

Hahnel, Carolin; Goldhammer, Frank; Kröhne, Ulf; Naumann, Johannes **Reading digital text involves working memory updating based on task characteristics and reader behavior**

formal und inhaltlich überarbeitete Version der Originalveröffentlichung in:

formally and content revised edition of the original source in:

Learning and individual differences 59 (2017), S. 149-157



Bitte verwenden Sie in der Quellenangabe folgende URN oder DOI /
Please use the following URN or DOI for reference:

urn:nbn:de:0111-pedocs-175584

10.25656/01:17558

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-175584>

<https://doi.org/10.25656/01:17558>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft

**Reading digital text involves working memory updating based on task characteristics
and reader behavior**

Carolin Hahnel ^{ab}, Frank Goldhammer ^{ab}, Ulf Kröhne ^a & Johannes Naumann ^c

^a German Institute for International Educational Research (DIPF), Frankfurt am Main, Germany

^b Centre for International Student Assessment (ZIB), Germany

^c Goethe University, Frankfurt am Main, Germany

accepted for publication in *Learning and Individual Differences*

DOI: 10.1016/j.lindif.2017.09.001

Correspondence concerning this article should be addressed to:

Carolin Hahnel, German Institute for International Educational Research (DIPF)

Schlossstraße 29, 60486 Frankfurt am Main, Germany

Phone: +49 (0) 69 24 708 727

Email: hahnel@dipf.de

Abstract

Receiving and using web-based information has become part of everyday life, but the non-linear presentation of information can make considerable demands on cognitive resources, affecting text comprehension. This study examined whether memory updating predicts students' comprehension of digital hypertext over and above skills in reading linearly structured text, and whether this association is affected by particular characteristics of reading tasks, the hypertext and individual reading behavior. Measures included reading comprehension as assessed via hypertext (digital reading) and linear text (linear reading) as well as memory updating among 15-year-old German students ($N = 288$). The number of nodes in a hypertext and cognitive reading operations required for task processing were regarded as task characteristics. Indicators of reader behavior were derived from log files. The results demonstrated a general effect of memory updating on digital reading over and above linear reading. This effect was not affected by the number of available nodes but by cognitive reading operations and individual reader behavior. Implications for students' cognitive processing of hypertexts are discussed.

Keywords

Digital reading; Hypertext comprehension; Memory updating; Task characteristics; Log file analyses.

Highlights

- Memory updating predicted digital reading over and above linear reading.
- The effect of memory updating was not affected by the number of available nodes.
- The effect vanished when students were instructed to reflect on and evaluate text.
- The memory updating effect increased as more target nodes were accessed.
- This effect increase disappeared when revisits to target nodes were taken into account.

In today's society, receiving and using information from the World Wide Web (WWW) has become integral part of many private, academic, and occupational activities (Leu, Kinzer, Coiro & Cammack, 2004). As a result, measures of reading web-based information have been included in international comparative studies like the Programme for International Student Assessment (PISA), which aims to evaluate the skills and knowledge of students at the end of compulsory education (OECD, 2011). Web-based information is frequently structured in the form of non-linearly organized text pieces ("nodes") that are associated with one another and accessible through hyperlinks. Hypertexts offer readers numerous ways of collecting and combining pieces of information for specific reading purposes. However, processing information that is not presented contiguously can seriously affect comprehension of a text (Coiro, 2011; Rouet, 2006), since individuals' cognitive resources are limited (Feldman Barrett, Tugade & Engle, 2004) and decision-making and navigation requirements add to the load on readers' working memory (WM; DeStefano & LeFevre, 2007; Foltz, 1996; Scheiter, Gerjets, Vollmann & Catrambone, 2009).

In the present study, we investigated interindividual differences in 15-year-old German PISA students' comprehension of hypertexts. We examined how such differences are related to memory updating – the individual skill of actively monitoring and manipulating WM content (e.g., Oberauer & Kliegl, 2006). We aimed to investigate (1) whether memory updating is predictive of students' hypertext comprehension over and above their general reading skills, and (2) whether such an association is affected by particular characteristics of reading tasks, the hypertext and reading behavior. Examining these research questions will provide evidence on the relation between hypertext comprehension and WM (e.g., Naumann, Richter, Christmann & Groeben, 2008; Pazzaglia, Toso & Cacciamani, 2008), and generate further insights on the nature of information processing from hypertext. In the following, we will refer to the skills of comprehending electronic hypertext and linearly structured text as digital reading and linear reading, respectively.

1.1 Working memory and digital reading

Reading is an individual process of receiving and processing written information, ranging from decoding and recognizing words up to higher processes of word-text integration and meaning making (Perfetti & Stafura, 2014). In both digital and linear text, information should be conveyed in a coherent form that enables readers to extract meaning and to form a mental representation of the text situation (Foltz, 1996; Kintsch, 1998). In this regard, WM generally plays an essential role since individuals need to integrate information retrieved from the text and information activated from their long-term memory (e.g., Daneman & Merikle, 1996; Hannon, 2012; Oakhill, Yuill & Garnham, 2011). Hypertexts, though, offer readers a great deal of freedom in terms of how they receive information by simultaneously providing fewer cues about what information to process next and where to find it (Foltz, 1996). Therefore, digital reading requires increased activation of cognitive resources to allow readers to deal appropriately with the non-linear text structure without getting lost (Coiro 2011; Gyselinck, Jamet & Dubois, 2008; Srivastava & Gray, 2012). Accordingly, visuospatial WM capacity was shown to be associated with the recognition of hypertext structures among sixth graders, whereas verbal recall predicted their semantic knowledge (Pazzaglia et al., 2008). These effects were not due to linear reading skills, prior knowledge or short term memory. Similar effects were found for university students. Readers with a low verbal WM capacity recalled noticeably less information from digital text than from linear text (Lee & Tedder, 2003), and low visuospatial WM capacity was associated with difficulties in recalling hypertext structures and keeping track of link hierarchies (Rouet, Vörös & Pléh, 2012).

Previous studies have mainly related digital reading processes to verbal and visuospatial WM subcomponents, but not to the domain-general WM functions of active information storage and processing. Conceptualizing WM as “a system for building, maintaining and rapidly updating arbitrary bindings” for goal-directed information processing (Wilhelm, Hildebrandt & Oberauer, 2013, p.3), the memory updating paradigm was found to be a good

representation of the individual skill to flexibly bind structures into mental WM representations (Schmiedek, Hildebrandt, Lövdén, Wilhelm & Lindenberger, 2009). In contrast to other WM theories (e.g., Engle, 2002; Miyake et al., 2000), WM capacity limits are assumed to arise from interference due to temporary bindings that limit the complexity of novel representations (Oberauer, 2009). Since digital reading requires making sense of text by simultaneously monitoring and flexibly manipulating representations of the text situation and spatial relations between nodes, it should be closely related to memory updating.

1.2 Task influences

In general, readers are sensitive to demands of reading tasks that influence the way of their cognitive information processing (cf. Kendeou, van den Broek, Helder & Karlsson, 2014; Naumann, 2015; McCrudden & Schraw, 2007; Rouet, 2006). Such demands are often described as sources of cognitive load in WM (DeStefano & LeFevre, 2007; Rouet, 2009; Scheiter et al., 2009). Higher cognitive load is associated with differences in learning performance across different text structures (Zumbach & Mohraz, 2008), navigational maps (Amadiou et al., 2009; Scott & Schwartz, 2007), and reading orders (Madrid, Van Oostendorp & Puerta Melguizo, 2009). Readers reported less cognitive load, for example, when they had high prior knowledge or positive attitudes towards the text content (Amadiou, van Gog, Paas, Tricot & Mariné, 2009; Scheiter et al., 2009).

In PISA (OECD, 2013, p.66), “mental strategies, approaches or purposes that readers use to negotiate their way into, around and between texts” are described as “reading aspects”. These include the facets *access and retrieve*, *integrate and interpret*, *reflect and evaluate* and – the digital reading-specific aspect – *complex*. Table 1 lists examples of each reading aspect as well as operations required for task processing. Illustrated tasks refer either to a hypertext detailing an email exchange between two girls looking for a sports club (“Sports Club”), or a social media-like language learning platform (“Language Learning”). The different methods of text processing invoked by these reading aspects (Table 1) might involve WM

representations being updated differently. When asked for explicit or implicit text information (i.e., *access and retrieve, integrate and interpret*), students have to interact with the hypertext in order to locate and connect information distributed over a hypertext. As a result, they will put effort into decisional and navigational actions when searching the hypertext for the requested information. Keeping text information active and updated while using representations of spatial relations for navigating might then especially draw upon memory updating skills. In contrast, when asked to articulate an opinion (i.e., *reflect and evaluate*), students will have to retrieve ideas, attitudes, and experiences with similar texts from their general knowledge in a similar way to demands in linear reading.

Another task-specific influence might concern structural conditions in hypertexts. In their review of hypertext reading, DeStefano and LeFevre (2007) proposed that an increasing amount of information in hypertext – or more precisely, the number of hyperlinks – increases cognitive load in readers. Although there is some evidence against this claim (Madrid et al., 2009), it can be argued that memory updating might become especially important when required information is widely distributed across the text. Since readers need to create bindings regarding the location, content and relations of nodes, they might have to evaluate and update their mental representation of a hypertext every time a new node is encountered. This should hold for the nodes that are necessary and germane for a specific task (target nodes; cf. McCrudden & Schraw, 2007), but also for the nodes that are completely irrelevant to it (irrelevant nodes).

The way in that readers interact with a hypertext structure (i.e., their navigation behavior) particularly influences how text is received, processed and comprehended (e.g., Hahnel, Goldhammer, Naumann & Kröhne, 2016; Madrid et al., 2009; Naumann & Salmerón, 2015). Navigation requires more controlled processing since readers need to simultaneously integrate information as well as anticipate and plan their reading progress as they read (Foltz, 1996; Naumann & Goldhammer, 2017). Therefore, navigation events as recorded in log files, for

example, are frequently used to shine a light on students' decisions and strategies of information access and use (cf. Scheiter et al., 2009). Empirically, navigation behavior has been found to partially mediate the effect of WM capacity on hypertext learning outcomes (Naumann et al., 2008), especially when it is characterized as the comparison of information from different perspectives (Kornmann et al., 2016). However, general reading skills have rarely been taken into account in previous studies. Less skilled readers might struggle to accurately choose relevant text sections or misjudge their level of comprehension (cf. Foltz, 1996; Salmerón, Cañas, Kintsch & Fajardo, 2005). According to Walczyk (2000), readers can apply controlled adjustments to their reading (e.g. rereading the text) in order to overcome deficits in decoding skills. Less skilled comprehenders, though, often believe they are able to answer questions without rereading text passages, decide more often against going back to a text and fail more often to give correct answers than skilled comprehenders (Vidal-Abarca, Mañá & Gil, 2010).

1.3 Rationale

In this study, we examined whether and how interindividual differences in memory updating as the individual skill to flexibly bind and unbind WM structures affect 15-year-olds' digital text comprehension. Memory updating is supposed to play a central role in simultaneously monitoring and manipulating representations of text and space and, therefore, should predict digital reading over and above joint processes with linear reading.

***H1.** The probability of solving a digital reading task correctly is predicted positively by both linear reading and memory updating.*

Assuming that an impact of memory updating does not result from individual differences rooted in shared processes of digital and linear reading (e.g., decoding or comprehension processes), its effect on digital reading is expected to differ depending on specific reading conditions concerning cognitive operations and hypertext structure.

H2. The effect of memory updating is more pronounced in tasks focusing on the text base (i.e., access and retrieve, integrate and interpret) than in tasks requiring knowledge-based judgments (i.e., reflect and evaluate).

H3a. The memory updating effect increases with the number of target nodes.

H3b. The memory updating effect increases with the number of irrelevant nodes.

When readers encounter new target nodes, they need to form bindings related to the new information and update existing mental representations. However, in order to consolidate relevant WM content, readers might adjust their reading by revisiting target information. A “refreshing” of WM bindings could alleviate differences due to individual memory updating skills and help prevent deficits in comprehension.

H4a. The effect of memory updating increases with the number of target nodes visited.

H4b. The effect of memory updating is reduced when students revisit target nodes.

Method

2.1 Sample

Data from 288 15-year-old students ($M=15.85$, $SD=0.29$) was used (53.47% female). These students participated in the PISA 2012 digital reading assessment and additionally in a German add-on study on computer-based assessment (CBA). The PISA sampling procedure included a two-stage sampling design. Across 212 PISA-eligible schools, 14 students were randomly drawn to participate in PISA CBA (OECD, 2014), which included the digital reading assessment. Another 14 students from 77 schools were drawn to participate in the add-on study. The intersection of these two subsamples formed the data basis.

2.2 Materials and measures

2.2.1 Digital reading

Digital reading was assessed via 19 items, clustered into six units (OECD, 2013). A unit provided a simulated hypertext and included two to four reading items (Figure 1). The

hypertext environment provided functional menus, tabs, buttons, and hyperlinks which students could use to navigate freely within the hypertext while answering a particular item. Navigation back to previous items was not possible. The hypertext content covered different reading situations (e.g., private, educational), text types (e.g., description, argumentation), and comprised between nine and 33 nodes in total. The number of target nodes (i.e., nodes with information required to solve a digital reading task correctly) ranged from one to five pages per item ($M=1.63$, $SD=1.26$). The fourth example task in Table 1, for instance, contained five target nodes: The email exchange that provided criteria on which students should base their answer (i.e., common date, sport preferences, low price) and the nodes of two gyms addressing these restrictions (one node for gym 1 and three for gym 2). Nodes that needed to be passed to reach a target node but did not contribute to task completion were considered neither target nor irrelevant. The number of irrelevant nodes (i.e., nodes which provide no helpful information for the task solution) ranged from four to 32 pages ($M=16.11$, $SD=8.08$).

Item response formats included multiple-choice questions, open text answers, and mixed formats. Open text responses were scored by trained coders with the aid of standardized coding guidelines (OECD, 2014). Students' responses were coded dichotomously. Six partial credit items were dichotomized to apply generalized linear mixed modeling (partial and full credit combined). Across items, the proportion of correct responses ranged from 17.84% to 94.47%. The items fit a Rasch model (Embretson & Reise, 2000) and appropriately covered the range of students' digital reading skills (Figure 2). Items for the reading aspect *access and retrieve* tended to be easier on average than items associated with other reading aspects, with *complex* items the most difficult. Cronbach's α (.77) and the reliability of expected a posteriori (EAP) scores (.74) were acceptable.

The reading aspects and the numbers of target and irrelevant nodes were regarded as task characteristics. To represent reader behavior, events recording visits to target nodes were extracted from students' log files. By averaging their visits across items, the number of target

nodes *visited* ($M=1.55$; $SD=0.45$) and the number of target node *visits* ($M=2.68$; $SD=1.17$) were derived for each student. Note that the first indicator counts visits only once whereas the second indicator counts visits and revisits. To examine the specificity of the effects of goal-relevant navigation, similar indicators were also derived for irrelevant node visits (number of irrelevant nodes visited: $M=0.74$, $SD=0.54$; number of irrelevant node visits: $M=1.09$, $SD=1.03$).

2.2.2 Memory updating

Students were asked to complete a numerical memory updating task by memorizing a sequence of numbers and mentally adding or subtracting numbers presented afterwards. Figure 3 illustrates the process for one item. The start sequence varied from two to four digits. The operators ranged from -8 to +8. All start, interim, and resulting numbers ranged from zero to nine. After stimulus presentation, students were asked to type in the result. Corrections were possible. Responses for 21 items were collected (Cronbach's $\alpha = .90$, EAP reliability = .88). The proportion of correct responses ranged from 7.90% to 71.07%. EAP scores, derived from a 2PL item response model (sample for scaling: $N = 639$), served as estimates of memory updating.

2.2.3 Linear reading

Linear reading was measured via 18 items from PISA 2009 (OECD, 2010). The items were clustered into units containing a linear text and three to four items. The texts included different formats (e.g., continuous, non-continuous) and types (e.g., description, narration). The items covered different reading aspects (*access and retrieve*, *integrate and interpret*, *reflect and evaluate*) and reading situations (e.g., public, educational). Examples can be retrieved from <http://www.oecd.org/pisa/38709396.pdf>. Units were administered either via computer or paper based; no effect of administration mode was found at the construct level (Kröhne, Hahnel, Schiepe-Tiska & Goldhammer, 2013). Response formats included multiple-choice and open text answers. Responses to 16 items were coded dichotomously; three were

coded with partial credits. The proportion of respondents receiving full credit on each item ranged from 8.02% to 78.35%. EAP scores, derived from a generalized partial credit model (Nering & Ostini, 2010; $N = 880$), served as estimates of linear reading (Cronbach's $\alpha = .83$, EAP reliability = .76).

2.3 Procedure

Students participated in groups of 14 on two days and were given instructions by trained test administrators. They were randomly assigned to all test conditions and received comprehensive tutorials. In the digital reading assessment, students received either three or all six digital reading units on the first day. Students participating in the add-on study, which took place within one week of the first testing day, received four linear reading units and then either (1) the other linear reading units, (2) the updating test, or (3) both.

2.4 Data analyses

Generalized linear mixed models (GLMMs; Bolker, et al., 2008; De Boeck et al., 2011) were used to test the hypotheses. Using a logit-link, these models can be used for regressions of the dichotomous digital reading scores on several predictor variables on both the student (i.e., memory updating, linear reading, indicators of node visits and nodes visited) and item level (i.e., reading aspects, number of target and irrelevant nodes). They also can take into account that digital reading scores are hierarchically nested within items, students and schools. In each model, items were modeled as fixed effects (i.e., item easiness), while random effects were modeled for students and schools (i.e., students' digital reading skill and the performance level of a school). Furthermore, each model included a fixed effect of linear reading to account for the effects of general comprehension (cf. Vidal-Abarca et al., 2010).

All analyses were carried out in R 3.1.3 (R Core Team, 2015) with the additional packages *TAM* (Kiefer, Robitzsch & Wu, 2014), *ltm* (Rizopoulos, 2006), *WrightMap* (Irribarra & Freund, 2014), and *lme4* (Bates, Mächler, Bolker & Walker, 2014). All tests were one-tailed, with a Type I error probability of 5%. All metric variables were z-standardized.

Thus, the regression coefficients are interpretable as predicted changes in the log odds of the probability of giving a correct response if a predictor increases by one standard deviation.

Results

The variances of students' digital reading skills and the performance level of schools in a baseline model without predictors were about 0.52 and 0.98, corresponding to intraclass correlations (ICC1; Skrondal & Rabe-Hesketh, 2004) of 0.14 and 0.23. Table 2 shows descriptive statistics and correlations between the student-specific variables.

3.1 General effect of memory updating

To test the general effect of memory updating (H1), digital reading was regressed on linear reading with and without memory updating. The first model showed that linear reading positively predicted students' success in digital reading tasks ($b_1=0.71$, $z=9.97$, $p<.001$) and explained a large amount of variance between students ($R^2=.35$). The explained variance increased by 7.55% when memory updating was added as a second predictor. Although the predictors were highly correlated (Table 2), memory updating still significantly predicted digital reading ($b_2=0.26$, $z=3.85$, $p<.001$) after controlling for linear reading ($b_1=0.59$, $z=7.39$, $p<.001$).

3.2 Variation across task characteristics

A model including interaction effects between reading aspects and memory updating (Table 3) was specified to test whether the memory updating effect on digital reading varies for different cognitive reading operations (H2). The memory updating effect was relatively high for *access and retrieve* items, indicating that students with efficient memory updating had an advantage in such tasks over students with less efficient memory updating. This advantage diminished significantly towards zero for *reflect and evaluate* items. Memory updating had no effect over and above linear reading in tasks requiring the evaluation of a text using one's own knowledge and personal experiences. The effects for *integrate and interpret*

as well as *complex* items did not differ significantly from the memory updating effect in *access and retrieve* tasks.

The numbers of target and irrelevant nodes were added as predictors to test whether the memory updating effect increases with the number of target nodes (H3a) and irrelevant nodes (H3b). Although the memory updating effect was still significant (Table 4), the interaction effects with the number of target and irrelevant nodes did not reach statistical significance. The memory updating effect did not increase with the number of target or irrelevant nodes.

3.3 Interaction with reader behavior

The indicators of reader behavior were positively correlated (Table 2). The means and the correlations between the unique visit and visit-revisit variables indicate that revisits of nodes seldom occurred for either target or irrelevant nodes. The correlations between the target and irrelevant node visit indicators suggest a tendency among students to explore the hypertext content, regardless of the nodes' relevance.

To test whether the memory updating effect is affected by the number of target nodes accessed (H4a) but diminishes when target nodes are revisited (H4b), two models were specified including unique node visits and repeated visits, respectively. To show that the effects of memory updating do not depend on general navigation behavior, the irrelevant page visits were included in the models (Table 5). The target node visit indicators exhibited a generally positive effect on digital reading. The more target nodes students located, the more likely they were to respond correctly to a digital reading item. In line with our hypotheses, the interaction in the first model showed a significant increase of the memory updating effect with the number of target nodes visited, whereas the interaction with the number of target node visits was not significant. In both models, there were no further effects of the irrelevant node visit indicators on the probability of a correct item response.

Discussion

Digital information resources provide great flexibility for readers to gather information quickly and efficiently. However, processing information that is not presented contiguously can affect comprehension by producing additional load on readers' WM (cf., DeStefano & LeFevre, 2007). Therefore, the present study investigated the effects of memory updating on digital reading. It examined whether the digital reading performance of 15-year-old German students would be predicted by their skill in memory updating. Moreover, this association was considered to be affected by the demands of tasks, hypertexts and actual reader behavior. In summary, the results showed that students benefited from efficient memory updating in their hypertext comprehension over and above linear reading skills. This general effect did not differ between students visiting different amounts of target or irrelevant nodes. It vanished, though, when the task required reflecting on and evaluating text rather than simply extracting text information explicitly or implicitly. Accessing nodes with target information was also shown to accelerate the effect of memory updating, but this interaction effect did not hold when revisits were taken into account.

4.1 Relationship between memory updating and digital reading

Taken together, the results contribute to a growing body of research on the role of WM for hypertext processing. Our findings showed the same relational pattern as studies using other methodological approaches (e.g., Gyselinck et al., 2008; Lee & Tedder, 2003; Naumann et al., 2008): When trying to comprehend hypertext, students benefit from efficient WM functions over and above their linear reading skills. In this respect, the relationship between memory updating and linear reading identified here was in line with previous research on reading and WM capacity (e.g., Dehn, 2008; Oakhill et al., 2011). Reading-related memory processes were taken into consideration in the analyses conducted, adding further evidence in support of the general assumption that decisional and navigational demands in hypertext draw upon WM resources.

Due to the GLMM approach, we were able to elaborate on the role of memory updating in digital reading in more detail by analyzing effects on the task and student levels. Concerning the cognitive reading operations of tasks, memory updating skills seemed to be particularly required when students were instructed to explicitly or implicitly extract information from a hypertext (i.e., *access and retrieve, integrate and interpret*) rather than retrieve and use knowledge about texts (i.e., *reflect and evaluate*). That does not mean that memory updating is unnecessary in *reflect and evaluate* tasks, but it is not required above shared processes with linear reading. The effect variation is also not a symptom of task difficulty, as some might suggest. Along with other items, the task difficulty of *reflect and evaluate* items covered an area of average ability in digital reading (Figure 2). We concluded that in information extraction tasks, memory updating skills serve the need to keep text bindings active while locating and evaluating further information in the digital space.

Concerning the number of nodes within a hypertext, there was no support for an interaction with memory updating. It might not be the mere quantity but the quality of information that is crucial. The ease of retaining WM representations can be affected by many other variables, like prior knowledge (Amadiou et al., 2009; Rouet, 2009) or mental integration processes (e.g., chunking or subvocalisation as a form of inner speech; Dehn, 2008). Demands on WM might be relieved, for instance, by maintaining a mental configuration that represents the hypertext structure (Pazzaglia et al., 2008) or establishing coherence between text parts (cf. Kintsch, 1998). Future research needs to address such micro processes of information management and integration in digital reading. Readers' actual interactions with nodes are another explanation for the lack of significant effects. Readers' node selection has previously been shown to interact with their prior knowledge and to affect hypertext comprehension (Salmerón et al., 2005). Our results also showed that the effect of memory updating increased with the number of target nodes accessed. The perception of cognitive load might be influenced by individual behavior to a greater extent than suggested

by DeStefano and LeFevre (2007). Regarding the interplay between hypertext and readers' use, the amount of information might affect readers' perception of cognitive load in a differential way depending on their strategies for dealing with a hypertext environment.

Furthermore, the finding that the additional memory updating effect was not found when revisits to target nodes were taken into account is of particular interest. It shows that the impact of an individual skill directly corresponds to specific behaviors observable in log files. It suggests that students with less efficient memory updating skills can compensate for their deficits by consolidating bindings in WM representations through revisits to particular nodes (cf. Walczyk, 2000). An implication might be that students could benefit from training in how to improve their skills in information management or their use of compensatory strategies. However, log files only allowed us to observe when students visited nodes repeatedly. We assumed that readers used revisits for rereading, but what exactly they did in terms of cognitive processing was not traceable. Other reasons may include deficient decoding or self-regulation skills that led to an improper processing of node contents (Foltz, 1996; Vidal-Abarca et al., 2010).

4.2 Limitations

There are at least four major limitations of the present study. First, the specificity of a memory updating effect as an effect of maintaining and manipulating WM content is not completely established. Memory updating tasks have previously been shown to be perfectly accounted for by a WM factor comprising various WM measures on a latent level (Schmiedek et al., 2009). Therefore, it represents a mixture of general WM capacity as well as the specific efficiency of executive processes (Wilhelm et al., 2013). Further investigations should try to avoid mono-operation biases and include other measures of verbal and visuospatial WM (e.g., Gyselinck et al., 2008). Applying a latent modeling approach can then be used to rule out task-specific variance and strengthen the effect interpretations of different WM components and their functions in digital reading.

Second, differences in the interaction effects between memory updating and the node visit indicators might be due to the closed environment of the digital reading units. The hypertexts are merely small, partial simulations of the WWW, and possible actions are limited to the functions provided (e.g., search boxes were visible but did not serve any function). The question arises as to whether revisiting nodes represents a behavior actually conducted in an open web space. Readers often just skim texts for information (Coiro, 2011), and revisits were generally rare events in our data. Closely related to this, the hypertexts also contained only a small number of nodes and did not vary systematically in their structural complexity (e.g., simple tree vs. complex network), or in the availability of maps or other orientation aids (cf. Amadieu et al., 2009). That might have limited the variability in observing revisits.

Third, information is highlighted as relevant through the instruction in tasks (McCrudden & Schraw, 2007). Therefore, the definition of relevance depends on the type of task and reader characteristics. Nodes will not be considered relevant if they do not meet readers' perception of a task that varies as a function of their comprehension skills and prior knowledge. Accordingly, different tasks can trigger readers to modify their navigation behavior and to visit nodes based on different intentions (e.g., rereading an isolated piece of information vs. reviewing for the purpose of integrating information; Rouet, Vidal-Abarca, Erboult & Millogo, 2001). An in-depth examination of node visits under different reading tasks will require further research that includes systematic variations of task types and relevance instructions.

Finally, logit-based GLMMs require dichotomous data that relates to a unidimensional skill. To use this approach, it was necessary to dichotomize six originally partial-credit scored items, even though dichotomization can result in an artificial reduction of variance and a loss of information on individual differences (MacCallum, Zhang, Preacher & Rucker, 2002). The dichotomous responses fit a Rasch model, but the results still raise questions about the unidimensionality of the digital reading construct. From a psychometric point of view, items

would be expected to tap into multidimensional skills if their processing required different component skills (e.g., the results of memory updating under different reading aspects).

Indeed, the effect pattern is consistent with the provided theoretical background and serves as an indicator of the complex nature of hypertext processing, but it calls the sensitivity of GLMMs into questions, meaning that further research is required.

4.3 Conclusions

Different implications might be drawn for learning and instructional purposes. For learning purposes, the results suggest that readers should learn from hypertext that is designed to be an appropriate fit to their cognitive skills (e.g., by providing optional opportunities for note-taking, scaffolding, or the repetition of central information). For the purpose of instruction, the difficulty of digital reading tasks should be increased gradually to stimulate new strategies and foster integration skills in readers. However, more research is needed to verify such conclusions. Analyzing the processes that lead to particular response outcomes should be a central focus here. Combining performance indicators derived from real web tasks with more fine-grained process data (e.g., eye-tracking or dual task approaches) can be used to validate interpretations about strategies and behaviors from log files.

Acknowledgments

This research was funded by the Centre for International Student Assessment (ZIB). We want to thank Florian Schmiedek for providing us with the memory updating test and Keri Hartman for proofreading our manuscript. Furthermore, we are thankful to two anonymous reviewers for their valuable insights and constructive comments.

References

- Amadiou, F., van Gog, T., Paas, F., Tricot, A., & Mariné, C. (2009). Effects of prior knowledge and concept-map structure on disorientation, cognitive load, and learning. *Learning and Instruction, 19*, 376–386. <https://doi.org/10.1016/j.learninstruc.2009.02.005>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. Retrieved from <http://CRAN.R-project.org/package=lme4>
- Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J.-S. S. (2008). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology and Evolution, 24*, 127–135. <https://doi.org/10.1016/j.tree.2008.10.008>
- Coiro, J. (2011). Predicting Reading Comprehension on the Internet: Contributions of Offline Reading Skills, Online Reading Skills, and Prior Knowledge. *Journal of Literacy Research : A Publication of the Literacy Research Association, 43*, 352–392. <https://doi.org/10.1177/1086296X11421979>
- Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*, 422–433.
- De Boeck, P., Bakker, M., Zwitser, R., Nivard, M., Hofman, A., Tuerlinckx, F., & Partchev, I. (2011). The estimation of item response models with the lmer function from the lme4 package in R. *Journal of Statistical Software, 39*, 1–28.
- Dehn, M. J. (2008). *Working memory and academic learning: assessment and intervention*. Hoboken, N.J: John Wiley & Sons, Inc.
- DeStefano, D., & LeFevre, J.-A. (2007). Cognitive load in hypertext reading: A review. *Computers in Human Behavior, 23*, 1616–1641. <https://doi.org/10.1016/j.chb.2005.08.012>
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, N.J: L. Erlbaum Associates.

- Engle, R. W. (2002). Working Memory Capacity as Executive Attention. *Current Directions in Psychological Science*, 11, 19–23. <https://doi.org/10.1111/1467-8721.00160>
- Feldman Barrett, L., Tugade, M. M., & Engle, R. W. (2004). Individual Differences in Working Memory Capacity and Dual-Process Theories of the Mind. *Psychological Bulletin*, 130, 553–573. <https://doi.org/10.1037/0033-2909.130.4.553>
- Foltz, P. (1996). Comprehension, Coherence, and Strategies in Hypertext and Linear Text. In J.-F. Rouet, J. J. Levonen, A. Dillon, & R. J. Spiro (Eds.), *Hypertext and Cognition* (pp. 109–136). Mahwah, NJ: Lawrence Erlbaum Associates.
- Gyselinck, V., Jamet, E., & Dubois, V. (2008). The role of working memory components in multimedia comprehension. *Applied Cognitive Psychology*, 22, 353–374. <https://doi.org/10.1002/acp.1411>
- Hahnel, C., Goldhammer, F., Naumann, J., & Kröhne, U. (2016). Effects of linear reading, basic computer skills, evaluating online information, and navigation on reading digital text. *Computers in Human Behavior*, 55, 486–500. <https://doi.org/10.1016/j.chb.2015.09.042>
- Hannon, B. (2012). Understanding the Relative Contributions of Lower-Level Word Processes, Higher-Level Processes, and Working Memory to Reading Comprehension Performance in Proficient Adult Readers. *Reading Research Quarterly*, 47, 125–152. <https://doi.org/10.1002/RRQ.013>
- Irribarra, D. T., & Freund, R. (2014). *Wright Map: IRT item-person map with ConQuest integration*. Retrieved from <http://github.com/david-ti/wrightmap>
- Kendeou, P., van den Broek, P., Helder, A., & Karlsson, J. (2014). A Cognitive View of Reading Comprehension: Implications for Reading Difficulties. *Learning Disabilities Research & Practice*, 29, 10–16.
- Kiefer, T., Robitzsch, A., & Wu, M. (2015). *TAM: Test Analysis Modules*. Retrieved from <http://CRAN.R-project.org/package=TAM>

Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. Cambridge University Press.

Kornmann, J., Kammerer, Y., Anjewierden, A., Zettler, I., Trautwein, U., & Gerjets, P.

(2016). How children navigate a multiperspective hypermedia environment: The role of spatial working memory capacity. *Computers in Human Behavior*, 55, 145–158.

<https://doi.org/10.1016/j.chb.2015.08.054>

Kröhne, U., Hahnel, C., Schiepe-Tiska, A., & Goldhammer, F. (2013, August). *Analyzing*

Mode Effects of PISA Print Reading Including a Comparison of Time-related

Information. Presented at the 15th Biennial Conference of the European Association for Research on Learning and Instruction (EARLI), Munich, Germany.

Lee, M. J., & Tedder, M. C. (2003). The effects of three different computer texts on readers'

recall: based on working memory capacity. *Computers in Human Behavior*, 19, 767–783.

[https://doi.org/10.1016/S0747-5632\(03\)00008-6](https://doi.org/10.1016/S0747-5632(03)00008-6)

Leu, D. J., Kinzer, C. K., Coiro, J. L., & Cammack, D. W. (2004). Toward a Theory of New

Literacies Emerging From the Internet and Other Information and Communication

Technologies. In R. B. Ruddell & N. Unrau (Eds.), *Theoretical Models and Processes of Reading* (5th ed., pp. 1568 – 1611). Newark: International Reading Association.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of

dichotomization of quantitative variables. *Psychological Methods*, 7, 19–40.

<https://doi.org/10.1037//1082-989X.7.1.19>

Madrid, R. I., Van Oostendorp, H., & Puerta Melguizo, M. C. (2009). The effects of the

number of links and navigation support on cognitive load and learning with hypertext:

The mediating role of reading order. *Computers in Human Behavior*, 25, 66–75.

<https://doi.org/10.1016/j.chb.2008.06.005>

McCrudden, M. T., & Schraw, G. (2007). Relevance and Goal-Focusing in Text Processing.

Educational Psychology Review, 19, 113–139. [https://doi.org/10.1007/s10648-006-9010-](https://doi.org/10.1007/s10648-006-9010-7)

- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cognitive Psychology*, 41, 49–100. <https://doi.org/10.1006/cogp.1999.0734>
- Naumann, J. (2015). A model of online reading engagement: Linking engagement, navigation, and performance in digital reading. *Computers in Human Behavior*, 53, 263–277. <https://doi.org/10.1016/j.chb.2015.06.051>
- Naumann, J., & Goldhammer, F. (2017). Time-on-task effects in digital reading are non-linear and moderated by persons' skills and tasks' demands. *Learning and Individual Differences*, 53, 1–16. <https://doi.org/10.1016/j.lindif.2016.10.002>
- Naumann, J., Richter, T., Christmann, U., & Groeben, N. (2008). Working memory capacity and reading skill moderate the effectiveness of strategy training in learning from hypertext. *Learning and Individual Differences*, 18, 197–213. <https://doi.org/10.1016/j.lindif.2007.08.007>
- Naumann, J. & Salmerón, L. (2016). Does navigation always predict performance? Effects of navigation on digital reading are moderated by comprehension skills. *The International Review of Research in Open and Distributed Learning*, 17, 42–59. <https://doi.org/10.19173/irrodl.v17i1.2113>
- Nering, M. L., & Ostini, R. (Eds.). (2010). *Handbook of polytomous item response theory models*. New York: Routledge.
- Oakhill, J., Yuill, N., & Garnham, A. (2011). The differential relations between verbal, numerical and spatial working memory abilities and children's reading comprehension. *International Electronic Journal of Elementary Education*, 4, 83–106.
- Oberauer, K. (2009). Design for a Working Memory. *Psychology of Learning and Motivation*, 51, 45–100. [https://doi.org/10.1016/S0079-7421\(09\)51002-X](https://doi.org/10.1016/S0079-7421(09)51002-X)

- Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language*, 55, 601–626.
<https://doi.org/10.1016/j.jml.2006.08.009>
- Organisation for Economic Co-operation and Development (OECD). (2010). *PISA 2009 Assessment Framework*. Paris: OECD Publishing.
- Organisation for Economic Co-operation and Development (OECD). (2011). *PISA 2009 Results: Students On Line*. Paris: OECD Publishing.
- Organisation for Economic Co-operation and Development (OECD). (2013). *PISA 2012 Assessment and Analytical Framework*. Paris: OECD Publishing.
- Organisation for Economic Co-operation and Development (OECD). (2014). *PISA 2012 Technical Report*. Paris: OECD Publishing.
- Pazzaglia, F., Toso, C., & Cacciamani, S. (2008). The specific involvement of verbal and visuospatial working memory in hypermedia learning. *British Journal of Educational Technology*, 39, 110–124. <https://doi.org/10.1111/j.1467-8535.2007.00741.x>
- Perfetti, C., & Stafura, J. (2014). Word Knowledge in a Theory of Reading Comprehension. *Scientific Studies of Reading*, 18, 22–37. <https://doi.org/10.1080/10888438.2013.827687>
- R Core Team. (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rizopoulos, D. (2006). ltm: An R package for latent variable modeling and item response theory analyses. *Journal of Statistical Software*, 17, 1–25.
- Rouet, J.-F. (2006). *The skills of document use: from text comprehension to Web-based learning*. Mahwah, NJ: Erlbaum.
- Rouet, J.-F. (2009). Managing cognitive load during document-based learning. *Learning and Instruction*, 19, 445–450. <https://doi.org/10.1016/j.learninstruc.2009.02.007>

- Rouet, J.-F., Vidal-Abarca, E., Erboul, A. B., & Millogo, V. (2001). Effects of Information Search Tasks on the Comprehension of Instructional Text. *Discourse Processes*, 31, 163-186.
- Rouet, J.-F., Vörös, Z., & Pléh, C. (2012). Incidental learning of links during navigation: the role of visuo-spatial capacity. *Behaviour & Information Technology*, 31, 71–81.
<https://doi.org/10.1080/0144929X.2011.604103>
- Salmeron, L., Canas, J. J., Kintsch, W., & Fajardo, I. (2005). Reading Strategies and Hypertext Comprehension. *Discourse Processes*, 40, 171–191.
https://doi.org/10.1207/s15326950dp4003_1
- Scheiter, K., Gerjets, P., Vollmann, B., & Catrambone, R. (2009). The impact of learner characteristics on information utilization strategies, cognitive load experienced, and performance in hypermedia learning. *Learning and Instruction*, 19, 387–401.
<https://doi.org/10.1016/j.learninstruc.2009.02.004>
- Schmiedek, F., Hildebrandt, A., Lövdén, M., Wilhelm, O., & Lindenberger, U. (2009). Complex span versus updating tasks of working memory: The gap is not that deep. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 1089–1096.
<https://doi.org/10.1037/a0015730>
- Scott, B. M., & Schwartz, N. H. (2007). Navigational spatial displays: The role of metacognition as cognitive load. *Learning and Instruction*, 17, 89–105.
<https://doi.org/10.1016/j.learninstruc.2006.11.008>
- Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: multilevel, longitudinal, and structural equation models*. Boca Raton: Chapman & Hall/CRC.
- Srivastava, P., & Gray, S. (2012). Computer-Based and Paper-Based Reading Comprehension in Adolescents With Typical Language Development and Language-Learning Disabilities. *Language Speech and Hearing Services in Schools*, 43, 424–437.

Vidal-Abarca, E., Mañá, A., & Gil, L. (2010). Individual differences for self-regulating task-oriented reading activities. *Journal of Educational Psychology, 102*, 817–826.

<https://doi.org/10.1037/a0020062>

Walczyk, J. J. (2000). The interplay between automatic and control processes in reading. *Reading Research Quarterly, 35*, 554–566.

Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology, 4*.

<https://doi.org/10.3389/fpsyg.2013.00433>

Zumbach, J., & Mohraz, M. (2008). Cognitive load in hypermedia reading comprehension: Influence of text type and linearity. *Computers in Human Behavior, 24*, 875–887.

<https://doi.org/10.1016/j.chb.2007.02.015>

MEMORY UPDATING IN READING DIGITAL TEXT

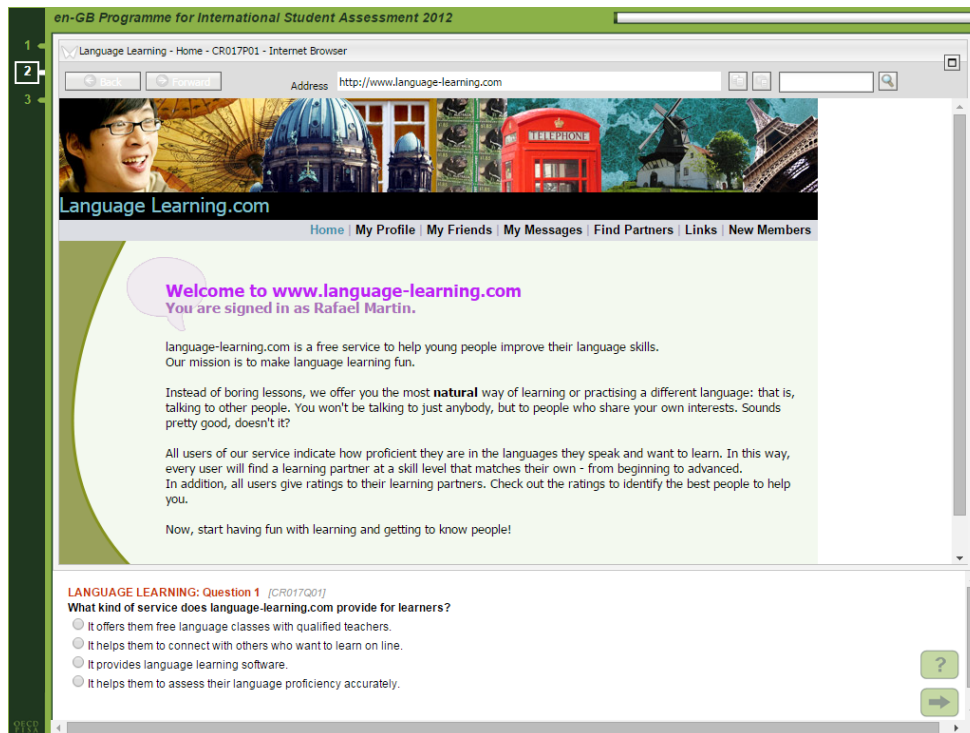


Figure 1. Screenshot of the digital reading hypertext “Language learning”. Further example tasks can be retrieved from the PISA 2012 Assessment Framework (OECD, 2013, p.72-78) as well as from the OECD website <https://www.oecd.org/pisa/pisaproducts/PISA%202009%20reading%20test%20items.pdf> (Annex A2, p. 233-247).

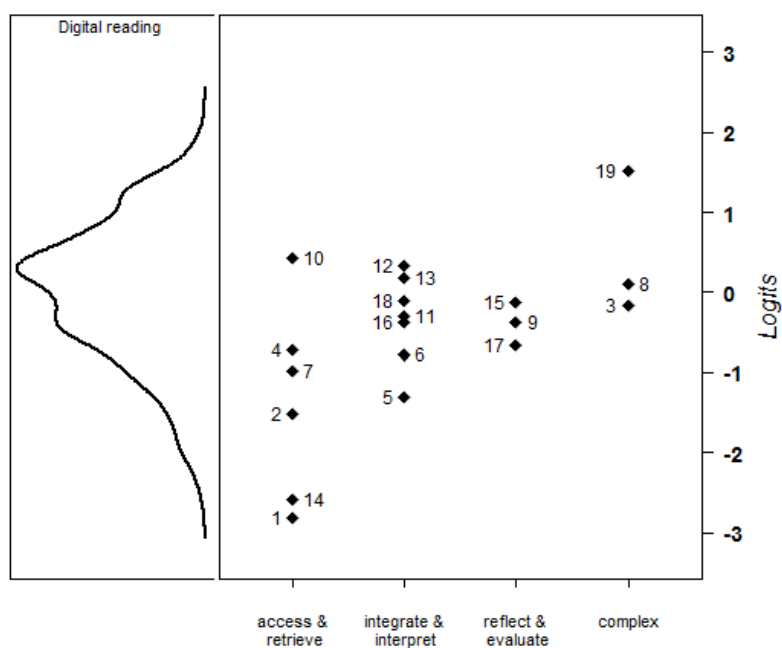


Figure 2. Wright map of the distribution of students' digital reading skills (left) mapped on the same scale as the difficulty of the digital reading items (right). Item difficulties are clustered according to each respective item's reading aspect (x-axis).

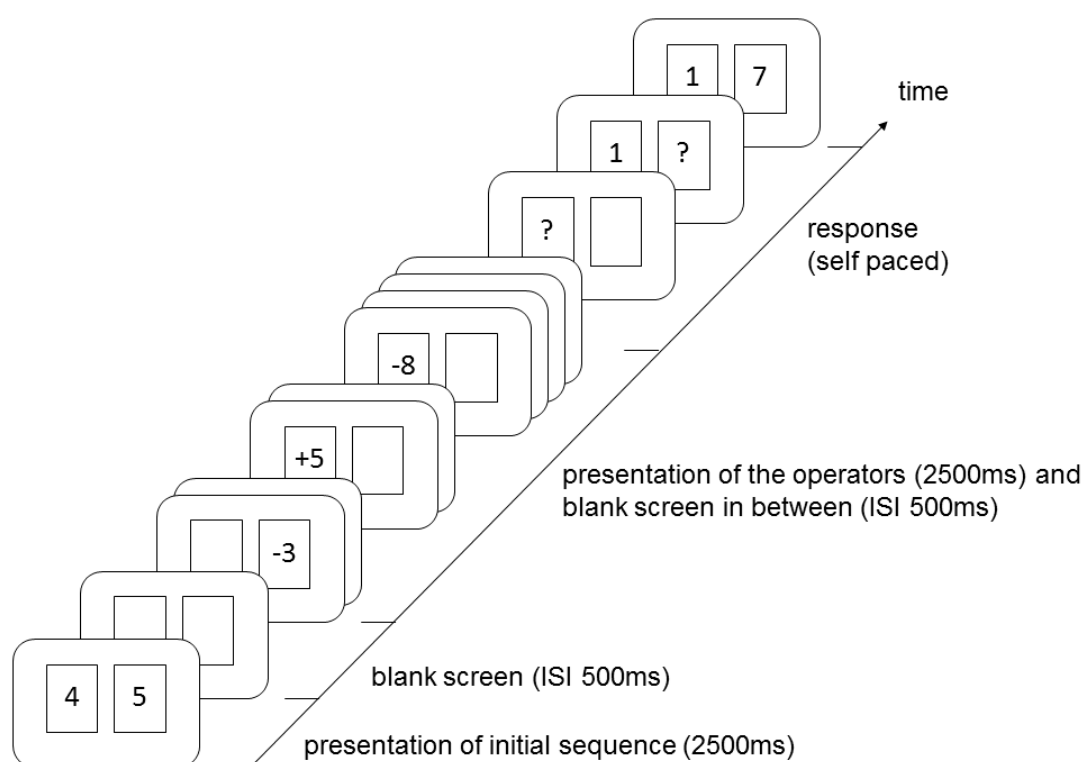


Figure 3. Example of the process of a memory updating item with two digits. The start sequence was presented for 2500ms. After interstimulus intervals (ISI) of 500ms, two operations per digit appeared successively and were presented for 2500ms. Adaptation according to Oberauer and Kliegl (2006, p. 603).

MEMORY UPDATING IN READING DIGITAL TEXT

Table 1

Examples of reading aspects in the digital reading items

Reading aspect	Description	Number of items	Instruction of an example task	Goal process in task example
<i>Access and retrieve</i>	finding, extracting and combining one or more pieces of information explicitly stated in the text	6	Unit <i>Sports Club</i> : Which sports club offers the cheapest monthly rates for 15-year-olds?	searching four websites to identify a match with a single specified criterion
<i>Integrate and interpret</i>	inferring on the basis of implicit assumptions, relations, or implications within the text to show a holistic understanding of the text	7	Unit <i>Language Learning</i> : What kind of service does language-learning.com provide for learners?	making inferences from text information on the function of a website
<i>Reflect and evaluate</i>	drawing upon one's own knowledge and experiences, and relating them to text content and form	3	Unit <i>Language Learning</i> : Look at "My Messages". Do you think Rafael should take up the VocabTrainer suggestion? Write Yes or No and give a reason for your answer.	evaluating the credibility and utility of an advertisement through the use of contextual information
<i>Complex</i>	providing reading tasks that are as realistic as possible (i.e. encompassing features of all the former aspects)	3	Unit <i>Sports Club</i> : Which sports club would suit Liz and Anna best? Write the name of the sports club and give two reasons for your answer.	(1) locating descriptions in several websites by following a series of links, (2) comparing a series of descriptions with a set of requirements retrievable from the e-mail exchange, (3) integrating information from several websites and forming an opinion consistent with the requirements stated in the e-mail exchange

Note. Note that these reading aspects are not intended to be mutually exclusive but emphasize particular ways of text processing. Information about the items is derived from the PISA coding guidelines.

MEMORY UPDATING IN READING DIGITAL TEXT

Table 2

Descriptive statistics for the individual independent variables and their bivariate correlations (Pearson's r)

Independent variable	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	1	2	3	4	5
1. memory updating	0.00	0.97	-1.68	2.15	-				
2. linear reading	-0.03	0.82	-1.92	1.84	.50***	-			
3. target nodes visited	2.65	1.05	0.47	5.44	.48***	.58***	-		
4. target node visits	1.57	0.42	0.00	2.33	.37***	.46***	.82***	-	
5. irrelevant nodes visited	0.89	0.76	0.47	4.22	.17**	.12*	.31***	.49***	-
6. irrelevant node visits	0.62	0.43	0.00	2.11	.13*	.05	.25***	.46***	.94***

Note. * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 3

Results of the model containing the interaction of memory updating and the reading aspects

Fixed effects	Estimate	SE	z	p	
linear reading	0.66	0.08	8.26	<.001	***
memory updating (MU) ¹	0.39	0.10	3.80	<.001	***
MU: integrate and interpret	-0.14	0.11	-1.34	.179	
MU: reflect and evaluate	-0.42	0.13	-3.30	<.001	***
MU: complex	-0.12	0.13	-0.93	.355	

Notes. ¹ Effect of memory updating for *access and retrieve* items. * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 4

Results of the model containing the interaction of memory updating with the number of target nodes and the number of irrelevant nodes

Fixed effects	Estimate	SE	z	p	
linear reading	0.59	0.08	7.40	<.001	***
memory updating (MU)	0.26	0.07	3.86	<.001	***
MU : No. of target nodes	0.05	0.05	0.93	.177	
MU : No. of irrelevant nodes	0.07	0.06	1.16	.123	

Note. * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 5

Results of the models containing the interaction of memory updating with the number of target and irrelevant nodes visited (left model), and the number of target and irrelevant node visits (right model)

Fixed effects	Model with unique visit indicators					Model with visit-revisit indicators				
	Est	SE	z	p		Est	SE	z	p	
linear reading	0.21	0.06	3.50	<.001	***	0.45	0.07	6.06	<.001	***
memory updating (MU)	0.12	0.05	2.34	.009	**	0.19	0.06	3.00	.001	**
visits on target nodes	0.98	0.07	14.58	<.001	***	0.52	0.07	7.25	<.001	***
visits on irrelevant nodes	-0.03	0.05	-0.70	.242		-0.06	0.06	-1.01	.156	
MU : visits on target nodes	0.11	0.05	2.17	.015	*	-0.06	0.06	-1.04	.148	
MU : visits on irrelevant nodes	0.07	0.05	1.41	.080		0.06	0.06	0.99	.160	

Note. * $p < .05$; ** $p < .01$; *** $p < .001$.