Like Goldman *et al.* (2012) the current investigation sought to better understand the dynamics of the reading and researching process. Therefore, qualitative analyses of individual cases were conducted using two better and two poorer learners.

Better learner 2

Participant 2 established a consistent pattern of reading sentence-by-sentence aloud and then commenting on each. The reader monitored comprehension

of each sentence by either paraphrasing the content in her own words in English as in ‘We are not gonna send any kids home’, or verbalizing the lack of understanding as in ‘I am not sure what the last part of that sentence is talking about’. Whenever she encountered a word she was not familiar with she used the placeholder ‘something’ trying to make meaning of the rest of the sentence as in ‘Something about giving people a good life in the country?’ At the same time she paraphrased the sentence as a question indicating that she was not completely sure about her interpretation.

This reader also frequently summarized the ideas of multiple paragraphs as well as at the end of each of the texts. Her summaries, however, were general and did not include any specific details that she had been able to understand or planned to use for the essay. After several paragraphs she verbalized ‘So this one is all about being good to immigrants’. She summarized an entire text saying ‘So this article is about like, some kind of thing that they tried where they integrated a couple of people’.

Overall, this better reader focused her attention on what she was able to comprehend from each sentence. While she was able to concretely state the content of each sentence her comprehension at the paragraph and text level were rather vague. It might very well be that she used her processing resources for sentence level comprehension and did not have any resources left to integrate the information into a more coherent mental model of each individual text. She also did not use many self-explanation strategies to elaborate on the content.

It seems that participant 2 was able to compensate for the lack of creating a coherent intertext model while reading by taking notes. Throughout the entire reading phase she took notes which she referred to constantly as she completed the essay task.

Better learner 20

Participant 20 showed the same pattern of reading sentence-by-sentence. As he paraphrased the content of the sentence in English he monitored comprehension. Whenever this reader struggled with understanding a sentence he reread the sentence in German. The following is the reaction after reading a sentence for the first time: ‘Ok. So. All right. This is a little [///] I’m tripping myself up here.’ After the second reading he was able to verbalize his satisfaction with comprehending the sentence as well as the content of the text ‘Ok. All right. So. They [/] they don’t want people to have to decide whether they want one identity or the other identity. They won’t be happy with like being *deutsch und* something else. Cool.’

This reader repeatedly applied his background knowledge to the reading task processing the content at a deeper level ‘Ahm, that’s interesting, because

I've heard a lot that says the exact opposite of that.' Likewise, the reader continuously employed the self-explanation strategy making evaluative statements about the content just read 'Ok. So they're gonna test for their *Sprachkenntnisse*. Ok. So knowing the language better helps you to integrate into the country. Ok. That's a given obviously.'

On multiple occasions participant 20 used the source evaluation strategy. He identified the source of the website he was currently reading expressing his curiosity about other groups' perspectives on immigration (political parties in this case) 'Ok. So this is interesting because I found out what [///] that this is the SPD's position on this so maybe I would wanna see what like the CDU has to say about this, because this is like their platform. Ok. Cool. Ah. Now I'm going to the next one.'

Finally, this reader showed his awareness of the recurrence of ideas within a text verbalizing 'They repeated that from the beginning'. He was able to make these surface connections within a single text but did not verbalize this strategy to indicate connections across texts. This reader also took notes while reading which he used for constructing the essay.

Poorer learner 3

For participant 3 reading was a slow and laborious process. He approached reading the individual websites by reading entire paragraphs or parts of paragraphs aloud before paraphrasing them. As he attempted to paraphrase multiple sentences he mainly focused on words he did not know. As he inserted the German words in his English paraphrase he tried to guess the meaning of words 'I'm thinking armen Menschen is ahm maybe military people'. Many of these guesses were based on graphemic or phonemic misinterpretations. In this example he incorrectly interprets the German word 'arme' which means poor as army. His focus on combining the translations of individual words rendered many of the paraphrases nonsensical 'Ah. *In Scharen nach Deutschland*. Ahm. [...] They are coming they are sharing Germanic land and trying to making things political and media like they are in other land. I think.'

This poorer reader did not seem to differentiate between the different sources and did not perceive the reading as a research task where he had to select information for the purpose of writing an essay on the policies of the German government on immigration. For two of the websites he skipped paraphrasing the titles. When he reached the end of the fifth website he seemed to summarize all five texts stating 'So most [/] most of these articles seem positive towards immigration, what they can do to help immigrants' and 'Let's see. [...] So they are very positive ah about the future with immigrants. [...] try to educate the immigrants. Make them a part of their community in positive ways.'

This reader also did not read all websites completely. In fact, he jumped between sites as he still had time in the end. At that point he reread some of the paragraphs he had read earlier or started reading the last paragraph he had not read during the first visit of the website. It seems that this poorer learner was searching for a paragraph he could comprehend better. Participant 3 did not take notes.

Poorer learner 13

Participant 13's main approach to reading the websites was by interspersing the reading of the German text with English translations 'Laut Studie sind let's see 54 Prozent of the people [der Personen] would ask this question [die gefragt wurde,] integrations politics [mit der Integrationspolitik] let's see [der Bundesregierung zufrieden. Die Studie zeigt deutlich das Umsteuern und der 2005 Jahre alt eingeleitete Kurs bei der Integration haben]. Ok wonderful. And across America and Deutschland. Let's see. [einen positiven Effekt.]' These German-English renderings of the text were furthermore interspersed with self-explanations and evaluations, such as 'Ok wonderful', and 'Oh great'. While these verbalizations indicate understanding, it is unclear whether the reader comprehended any details of the texts. In fact, he tended to summarize one entire website in one sentence 'Well, let's see. So this one basically says the government is doing great in integrating people in.' Unlike other readers this reader did not once question his understanding of a text or indicate that he was using the self-monitoring strategy.

Additionally, this reader interspersed his summaries with irrelevant comments that did not contribute to the comprehension of the website. He referred to his background knowledge about the United States stating 'Sounds like what we do in the south. Or actually we do that all around for tax breaks. Sadly.' In another instance he focused on a name 'Lets see. EU commissioner Lasslo less-low. Nice name.' or judged the author of a text 'I'm pretty sure he is probably exaggerating the situation'. Likewise, he engaged in a meta-discourse about his verbalizations. He mentioned repeatedly that he took shortcuts to verbalizing his thought process 'integrating. Ok. Integrations is pretty much the same in terms of spelling. Let's have girls, women, I'm just gonna go with Mädchen, which is the plural because I'm lazy. 18 and older. To take integration classes. Oh, become a citizen.' This reader was seemingly more interested in commenting about the content of a text than gaining a detailed understanding of the ideas in the texts.

Nevertheless, he was the only reader who mentioned that he was aware of the writing assignment 'Ok yeah I supposed I could use it as a loose opener maybe. Hm.' He took some notes which he used during the writing task.

Discussion

The current investigation was a conceptual replication study of Goldman *et al.* (2012) who explored the research skills of better and poorer learners. Our aim was to determine whether Goldman *et al.*'s findings, which were based on college students researching Internet sources in their first language, are generalizable for the L2 reading and researching context. Goldman *et al.*'s study was motivated by the need to better understand whether students have the literacy skills to effectively conduct research on the Internet. Since not one text or online resource can provide all the necessary information about a topic, research skills need to include the locating, synthesizing, evaluating and integrating of sources (e.g., Goldman, 2004). The Internet has also catalyzed research among L2 learners. Oftentimes Internet based research assignments are an integral part in culture focused classes (e.g., Arnold, Ducate, & Kost, 2009; Rott & Weber, 2013). Therefore, the question whether L2 learners have effective research skills and the ability to integrate information from multiple sources is vital.

The current replication study adapted the research methodology of Goldman *et al.* (2012). Think-aloud methodology was used as a vehicle for understanding the text processing behavior of better and poorer L2 learners when conducting a research task on the Internet. Current findings only partially replicated findings from the original study. The original study identified that better and poorer learners used overall the same set of strategies during the reading and researching phase. Yet, these two groups of learners differed in *when* they used specific strategies. The description of the case studies highlighted the dynamic interaction between the readers and the text as they were developing a coherent intertext model to explain the eruption of Mt St Helens. Better learners used strategies in combination and used them strategically in a certain order and in specific contexts. In particular, better learners employed self-explanation and monitoring strategies to navigate through reliable and unreliable websites, spending more time on reliable sites. Moreover, better learners' explanation-driven processing in combination with an awareness and evaluation of knowledge gain was essential to making decisions about navigation and eventually the selection of information.

The analysis of the relative frequencies of reading behaviors by L2 readers in the current study replicated the original findings. Better and poorer L2 learners generally did not differ in their use of reading strategies when approaching texts that contained information to complete the research task (reliable) and texts that did not contain such information (unreliable). In fact, better and poorer learners were equally successful in identifying the texts which contained information for the research task. Likewise, they demonstrated an equal amount of learning about the German government's policy

on immigration, which was assessed on the content knowledge post-test. Yet, the qualitative analysis in conjunction with the quantitative tally of strategy use showed that even though poorer learners were able to identify the more reliable text they were not able to effectively apply strategies. Better learners used more strategies overall. In particular, self-explanation, and monitoring strategies were associated with better research; elaborating on the focal content of a sentence or paragraph, retrieving associated knowledge from long-term memory, or raising questions about the content enhanced learning from the multiple textual sources. Likewise, strategic monitoring by questioning comprehension seemed to result in a more flexible textual model that learners were able to modify as they read along. These results replicated Goldman *et al.*'s (2012) findings.

In addition, better L2 learners were better able to handle unfamiliar words. Better learners focused on what they comprehended and less on what they did not understand. That way better learners maintained a more fluent reading process than poorer learners who got stuck by focusing on unfamiliar words and engaged in a word-for-word reading behavior. Graphemic misinterpretation of words and the retrieval of irrelevant background knowledge resulted in construction of incoherent textual models. Consequently, poorer learners lacked comprehension of details verbalizing a mostly vague and general understanding of ideas.

Additionally, better learners spent significantly more time on reliable texts than poorer learners. As a result they were able to outline significantly more policies of the government in their essays than poorer learners. Nevertheless, participants in both groups outlined only a small fraction of the possible perspectives. While the reliable texts mentioned 18 governmental policies, better readers only explained between four and seven on average. Poorer learners mentioned only three positions on average.

Similarly, like in the study by Wiley *et al.* (2009), both groups of learners mentioned a roughly equal number of unreliable (here non-governmental) positions. That is, the L2 readers did not make a distinction between reliable and unreliable sources, but rather wrote about ideas they had comprehended and considered noteworthy. In fact, the case studies confirmed that, in particular, poorer L2 readers dealt with each text as a separate entity. They did not verbalize the goal of the assignment, which was to outline the government's policies on immigration, and they did not select any specific information from the texts to accomplish the task. Instead, it seemed that these poorer L2 readers moved through the texts by making meaning of individual ideas without establishing a coherent single-text model. Even though the combined length of the texts amounted to only 778 words, it might be that the L2 readers were overwhelmed with the task of reading five authentic texts

in one sitting. Likewise, it is possible that students' processing resources were depleted during word and sentence level processing of the texts – not leaving sufficient resources for the integration of perspectives from a single text, let alone the integration of perspectives from multiple textual sources.

Conclusion

Overall, this study has shown that better research assignments are produced by L2 learners who use more strategies and spend more time with texts containing relevant information. In any research task the goal is to learn more information. Naturally students' background knowledge about the topic will be low. In the current study, better learners were able to overcome this limitation by elaborating on the content through interpretative questions in combination with monitoring comprehension. Better learners' approach to the research task was further marked by a focus on familiar vocabulary and skipping unfamiliar words; that way unfamiliar words did not interfere with fluent reading.

Nevertheless, the study indicates that L2 learners enrolled in 2nd through 4th year language and culture courses might not have the L2 literacy skills required to independently conduct research on the Internet. In fact, language students might require scaffolding of reading and research task well into advanced learning stages. Likewise, students might benefit from practicing sub-skills to effectively read online, such as identifying the reliability of sources and the use of online dictionary extensions that allow the look-up of words via a simple mouse click (e.g., Google Translate extension for Chrome). Yet, as the review of literature on lexical text enhancements has shown, when L2 learners overuse these resources text comprehension can be negatively affected. Therefore, students might benefit from instruction on how to effectively take advantage of dictionary extensions. Similarly, better learners continuously make decisions about how new information affects their understanding of the topic and influences their navigation decisions within and between texts (Goldman *et al.*, 2012). Consequently, L2 readers will advance their research skills by practicing goal directed note-taking in conjunction with rereading texts or individual passages. Therefore, future research needs to investigate the effectiveness of practicing sub-skills for multiple source comprehension and conducting online research.

Limitations

The findings of the current investigation are limited in that we had only 19 participants which were divided into two groups based on their completion of the research assignment written in German. The standard deviations were

quite large indicating that the reading behavior in each group was not homogeneous. Future investigations with more participants and additional measures to distinguish between better and poorer learners, such as groupings according to the ACTFL or CEFR scores, might provide additional insights. Moreover, it would be interesting to explore different research assignments that have a clearer focus and a limited number of information readers need to select from texts. Finally, it might be useful to determine how the writing assignment affects the research behavior and the selection of information from texts. An assignment that requires L2 readers to outline advantages and disadvantages might affect their reading and researching behavior differently than an assignment that asks them to outline and describe policies.

About the authors

Susanne Rott (Ph.D., University of Illinois at Urbana-Champaign) is Associate Professor of Germanic Studies and Linguistics at the University of Illinois at Chicago. Her research focuses on second language lexical development in an instructed learning setting. She investigates how form-meaning mappings develop across interrelated continua that mark partial to complete, weak to robust and non-target-like to target-like continua.

Bianca Gavin (M.A., University of Illinois at Chicago) is a Ph.D. student at Penn State University in German Linguistics and Applied Linguistics.

References

- Akyel, A. & Erçetin, G. (2009). Hypermedia reading strategies employed by advanced learners of English. *System*, 37 (1), 136–152. <http://dx.doi.org/10.1016/j.system.2008.05.002>
- Anderson, N. J. (1991). Individual differences in strategy use in second language reading and testing. *The Modern Language Journal*, 75 (4), 460–472. <http://dx.doi.org/10.1111/j.1540-4781.1991.tb05384.x>
- Ariew, R. & Erçetin, G. (2004). Exploring the potential of hypermedia annotations for second language reading. *Computer Assisted Language Learning*, 17 (2), 237–259. <http://dx.doi.org/10.1080/0958822042000334253>
- Arnold, N., Ducate, L., & Kost, C. (2009). Collaborative writing in wikis: Insights from culture projects in intermediate German classes. In L. Lomicka & G. Lord (Eds), *The next generation: Social networking and online collaboration in language learning*, 115–144. San Marcos, TX: CALICO.
- Aydin, Z. & Yildiz, S. (2014). Using wikis to promote collaborative EFL writing. *Language Learning & Technology* 18 (1), 160–180.
- Bowles, M. (2010). *The think-aloud controversy in second language research*. New York: Routledge.

- Britt, M. A. & Aglinskas, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction*, 20 (4), 485–522. http://dx.doi.org/10.1207/S1532690XCI2004_2
- Britt, M. A. & Rouet, J. F. (2012). Learning with multiple documents: Component skills and their acquisition. In M. J. Lawson & J. R. Kirby (Eds), *The quality of learning: dispositions, instruction, and mental structures*, 276–314. Cambridge: Cambridge University Press.
- Carrell, P. L. (1998). Metacognitive awareness and second language reading. *The Modern Language Journal*, 73 (2), 121–134. <http://dx.doi.org/10.1111/j.1540-4781.1989.tb02534.x>
- Carrell, P. L., Pharis, B. G., & Liberto, J. C. (1989). Metacognitive strategy training for ESL reading. *TESOL Quarterly*, 23 (4), 647–678. <http://dx.doi.org/10.2307/3587536>
- Chun, D. (2001). L2 reading on the web: Strategies for accessing information in hypermedia. *Computer Assisted Language Learning*, 14 (5), 367–403. <http://dx.doi.org/10.1076/call.14.5.367.5775>
- Chun, D. (2012). Review article: Replication studies in CALL research. *CALICO Journal*, 29 (4), 591–600. <http://dx.doi.org/10.11139/cj.29.4.591-600>
- Ducate, L., Anderson, L., & Moreno, N. (2011). Wading through the world of wikis: An analysis of three wiki projects. *Foreign Language Annals*, 44 (3), 495–524. <http://dx.doi.org/10.1111/j.1944-9720.2011.01144.x>
- Goldman, S. R. (2004). Cognitive aspects of constructing meaning through and across multiple texts. In N. Shuart-Faris & D. Bloome (Eds), *Uses of intertextuality in classroom and educational research*, 317–352. Greenwich, CT: Information Age.
- Goldman, S., Braasch, J., Wiley, J., Graesser, A., & Brodwinska, K. (2012). Comprehending and learning from internet sources: Processing patterns of better and poorer learners. *Reading Research Quarterly*, 47 (4), 356–381.
- Kessler, G., Bikowski, D., & Boggs, J. (2012). Collaborative writing among second language learners in academic web-based Projects. *Language Learning & Technology*, 16 (1), 91–109.
- Konishi, M. (2003). Strategies for reading hypertext by Japanese ESL learners. *The Reading Matrix*, 3 (3), 97–117.
- Kost, C. R. (2011). Investigating writing strategies and revision behavior in collaborative wiki projects. *CALICO Journal*, 28 (3), 606–620. <http://dx.doi.org/10.11139/cj.28.3.606-620>
- Lawless, K. A., Goldman, S. R., Gomez, K., Manning, F., & Braasch, J. (2012). Assessing multiple source comprehension through evidence centered design. In J. P. Sabatini, E. Albro, & T. O'Reilly (Eds), *Reaching an understanding: Innovations in how we view reading assessment*. Lanham, MD: R&L Education.
- Perfetti, C. A., Rouet, J., & Britt, M. A. (1999). Toward a theory of documents representation. In H. van Oostendorp & S. R. Goldman (Eds), *The construction of mental representations during reading*, 88–108. Mahwah, NJ: Erlbaum.
- Rott, S. & Weber, E. D. (2013). Preparing students to use wiki software as a collabora-

- tive learning tool. *CALICO Journal*, 30 (2), 179–203. <http://dx.doi.org/10.11139/cj.30.2.179-203>
- Rouet, J. & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds), *Text relevance and learning from text*, 19–52. Charlotte, NC: Information Age.
- Sakar, A. & Erçetin, G. (2005). Effectiveness of hypermedia annotations for foreign language reading. *Journal of Computer Assisted Learning*, 21 (1), 28–38. <http://dx.doi.org/10.1111/j.1365-2729.2005.00108.x>
- Sanz, C., Lin, H. J., Lado, B., Bowden, H. W., & Stafford, C. A. (2009). Concurrent verbalizations, pedagogical conditions, and reactivity: Two CALL studies. *Language Learning*, 59 (1), 33–71. <http://dx.doi.org/10.1111/j.1467-9922.2009.00500.x>
- Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Source evaluation, comprehension, and learning in Internet science inquiry tasks. *American Educational Research Journal*, 46 (4), 1060–1106. <http://dx.doi.org/10.3102/0002831209333183>

Appendix A

Think-Aloud instructions

We are interested in how you do your research for your projects in German. While reading the texts we want you to say out loud what is on your mind. Anything that comes to your mind while reading about the text you should say aloud. This is called a Think-Aloud. During the research phase you will do the Think-Aloud the entire time. You should make comments whenever you like, but try to comment on each sentence you read. Basically, you will speak all the time. If, at any time you stop to comment or become quiet I will prompt you by saying ‘What are you thinking?’. Let’s give the Think-Aloud a try. I will read the first paragraph on the text about Hochschule to go (Appendix B), then you will read the second paragraph to practice thinking aloud. This text is just to practice; it has nothing to do with the text that you will read during the research phase.

Appendix B

Think-Aloud Example

The following text for the Think-Aloud example came from the website Sueddeutsche.de; <http://jetzt.sueddeutsche.de/texte/anzeigen/579290/Hochschule-to-go>

Hochschule to go [online seminars and degrees are already pretty popular in America]

Text: Charlotte Haunhorst

Mit dem Unistart sind in Deutschland auf der Plattform iversity.org auch die ersten MOOCs - also Onlinevorlesungen für alle – an den Start gegangen. [mmhhh, I didn’t know that there are no online degrees in Germany yet.] Über 100.000 Menschen haben sich dort bereits angemeldet. [That’s a lot of students that have signed up, even though this is the first semester that they are offering online classes.] Die klassischen Unis werden zukünftig umdenken müssen. [I wonder how will this affect the students? Or even the universities themselves if online classes become popular and students prefer those over regular classroom classes?]

Now you try. Read each sentence out loud and comment after it.

Die Anmeldung dauert keine Minute: Mailadresse angeben, Passwort reinhacken und schwups - schon ist man Student. Abiturnote? Höchster Bildungsabschluss? Semesterbeitrag? Interessiert hier niemanden. Denn anders als Präsenzunis ist iversity.org eine Online-Bildungseinrichtung für die Massen. Das Angebot ist gratis, die Anmeldung funktioniert auch über Facebook.