Do a simultaneous rather than sequential presentation of multiple documents and the possibility to highlight text foster multiple document comprehension?

The roles of reading interactions in the effects of document presentation and text-highlighting.

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Caroline Leroy, M.Sc.

aus Karlsruhe

Tübingen

2021

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	27.07.2021
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter:	Prof. Dr. Peter Gerjets
2. Berichterstatter:	Prof. Dr. Martin Butz

Acknowledgements

Of course, this work would not have been possible without the many supports I received throughout the almost four years of my dissertation. First and foremost, I want to thank my supervisor, Yvonne Kammerer, for all the time and effort she put into this work. Yvonne, your never tiring support by answering and discussing the myriad of questions regarding methodology, analyses, and structure of papers, your detailed work in reviewing my doings, and the countless iterations through paper manuscripts can not be thanked for enough. Thanks to your hard and so much appreciated work, my stubborn attempts in investigating the role of reading interactions in the respective effects have taken structure and could finally ‘be put out into the world to see’ and contribute to the research field of multiple document comprehension. In relation to the manuscripts, I also especially want to thank Peter Gerjets and Uwe Oestermeier for their very detailed reviews, regardless of how short of a notice they were given. Other than the manuscripts, also, the support you two gave by guiding my work (may it be the WCT project or the dissertation) needs special mention.

Furthermore, I want to shout out a big thank you to everybody (else) involved in how welcomed and integrated I felt at IWM from the very beginning. Steffen and Franz, sharing our trampoline-equipped office and exchanging statistical, programming, and worldly wisdoms with you was educational and joyful at the same time. Everybody in the working group, thank you for making me feel safe in sharing my thoughts concerning my work, for making me feel like a valuable part of the team by reaching out to me concerning your work (e.g., “Hey Caro, you’re a cognitive scientist, are you?”), and by making the working atmosphere as pleasant as it was. Also, I want to especially mention my two co-working PhD speakers, Leonie and Lotte, who have supported me not only in attempting to improve the conditions for our fellow PhDs.

Naturally, I also want to thank everybody who contributed to this work by being a part of my “work-LIFE-balance” throughout the time of my dissertation. All the IWM and WCT PhDs whom I shared valuable coffee breaks and other moments with, inside and outside of work, everybody in our pandemic-initiated online PhD-writing-sessions, who have kept my spirits high, and my body stretched, my Spikeballers, and other after-work-companions. Last but not least, my beloved friends and family members who found countless ways of support even throughout the pandemic.

All of you: THANK YOU FOR YOUR SUPPORT!

Table of Contents

SUMMARY VI

ZUSAMMENFASSUNG VII

1. INTRODUCTION 1

1.1 Theoretical Models of Multiple Document Comprehension 3

1.1.1 The Documents Model: Mental Representation of Multiple Documents and Assessment of Multiple Document Comprehension 4

1.1.2 MD-TRACE Model 5

1.1.3 Extensions of MD-TRACE that Emphasize Reading Actions 8

1.2 The Relation Between Strategical Task Processing and Multiple Document Comprehension 10

1.3 Characteristics of a Digital Reading Environment to Foster Multiple Document Comprehension 12

1.3.1 Simultaneous vs. Sequential Document Presentation 12

1.3.2 Text-Highlighting Tool 14

1.4 Reading Interactions with Documents Reflecting Strategical Processing 26

1.4.1 Grouping Documents as Specific Form of Re-Organization 26

1.4.2 Re-Reading Documents 28

2. SUMMARY AND OVERVIEW OF STUDIES 31

2.1 Research Questions and Hypotheses 31

2.1.1 Additional Contributions of Studies 2 and 3 Based on Findings in Study 1. 37

2.2 Methodological Similarities and Differences Across Studies 38

2.2.1 Documents 38

2.2.2 Document Presentation and Interaction Possibilities 39

2.2.3 Control Measures and Assessment of Multiple Document Comprehension 41

3. *STUDY I*: SIMULTANEOUS PRESENTATION OF MULTIPLE DOCUMENTS AND TEXT-HIGHLIGHTING: ONLINE INTEGRATIVE PROCESSES AND OFFLINE INTEGRATED UNDERSTANDING 44

Abstract 45

3.1 Introduction 46

3.1.1 The role of characteristics of the reading environment for integrated understanding 47

3.1.2 The relation between online integrative processes and integrated understanding 49

3.1.3 Present study 50

3.2 Method 52

3.2.1 Participants and experimental design 52

3.2.2 Documents 52

3.2.3 Documents presentation and text-highlighting 53

3.2.4 Dependent measures 53

3.2.5 Procedure 55

3.3 Results 55

3.3.1 Comparability of experimental conditions regarding control variables 55

3.3.2 Task processing and (implicit) relevance assessments 56

3.3.3 Effects of documents presentation and possibility of text-highlighting on intertextual connections 56

3.3.4 Relations of intertextual connections to spatially organizing documents and number of revisits 57

3.4 Discussion 59

3.5 Supplementary Material 63

4. STUDY II: INVESTIGATING THE ROLES OF DOCUMENT PRESENTATION AND READING INTERACTIONS ON DIFFERENT ASPECTS OF MULTIPLE DOCUMENT COMPREHENSION. 73

Abstract 74

4.1 Introduction 75

4.2 Theoretical and Empirical Background 76

4.2.1 Readers' Mental Representation of Multiple Documents 76

4.2.2 Effects of a Simultaneous Presentation of Documents on Multiple Document Comprehension 78

4.2.3 The Role of Reading Interactions in Forming an Understanding of Multiple Documents 80

4.2.4 Present study 82

4.3 Method 83

4.3.1 Participants and Experimental Design 83

4.3.2 Task and Documents 83

4.3.3 Documents Presentation 84

4.3.4 Multiple Document Comprehension Measures 85

4.3.5 Reading Interactions 87

4.3.6 Control Variables 88

4.3.7 Procedure 89

4.4. Results 89

4.4.1 Differences Regarding Control Variables and Multiple Document Comprehension (RQ1) Between Experimental Conditions 90

4.4.2 The Role of Grouping Documents During Reading for Multiple Document Comprehension (RQ2) 91

4.4.3 The Role of Number of Document Revisits for Multiple Document Comprehension (RQ3) 93

4.5 Discussion 94

4.5.1 Multiple Document Comprehension When Reading Multiple Documents in a Reading Environment (not) Enabling Simultaneous Presentation 95

4.5.2 The Role of Grouping Documents for Multiple Document Comprehension in a Reading Environment Enabling a Simultaneous Presentation of Documents .. 96

4.5.3 The Role of Document Revisits for Multiple Document Comprehension in a Reading Environment Imposing a Sequential Presentation of Documents 97

4.5.4 Conclusion 97

4.4.5 Limitations and Future Work 98

5. *STUDY III*: READING MULTIPLE DOCUMENTS ON A HEALTH-RELATED ISSUE: THE ROLES OF A TEXT-HIGHLIGHTING TOOL AND RE-READING BEHAVIOR ON INTEGRATED UNDERSTANDING 100

Abstract 101
5.1 Introduction 102
5.2 Theoretical and Empirical Background 103
5.2.1 Readers' Integrated Understanding of Multiple Documents 103
5.2.2 Mechanisms of Text-Highlighting 104
5.2.3 The Effects of Text-Highlighting on Integrated Understanding 105
5.2.4 Present Study 107
5.3. Method 109
5.3.1 Participants and Experimental Design 109
5.3.2 Task and Documents 109
5.3.3 Documents Presentation 111
5.3.4 Measures 113
5.3.5 Procedure 119
5.4. Results 120
5.4.1 Comparability of Conditions 120
5.4.2 Number of Fixations on Documents During Short Re-Accesses 123
5.4.3 Interaction Effects Between the Possibility to Highlight Text and the Number of Re-Readings on Integrated Understanding Measures (RQ1) 123
5.4.4 Effect of the Possibility to Highlight Text on Initial Reading and Re-Reading Behaviour (RQ2) 126
5.5 Discussion 130
5.5.1 The Moderating Role of Re-Reading in the Effect of the Text-Highlighting Tool on Intertextual Integration and the Role of Re-Reading in Source-Content Integration (RQ1) 131
5.5.2 Differences in (Re-)Reading Behaviour With Versus Without Text-Highlighting (RQ2) 133
5.5.4 Limitations and Future Work 134
5.5.5 Conclusion 135
5. 6 Supplementary Material 137

6. GENERAL DISCUSSION 147

6.1 The Effect of Document Presentation on Multiple Document Comprehension (RQ1), and the Role of Readers' Spontaneous Engagement in Grouping Documents Therein (RQ2) 148

6.2 The Effect of the Possibility to Highlight Text on Multiple Document Comprehension (RQ3), and the Role of Readers' Engagement in Re-Reading Documents Therein (RQ4) 153

6.2.1 Readers' Visual Attention to Source Information Versus Text During Re-Reading. Additional Analyses of Eye-Tracking Data in Study 3 156

6.2.2 Is the Definition of Revisits in Studies 1 and 2 Valid? Insights from Eye-Tracking in Study 3 160

6.3 Discussion of Discrepant Results Across Studies 161

6.5 Theoretical and Practical Implications.. 165

6.6 Limitations, Strengths, and Future Work 167

6.7 Conclusion 170

REFERENCES 172

LIST OF TABLES 194

LIST OF FIGURES 198

Summary

Gaining a comprehensive understanding of a complex subject matter requires readers to consult multiple documents, since single documents might only provide one-sided information. These documents might furthermore provide complementary or conflicting information, which requires readers to compare, contrast, evaluate, and integrate information across documents. As previous work has shown, however, this is a challenging task even for advanced readers. Research in the field of multiple document comprehension therefore aims to identify tools to support readers in comparing, contrasting, and integrating information across documents, as well as to understand what reading processes (i.e., cognitive or behavioral) are positively related to multiple document comprehension.

The present dissertation aimed to contribute to this line of research by (a) investigating the effects of two characteristics of a digital reading environment on readers' comprehension of multiple, partly conflicting documents, and (b) by investigating reading interactions (i.e., interactions with the documents during reading) assumed to reflect cross-document information comparison as potential rationales for the hypothesized effects of the reading environment. Specifically, the two characteristics of a reading environment examined were whether it enabled a simultaneous rather than sequential presentation of documents, and whether it enabled text-highlighting. As reading interactions, whether participants had grouped the partly conflicting documents during reading as well as the number of revisits to documents were assessed. Of note, both of the characteristics of a reading environment have previously only scarcely been investigated in relation to multiple document comprehension.

Across the three studies conducted as part of this dissertation, findings showed mixed results regarding the main effects of document presentation and text-highlighting on multiple document comprehension – possibly due to methodological differences across studies. However, findings consistently indicate important roles of both reading interactions examined for (specific aspects of) multiple document comprehension, which were partly dependent on the reading environment. For instance, findings of the present work suggest that the overall inconclusive findings regarding the effectiveness of text-highlighting might be due to differences in participants' revisiting behavior across studies.

Overall, findings are supportive of the assumption of a complex interplay between reading environment, reading (inter-)actions and multiple document comprehension which was made by a theoretical model in the field.

Zusammenfassung

Um ein umfassendes Verständnis eines komplexen Themas zu erlangen, müssen multiple Dokumente gelesen werden, da einzelne Dokumente möglicherweise nur einseitige Informationen bereitstellen. Da Dokumente darüber hinaus ergänzende oder widersprüchliche Informationen beinhalten können, müssen Lesende die Informationen in den verschiedenen Dokumenten vergleichen, kontrastieren, bewerten und integrieren. Frühere Arbeiten haben jedoch gezeigt, dass dies selbst für fortgeschrittene Leser eine anspruchsvolle Aufgabe ist. Die Forschung auf dem Gebiet des Verstehens multipler Dokumente versucht daher, Hilfsmittel zu identifizieren, die Lesende beim Vergleichen, Kontrastieren und Integrieren von Informationen multipler Dokumente unterstützen, sowie zu verstehen, welche Leseprozesse (d.h. kognitive oder verhaltensbezogene) positiv mit dem Verstehen multipler Dokumente zusammenhängen.

Die vorliegende Dissertation leistet einen Beitrag zu dieser Forschungsrichtung, indem sie (a) die Auswirkungen von zwei Eigenschaften einer digitalen Leseumgebung auf das Verständnis multipler, teils widersprüchlicher Dokumente untersuchte und (b) Leseinteraktionen mit den Dokumenten, die dokumentenübergreifenden Informationsvergleich widerspiegeln, als potenzielle Ursache für die angenommenen Effekte der Leseumgebung untersuchte. Die beiden untersuchten Charakteristika einer Leseumgebung – beide in bisheriger Forschung zum Verstehen multipler Dokumente nur wenig untersucht – waren dabei, ob eine simultane versus sequentielle Ansicht der Dokumente oder Text-Highlighting ermöglicht wird Als Leseinteraktionen wurde erhoben, ob die teils widersprüchlichen Dokumente während dem Lesen gruppiert wurden, und wie oft Dokumente nochmals geöffnet wurden.

Die Ergebnisse der drei im Rahmen dieser Dissertation durchgeführt wurden Studien sind hinsichtlich der Haupteffekte von Dokumentenpräsentation und Text-Highlighting auf das Verstehen mehrerer Dokumente uneindeutig - möglicherweise aufgrund methodischer Unterschiede zwischen den Studien. Die Befunde deuten jedoch durchgängig auf wichtige Rollen der beiden untersuchten Leseinteraktionen für das Verständniss multipler Dokumente hin, die teilweise von der Leseumgebung abhängig waren. Beispielsweise deuten die Ergebnisse der vorliegenden Arbeit darauf hin, dass die insgesamt nicht schlüssigen Befunde früherer Forschung zur Effektivität von Text-Highlighting auf Unterschiede der Häufigkeit, mit der Dokumente erneuten betrachtet wurden, zwischen den Studien zurückzuführen sein könnten.

Insgesamt stützen die Befunde die Annahme eines komplexen Zusammenspiels von Leseumgebung, Leseinteraktionen und dem Verstehen multipler Dokumente, die von einem theoretischen Modell im Feld aufgestellt wurde.

1. Introduction

What is known as the *new literacies*, that is, the new types of literacies that have emerged through new technologies such as the Internet (Peterson et al., 2010), makes a vast amount of information easily available to information seekers. For example, the Internet is commonly used to inform oneself about health-related issues (AlGhamdi & Moussa, 2012) such as the health effects of ultraviolet (UV) radiation. Especially for such rather complex subject matters, the opulence of documents available necessitates information seekers to read multiple documents, since single documents may provide only general, incomplete, or one-sided information. However, since different documents might provide conflicting or complementary information, this poses additional demands on readers compared to when reading single documents (Mahlow et al., 2020), namely to compare, evaluate, and integrate information across documents (see also Britt & Rouet, 2012; Primor & Katzir, 2018; Wineburg, 1991). Correspondingly, previous work has found that readers often struggle with multiple document reading tasks and, as a result of not sufficiently comparing and integrating information across documents, often end up with a non-comprehensive, or even one-sided understanding of the issue at hand (e.g., Bråten, Ferguson, et al., 2014; Bråten & Strømsø, 2011; Britt & Aglinskas, 2002; Ferguson et al., 2012; Kiili & Leu, 2019; List, Du, et al., 2019; List & Du, 2021; Wiley et al., 2009). Determining how readers can be supported in building a comprehensive understanding of multiple documents, for instance by augmenting digital reading environments with tools to support cross-document information integration (i.e., support tools; cf. Barzilai et al., 2018), has thus become an objective of great interest in the area of multiple document comprehension (e.g., Britt et al., 2004; Delgado et al., 2020; Haber et al., 2014; Kobayashi, 2009; Lombard et al., 2021; Margolin et al., 2013; Olive et al., 2008; Salmerón et al., 2018a, 2009, 2010; Salmerón & Llorens, 2019).

Two characteristics of a reading environment that have previously been assumed to support readers' text comprehension are whether it allows for documents to be presented simultaneously rather than sequentially, and whether it allows for text to be highlighted. However, previous work in the area of multiple document comprehension investigating the effect of document presentation has used an only partly simultaneous presentation of multiple documents (Lombard et al., 2021; Olive et al., 2008; Wiley, 2001), or used only two documents (Kobayashi, 2016) as reading material. By using a large display (i.e., a multi-touch table), Studies 1 and 2 of the present dissertation were the first to investigate the effect of a reading environment enabling a *fully* simultaneous presentation of all (i.e., 6 and 5) documents provided on readers' multiple document comprehension. Furthermore, previous work investigating the effect of text-highlighting on text comprehension overall, or multiple document comprehension in specific, yielded inconclusive findings (for reviews, see Dunlosky et al., 2013; Miyatsu et al., 2018) – especially when considering research that found positive effect of text-highlighting only for printed, but not for digital documents (e.g., Ben-Yehudah & Eshet-Alkalai, 2018, 2020). As will be argued in Section 1.3.2, this might be due to the fact that highlighting text digitally with a computer mouse is more effortful than highlighting printed text with a pen (Goodwin et al., 2020). Therefore, the present work implemented touch-based digital text-highlighting, which can be assumed to be more similar to highlighting text with a pen. As such, Studies 1 and 3 of the present work were the first to investigate the effect of a touch-based digital highlighting tool on readers' multiple document comprehension. Furthermore, previous work investigating the effects of document presentation and text-highlighting on multiple document comprehension have primarily assessed readers' *qualitative* comprehension rather than specific multiple document comprehension measures (cf. Section 1.1.1). Whether these characteristics of a reading environment affect different aspects of readers' multiple document comprehension differentially, thus, remains an open question from previous work, but was addressed as a first major goal in the studies of the present dissertation.

The second major goal of the present dissertation was to examine potential rationales for the effects of document presentation and text-highlighting in more detail, which have not been investigated in previous work. That is, since comparing, re-evaluating and integrating information across documents is essential for multiple document comprehension (Britt & Rouet, 2012; Mahlow et al., 2020; Primor & Katzir, 2018; Wineburg, 1991), potential effects of a reading environment on multiple document comprehension might be due to differences in reading behaviors that reflect these processes. In order to investigate this assumption, one kind of readers' interaction with documents that can be assumed to reflect the degree of their cross-document information comparison and re-evaluation was assessed in each experimental condition (see Sections 1.4.1 and 1.4.2). Specifically, it was investigated whether *systematically re-arranging* the partly conflicting documents according to their overall stance (i.e., grouping them) during reading would play a role in the effect of document presentation on multiple document comprehension. This was assumed based on previous observatory studies suggesting that readers of printed documents commonly spread out documents (Haber et al., 2014; O'Hara et al., 2002; O'Hara & Sellen, 1997) and re-arrange them systematically (e.g., Bi & Balakrishnan, 2009; O'Hara et al., 2002), and based on previous findings indicating that having documents presented side by side can support cross-document information comparison (Olive et al., 2008). Furthermore, it was investigated whether participants' *re-reading* behavior played a role in the potential effect of text-highlighting on multiple document comprehension. This was assumed based on previous work suggesting that highlights can save readers time and effort in re-processing potentially relevant pieces of information (Yeari et al., 2017). Of note, when readers' engagement in re-reading qualifies the effect of text-highlighting, this might furthermore be one potential reason for the inconclusive findings in previous work on the effect of text-highlighting on text comprehension.

As such, the three studies conducted as part of the present dissertation contribute to previous research by investigating effects of the aforementioned two characteristics of the reading environment on different measures of readers' multiple document comprehension (rather than their overall quality of multiple document comprehension), and by especially taking into account the respective role of reading interactions therein. In the remainder of the introduction, first, theoretical models of multiple document comprehension are outlined in order to motivate both, the multiple document comprehension measures and reading interactions assessed. Subsequently, the importance of comparison and re-evaluation processes during reading, as well as previous work regarding document presentation, text-highlighting, and reading interactions examined are outlined.

1.1 Theoretical Models of Multiple Document Comprehension

Theoretical models of readers' multiple document comprehension build on Kintsch's Construction-Integration model (1988, 1998), which describes the mental representation readers of single texts construct. Namely, the mental representation of single texts comprises the *textbase*, which represents the meaning of the text without any additional information not stated in the text, and the *situation model*, which represents a reader's interpretation of the information built by drawing inferences between parts of the text and their prior knowledge. However, reading multiple documents on the same topic is more complex in that different documents are usually written by different authors covering differing perspectives, and potentially pursuing different intentions (Bråten, Anmarkrud, et al., 2014). Hence, multiple documents on the same subject matter can support, complement, or contradict each other (in parts), which, in return, makes a coherent mental representation of multiple documents more complex than that of single documents (see also Mahlow et al., 2020). Of note, in order to capture realistic multiple documents reading tasks, the document set used in multiple document comprehension research is typically partly conflicting (see Primor & Katzir, 2018) – as was in the three studies of the present dissertation.

In the following, one theoretical model that describes the mental representation of multiple documents as well as three theoretical models that further describe the processes involved in reading multiple documents are outlined.

1.1.1 The Documents Model: Mental Representation of Multiple Documents and Assessment of Multiple Document Comprehension

As the first theoretical model on multiple document comprehension posited, the Documents Model Framework (DMF; Perfetti et al., 1999) describes the mental representation that readers of multiple documents (ideally) construct – the *documents model*. In comparison to the mental representation of single documents, the documents model comprises two additional layers, namely the *intertext model* and the *integrated mental model* (or, how it was initially called in reference to Kintsch's situation model, the s*ituations model*, cf. Britt & Rouet, 2012). The intertext model contains one *document node* for each document, holding, for instance, the main ideas provided by a document, information about the document itself such as its type (e.g., newspaper article, blog post), and information about the source (i.e., the information provider) such as its name (e.g., Frankfurter Allgemeine Zeitung or the name of a journalist). The *integrated mental model* is a reader's global understanding, or interpretation, of the issue addressed across documents. A fully integrated mental model would contain (a) links between each source and the respective information provided by it such that a reader would be able to distinctively reflect upon *who said what* (i.e., *source-content links*), (b) links between document nodes such that a reader would be able to report how documents as whole are related (i.e., whether they complement, support, or contradict each other; i.e., *source-source links*), and (c) links between specific pieces of information such that readers would be able to report how they relate to one another (i.e., *content-content links*). Figure 1.1 (top) schematically represents the documents model.

By the different components of a documents model described, the DMF also provides a basis for assessing the quality of readers' comprehension of multiple documents. That is, readers' understanding of multiple documents can be assessed by tasks that address the different contents and links represented in a reader's documents model (for reviews of multiple document comprehension measures used in previous research, see Barzilai et al., 2018; Primor & Katzir, 2018). Table 1.1 provides examples of tasks and measures used in previous research to assess readers' multiple document comprehension.

1.1.2 MD-TRACE Model

As an extension of the DMF, the MD-TRACE (Multiple Documents Task-based Relevance Assessment and Content Extraction) model (Britt & Rouet, 2012) furthermore considers that readers are active entities with internal resources such as their reading and self-regulation skills, prior topic knowledge, topic interest, and working memory capacity, to whom external resources such as the reading task, the documents, and tools in the environment are available. Given the reading task and a reader's understanding of it, according to the MD-TRACE model, a reader will iteratively go through five processing steps and will have to make several decisions along the reading process (see Figure 1.1). That is, in a first step, a reader will build an understanding of the task itself and define a reading goal. In a second step, the reader will examine the further information need on the basis of prior knowledge (and the previously read information). When deciding to read further, in a third step, the reader has to select, process, and integrate the information just read in the mental representation – that is, the mental model will be updated for each newly income information that is evaluated as task-relevant. In the fourth step, the task product is updated, and, in a fifth step, its task fulfillment is evaluated. If the task product does not meet the task goal, the reader will continue with this process until satisfied with the task product.

Figure 1.1

Schematic illustration of the documents model (oval with grey background on the top; Perfetti et al., 1999) a reader of two documents might construct, and the iterative reading process suggested by the MD-TRACE model (bottom; Britt & Rouet, 2012). Documents Model: circles with white background reflect one source's information represented in the document node; circles with black background ("i") reflect one piece of information (or the reader's interpretation of it) provided by the respective document it is linked to; arrows reflect source-content, source-source, and content-content links, respectively. MD-TRACE model: Rhombuses reflect decision steps; arrows reflect possible further steps after a decision has been made, with solid lines reflecting the "yes" path and dashed lines reflecting the "no" path. Please note that this representation is not necessarily exhaustive, but focuses on the aspects of the respective models that are relevant for the purposes of the present dissertation.

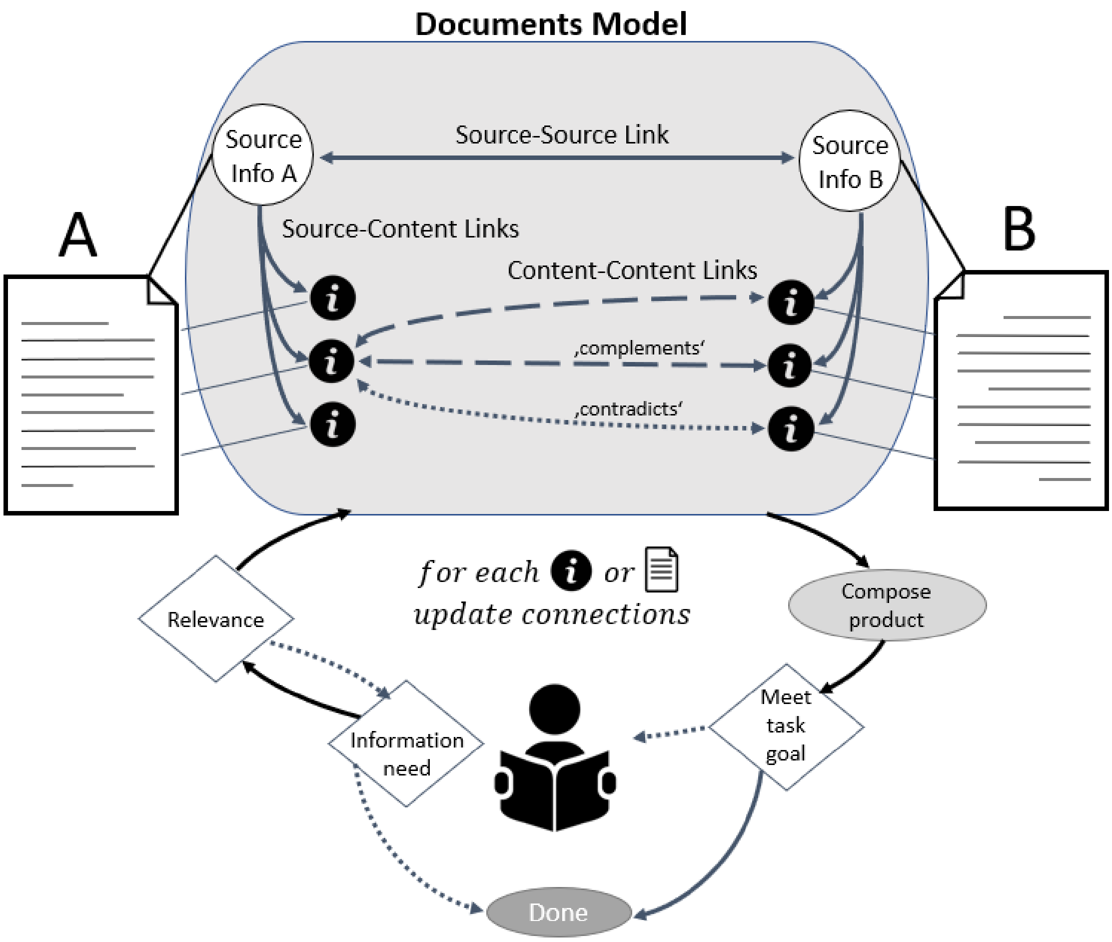


Table 1.1

Examples of tasks and measures used to assess multiple document comprehension, as well as the component of the documents model that the respective task addresses, and examples of previous research that used these tasks. Measures used in the studies of the present dissertation are printed in bold.

Component of the documents model	Task	Measure	Previous research using the respective measure
Source information	Source names recall	**# Sources correctly recalled**	e.g., Braasch et al. (2012); Rouet et al. (2016)
	Argumentative essay	**# Source citations**	e.g., Bråten, Ferguson, et al. (2014); Kammerer, Kalbfell, et al. (2016); Salmerón et al. (2018a)
Source-content links	Source-content mapping task	**# Correct mappings**	e.g., Delgado et al. (2020); Kammerer, Meier, et al. (2016); Stang Lund et al. (2019); Strømsø et al. (2010)
	Argumentative essay	# Source citations with content provided by that source	e.g., Bråten, Ferguson, et al. (2014); Britt & Aglinskas (2002); List (2019)
Content-content links (cross-document)	Intertextual inference verification task	# Correctly answered questions	e.g., Salmerón et al. (2010); Strømsø et al. (2010); Wiley (2001)
	Argumentative essay	**# Intertextual connections**	e.g., Kobayashi (2009); Linderholm & Van den Broek (2002); List (2019); Salmerón et al. (2018a)
Overall quality of integration	Argumentative essay	Overall level of integration (cf. (Reznitskaya et al., 2009)	e.g., Anmarkrud et al. (2013); Bråten, Ferguson, Strømsø, et al. (2013); List et al. (in press)

1.1.3 Extensions of MD-TRACE that Emphasize Reading Actions

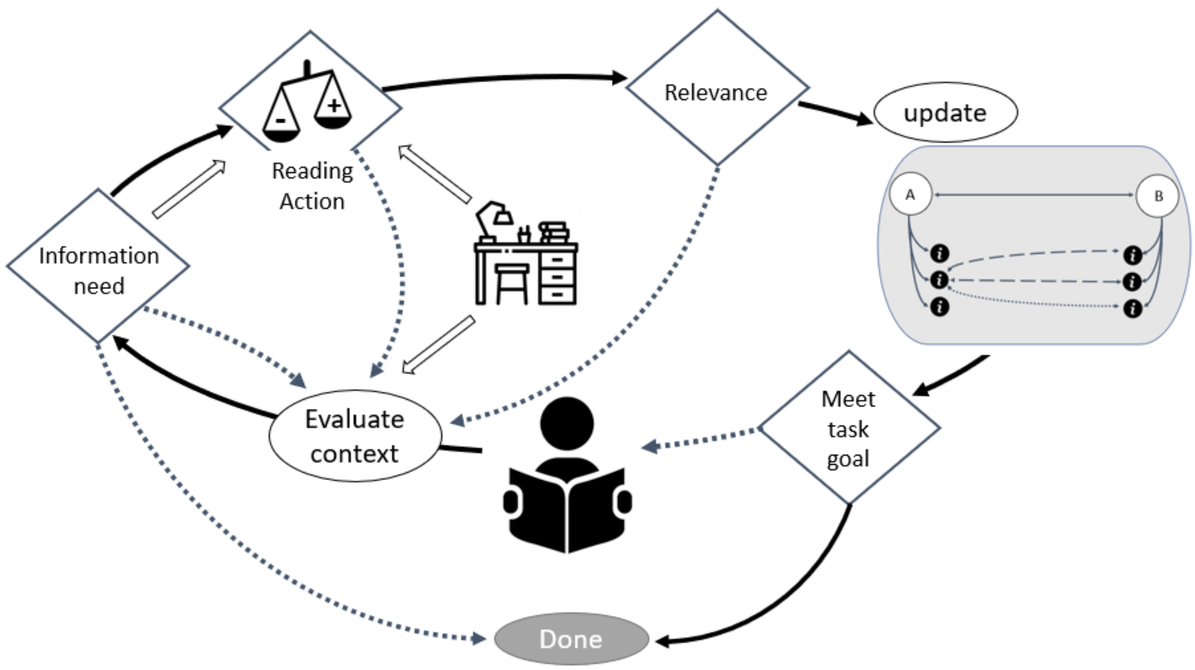
Two theoretical models that extend the DMF by, amongst others, emphasizing reading actions taken, are the Cognitive Affective Engagement Model (CAEM; List & Alexander, 2017a) and the REading as problem SOLVing model (RESOLV; Britt et al., 2018; Rouet et al., 2017). Both models take into account that the reading process is not always as static and recursive as described by the MD-TRACE model, but that factors such as the task and individual factors such as readers' prior topic knowledge, topic interest, or motivation can affect the reading process – and, in turn, multiple document comprehension.

Specifically, the CAEM (List & Alexander, 2017a) posits that readers' engagement with the task determines their multiple document comprehension after reading. Thereby, readers' engagement with the task will depend on several factors of the reading context and reader characteristics. Furthermore, readers' engagement with the task can be reflected in observable reading actions such as the navigation through documents provided. For instance, the CAEM states that readers who are engaged in the task might follow a nonlinear rather than a linear navigation through the documents provided, pay more attention to source information, and corroborate information to a greater extent than those who only engage in the task little. Finally, according to the CAEM, readers' task engagement – as can be reflected in their reading actions (or behaviors) – might predict multiple document comprehension.

Quite similar to the CAEM (List & Alexander, 2017a), the RESOLV model (Britt et al., 2018; Rouet et al., 2017) emphasizes the physical, social, and cognitive context readers find themselves in. Thereby, the RESOLV model (see Figure 1.2) more precisely defines the reading context as all of the physical and social factors to do with the reading task, including, for instance, the task itself, and readers' engagement (which is defined by their goals, perspectives, and motivation), as well as the document set and tools provided by the environment. More specifically, the RESOLV model further categorizes such tools provided by the reading environment (i.e., the physical part of the reading context) into *supports* and *obstacles* – which directly implies that the reading environment can have a positive or negative effect on readers' comprehension of multiple documents (see Section 1.2). Another distinction between the CAEM and RESOLV model lays in the benefit-cost assumption regarding reading actions that is posited by the RESOLV model. Specifically, the benefit-cost assumption states that before and during reading, readers evaluate the benefits (in relation to, for instance, the task, their prior knowledge, or motivation; see CAEM) and costs (in relation, for instance, to the reading environment) of taking reading actions, and only execute them when the benefit-cost ratio exceeds a certain threshold (which, again, is individual to the reader). Thereby, reading actions

Figure 1.2

Schematic representation of the RESOLV model (Britt et al., 2018; Rouet et al., 2017). Rhombuses represent decision steps which, in comparison to the MD-TRACE model, include one regarding whether the respective reading action should be taken based on the benefit-cost assessment. Both, the evaluation of the reading context and the evaluation of a reading action's costs can be affected by the reading environment (wide arrows with white filling). Furthermore, the reader's information need can influence the benefits of a reading action. As in the MD-TRACE model, solid arrows represent the "yes" path and dashed lines reflect the "no" path. Please note that this representation is not necessarily exhaustive, but focuses on the aspects of the RESOLV model that are relevant for the purposes of the present dissertation.



can either be covert mental processes such as comparing or (re-)evaluating information, or observable reading actions such as looking up more reading materials, re-organizing documents or pieces of information, or re-reading documents. For example, the benefit of the reading interaction 're-reading information in another document' might depend on the reader's previous understanding, relevance evaluation, and current memory of the information provided by that document, whereas its costs might depend on the effort it will take to re-access the information (i.e., the document or the specific portion of text) in the given reading environment. As such, the RESOLV model highlights a potential interplay between the reading environment, reading actions, and readers' resulting multiple document comprehension. The studies of the present dissertation investigated this interplay in different reading environments by assessing readers' interactions with documents (i.e., their reading interactions) and several aspects of their multiple document comprehension.

1.2 The Relation Between Strategical Task Processing and Multiple Document Comprehension

The MD-TRACE model, CAEM, and RESOLV model suggest a positive relation between readers' multiple document comprehension and relevance assessment, task engagement, and reading actions, respectively – and, therefore, with readers' task processing. In a similar fashion, other theoretical work suggests that readers' strategical (mental) processing of information is related to their multiple document comprehension (e.g., Afflerbach & Cho, 2009; Braasch & Bråten, 2017; Cho & Afflerbach, 2017; Rouet et al., 2017; Rouet & Britt, 2011; Stadtler & Bromme, 2014). More specifically, for instance, the taxonomy provided by Cho and Afflerbach (2017) suggests that, amongst other strategies, comparing, contrasting, re-reading and linking information, as well as "perceiving that multiple texts related to the same topic can provide diverse and contrasting views about the topic, complementary information about the topic, or both" (p. 120) are required for multiple document comprehension.

This positive relation between readers' strategical processing and multiple document comprehension was, for instance, suggested in (Wineburg, 1991) seminal work in which historians (i.e., experts in the field) and students (i.e., less experienced in the field) were asked to think aloud while examining eight text and three graphical documents on a historical event. An assessment of thinking-aloud protocols showed that historians had engaged in strategical comparison of information across documents and evaluation of information via source information more than students. Building on the work by Wineburg (1991), a large body of research found a positive relation between multiple document comprehension and readers' engagement in comparison, corroboration, and evaluation of information across documents during reading (Anmarkrud et al., 2014; Bråten, Ferguson, Strømsø, et al., 2013; Bråten & Strømsø, 2011; Ferguson et al., 2013; Ferguson & Bråten, 2013; Goldman, Braasch, et al., 2012; Hagen et al., 2014; Strømsø et al., 2003; Wolfe & Goldman, 2005). Such findings suggest that the positive relation between readers' task engagement or reading actions and multiple document comprehension posited by the CAEM and RESOLV model might especially be true for task engagement and reading actions related to comparison, corroboration, and (re-)evaluation of information across documents.

Moreover, previous findings imply that (cross-document) information comparison and (re-)evaluation are *strategical* processes since, for instance, readers' engagement in these processes was found to be greater with task-relevant rather than irrelevant information (e.g., Anmarkrud et al., 2014; Kaakinen & Hyönä, 2007; Matthew T. McCrudden et al., 2010; Matthew T. McCrudden & Schraw, 2007; Rouet et al., 2001) and with conflicting rather than non-conflicting information (Rinck et al., 2003; Rouet et al., 2016). Specifically, with conflicting information, such increased engagement in strategical processing has been suggested to reflect readers' attempts to re-establish coherence (cf. Braasch & Bråten, 2017; Richter & Maier, 2017; Stadtler & Bromme, 2014) through additional information linking (Ferguson & Bråten, 2013), or linking of the respective pieces of information to its source (Braasch et al., 2012; Braasch & Bråten, 2017; Ferguson & Bråten, 2013; Kammerer, Kalbfell, et al., 2016; Rouet et al., 2016; Stang Lund et al., 2019). In turn, readers have been found to spend more time processing conflicting than non-conflicting information (Braasch et al., 2012), to return to previously read conflicting information more often than to non-conflicting information (Rinck et al., 2003; Rouet et al., 2016), and to (re)turn to source information to a greater extent when information is conflicting rather than non-conflicting (Braasch et al., 2012). Accordingly, the observable measures of readers' task engagement (i.e., reading actions) such as the time spent processing information (e.g., Bråten et al., 2018; Bråten, Anmarkrud, et al., 2014; Goldhammer et al., 2014; Latini et al., 2019; List, 2019), or returns to previously read information (Ariasi & Mason, 2011; List & Alexander, 2018; Yeari et al., 2017), have been found to be positively related to multiple document comprehension.

However, despite the fact that previous work found readers to engage in strategic processing or reading actions, students have been reported to rarely pay attention to source information (Britt & Aglinskas, 2002; Wineburg, 1991) and to differ greatly in their degree of linking information across documents (Anmarkrud et al., 2014; Cerdán & Vidal-Abarca, 2008; List et al., in press; Stadtler & Bromme, 2007; Wolfe & Goldman, 2005). Therefore, empirical research is required to determine how readers can best be supported in building a comprehensive understanding of multiple document – for instance by providing tools in the reading environment that might support their cross-document information comparison.

1.3 Characteristics of a Digital Reading Environment to Foster Multiple Document Comprehension

Reading environments provide readers with tools that can affect their reading process and, in turn, their multiple document comprehension. As a first major goal, the present dissertation investigated the effects of two characteristics of a digital reading environment on multiple document comprehension. Specifically, the effect of a reading environment enabling a simultaneous as compared to imposing a sequential presentation of multiple documents and a reading environment either providing a text-highlighting tool or not were investigated. Based on previous findings, it was assumed that both can support readers' strategical processing of multiple, partly conflicting documents by supporting cross-document information comparison. Why and how these characteristics of a reading environment were assumed to support these processes as well as questions remaining open from previous work are outlined in the following two subsections.

1.3.1 Simultaneous vs. Sequential Document Presentation

When reading digitally, which is usually done on a computer screen or tablet, screen space is limited. This spatial restriction imposes a sequential presentation of multiple documents for reading. However, research in the area of human-computer interaction and multiple document comprehension suggests that a reading environment that enables a simultaneous rather than imposes a sequential presentation of multiple documents can support information integration across computer applications (Andrews et al., 2010; Ball & North, 2005; Czerwinski et al., 2003; Hutchings et al., 2004; Hutchings & Stasko, 2004; Robertson et al., 1998) and across documents (Jang et al., 2011, 2012; Jang & Schunn, 2012, 2014; Olive et al., 2008; Wiley, 2001). Overall, it is argued that a simultaneous presentation of documents reduces readers' (or users') cognitive load in integrating information across documents by reducing the effort required to (re-)access documents (Andrews et al., 2010; Bi & Balakrishnan, 2009; Haber et al., 2014; Jang et al., 2012; Jang & Schunn, 2012; O'Hara et al., 2002; O'Hara & Sellen, 1997; Takano et al., 2015), which might, in turn, support cross-document information comparison (cf. Ginns, 2006).

That a simultaneous presentation of multiple documents can reduce readers' cognitive load and effort required to re-access documents is especially indicated by one study in the area of multiple document comprehension by Olive et al. (2008). In the two experiments of that study, participants were asked to read and summarize three documents. Whereas the documents were presented in a sequential manner in a document viewer in both conditions, the notepad (for summary composition) was either displayed next to the document viewer (simultaneous condition) or alternating with the document viewer (sequential condition). Participants furthermore worked on a secondary task from which their mental load during reading and writing was assessed. In both experiments, participants in the simultaneous condition showed lower cognitive load during writing (see also Jang et al., 2011) and made more switches (see also Jang et al., 2012) between the document viewer and notepad than participants in the sequential condition. In the first experiment, in which participants were asked to write summaries in their own words, the quality of summaries was comparable across conditions. Most interestingly, however, in the second experiment, in which participants were asked to write summaries by typing 'copies' of parts of the text in the source documents, the lexical proximity between source documents and summaries was closer for participants in the simultaneous condition. These results, especially of the second study, imply that comparison processes between notepad and source documents was more frequent and more successful in the simultaneous than in the sequential condition. Hence, a simultaneous presentation of multiple documents might not only reduce the costs of switching between documents, but also support cross-document information comparison and, in turn, integration of documents.

In another study by Wiley (2001), participants were asked to read ten documents on a historical topic on a regular computer screen either in a partly simultaneous condition presenting two documents next to each other at a time (i.e., five of the documents were presented sequentially on the left and right side, respectively) or in a sequential condition with only one document presented at a time. After reading, participants had to write an argumentative essay on the topic without having the documents available. Analyses of the essays with regard to the level of integration they reflected indicated that the partly simultaneous presentation of documents can support integration processes – but only when an instruction is given on why documents are presented side by side. Hence, of note, merely providing readers with such a reading environment may not be sufficient to support them in building an integrated understanding of multiple documents. Rather, they need to actually use the opportunities provided by it for corroborating and integrating information across documents (for a more detailed discussion of this aspect, see Section 1.4.1). However, notwithstanding Wiley's (2001) finding that the partly-simultaneous presentation of documents was only beneficial when readers received an instruction on why, both the studies by Wiley (2001) and Olive et al. (2008) provide first indications for a beneficial effect of reading environments that enable a simultaneous rather than imposing a sequential presentation of documents.

On a more critical note, in another (and last) study in the field of multiple document comprehension that investigated the effect of a simultaneous as compared to a sequential presentation of multiple documents, Kobayashi (2016) did not find any difference in the degree to which participants had stated to compare and contrast pieces of information across documents during reading. However, Kobayashi used only two documents in this study, and those documents were rather short (around 70 Japanese characters each). Hence, participants in the sequential condition may have been able to compare and contrast documents from memory just as well as participants in the simultaneous condition who may have relied on memory less, but rather went back and forth between documents to do so.

In conclusion, previous work investigating the effect of a reading environment enabling a simultaneous rather than imposing a sequential presentation of multiple documents is indicative of a simultaneous presentation yielding easier re-accessing of documents, which might support comparison and integration of information across documents and might thus foster readers' multiple document comprehension. However, previous research on this matter in the area of multiple document comprehension is scarce, compared a reading environment presenting two rather than all documents simultaneously with a sequential presentation of multiple documents, and only assessed the *quality* of participants' integrated understanding rather than the several components that constitute readers' documents models. By using a multi-touch table which, in the simultaneous condition, enabled a simultaneous presentation of all documents, two studies of the present dissertation thus were the first to investigate the effect of a reading environment enabling a fully simultaneous rather than imposing a sequential presentation of multiple documents on several distinct measures of multiple document comprehension.

1.3.2 Text-Highlighting Tool

Highlighting text during reading is a commonly used reading strategy (e.g., Bell & Limber, 2010; Gurung et al., 2010; Hartwig & Dunlosky, 2012; Miyatsu et al., 2018; Peterson, 1991; Yan et al., 2014; Yik et al., 2018). Accordingly, the belief that text-highlighting supports text comprehension is pervasive. Yet, overall, research findings on its effectiveness on retention as well as on (specific measures of) multiple document comprehension are inconclusive (for reviews see Dunlosky et al., 2013; Miyatsu et al., 2018; see also Table 1.2).

Nonetheless, previous findings on the effectiveness of text-highlighting have led to the formulation of two hypotheses on why text-highlighting might support text comprehension. These hypotheses were derived from findings such as those of the seminal work by Fowler and

Barker (1974), who investigated how being able to actively highlight text versus being provided with previously highlighted text (i.e., either by an experimenter or by another participant), or reading text without highlights would affect readers' memory for text. Results showed that the likelihood to give a correct answer to a retention item for which the answer had been highlighted in the text was greater for participants who had *actively* highlighted text than for participants who had received previously highlighted text. Such benefit of active as compared to passive highlighting (see also Rickards & August, 1975; Schnell & Rocchio, 1978) led to assuming that active highlighting might be beneficial for text comprehension (or retention) since it entails an *encoding mechanism* (see below). Moreover, Fowler and Barker (1974) found that with highlights provided or actively made (i.e., irrespective of who had set the highlights), the likelihood of participants answering retention items correctly was greater when the respective answer had been highlighted in the text than when the answer had not been highlighted. Such findings (see also Crouse & Idstein, 1972; Hartley et al., 1980; Peterson, 1991; Ponce & Mayer, 2014; Yeari et al., 2017) led to the assumption that highlights might support *retrieval mechanisms* (see below). In accordance with these findings, Peterson (1991) reported that for 65% of her participants text-highlighting served both an encoding and retrieval function, for 26% it mostly served a retrieval function, and for 7%, it mostly served an encoding function.

Before (the very limited) previous research investigating the effect of text-highlighting on multiple document comprehension is outlined, first, the encoding and retrieval hypotheses are described by furthermore giving a brief overview of previous findings that support and challenge them, respectively. Table 1.2 provides a non-exhaustive overview of previous findings on the effectiveness of text-highlighting as a function of text comprehension measure assessed, by also providing additional information regarding the experimental methodology that might furthermore be relevant for the effectiveness of text-highlighting (however, the latter aspect will be discussed only partly in the following).

The encoding function of active text-highlighting. Actively highlighting text during reading requires readers to make decisions regarding the relevance of the to-be-highlighted piece of text. Such active decisions can yield deeper processing and, in turn, yield better memory of the respective pieces of information (cf. Craik & Lockhart, 1972). Other than the studies listed in Table 1.2 that found a beneficial effect of active in contrast to passive highlights, one study by Gier et al. (Gier et al., 2011) provides findings in support of this assumption. In this study, participants were asked to read inappropriately pre-highlighted text (i.e., the pre-highlighted information was not relevant for the task), and to use an electronic highlighter that only stored the highlighted information internally without participants being able to review it. In line with the assumption of the encoding function of actively highlighting text, the mere act of *pretending* to highlight text in that study was found to counteract the detrimental effect of inappropriate highlights (e.g., Lorch et al., 1995; Nist & Hogrebe, 1987; Yue et al., 2014). However, findings of another study by Gier et al. (Gier et al., 2010) somewhat challenges these findings. In that study, the detrimental effects of inappropriate highlights were only counteracted when participants could highlight text in another color than the pre-existing highlights, but not when they used their finger to 'highlight' text (i.e., not leaving an actual highlight). Furthermore challenging the assumption of an encoding function of actively highlighting text, Nist and Hogrebe (1987) found a beneficial effect of active highlighting as compared to experimenter-generated highlights for retention of cued information *only* when the experimenter-highlighted information were task-irrelevant, but not when they were task-relevant. Moreover, for overall retention (i.e., in contrast to retention of the respective highlighted information), a vast amount of research has found no beneficial effect of active as compared to passive highlighting (Annis & Davis, 1978; Ben-Yehudah & Eshet-Alkalai, 2018; Idstein & Jenkins, 1972; Kobayashi, 2009; Li et al., 2016; Peterson, 1991; Ponce & Mayer, 2014).

The retrieval function of text-highlights. In contrast to the encoding hypothesis which assumes a beneficial effect of *active* text-highlighting through the deeper processing of to-be-highlighted information, the retrieval hypothesis assumes a beneficial effect of highlights in general (i.e., may they be set by the reader or by someone else) through the fact that highlighted text stands out (e.g., Strobelt et al., 2016). This can be explained by the Gestalt principle of similarity (cf. Wertheimer, 1925): Since highlighted text has different characteristics (i.e., background color) than its surrounding, it will be perceived as distinct from its surrounding. As such, highlights (as well as other typographical cues such as when text is underlined or capitalized) direct readers' attention (Gaddy et al., 2001; Lorch, 1989; Mayer & Fiorella, 2014) and the respective highlighted text can be located easier than non-highlighted text (cf. Chi et al., 2007; Yeari et al., 2017). In line with this 'attention drawing' assumption, it has previously been shown that readers attribute more (re-)processing time and visual attention to cued than non-cued parts of the text (Lorch et al., 1995, experiment 2; Yeari et al., 2017). In turn, (passively) highlighted information might be recalled better than non-highlighted information (*von Restorff effect*; von Restorff, 1933). On a critical note, however, the increased re-processing time and visual attention to cued as compared to non-cued parts of the text have been found irrespective of the quality of cued text (i.e., regarding relevance or appropriateness; Lorch et al., 1995, experiment 2; Yeari et al., 2017). In turn, previous work also found that a low quality of self-generated (Winchell et al., 2020; Yang et al., 2020) and experimenter-generated (Gier et al., 2009; Silvers & Kreiner, 1997; So & Chan, 2009) highlights can impede text comprehension. Furthermore, some previous work found no difference in participants' retention of information when it had been highlighted versus not (Coles & Foster, 1975, experiments 1 and 2; Crouse & Idstein, 1972).

The effect of text-highlighting on multiple document comprehension. While some of the aforementioned previous work investigating the effect of text-highlighting used two or more texts, only few studies have assessed the more typical measures of multiple document comprehension described in Section 1.1 rather than overall retention or retention of cued text (see Table 1.2). First, in the study by Ben-Yehudah and Eshet-Alkalai (2018), the possibility to highlight text in 4 pages of a textbook yielded no difference in overall retention, but yielded a positive effect of highlighting on inference compared to when highlighting was not possible – however only when text was provided as print, but not when texts were provided digitally. Similar results were reported in a study by Kobayashi (2009) in which participants were either allowed to annotate (i.e., highlight, underline, or comment on) six partly conflicting texts or not. He found that the number of intertextual connections in essays composed after reading was higher for those participants who were allowed to annotate text, whereas the number of information correctly recalled from texts did not differ across groups. The findings of these two studies are especially interesting since they both reported no effect on the retention measures (thus also challenging the encoding hypothesis), but a positive effect on readers' inferential and intertextual-linking processing – yet this positive effect was restricted to reading *printed* documents. Somewhat corroborating the finding by Ben-Yehudah and Eshet-Alkalai (2018) that highlighting does not affect multiple document comprehension when documents are read digitally, Li et al. (2016) found that being able to highlight text in a hypertext environment containing 23 documents did not support readers' retention or inferential knowledge compared to when highlighting was not possible. Furthermore, in a fourth study in the area, List and Alexander (2020) let participants read four partly conflicting documents whilst highlighting information according to one of three instructions: Participants were either asked (a) to highlight information relevant for comprehension of each single text (i.e., comprehension purpose), (b) to highlight pieces of information that are related to information provided by another document (i.e., intertextual-linking purpose), or (c) to highlighting text that provides information they did not know before (i.e., meta-comprehensive purpose). Results showed that participants had actually set their highlights distinctive as per their instruction, however, the level of integration reflected in essays did not differ across instruction groups. Hence, the mere act of highlighting for an intertextual-linking purpose did not yield deeper processing of intertextual relations.

In conclusion, research findings regarding the effect of highlighting on both single text comprehension as well as on multiple document comprehension are inconclusive. Of note, however, in the digital conditions of the abovementioned studies investigating the effect of highlighting on multiple document comprehension, highlighting was initiated with a computer mouse. Compared to highlighting with a pen, this might be more effortful for readers (Goodwin et al., 2020) and might thus affect their task processing (e.g., by highlighting less or by interrupting the reading process to a higher extent) and, in turn, multiple document comprehension. Conversely, when digital text-highlighting is less effortful, such as when it is implemented via touch, it might support readers in building an understanding of multiple documents. Of note, the studies of the present dissertation that investigated the effect of text-highlighting on multiple document comprehension did not contrast digital text-highlighting via touch versus by using a computer mouse. Rather, the present work aimed to provide a first basis for whether digital text-highlighting via touch *can* foster multiple document comprehension. Moreover, as is suggested by the assumption of the retrieval function of text-highlights, highlighting might especially be beneficial for text comprehension (of multiple document comprehension) when readers engage in re-reading. However, none of the abovementioned previous studies has reported on participants' re-reading behavior. Section 1.4.2 elaborates on this assumption in more detail in order to motivate Studies 1 and 3 of the present dissertation in which the role of re-reading in the effect of text-highlighting on multiple document comprehension was investigated.

Table 1.2

Non-comprehensive summary of previous work investigating the effect of experimenter-generated (Exp.) or reader-generated (Self) text-highlighting by measure category assessed. Results regarding the beneficial effect of the respective manipulation are represented in the last three columns, whereas '0' reflects no beneficial effect, '+' reflects a beneficial effect, and '—' reflects a detrimental effect. Studies are provided in alphabetical order per measure category.

Measure category // Description		Study	Reading material	Type of cueing // other manipulations [a]	Manipulation of cueing		
					Exp. vs. none	Self vs. Exp.	Self vs. none
1	Retention overall	Annis and Davis (1978) *(see also category 3)*	1 text (1.525 words) (only immediate test reported, not 1 week delay)	Underlining // Additional group was instructed to take notes during reading			0
		Ben-Yehudah and Eshet-Alkalai (2018) *(see also category 5)*	4 pages text (total of 858 words; incl. 2 figures)	Highlighting // print vs. digital			0
		Fass and Schumacher (1978)	1 text; low reading difficulty vs. high reading difficulty (about 1.000 words)	Underlining // Also manipulated compensation for participation (**high** vs. low motivation) and whether text was easy or **difficult** to read			+ (conditions in bold; others 0)
		Fowler and Barker (1974) *(see also category 2)*	2 texts (in total about 8.000 words)	Experiment 1: Highlighting	0	0	0
				Experiment 2: Experimenter-generated cueing either Highlighting or Underlining	0		
		Hartley et al. (1980) *(see also category 2)*	1 text (about 300 words)	Underlining	0		

Hershberger (1964)	2 texts	Underlining // Also manipulated text length (about 3.500 and 2.000 words vs. 1.500 and 850 words)	0		
Idstein and Jenkins (1972)	1 text (1.200 words)	Underlining // Group without underlining was instructed to read repetitively			0
Johnson (1988) *(see also category 2)*	1 text; 20 passages (1844 words)	Underlining of most important sentence in each passage // Exp. 1: Reviewing not allowed Exp. 2: Reviewing allowed			 0 0
Kobayashi (2009) *(see also category 6)*	6 partly conflicting texts (*M* = 305 Japanese letters)	External strategies (i.e., highlighting, underlining, annotating, etc.)			0
Li et al. (2016) *(see also category 6)*	23 hypertext documents	Highlighting // In one additional condition, participants could highlight and re-organize text			0 (only HL) -- (with re-organizing)
Peterson (1991) *(see also categories 2+5)*	1 text (10.000 words)	Underlining // Group "self" either reviewed annotated text or blank text			0
Ponce and Mayer (2014) *(see also category 5)*	1 partly conflicting text (123 words)	Highlighting // Additionally, "graphical organizer" groups, with either experimenter- or self-generated	+	--	+

Ponce et al. (2018) *(see also category 3)*	1 text (200 words) including 2 conflicting positions	Highlighting // All experimental groups: read-only, active highlighting (HL), notetaking (N), graphic organizer (GO), HL + N, HL + GO			0
Rickards and August (1975) *(see also category 2)*	1 text of 16 passages (4 passages presented in block on one page)	Underlining // Group "self" was furthermore instructed to underline most vs. least important information, or to underline according to their own choosing; experimenter-generated underlines were either important or not.	0	+	+
Schnell and Rocchio (1978)	1 text (1.100 words)	Underlining // Group "self" either instructed on how to underline or not	0	+ (only with instruction; without 0)	+
Skuballa et al. (2018) *(see also category 3)*	1 biology textbook chapter	Highlighting // Exp. 1: Both groups were furthermore explicitly instructed to think about the key concepts Exp. 2: Third group without explicit instruction to think about key concepts and without being able to highlight text			-- --
Winchell et al. (2020) *(see also category 1)*	3 passages of a biology textbook presented sequentially, each for 5 min	Highlighting			+
Yeari et al. (2017) *(see also category 2)*	9 texts (between 314 and 471 words, each)	Highlighting // Experimenter-generated cues were either relevant or not	0		

		Yik et al. (2018)	1 text	Highlighting	0		
		Yue et al. (2014)	1 text (856 words)	Highlighting // Additionally, participants re-read text either immediately or after 30 min; test occurred after the 2nd reading			+ (especially without delay between readings)
2	Retention of respective cued information	Cashen and Leicht (1970)	3 texts for a General Psychology class	Underlining // Also assessed retention of text adjacent to cued – same results	+		
		Coles and Foster (1975)	1 text (1.050 words) Experiments 1 + 2: no reading strategy instruction	Underlining;	0		
			1 text (1.050 words) Experiment 3: incl. reading strategy instruction	Underlining;	+		
		Crouse and Idstein (1972)	Experiment 1: Short passages, each about 210 words	Underlining	0		
			Experiment 2: 1 text (6.000 words)	Underlining	+		
		Fowler and Barker (1974) *(see also category 1)*	2 texts (in total about 8.000 words)	Highlighting; Experiment 1	+	+	+
		Hartley et al. (1980) *(see also category 1)*	One text (about 300 words)	Underlining	+		

Johnson (1988) *(see also category 1)*	1 text; 20 passages (1844 words)	Underlining of most important sentence in each passage // Exp. 1: Reviewing not allowed Exp. 2: Reviewing allowed			 0 0
Lorch et al. (1995)	1 text (about 2.400 words; across 4 pages in experiment 1 and one sentence per screen in experiment 2)	Exp. 1: Underlining; Exp. 2: Capitalization // Experimenter-generated cues either only for relevant or also irrelevant information	+		
Nist and Hogrebe (1987)	1 text (about 2.200 words)	Underlining // Experimenter-generated cues either only for relevant or also irrelevant information		0	
Peterson (1991) *(see also categories 1+5)*	1 text (10.000 words)	Underlining // Group “self” either reviewed annotated text or blank text			+
Rickards and August (1975) *(see also category 1)*	1 text of 16 passages (4 passages presented in block on one page)	Underlining // Reviewing not allowed		+	+
Winchell et al. (2020) *(see also category 1)*	3 passages of a biology text book presented sequentially, each for 5 min	Highlighting			+
Yeari et al. (2017) *(see also category 1)*	9 texts (between 314 and 471 words, each)	Highlighting // Experimenter-generated cues either only for relevant or also irrelevant information	+		

3	Comprehension	Amer (1994) *(see also category 5)*	6 passages (length not mentioned)	Underlining // Additional group was instructed to draw concept maps during reading		+
		Annis and Davis (1978) *(see also category 1)*	1 text (1.525 words) (only immediate test, not 1week delay)	Underlining // Additional group was instructed to take notes during reading		+
		Ponce et al. (2018) *(see also category 1)*	1 text (200 words) including 2 conflicting positions	Highlighting // All experimental groups: read-only, active highlighting (HL), notetaking (N), graphic organizer (GO), HL + N, HL + GO		0
		Skuballa et al. (2018)) (measure = transfer) *(see also category 1)*	1 biology textbook chapter	Highlighting // Experiment 1: Both groups were furthermore explicitly instructed to think about the key concepts		0
				Experiment 2: Third group only read without explicit instruction to think about key concepts and without being able to highlight text.		--
4	Comprehension + inference (combined)	Stordahl and Christensen (1956)	2 texts (each about 2.600 and 3.800 words)	Underlining		0
		Shi et al. (2020)	3 texts (each between 1.200 and 1.300 words)	Underlining and bold print // Texts were printed or digital; One additional group received texts with cues and an organization map	+ (printed and digital)	

5	Inference / Integration (not specifically intertextual)	Amer (1994) *(see also category 3)*	6 passages (length not mentioned)	Underlining // Additional group was instructed to draw concept maps during reading			+ (underlining and mapping)
		Ben-Yehudah and Eshet-Alkalai (2018) *(see also category 1)*	4 pages text (total of 858 words; incl. 2 figures)	Highlighting // Documents were either print or digital			+ (only print; digital 0)
		Peterson (1991) *(see also categories 1+2)*	1 text (10.000 words)	Underlining // Group "self" either reviewed annotated text or blank text			0 / -- (review of annotated detrimental)
		Ponce and Mayer (2014) *(see also measure 1)*	1 partly conflicting text (123 words)	Highlighting // Additionally, "graphical organizer" (GO) groups, with either experimenter- or self-generated	0	0	0
6	Intertextual integration	Kobayashi (2009) *(see also category 1)*	6 partly conflicting texts (*M* = 305 Japanese letters)	External strategies (i.e., highlighting, underlining, annotating, etc.)			+
		Li et al. (2016) *(see also category 1)*	23 hypertext documents	Highlighting // In one additional condition, participants could highlight and re-organize text			0 (only HL) -- (with organizing)

[a] Please note that only results regarding the manipulation of cueing (i.e., highlighting, underlining, or capitalization, respectively) are reported in this table. Results regarding other manipulations relevant to the studies of the present dissertation (e.g., review or re-organization of the text) are provided in Section 1.4.

1.4 Reading Interactions with Documents Reflecting Strategical Processing

As outlined in Section 1.2, readers' strategical processing of multiple documents can be reflected in reading actions related to cross-document information comparison. In turn, multiple document comprehension might be related to such reading actions. Thereby, the extent to which readers engage in such reading actions can, on the one hand, reflect task engagement (see CAEM; (List & Alexander, 2017a), and, on the other hand, depend on the reading environment (see benefit-cost assumption of the RESOLV model; Britt et al., 2018; Rouet et al., 2017).

In the taxonomy of reading strategies provided by Cho and Afflerbach (2017), it was suggested that rereading and linking information, as well as "perceiving that multiple texts related to the same topic can provide diverse and contrasting views about the topic, complementary information about the topic, or both" (p. 120) are required for multiple document comprehension. In accordance with this suggestion, the two reading interactions (i.e., interactions with the documents during reading) assessed in the studies of the present dissertation were the number of revisits to documents and whether participants had grouped the partly conflicting documents according to their stance taken on the health effects of UV radiation (i.e., reflecting whether the complementary and contrasting views, respectively, had been realized). The following two subsections outline why these reading interactions can be assumed to reflect cross-document information comparison.

1.4.1 Grouping Documents as Specific Form of Re-Organization

In their much cited observative study, O'Hara and Sellen (1997) found that readers of printed documents who were asked to summarize a four paged article laid out the documents in front of them to "gain a sense of overall structure" and to "check on or to relate specific pieces of information across documents" (p. 346; see also Haber et al., 2014; O'Hara et al., 2002; Shibata et al., 2013; Takano et al., 2015). Conversely, those who read the same documents digitally (i.e., in a spatially limited reading environment) made remarks about the lack of flexibility on screen. These findings suggest that readers use a reading environment enabling flexible re-arrangement of multiple documents (i.e., might it be non-digital) not simply to spread out documents (i.e., potentially into a simultaneous presentation), but to do so in a strategical manner. Thereby, when done strategically, the resulting layout might be an externalization of readers' understanding of the relations across documents (Skuballa et al., 2018), and thus be a result of their cross-document information comparison.

Of note, such strategical re-organization of documents might also be done in a spatially restricted reading environment that imposes a sequential presentation of multiple documents. For example, when reading multiple partly conflicting documents in a reading environment that presents one document at a time but allows for re-arrangement of documents, readers might re-arrange documents taking the same stance to the left, and documents taking another stance to the right – for example when piling conflicting documents according to their overall stance on a desk. When re-accessing a document, its location, then, bears contextual information regarding the overall stance taken in that document, which can reduce a reader's cognitive load (Skuballa et al., 2018) compared to when no such re-arrangement is possible. Still, strategical re-arrangement of documents in the form of grouping them according to their overall stance might be more beneficial when the reading environment enables a simultaneous rather than imposes a sequential presentation of documents. This is because in a simultaneous presentation of multiple documents, such strategical re-arrangement of documents might furthermore reduce readers' effort in cross-document comparison and integration by having related documents (even) more easily and, most importantly, simultaneously accessible (cf. Andrews et al., 2010; Ball & North, 2005), which might further support cross-document information comparison (Ginns, 2006; Jang & Schunn, 2012; Olive et al., 2008).

Accordingly, as one form of strategical re-arrangement of documents, in the studies of the present dissertation (i.e., Studies 1 and 2), it was assessed whether participants had *grouped* documents according to their overall stance taken regarding the health effects UV radiation. Of note, this measure has not been assessed in previous work. Nonetheless, the relation between grouping documents during reading and multiple document comprehension can be assumed based on previous work examining the effect of scaffolds such as concept maps or organizational aids. In such scaffolds, information from source documents is re-organized in a structured way such that both the key concepts and the relations between them become more salient (Barzilai et al., 2018; Lee & Baylor, 2006; Salmerón & Garcia, 2012). Accordingly, previous work found a positive effect of such scaffolds for retention (Ponce et al., 2018; Ponce & Mayer, 2014; Shi et al., 2020), for overall comprehension (Amer, 1994; Ponce et al., 2018; Ponce & Mayer, 2014), for inferential knowledge (Ponce et al., 2018; Praveen & Rajan, 2013; Shi et al., 2020), and for cross-document information integration (Barzilai & Ka'adan, 2017; Renkl & Scheiter, 2017) compared to when participants read documents without such scaffolds (for a review, see Barzilai et al., 2018). On a critical note, participants in these studies were either instructed to fill in scaffolds or were provided with already filled-in scaffolds. In relation to this note, Lombard et al. (2021) found that only those participants who were explicitly *instructed* to extract and re-organize information from multiple documents wrote more integrated essays than those not able to use such scaffolding during reading. Those who could use this scaffolding during reading, but were not explicitly instructed to use it, did not write more integrated essays than those who were not able to use the scaffolding. However, Lombard et al. (2021) neither reported on the extent to which those (not) explicitly instructed actually used the possibility to re-organize information, nor did they report on whether those who used it without being instructed to still wrote more integrated essays than those who did not use it. Still, these findings (i.e., the positive effect of the re-organization tool when instructed to use it) indicate that re-organization of information *can* support multiple document comprehension.

In conclusion, notwithstanding the fact that participants of the abovementioned studies were instructed to use the respective information re-organization procedures, they suggest that interrelations can become more salient through re-organization, which, in turn, can support text comprehension. Likewise, readers who spontaneously (i.e., without being instructed to) engage in strategically re-organizing documents (Haber et al., 2014; O'Hara & Sellen, 1997; Shibata et al., 2013; Takano et al., 2015) – in the form of grouping documents according to their overall stance during reading – might also show better multiple document comprehension, and specifically better intertextual integration. This assumption was addressed in Studies 1 and 2 of the present dissertation.

1.4.2 Re-Reading Documents

Re-reading is a commonly used study strategy (Carrier, 2003; Hartwig & Dunlosky, 2012; Kornell & Bjork, 2007) that has mainly been investigated with single texts by instructing participants to re-read text or not. Overall, while research findings on the effectiveness of re-reading are somewhat inconclusive (Dunlosky et al., 2013; Miyatsu et al., 2018), two hypotheses regarding the reasons for why re-reading might foster text comprehension are discussed in the literature (see Bromage & Mayer, 1986; Dunlosky et al., 2013; Kiewra et al., 1991). The *quantitative* hypothesis states that re-reading supports retention of the respective re-read information through increased processing (i.e., re-processing) of information. This hypothesis is supported by several studies that found a positive effect on overall retention when participants were instructed to read text twice (or several times) compared to only reading it once (e.g., Karpicke & Blunt, 2011; Rawson & Kintsch, 2005; Rothkopf, 1968). However, some studies found no effect of re-reading on retention (e.g., Callender & McDaniel, 2009; Griffin et al., 2008; for reviews see Dunlosky et al., 2013; Miyatsu et al., 2018) or a beneficial effect of re-reading on comprehension rather than on retention (e.g., Bromage & Mayer, 1986; Mayer, 1983). The possibility that re-reading might rather support comprehension than overall retention is addressed in the *qualitative* hypothesis which states that readers shift their attention between the initial and following reading such that during initial reading, readers focus on the facts, whereas during re-reading, they rather focus on a conceptual organization (i.e., mentally) of the provided information (e.g., Kiewra et al., 1991). This hypothesis is supported by studies suggesting that readers allocate more attentional resources to information relevant for global understanding and for relation-building during re-reading than during initial reading (e.g., Millis et al., 1998; Millis & King, 2001; Yeari et al., 2017). In further support of the qualitative hypothesis, some studies showed that participants who were allowed to re-read also scored higher in inference questions (i.e., requiring information connection) than those who were not allowed to re-read text (e.g., Karpicke & Blunt, 2011; Rawson & Kintsch, 2005). Nonetheless, in their review, Dunlosky et al. (2013) state that overall "results have been somewhat mixed, but the evidence appears to favor the qualitative hypothesis" (p. 27).

In the context of multiple document comprehension, re-reading information in a document (often) requires re-accessing it. Especially when documents are partly conflicting, it is argued that readers might re-access documents in order to corroborate information across documents or to re-evaluate them in the light of information provided in another document (Goldman, Braasch, et al., 2012; List & Alexander, 2018; Wiley et al., 2009; Wineburg, 1991). In this regard, List and Alexander (2017a) wrote in their derivation of the CAEM that multiple document comprehension might be related to readers' document navigation (e.g., Salmerón & García, 2011; Wiley et al., 2009) in that readers who follow a more evaluative or critical approach might re-access documents more often (i.e., in order to compare and re-evaluate information across documents) than disengaged readers. For example, Goldman, Braasch, et al. (2012) fond that participants who learned more from reading (i.e., as measured in the difference between a pre- and posttest) also had a higher number of document re-accesses during reading. Furthermore, the Discrepancy-Induced Source Comprehension (D-ISC) model (Braasch & Bråten, 2017) suggests that readers return to previously read information at conflict detection in order to re-evaluate the conflicting information in the light of the sources (i.e., their potential perspectives or intentions) and to, eventually, re-establish coherence. In case the conflict is located in different documents, hence, readers would have to re-access a previously read document in order to link the information provided to its source to re-establish coherence. In line with this assumption, (List & Alexander, 2018) found that participants who had re-accessed at least one document in the reading process were more likely to include source information in their essays than participants who had not re-accessed any document (see also List, Stephens, et al., 2019). Furthermore, in a recent study by Delgado et al. (2020), participants who re-accessed pop-up windows providing information about conflicts across documents more often, showed better source-content integration than those who had re-accessed these windows less often.

However, previous work examining the relation between re-accessing documents and multiple document comprehension is scarce and only assessed single aspects of multiple document comprehension in isolation. One contribution of the present dissertation for research on the relation between re-reading (or re-accessing documents) and multiple document comprehension lays in the fact that several specific measures of multiple document comprehension were assessed at once, which provides a more differentiable understanding of the relation between re-reading and multiple document comprehension than previous work. The main contribution of the present work, however, lays in the investigation of the role of re-reading (or re-accessing documents) in the effect that text-highlighting has on readers' multiple document comprehension.

Re-reading with versus without text-highlighting. As mentioned in Section 1.3.2, re-reading might play an important role in the effect of text-highlighting on multiple document comprehension. O'Hara et al. (2002) noted that without text-highlighting, readers often have "trouble knowing where they are in a text and finding their way to desired locations within the text" (p. 273). Hence, when readers re-access documents in an attempt to re-evaluate and link information across documents, such a hurdle of re-locating information in a document might interrupt strategic re-reading and, thus, (cross-document) information corroboration. Conversely, with highlights, the presumably relevant portions of text (i.e., those that have been highlighted) can be re-located more effectively. For example, Yeari et al. (2017) found indications that readers are more likely to re-process pre-highlighted than non-highlighted text – irrespective of the relevance of the highlighted information – and concluded that highlights can "save time and effort in reprocessing the textual content" (p.2). In turn, such improved re-processing of information that is deemed relevant by readers (i.e., especially when readers are able to highlight text themselves) due to better re-location might support readers' corroboration processes.

2. Summary and Overview of Studies

As suggested by the RESOLV model (Britt et al., 2018; Rouet et al., 2017), the extent to which readers engage in reading actions in a given reading environment might take an integral role in the effect the respective reading environment has on their multiple document comprehension. The overall goal of the present dissertation was to investigate the effects of two characteristics of a digital reading environment on several aspects of readers' multiple document comprehension, and to examine the role of specific reading interactions (i.e., interactions with documents during reading) in these effects. As outlined in Section 1.3, a reading environment enabling a simultaneous as compared to imposing a sequential presentation of multiple documents as well as the provision of a text-highlighting tool are two characteristics of a reading environment that bear great potential in supporting readers in their understanding of multiple documents. However, some questions remain unanswered by previous work. The following subsection explicates the respective research questions deduced from the literature that the present dissertation aimed to answer and summarizes the literature outlined in the introduction in order to motivate the hypotheses. Table 2.1 gives a comparative overview of the studies.

2.1 Research Questions and Hypotheses

First, previous work in the area of human-computer interaction and multiple document comprehension suggests a beneficial effect of a simultaneous rather than a sequential presentation of multiple documents for cross-document information integration (Andrews et al., 2010; Ball & North, 2005; Czerwinski et al., 2003; Hutchings et al., 2004; Hutchings & Stasko, 2004; Jang et al., 2011, 2012; Jang & Schunn, 2012, 2014; Olive et al., 2008; Robertson et al., 1998; Wiley, 2001) through easier (re-)accessing of documents (Andrews et al., 2010; Bi & Balakrishnan, 2009; Haber et al., 2014; Jang et al., 2012; Jang & Schunn, 2012; O'Hara et al., 2002; O'Hara & Sellen, 1997; Takano et al., 2015). However, in the field of multiple document comprehension, the rather scarce previous work on this matter has used side-by-side views, and thus an only partial-simultaneous presentation of documents (Kobayashi, 2009; Olive et al., 2008; Wiley, 2001). Moreover, the only one of these studies that investigated this effect on readers' post-reading multiple document comprehension (i.e., in contrast to the quality of summaries composed with documents available; Olive et al., 2008) and used more than two documents (Wiley, 2001; i.e., in contrast to Kobayashi (2009) who used only two documents) assessed the *quality* of participants' argumentative essays written after reading as measure of their multiple document comprehension. Hence, the effect of a reading environment enabling a *fully* simultaneous as compared to imposing a sequential presentation of multiple documents on *specific aspects* of readers' multiple document comprehension (cf. Perfetti et al., 1999) has not been investigated in previous research. Based on findings and suggestions of previous observational and experimental work in the field of human-computer interaction (Andrews et al., 2010; Ball & North, 2005; Czerwinski et al., 2003; Hutchings et al., 2004; Hutchings & Stasko, 2004; Jang et al., 2011, 2012; Jang & Schunn, 2012, 2014; Robertson et al., 1998), the simultaneous presentation of multiple documents was expected to support cross-document information corroboration, and, thus, an effect of document presentation especially on measures related to linking information was expected.

Research question 1: Does a reading environment enabling a *fully* simultaneous rather than imposing a sequential presentation of multiple documents foster readers' multiple document comprehension? (Studies 1 and 2)

Hypothesis 1: A reading environment enabling a simultaneous as compared to imposing a sequential presentation of multiple documents fosters readers' intertextual integration. (Studies 1 and 2)

+ exploratory analysis of the potential effect of document presentation on source-content integration and memory for source names. (Study 2)

Second, when possible in the respective reading environment, readers often spontaneously engage in re-arrangement of documents (Haber et al., 2014; O'Hara et al., 2002; O'Hara & Sellen, 1997; Shibata et al., 2013; Takano et al., 2015) – and this spontaneous re-arrangement can be assumed to emerge as a reading strategy, that is, in order to have related or to-be-compared documents close to each other (O'Hara et al., 2002). The resulting spatial proximity of documents can reduce the effort required for cross-document information comparison and integration (Andrews et al., 2010; Ball & North, 2005), and thus support these cognitive processes (Ginns, 2006; Jang & Schunn, 2012; Olive et al., 2008). It might thus be that in a reading environment that enables a simultaneous presentation of documents, especially those readers who engage in strategical re-arrangement of documents show a better multiple document comprehension after reading. In relation to this assumption, previous work found that strategical re-organization of information (i.e., on the basis of pieces of information rather than on the basis of full documents) can support comprehension of the respective relations between them (Amer, 1994; Barzilai & Ka'adan, 2017; Lee & Baylor, 2006; Ponce et al., 2018; Ponce & Mayer, 2014; Praveen & Rajan, 2013; Renkl & Scheiter, 2017; Salmerón & Garcia, 2012; Shi et al., 2020). Based on these findings and the observations that readers spontaneously engage in re-arrangement of documents (O'Hara et al., 2002), it was assumed that (at least some) readers of partly conflicting documents might also re-arrange documents during reading such that groups of documents taking the same stance will be created. However, no previous study had examined readers' spontaneous strategical re-arrangement of full documents (i.e., grouping documents) during reading.

Research question 2: What is the role of readers' spontaneous strategical re-arrangement (i.e., grouping) of partly conflicting documents during reading in the effect of document presentation on multiple document comprehension? (Studies 1 and 2)

Hypothesis 2.1: A reading environment enabling simultaneous presentation of multiple documents bears greater affordance to strategically re-arrange documents than one that imposes a sequential presentation of documents. Hence, more participants in the simultaneous than in the sequential condition were expected to group documents during reading. (Study 1)

Hypothesis 2.2: Readers who group the partly conflicting documents during reading according to their overall stance taken were expected to show better intertextual integration than those who do not. (Studies 1 and 2)

+ exploratory analysis of a potential relation between readers' spontaneous grouping of documents and source-content integration, as well as their memory for source names. (Study 2)

Hypothesis 2.3: Readers' spontaneous engagement in grouping documents during reading mediates the effect of document presentation on multiple document comprehension (i.e., specifically on intertextual integration; Study 1).

Third, previous research findings regarding the effectiveness of text-highlighting is rather inconclusive (Dunlosky et al., 2013; Miyatsu et al., 2018), but suggests that it might be more beneficial for tasks that require integration of information than for retention (Ben-Yehudah & Eshet-Alkalai, 2018; Kobayashi, 2009). However, previous work investigating the effect of text-highlighting on multiple document comprehension is scarce. Moreover, Ben-Yehudah and Eshet-Alkalai (2018) found a positive effect of active text-highlighting on inference questions only with printed but not with digital documents, and Kobayashi (2009) used only printed documents. It thus especially remains an open question, whether a *digital* text-highlighting tool can support *multiple document comprehension*. Previous work suggests that highlighting digital documents with a mouse might be more difficult than highlighting paper with a pen, resulting in a different highlighting behavior (Goodwin et al., 2020; O'Hara & Sellen, 1997; Schugar et al., 2011). The present work therefore used a digital highlighting technique more similar to highlighting printed documents with a pen, namely touch-based highlighting.

Research question 3: Does the provision of a digital touch-based text-highlighting tool support multiple document comprehension? (Studies 1 and 3)

Hypothesis 3: The provision of a digital touch-based text-highlighting tool supports readers' intertextual integration. (Studies 1 and 3)

+ exploratory analysis of the effect of the provision of the text-highlighting tool on source-content integration (Study 3).

Fourth, the retrieval function of text-highlighting (see Section 1.3.2) suggests that highlighted information can be retrieved, that is, re-located, easier than non-highlighted information. In line with this hypothesis, Yeari et al. (2017) found indications for highlights to "save time and effort in reprocessing the textual content" (p.2). Deduced from these findings, it can be assumed that re-reading is more beneficial with than without text-highlighting. Likewise, since readers have been suggested to rather focus on facts during initial reading and on information-linking during re-reading (Millis et al., 1998; Millis & King, 2001; Yeari et al., 2017), text-highlighting might especially have a beneficial effect on readers' multiple document comprehension when they engage in re-reading – and this might especially be true for their intertextual integration. However, no previous study has investigated the role of readers' spontaneous re-reading behavior in the effect of text-highlighting on comprehension – yet on multiple document comprehension. Furthermore, when readers of highlighted text re-read text, it can be assumed that their re-reading process will be guided by their highlights (Shi et al., 2020; Yeari et al., 2017), resulting in a smaller proportion of text re-read with compared to without highlights. This, however, has also not been examined in previous research but was investigated in Study 3 of the present dissertation by means of eye-tracking methodology. In accordance with previous work (Goldman, Braasch, et al., 2012; List & Alexander, 2018; Wiley et al., 2009; Wineburg, 1991), the number of re-accesses to documents was assessed as measure of readers' re-reading behavior in the studies of the present dissertation.

Research question 4: What is the role of re-reading (i.e., re-accessing documents) in the effect of text-highlighting on multiple document comprehension? (Studies 1 and 3)

Hypothesis 4.1: Text-highlighting is especially beneficial for intertextual integration for readers who engage in re-reading (i.e., those with a high number of document re-accesses). (Studies 1 and 3)

Hypothesis 4.2: The positive relation between the number of re-readings and intertextual integration is greater for readers who have the possibility to highlight text than for those who do not have the possibility to highlight text. (Studies 1 and 3)

+ exploratory analysis of the role of re-reading in the effect of text-highlighting (i.e., hypotheses 4.1 and 4.2) on source-content integration (Study 3).

Hypothesis 4.3: Readers who have the possibility to highlight text re-read a smaller proportion of text than readers who do not have the possibility to highlight text. (Study 3)

Table 2.1

Summary of the three studies conducted as part of the present dissertation with respect to the characteristics of the reading environment manipulated as well as reading interactions, multiple document comprehension measures, and additional behavioral measures assessed.

Category	Specifics	Study 1 (multi-touch table; *N* = 126)	Study 2 (multi-touch table; *N* = 108)	Study 3 (touch display; *N* = 95)
Reading environment	Simultaneous vs. sequential presentation	✓	✓ (without highlighting)	
	With vs. without text-highlighting tool	✓		✓ (sequential)
Reading interaction	# Revisits to previously read documents	✓ (only sequential)		✓ (re-readings)
	Grouping documents according to overall stance	✓ (simultaneous and sequential)	✓ (only simultaneous)	
Integrated understanding	# Intertextual connections in essays	✓	✓	✓
	# source citations in essays	✓	✓	
	Source-content mapping		✓	✓
	Source names recall		✓	
Additional behavioral measures	Overall reading time	✓		✓
	Initial reading versus re-reading time			✓
	Visual distribution during re-readings			✓

2.1.1 Additional Contributions of Studies 2 and 3 Based on Findings in Study 1.

The findings of Study 1, in which both the document presentation and the possibility to highlight text were manipulated, raised some questions that were addressed in Studies 2 and 3. First, in Study 1, participants' intertextual integration was assessed as the sole measure of multiple document comprehension. While the characteristics of the reading environment investigated with respect to their effects on multiple document comprehension were assumed to specifically foster cross-document information comparison, and, in turn, readers' intertextual integration, whether other aspects of readers' multiple document comprehension were affected was left unaddressed in Study 1. Thus, in addition to intertextual integration, Study 2 investigated potential effects of document presentation on source-content integration (as well as the role of grouping documents during reading therein) and on source name recall, and Study 3 additionally investigated a potential effect of text-highlighting on source-content integration (as well as the role of participants' re-reading behavior therein). Second, as is described in more detail in Section 2.2.2, the sequential condition in Study 1 differed from the rather typical sequential reading environment of a computer screen or tablet in that documents could also be re-arranged and, at re-access, appeared in the respective position it was last moved to. In the sequential condition of Study 2, documents could not be re-arranged, thus allowing an investigation of the effect of document presentation on multiple document comprehension with a more realistic implementation of a sequential presentation of multiple documents. Third, in order to assess the number of document re-accesses, and, more specifically, in order to not count in accidental re-accesses, in the sequential conditions of Studies 1 and 2, a *revisit* was defined for each re-access (i.e., starting from the second access) of a document that lasted longer than 1 sec. However, some readers adopt a 'sampling strategy' (List & Alexander, 2017a; Reader & Payne, 2007) and first only read a small portion of a document before fully reading it. In turn, a definition of revisits as re-accesses that last at least 1 sec might be biased by some behavioral artefacts and result in an overestimation of the number of re-accesses made in order to *re*-read information in a document since the respective document might not have been fully processed during its initial access. By means of eye-tracking technology, this issue was addressed in Study 3 in that the more fine-grained measure of *re-readings* was defined as the number of revisits after the document had been *fully* read for the first time. Specifically, this was achieved by defining an initial reading of a document as including all (re-)accesses of that document until the (accumulated) fixation heatmap revealed full coverage of the text (for more detailed description, see Section 5.3.4.2.1). Eye-tracking methodology furthermore was used to gain insights into how much information intake had occurred (i.e., via the number of fixations in a document) during short re-accesses lasting between below 1 sec, between 1 and 2 sec, between 2 and 3 sec, between 3 and 4 sec, and between 4 and 5 sec, as well as to assess the proportion of text fixated during re-reading when text-highlighting was possible versus not.

2.2 Methodological Similarities and Differences Across Studies

2.2.1 Documents

The opulence of documents on any subject matter available, for example on the Internet, almost makes it inevitable that readers come across (partly) conflicting information across documents. Hence, building an understanding of how pieces of information provided by different documents are related is an integral part of multiple document comprehension (Britt & Rouet, 2012; Perfetti et al., 1999), and the document set used for multiple document comprehension research is typically partly conflicting (see Primor & Katzir, 2018). The documents used in all three studies of the present dissertation discussed the health effects of UV radiation, with two documents providing arguments for positive effects, two documents providing arguments for negative effects, and one document providing neutral facts (e.g., the types of UV radiation). Tables 3.5.1 and 5.6.1 provide a more detailed description of the documents.

Documents were translated and adapted versions of documents used in previous research (e.g., Bråten, Anmarkrud, et al., 2014; Ferguson & Bråten, 2013; Lombard et al., 2021; Strømsø et al., 2016). All documents were comparable in readability and length (see Tables 3.5.1 and 5.6.1) and had one colored source logo at the top of the page. In all three studies, prior to reading, participants were informed that they were provided with documents from the Internet that discuss the health effects of UV radiation, and that their task was to write an argumentative essay about the topic later without having the documents available.

Main difference across studies. Study 1, in which both document presentation and text-highlighting were manipulated, furthermore tested the assumption that readers would differentially interact with a task-irrelevant document than with task-relevant documents. For this purpose, a sixth, task-irrelevant document providing information on the health effects of mobile phone radiation, was provided in addition to the five documents on the topic of UV radiation that were used in all three studies (though slightly adapted for Studies 2 and 3). Specifically, based on previous work suggesting that readers allocate more attentional resources to task-relevant than task-irrelevant information (e.g., Millis et al., 1998; Millis & King, 2001; Yeari et al., 2017), it was hypothesized that compared to the task-relevant documents, the task-irrelevant document would receive less re-accesses and highlights (i.e., from participants who were able to highlight text), and that participants would sort out the task-irrelevant document by moving it away from the task-relevant documents. These hypotheses were confirmed in Study 1 of the present dissertation (see Section 3.3.2). For the following two studies, therefore, only the five task-relevant documents were used.

2.2.2 Document Presentation and Interaction Possibilities

To enable an intuitive interaction with documents (especially for text-highlighting, which has been reported to be rather cumbersome with a computer-mouse; cf. Goodwin et al., 2020; O'Hara & Sellen, 1997; Schugar et al., 2011), as well as to assess reading interactions from log files, documents were presented on a touch display in all three studies. That is, in Studies 1 and 2, in order to enable a simultaneous presentation of all documents, a multi-touch table was used. In Study 3, in which documents were only presented sequentially, a more regular-sized touch display was used so that participants would not be puzzled about the free, yet un-usable space on screen.

In the sequential presentation conditions (i.e., in Studies 1, 2, and 3), only one document was presented at a time in the horizontal center of the respective screen. Documents could be opened from a menu just below the currently opened document, and all documents initially appeared in a size slightly larger than A4, and could be resized freely through a pinching gesture. In the simultaneous presentation conditions (i.e., in Studies 1 and 2), documents were also initially presented in a size slightly larger than A4 (i.e., the same size as the initial size of documents in the sequential condition) and could be resized freely by participants.

Main differences across studies. First, in Study 1, the menu from which documents could be opened in the sequential condition contained miniature views of the documents as well as labels (A – F; see Figure 3.2). In these miniature views, the colored source logos of the documents were not readable (i.e., they were too small), yet participants might have used their color and overall structure (e.g., the length) as landmarks when attempting to re-access a specific document. The layout of the menu was thus adapted for Studies 2 and 3 in that it no longer contained the miniature views of documents but only labels (A – E; see Figure 4.1). Second, in Study 1, the order in which documents were presented was randomized for each participant. Hence, the order of documents could potentially have been blocked (i.e., presenting the positive and negative documents in direct succession, respectively) or alternating (i.e., presenting a positive document after a negative document or vice versa) for some, but most likely not for all participants. This methodological decision was made to eliminate any potential order effects, since previous work has shown that whether partly conflicting documents are presented in a blocked or alternating order can affect cross-document conflict detection and, thus, multiple document comprehension (Braasch et al., 2021; Richter & Maier, 2017). Still, for Studies 2 and 3, two document orders (with positive and negative documents presented in an alternating fashion and the neutral document presented as third) were defined and counterbalanced between participants to somewhat gain more control over participants' initial information encounter. Please note, however, that participants were not instructed to use the provided document order when reading the documents, but they could switch between documents freely. Third, in the simultaneous condition of Study 1, documents were initially presented in a grid (2 rows, 3 columns; see Figure 3.2) without overlap – that is, documents were initially presented simultaneously. Especially in combination with the randomized order of documents, however, this could potentially have generated a layout for some participants in the simultaneous condition in which the documents taking the same stance were presented close to each other, hence rendering unnecessary actively grouping documents. Therefore, in the simultaneous condition of Study 2, documents were initially not presented simultaneously, but on a stack such that only the text of the respective topmost document was visible (see Figure 4.1), and participants had to actively create a simultaneous presentation (as well as groups) of documents themselves. Lastly, regarding interaction possibilities, re-arrangement of documents in the *sequential* condition was only possible in Study 1. This was done in order to investigate the likelihood to which participants in the sequential versus simultaneous condition would engage in grouping documents (i.e., moving them to similar locations) during reading. However, the possibility to re-arrange documents is not given in typical digital reading environments that present documents sequentially, such as computer screens or tablets (that is, the tablet itself can be moved, but the spatial relation between single documents can not be manipulated, which, however, was possible in Study 1 by using a multi-touch table). In Study 2, the sequential condition was thus made more comparable to a more typical sequential reading environment by not allowing re-arrangement of documents. Please note that documents in the sequential condition of Study 2 were still presented on a multi-touch table even though the space provided by it was no longer usable without the possibility to re-arrange documents. This was decided for – instead of presenting documents on a more regularly sized touch screen in the sequential condition – in order to preclude potential artefacts of the device used between presentation conditions. In contrast, in Study 3, in which documents were only presented sequentially, a more regularly sized touch screen was used.

2.2.3 Control Measures and Assessment of Multiple Document Comprehension

Both readers' text comprehension as well as their engagement in reading actions are related to their individual characteristics such as working memory capacity, prior knowledge, topic interest, or prior topic beliefs (Barzilai & Strømsø, 2018; Britt & Rouet, 2012; List & Alexander, 2017a; Maier & Richter, 2013; Perfetti, 1997; Rouet & Britt, 2011). To control for potential differences across experimental conditions, these characteristics were assessed as control variables in all three studies of the present dissertation. Specifically, participants' working memory capacity was assessed in a reading span task (Kane et al., 2004), their prior topic knowledge was assessed as a self-rated measure in Study 1 and as the number of correct facts and arguments in an essay written about the health effects of UV radiation prior to reading the documents in Studies 2 and 3, and their prior topic interest (1 item) and prior topic beliefs (2 items, one addressing the positive and negative beliefs, respectively) were assessed in self-rated measures on a 7-point Likert scale in all studies.

As measures of multiple document comprehension, in all studies, participants' intertextual integration was assessed in the number of intertextual connections in essays written after reading the documents without having the documents available. That is, each connection of two or more pieces of information that clearly stemmed from two different documents was counted as one intertextual connection (see also Kobayashi, 2009; Linderholm & Van den Broek, 2002; List, 2019; List et al., in press; Salmerón et al., 2018a). According to the set of documents used in the studies, Intertextual connections could combine information from two documents taking the same stance, from two documents taking opposite stances, or from the neutral document and one positive or negative document, respectively. Tables 3.5.2 and 5.6.2 provide examples of such intertextual connections.

Furthermore, in Study 1, the number of source citations included in participants' essays was assessed as measure of multiple document comprehension (see, for instance, Bråten, Ferguson, et al., 2014; Kammerer, Kalbfell, et al., 2016; List & Alexander, 2018a; Salmerón et al., 2018a). However, in line with previous work reporting that even advanced readers such as university students only rarely engage in citing sources spontaneously, only 21 out of the 126 participants in Study 1 cited at least one source. While there was a main effect of document presentation on the likelihood to cite sources in essays – with participants in the simultaneous condition being more likely to have included at least one citation in essays (namely 16 out of 62 participants in the simultaneous versus 5 out of 64 participants in the sequential presentation condition) – a different measure of participants' mental representation of source names was used in Study 2. That is, rather than their *spontaneous* source name citations, a source name recall task was used that specifically asked participants to name all document sources they recalled. In Study 3, however, no such measure was used, since text-highlighting was not expected to affect readers' memory for source names.

Moreover, to further investigate whether the characteristics of the reading environment, or the relation between the reading environment and reading interactions would be differentially related to different aspects of multiple document comprehension, a source-content mapping task was used as an additional measure of multiple document comprehension in Studies 2 and 3 (see Figure 4.2; e.g., Delgado et al., 2020; Kammerer, Meier, et al., 2016; Stang Lund et al., 2019; Strømsø et al., 2010). In this task, participants were provided with the source logos of documents in one column and the main statements of documents in a second column. Their task was to re-arrange the main statements such that they were positioned next to the logo of the document that provided the respective information. As such, this measure required integration of the source logo and information provided by the respective document. Of note, however, a relatively superficial source-content integration in the way that represented which document provided positive versus negative information was not sufficient to solve this task, since two documents provided positive and negative information, respectively. Hence, to solve this task, participants also had to have engaged in cross-document information comparison in order to differentiate between the documents taking the same stance. Please note that while the possibility to highlight text was not expected to affect source name recall, text-highlighting might still have affected source-content integration since the key information provided by each document was likely to be highlighted when possible. Therefore, the source-content mapping task was used in Study 3 as additional measure of multiple document comprehension compared to Study 1, but not the source name recall task.

3. *Study I*: Simultaneous Presentation of Multiple Documents and Text-Highlighting: Online Integrative Processes and Offline Integrated Understanding

The content of the following chapter has been published in *Scientific Studies of Reading*. The proportional contributions of all co-authors to the manuscript are presented in the subsequent table. This article may not exactly replicate the final version published in the journal. It is not the copy of record.

Author	Author position	Scientific ideas	Data generation	Analysis interpretation	Paper writing
Caroline Leroy	1st author	50 %	80 %	65 %	55 %
Peter Gerjets	2nd author	15 %	0 %	5 %	5 %
Uwe Oestermeier	3rd author	10 %	0 %	0 %	5 %
Yvonne Kammerer	4th author	25 %	20 %	30 %	35 %

Title of paper:	Simultaneous Presentation of Multiple Documents and Text-Highlighting: Online Integrative Processes and Offline Integrated Understanding
Status in publication process:	Published: *Scientific Studies of Reading, 25(2)*, 179-192 https://doi.org/10.1080/10888438.2020.1784903

Leroy, C., Gerjets, P., Oestermeier, U., & Kammerer, Y. (2021). Simultaneous presentation of multiple documents and text-highlighting: Online integrative processes and offline integrated understanding. *Scientific Studies of Reading, 25(2)*, 179-192. https://doi.org/10.1080/10888438.2020.1784903

Abstract

This study examined the effect of the reading environment, i.e. documents presentation and possibility of text-highlighting, on readers' integrated understanding, as well as the interplay between the reading environment and overt reading processes (i.e., online integrative processes) in forming intertextual connections. University students (*N* = 126) read six partly conflicting documents presented on a large multi-touch table, which were presented simultaneously or sequentially, with or without the possibility of text-highlighting. The simultaneous presentation yielded better integrated understanding, with this effect being fully mediated by an increased likelihood to spatially organize documents during reading. Furthermore, the possibility of text-highlighting also fostered readers' integrated understanding. Additional analyses (sequential presentation only), however, indicated that this was not the case for participants who only infrequently revisited documents. We discuss this complex interplay between the reading environment and readers' online integrative processes in the light of the RESOLV model (Britt et al., 2018; Rouet et al., 2017)

Keywords: multiple documents comprehension, integrated understanding, online integrative processes, documents presentation, text-highlighting

3.1 Introduction

When informing oneself about a socio-scientific issue such as health effects of UV (ultraviolet) radiation, readers are often confronted with multiple documents providing partly conflicting perspectives (Britt et al., 2018). Therefore, in order to gain an overarching *integrated understanding* of the issue addressed in multiple documents, readers do not only need to construct an understanding of the main ideas and arguments conveyed by individual documents, but also to form and represent connections *across* documents (e.g., Britt & Rouet, 2012; Perfetti, Rouet, & Britt, 1999; Salmerón, Gil, & Bråten, 2018). This is a quite challenging task for most readers (List, Du, Wang, & Lee, 2019), because different documents typically do not explicitly refer to each other as it is usually the case for different arguments conveyed within a single document. Furthermore, "switching between documents comes with a cost to attentional resources and hence jeopardizes the comprehender's ability to attend to related information simultaneously and, thereby, to infer meaningful connections" (Van Den Broek & Kendeou, 2015, p. 108).

The Documents Model framework (e.g., Britt, Perfetti, Sandak, & Rouet, 1999; Britt & Rouet, 2012; Perfetti et al., 1999) proposes that an adequate mental representation that readers construct from multiple documents comprises synthesized core information from single documents together with information on how contents from different documents relate (e.g., whether they contradict, support, or complement each other) and which contents stem from which source (i.e., information provider or author). Accordingly, readers' *offline* integrated understanding can, for instance, be assessed by using verification items requiring (cross-document) integration (e.g., Bråten & Strømsø, 2011; Wiley, 2001) or based on written products composed after reading (for a recent review, see Primor & Katzir, 2018), which allow to analyze the quality of integrative argumentation (e.g., Anmarkrud, McCrudden, Bråten, & Strømsø, 2013; Ferguson, Bråten, Strømsø, & Anmarkrud, 2013), the number of statements that combine information from two documents (i.e., intertextual connections; e.g., Kobayashi, 2009; List et al., 2019; Salmerón et al., 2018), or the number of source citations (e.g., Anmarkrud et al., 2014; List et al., 2019). .

While the Documents Model framework (e.g., Britt et al., 1999) mainly focuses on individuals' *mental representations* that account for an integrated understanding, other work has addressed the *cognitive processes* crucial for gaining an integrated understanding from multiple documents (e.g., Afflerbach & Cho, 2009; Bråten, Britt, Strømsø, & Rouet, 2011; Richter & Maier, 2017; Van den Broek & Kendeou, 2015; for a recent review, see Barzilai, Zohar, & Mor-Hagani, 2018). Typically, readers are expected to assess the relevance of documents or specific information therein for their current information need, identify discrepancies or consistencies, and (mentally) connect and combine information across documents (e.g., Rouet & Britt, 2011). For instance, conflict as compared to agreement between documents has been shown to yield better integrated understanding (Ferguson et al., 2013) and increased consideration and memory of source information (e.g., Braasch, Rouet, Vibert, & Britt, 2012; Stadtler & Bromme, 2014). Accordingly, the RESOLV (REading as problem SOLVing) model (Britt et al., 2018; Rouet et al., 2017) suggests a complex interplay between the *reading context* (e.g., the documents set and the environment including support tools available), the *reading processes* (i.e., cognitive processes such as mentally comparing information, or overt actions such as revisiting previously read documents), and readers' integrated understanding. Specifically, according to the RESOLV model, readers adaptively decide to perform specific reading actions based on their evaluation of respective benefits and (physical, cognitive, or emotional) costs (i.e., actions are performed when their perceived benefit-cost ratio exceeds a certain threshold). Noteworthy, the reading environment can affect this benefit-cost ratio (Rouet et al., 2017). The overarching goal of the present study was to investigate the interplay between characteristics of the reading environment, individuals' overt reading processes related to the formation of intertextual connections (referred to as *online integrative processes*), and the resulting offline integrated understanding of multiple documents.

3.1.1 The role of characteristics of the reading environment for integrated understanding

One characteristic of the reading environment that might affect integrated understanding is how multiple documents are presented for reading. Wiley (2001) assumed that presenting multiple documents simultaneously rather than sequentially may "[promote] understanding by supporting the comparison and integration of the individual [documents]" (p. 376). In support of her assumption, a partly simultaneous presentation of multiple documents (i.e., two out of 10 documents presented simultaneously in a split-screen) resulted in more integrated essays of undergraduate students than a sequential presentation.

Furthermore, Olive, Rouet, François, and Zampa (2008) argued that the costs for switching between documents, which requires attentional resources (Van den Broek & Kendeou, 2015), might be reduced in a simultaneous as compared to a sequential presentation (see also Jang, Schunn, & Nokes, 2011). In two studies Olive et al. (2008) found university and postdoctoral students' cognitive load in summarizing three documents to be lower when the document viewer and notepad were displayed simultaneously instead of alternatingly.

However, it should also be noted that Kobayashi (2007) found no difference in university students' self-reported degree of having mentally linked information across two documents during reading when documents were presented simultaneously versus sequentially. Yet, the documents used in that study were quite short (approximately 170 Japanese characters each), and, thus, drawing intertextual connections might not have been as resource demanding.

In sum, previous research indicates that a simultaneous presentation of multiple documents might foster readers' integrated understanding. Yet, to date this has only been investigated with a rather restrictive definition of a simultaneous presentation (Olive et al., 2008; Wiley, 2001), or with only two documents (Kobayashi, 2007). One major goal of the present study, thus, was to investigate the effect of a fully simultaneous as compared to a sequential presentation of more than two documents on readers' integrated understanding.

A second characteristic of the reading environment that might affect integrated understanding is whether text-highlighting is possible during reading. Since the early work of Fowler and Barker (1974), who found a benefit for active text-highlighting over reading pre-highlighted text in university students' one-week delayed memory test, it has been argued that text-highlighting supports reading comprehension because it requires deep information processing in order to decide which information to highlight. Overall, however, empirical evidence is rather diverse, indicating that the benefit of text-highlighting may depend on particular characteristics of the reading environment (for a recent review, see Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013). For instance, studies with undergraduate students found text-highlighting to be more beneficial with difficult rather than easy text (Fass & Schumacher, 1978) and with printed rather than digital text (Ben-Yehudah & Eshet-Alkalai, 2018). A potential reason for the latter finding might be that highlighting is more convenient and natural in printed than in digital text (Ben-Yehudah & Eshet-Alkalai, 2018).

Furthermore, because highlighted text "pops out" (Strobelt, Oelke, Kwon, Schreck, & Pfister, 2016), it might facilitate rereading, comparing, and integrating information across documents. Kobayashi (2009) investigated the effect of using external strategies that make information "pop out" (such as highlighting and annotating) on readers' integrated understanding of multiple documents. Text-highlighting turned out to be the most frequently used external strategy, and with the possibility to use external strategies readers achieved a better integrated understanding than without. Yet, to our knowledge no prior research has exclusively examined the effect of text-highlighting (as the only external strategy) on integrated understanding of multiple documents, which was a second major goal of the present study.

3.1.2 The relation between online integrative processes and integrated understanding

Since the seminal work of Wineburg (1991), several studies have shown a positive relation between readers' offline integrated understanding and their online integrative processes (e.g., Goldman, Braasch, Wiley, Graesser, & Brodowinska, 2012; Wiley et al., 2009). For instance, the degree to which undergraduates correctly identified relevant information in documents during reading (Anmarkrud et al., 2013), or the degree of comparison or corroboration of information within or across documents during reading (Anmarkrud et al., 2014) have been found to be positively related to offline integrated understanding. Besides, overt reading actions such as spatially organizing or revisiting documents, might also reflect online integrative processes. Previous qualitative studies found that university students and staff members, when reading multiple printed documents to write a report or summary, frequently organized documents spatially (Haber et al., 2014), for instance, to draw relations between documents (O'Hara et al., 2002). Furthermore, asking readers to create a concept map (i.e., a graphical organization of content from multiple documents; Barzilai et al., 2018), or an integration map (i.e., indicating agreements and disagreements between documents; Barzilai & Ka'adan, 2017) have been shown to foster offline integrated understanding.

Additionally, revisiting documents (i.e., returning to previously read documents), which might reflect re-evaluation of information in the light of other documents (List & Alexander, 2018; Wineburg, 1991), was also found to be positively related to readers' integrated understanding (e.g., Wiley et al., 2009). For instance, a study with high-school students by Goldman et al. (2012) showed that better learners (i.e., students with greater knowledge gain from pre- to post-test) revisited relevant documents more often than poorer learners, and, moreover, were more likely to include intertextual connections in their essays. Furthermore, List and Alexander (2018) found that the likelihood to revisit documents was associated with a greater number of source citations in undergraduates' essays.

However, to our knowledge, previous studies have neither examined the relation between readers' spontaneous, *non-instructed*, spatial organization of multiple documents and their integrated understanding, nor the interplay between the reading environment and overt reading actions (i.e., spatially organizing and revisiting documents) in forming intertextual connections.

3.1.3 Present study

The goal of the present study was to investigate whether a simultaneous presentation of multiple documents and the possibility of text-highlighting as two characteristics of the reading environment would foster readers' integrated understanding. Furthermore, we examined the interplay between the reading environment, online integrative processes, i.e., spatially organizing and revisiting documents, and readers' offline integrated understanding reflected in written argumentative essays. To this end, we presented six documents on a large multi-touch table (i.e., a large horizontally oriented touch-display, see Figure 3.1). This made it possible to (1) simultaneously present multiple documents without overlap, (2) easily and naturally arrange documents on screen similar to when interacting with printed documents (O'Hara et al., 2002), (3) conveniently highlight text with the finger (see Ben-Yehudah & Eshet-Alkalai, 2018), and (4) assess participants' spatial organization of documents, and, in the sequential presentation[1], also their revisits to documents.

Based on our theoretical and empirical background analysis, we derived the following eight hypotheses:

H1: Participants will engage less with a task-irrelevant document than with a task-relevant document (e.g., Anmarkrud et al., 2013).

H2 and H3: A simultaneous as compared to a sequential presentation of multiple partly conflicting documents (H2; Wiley, 2001) as well as the possibility of text-highlighting (H3; Kobayashi, 2009) will facilitate cross-document information comparison and corroboration and, thus, will foster readers' integrated understanding.

H4: Because in a simultaneous presentation all documents are visible at the same time, which should result in a greater benefit of spatially organizing documents, participants in this condition will be more likely to spatially organize documents during reading than those in the sequential condition.

H5: Participants who spatially organize documents (i.e., potentially in an attempt to draw relations between them; O'Hara et al., 2002) will also have a better offline integrated understanding (Barzilai et al., 2018).

[1] In the simultaneous presentation it was not possible to analyze revisits, because document revisits (i.e., re-inspections) can also occur without overt interactions. Accordingly, H7 and H8 (see below) can only be tested for the sequential presentation.

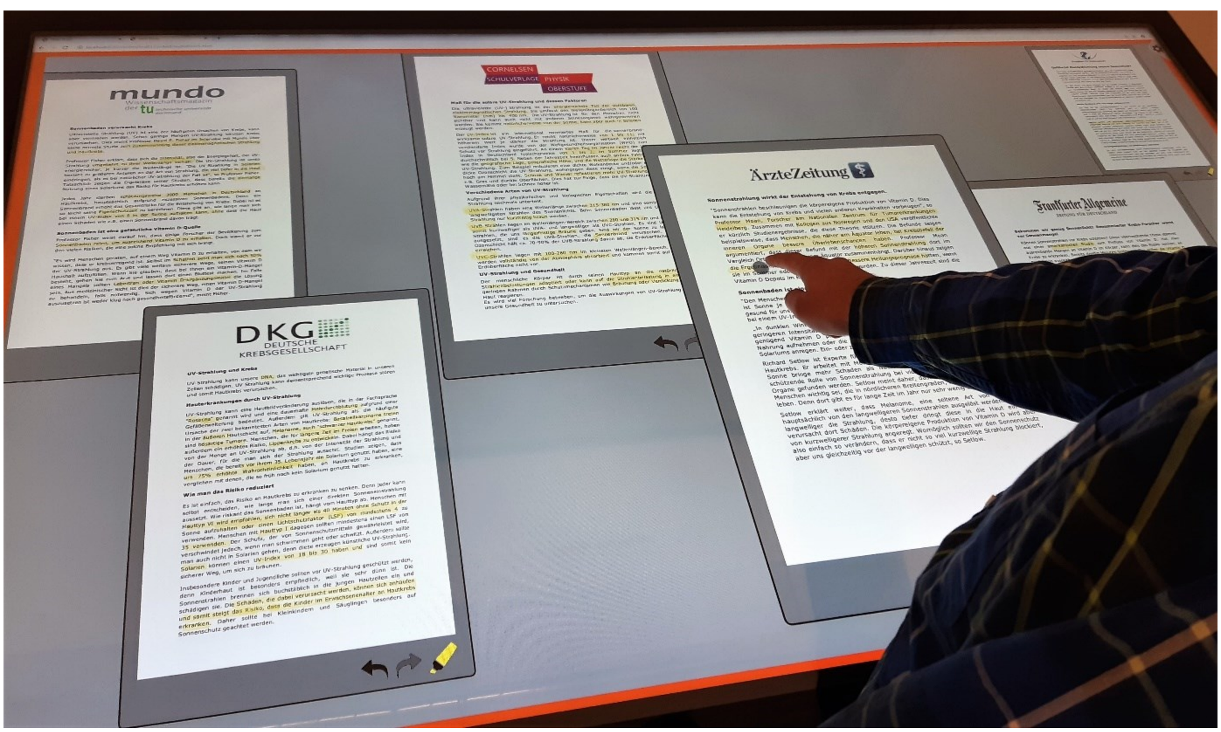
H6: The likelihood to spatially organize documents will mediate the effect of documents presentation on integrated understanding.

H7: Since revisiting previously read documents might reflect an attempt to re-evaluate information (List & Alexander, 2018), readers' integrated understanding will be positively related to the number of revisits to relevant documents during reading (Goldman et al., 2012).

H8: Because highlighted text "pops out" and can thus be assumed to be spotted more easily (for cross-document information comparison), the relationship between readers' integrated understanding and the number of revisits will be stronger with than without highlighting.

Figure 3.1

Reenactment of a participant conducting the reading task on the experimental multi-touch-table user interface in the simultaneous with-highlighting condition, currently highlighting text.



3.2 Method

3.2.1 Participants and experimental design

Participants were one-hundred twenty-six German-speaking university students (74.60 % female, *M* = 23.83 years, *SD* = 3.54) from different majors and semesters (*M* = 6.29 semesters, *SD* = 4.82). Students from fields related to medicine or biology were not invited because we wanted to examine laypersons regarding the topic addressed in the documents (i.e., health effects of UV radiation). The study was approved by the local ethics committee and conducted in accordance to the Declaration of Helsinki. Participants gave their informed consent at the beginning of the study and were rewarded with 12 € after the study.

As independent variables, documents presentation (DP: simultaneous vs. sequential) and the possibility of text-highlighting (HL: with vs. without) were varied between participants. Participants were randomly assigned to one of the four experimental conditions, with 31 or 32 participants serving in each condition.

3.2.2 Documents

As we aimed to examine the number of intertextual connections in participants' argumentative essays as an indication of their integrated cross-document understanding (see also Kobayashi, 2009; List et al., 2019; Salmerón et al., 2018), we used five partly conflicting documents presenting different perspectives on health effects of UV radiation (see Supplementary Tables 3.5.1 and 3.5.2, for a detailed description of the documents and for examples of intertextual connections). The documents were translated and slightly adapted versions of documents used in several previous studies (Ferguson et al., 2013; Strømsø et al., 2016). Two documents stated positive and two other documents stated negative health effects of UV radiation. A fifth document provided general information about UV radiation without taking a stance. In addition, to investigate participants' (implicit) relevance assessments during reading, we presented a sixth task-irrelevant document dealing with health effects of mobile phone radiation (also see Supplementary Table 3.5.1). The six one-paged documents were comparable in length (*M* = 366.67 words, *SD* = 24.78) and readability (*M* = 51.3, *SD* = 2.80; i.e., rather high text difficulty; Björnsson, 1968). All documents contained two to three subheadings and a source logo at the top of the page. Documents' sources were German equivalents of the sources used in previous studies. Trustworthiness of the six sources was indicated in an independent pilot study with 27 university students (see Supplementary Table 3.5.1).

3.2.3 Documents presentation and text-highlighting

The six documents were initially displayed in A4 size in randomized order on an AVR EDGE 55” 4K multi-touch table (resolution: 3840x2160 px, 1.88x1.06 meter; see Figure 3.1), either simultaneously or sequentially (see Figure 3.2). When presented simultaneously, all documents were initially displayed in two rows; when presented sequentially, documents could be opened from a menu containing miniature images of the documents labeled A–F, and were initially presented in the middle of the lower row. In both presentation conditions, participants could freely re-position and resize documents. When documents were re-opened in the sequential presentation, they were displayed in the previously manipulated position and size. In the with-highlighting conditions, text-highlighting could be activated (and de-activated) with a button provided in the respective document’s frame. Text-highlighting could be done directly with the finger. Screen-recordings and interaction logs of the task processing were recorded for each participant.

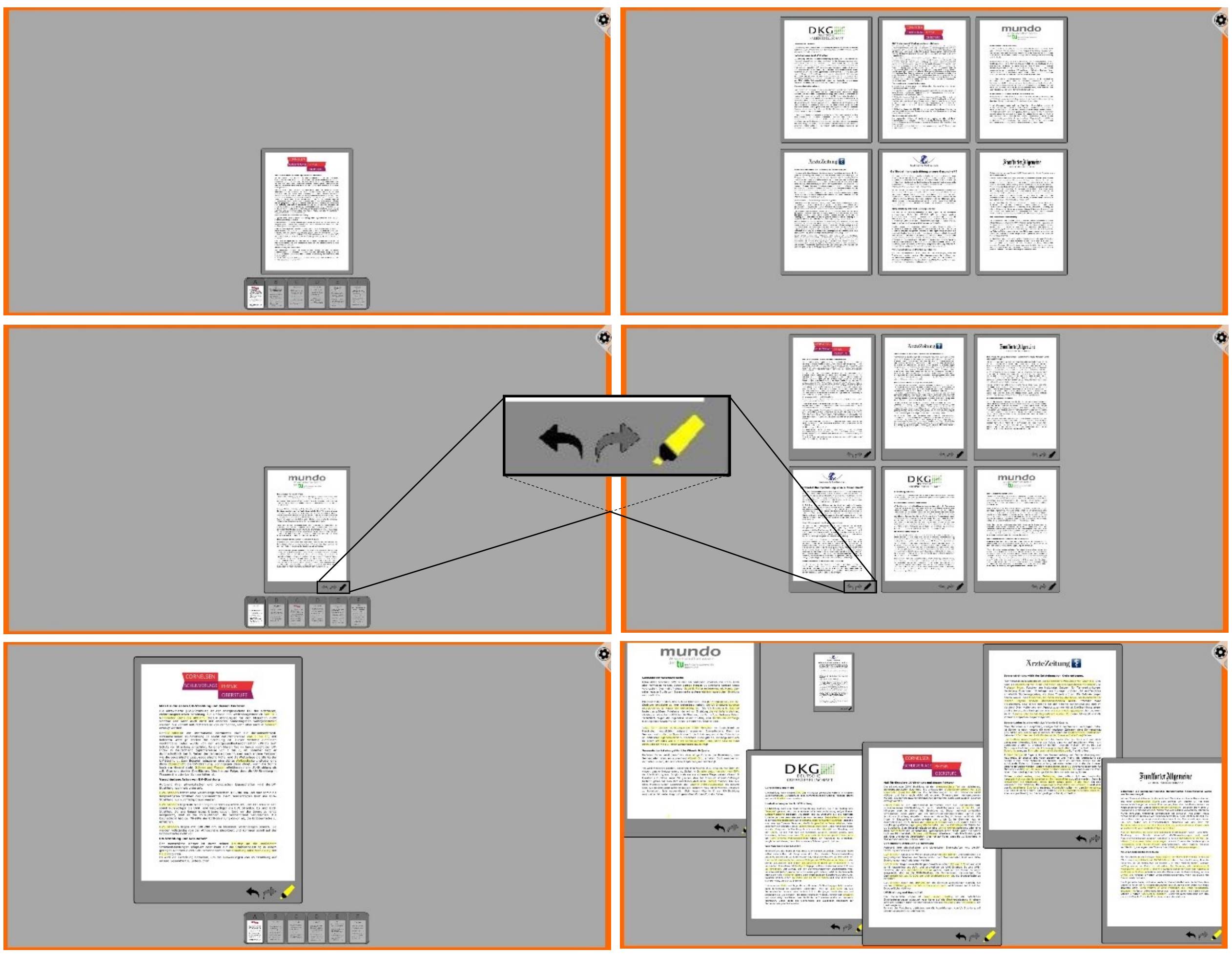
3.2.4 Dependent measures

As a measure of participants’ integrated understanding, we assessed the *number of correct intertextual connections* between relevant documents included in their argumentative essays. Intertextual connections could either connect document-specific information or reflect the global dispute (see Supplementary Table 3.5.2 for examples). Two raters independently coded a random selection of 25 essays for intertextual connections, achieving an interrater-agreement (based on a sentence-level) of 82.67%. Disagreements were resolved through discussion. The remaining 101 essays were coded by one rater and subsequently discussed for confirmation. Besides, two raters coded all essays regarding whether participants had cited any document sources or embedded sources (interrater-agreement: 100% and 98.41%, respectively). We also analyzed essay length, which was previously suggested to reflect task engagement (e.g., Latini, Bråten, Anmarkrud, & Salmerón, 2019).

The 126 screen recordings (displaying participants’ interactions on the multi-touch table) were rated by two independent raters regarding the *spatial organization of relevant documents* as well as regarding whether the task-irrelevant document was spatially separated from the relevant documents. Specifically, spatial organization of documents was coded when congruent or contradicting documents were grouped (simultaneous presentation) or moved to similar locations (sequential presentation) on the multi-touch table. Interrater agreement for spatial organization and sorting out the irrelevant document was 89.06% and 82.81% in the sequential presentation, and 72.58% and 69.36% in the simultaneous presentation, respectively. All disagreements were resolved through discussion.

Figure 3.2

Multi-touch table interface: Initial positions of documents in the sequential (left column) and simultaneous (right column) conditions, without (top row) and with highlighting (middle row). The bottom row shows an exemplary manipulation of documents in the respective documents presentation condition with highlighting.



Furthermore, the logfile recordings of participants in the sequential conditions were analyzed regarding the *number of revisits* to (i.e., re-openings of) the five relevant documents. To exclude unintentional re-openings, only revisits longer than one second were considered.

Control variables. To ascertain comparability across experimental conditions, we assessed participants' self-reported prior knowledge (1 item), interest (1 item), and topic-beliefs (2 items, one addressing positive and another addressing negative beliefs) regarding health effects of UV radiation using 7-point-Likert scales (see Supplementary Table 3.5.3), as well as participants' working memory capacity (15-item computer-based reading span task; cf. Kane et al., 2004). Topic-beliefs were assessed once again after the writing of the essay.

3.2.5 Procedure

Participants were tested in single sessions (of about 1.5h) in the lab. After assessing the control variables, participants received a written explanation of the interaction possibilities on the multi-touch table interface (i.e., freely re-positioning and zooming the documents, and, in the highlighting conditions, text-highlighting), without being prompted to any specific interactions. Subsequently, participants practiced interaction with six blank documents. When they felt comfortable in utilizing the interface, participants were told that they would be provided with six documents from the Internet, which they should read within 25 minutes in order to write an argumentative essay about health effects of UV radiation without having the documents available. After reading, participants were asked to write their argumentative essay on a laptop within 15 minutes, by justifying their statements with arguments and findings from the documents they had read.

3.3 Results

3.3.1 Comparability of experimental conditions regarding control variables

Overall, participants had moderately high working memory ($M = 10.36$, $SD = 2.70$), rather low prior topic knowledge ($M = 3.04$, $SD = 1.23$), moderate topic interest ($M = 4.52$, $SD = 1.16$), and rather negative topic beliefs (agreement with negative statement: $M = 5.28$, $SD = 1.09$; agreement with positive statement: $M = 2.63$, $SD = 1.30$). Experimental conditions did not differ regarding any of these five variables (see Supplementary Table 3.5.4), or regarding their change in topic-beliefs from pre- to post-reading (see Supplementary Table 3.5.5).

3.3.2 Task processing and (implicit) relevance assessments

An ANOVA showed that with the possibility of text-highlighting participants spent more time reading the documents than without highlighting, $F(1,122) = 19.51$, $p < .001$, $\eta_p^2 = .14$ (DP: $F < 1$; interaction: $F(1, 122) = 1.56$, $p = .215$; see Table 3.1 for means and standard deviations). Only six participants (four in the sequential and two in the simultaneous condition) did not make any highlights. On average, participants made 76.16% (*SD* = 33.23) of their highlights during their first interaction with the documents (sequential presentation: *M* = 78.94%, *SD* = 35.32; simultaneous presentation: *M* = 73.47%, *SD* = 31.47). Regarding participants' (implicit) relevance assessments of the documents during reading, repeated-measures ANOVAs using document relevance as within-subjects factor revealed that participants in the sequential conditions made more revisits to a relevant than to the irrelevant document, and participants in the with-highlighting conditions, on average, made more highlights in a relevant than in the irrelevant document, thus, confirming H1 (for means, standard deviations, and inferential statistics, see Supplementary Tables 3.5.6 and 3.5.7).

3.3.3 Effects of documents presentation and possibility of text-highlighting on intertextual connections

For the number of intertextual connections in participants' essays, an ANOVA showed main effects of DP, $F(1, 122) = 5.07$, $p = .026$, $\eta_p^2 = .04$, and HL, $F(1, 122) = 7.33$, $p = .008$, $\eta_p^2 = .05$ (interaction: $F < 1$). As expected in H2 and H3, participants' essays contained more intertextual connections when documents were presented simultaneously rather than sequentially, and when text-highlighting was possible rather than not (for means and standard deviations, see Table 3.1; for a more detailed description of the intertextual connections made, see Supplementary Table 3.5.8; for an analysis of participants' source citations, see Supplementary Table 3.5.9). As mentioned above, time on task was longer with than without highlighting. However, including time on task as a covariate into our model did not change the aforementioned results regarding number of intertextual connections (DP: $F(1, 121) = 4.796$, $p = .031$; HL: $F(1, 121) = 7.65$, $p = .007$; interaction: $F(1, 121) = 0.33$, $p = .568$; time on task: $F(1, 121) = 1.38$, $p = .242$).

Table 3.1

Means (and standard deviations) of reading time, number of highlights, and dependent measures as a function of documents presentation and highlighting.

Documents presentation	Sequential		Simultaneous	
Highlighting	Without	With	Without	With
Reading time [min]	14.39 (4.72)	17.00 (4.76)	13.60 (3.99)	18.25 (4.85)
# Highlights in all relevant documents	-	87.53 (67.77)	-	84.52 (57.08)
# Revisits to all relevant documents (> 1 sec)	11.50 (8.67)	8.19 (8.24)	-	-
# Intertextual connections	2.50 (1.37)	3.31 (1.66)	3.16 (1.55)	3.84 (1.34)
Spatial organization [% participants]	25.00	15.63	41.94	48.39
% Participants who sorted out the irrelevant document amongst those who spatially organized documents (1), or not (0)	1: 25.00 0: 4.17	1: 40.00 0: 7.41	1: 41.67 0: 10.53	1: 60.00 0: 0.00
# Intertextual connections in essays of participants who spatially organized documents (1), or not (0)	1: 2.75 (1.39) 0: 2.42 (1.38)	1: 5.00 (1.87) 0: 3.04 (1.43)	1: 3.77 (1.74) 0: 2.78 (1.26)	1: 3.93 (1.28) 0: 3.69 (1.49)

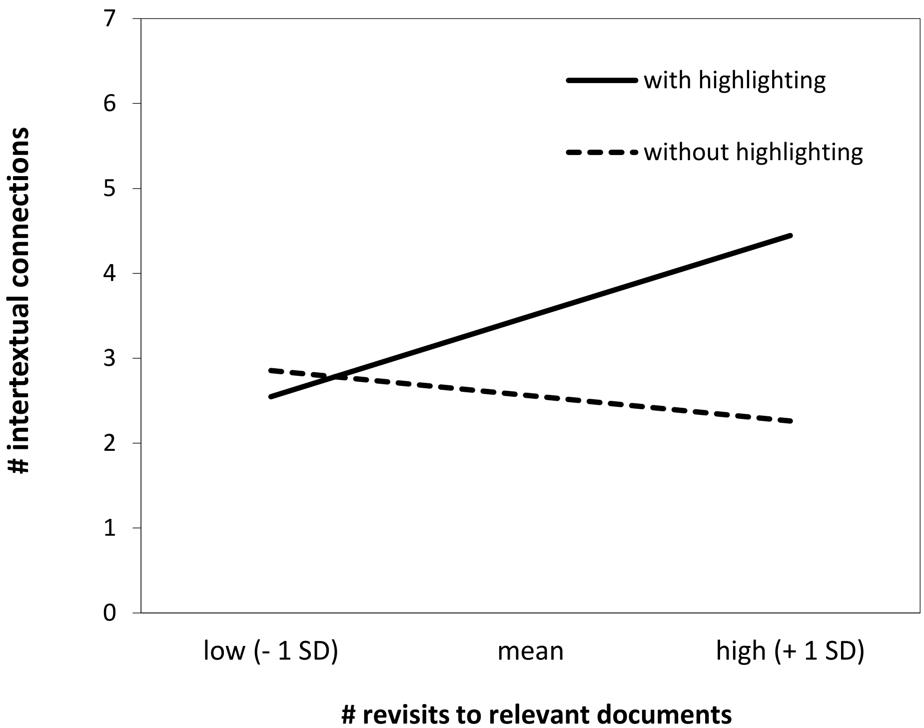
3.3.4 Relations of intertextual connections to spatially organizing documents and number of revisits

In line with H4, a logistic regression analysis showed that a simultaneous presentation yielded a greater likelihood to spatially organize documents during reading than a sequential presentation (45.16% vs. 20.31% of participants), Wald's χ^2 = 8.62, p = .003 (HL: Wald's χ^2 = 0.16, p = .689; interaction: Wald's χ^2 = 1.08, p = .298). Participants who spatially organized documents during reading were also more likely to sort out the irrelevant document during reading than those who did not (51.85% vs. 5.71% in the simultaneous and 30.77% vs. 5.88% in the sequential conditions; see also Table 3.1). Confirming H5, a Welch's t-test showed that participants who spatially organized documents had included significantly more intertextual connections in their essays than those who did not, $t(70.34) = 2.80$, $p = .006$ (also see Table 3.1). In turn, this pattern of findings enabled us to conduct a mediation analysis using a bootstrapping-based procedure by Hayes (2013) with 1,000 bootstraps. The model accounted for a statistically significant portion of the variance, $R^2 = .08$, $F(2, 123) = 5.57$, $p = .005$. In line with H6, the likelihood to spatially organize documents fully mediated the effect of documents presentation on number of intertextual connections (direct effect: $B = 0.20$, $SE = 0.14$, $\beta = .13$, $p = .146$; indirect effect: $B = 0.09$, $SE = 0.05$, $\beta = .06$, $CI_{95\%} = [0.013, 0.42]$, $p = .049$).

Finally, for participants in the sequential conditions, a linear regression analysis with number of revisits to task-relevant documents as continuous predictor (z-standardized) and highlighting as dichotomous predictor as well as the interaction term, and number of intertextual connections as the dependent variable showed a significant interaction between predictors, $\beta = .40$, $t(60) = 3.56$, $p = .001$ (HL: $\beta = .31$, $t(60) = 2.79$, $p = .007$; revisits: $\beta = .21$, $t(60) = 1.87$, $p = .066$). Correlation analyses showed a positive correlation between number of revisits and number of intertextual connections with, $r = .56$, $p = .001$, but not without highlighting, $r = -.22$, $p = .226$. Simple comparison analyses according to Aiken and West (1991) further revealed a beneficial effect of text-highlighting on intertextual connections for participants with a high (i.e., 1 *SD* above the mean; $\beta = .72$, $t(60) = 4.47$, $p < .001$), or medium number of revisits (see main effect; Figure 3.3), but not for participants with a low number of revisits (i.e., 1 *SD* below the mean; $\beta = -.14$, $t(60) = -0.57$, $p = .570$. Taken together, these results do not support H7, because there was no overall positive correlation between number of revisits and number of intertextual connections, but they support H8, which predicted a stronger correlation with than without text-highlighting. Moreover, these results qualify our results regarding H3 (i.e., the effect of highlighting on the number of intertextual connections), such that for the sequential presentation, H3 was not confirmed for participants with few revisits.

Figure 3.3

Interaction of highlighting condition and number of revisits to relevant documents (z-standardized) with regard to the number of intertextual connections in the sequential conditions.



3.4 Discussion

In the present study, we investigated the effects of two characteristics of the reading environment, namely a simultaneous presentation of multiple documents and the possibility of text-highlighting, on readers' offline integrated understanding. Integrated understanding was operationalized as the number of intertextual connections that participants comprised in their argumentative essays after the reading of six partly conflicting documents on a socio-scientific issue. Furthermore, we examined the interplay between the reading environment, overt online integrative processes, i.e., spatially organizing or revisiting documents, and readers' integrated understanding. In the following, we will discuss our results in the light of the RESOLV model (Britt et al., 2018; Rouet et al., 2017), which suggests that readers' integrated understanding is related not only to the reading environment, but also to reading processes such as overt reading actions that readers (decide to) execute when their perceived "benefit-cost ratio exceeds some threshold value" (Rouet et al., 2017, p. 203).

First, participants in the present study engaged less with the task-irrelevant than with task-relevant documents, which might indicate that participants assessed information relevance during reading. In line with RESOLV's assumptions, the beneficial value of the task-irrelevant document might have been perceived as rather low, which resulted in a smaller likelihood of interacting with it.

Second, in line with previous research (e.g., Wiley, 2001), the simultaneous presentation of multiple documents fostered readers' integrated understanding. In the present study, this effect was fully mediated by readers' engagement in spatially organizing the partly conflicting documents during reading. In line with RESOLV's assumptions, this pattern of results might indicate that (a) with documents being presented simultaneously participants have evaluated the cost of spatially organizing documents during reading lower than when presented sequentially, and (b) a spatial arrangement of simultaneously presented documents has lowered participants' perceived costs of comparing and corroborating information across documents, with lower perceived costs resulting in a greater likelihood to execute the respective actions, and, in turn, supporting information integration (see also Jang et al., 2011).

Third, the possibility of text-highlighting also resulted in a better integrated understanding, thus, corroborating and expanding previous research that examined several external strategies (i.e., text-highlighting, annotating, note-taking; Kobayashi, 2009). Note that in line with previous research (Ben-Yehudah & Eshet-Alkalai, 2018), the possibility of text-highlighting also resulted in longer reading times, which might indicate increased levels of engagement (Latini et al., 2019). Reading time, however, did not account for the effect of text-highlighting on integrated understanding. Furthermore, while the possibility of text-highlighting did not significantly affect the number of revisits to sequentially presented documents, additional analyses revealed that the benefit of text-highlighting was moderated by the degree of revisiting documents, such that for participants with few revisits no benefit was shown. This indicates that with text-highlighting, allocating one's attention during revisits (i.e., re-inspecting highlighted information) might have been eased, resulting in a more effective comparison and corroboration than without text-highlighting, particularly because three-fourth of the highlights, on average, were made during first visits. While we acknowledge that this conclusion is only speculative, it supports the notion of a "pop out" effect of highlighted text (Strobelt et al., 2016) rather than an effect of deeper processing through the act of text-highlighting per se.

Yet, the finding that the number of revisits to relevant documents and integrated understanding were uncorrelated without the possibility of text-highlighting contradicts previous research in which a general positive relation was found between the number of revisits and participants' integrated understanding (Goldman, Braasch, et al., 2012; List & Alexander, 2018). However, it should be noted that whereas in the present study document previews only displayed a source logo and a letter (A-F), in previous research documents could be (re-)accessed via Google search results pages presenting the titles and short abstracts of the documents. Such preview information might have facilitated intertextual integration on a coarse level early on. Thus, future research is needed to test effects of preview information on readers' revisiting behavior and its relation to integrated understanding.

We acknowledge that our research comes with certain limitations. First, assessing the number of revisits was only possible in the sequential presentation. Furthermore, even though we only considered revisits above one second, revisits might not always reflect readers' attempts of comparing or linking information across documents, but could also be initiated in order to search for a document or information. For future research on readers' revisiting behavior (also in a simultaneous presentation), we thus suggest the use of eye-tracking methodology, potentially in combination with cued retrospective verbal protocols (Van Gog et al., 2005). Second, whether participants had spatially organized documents was determined by two human raters based on screen recordings. In order to cross-validate the codings with participants' intentions, verbal protocols would also be a useful research tool. Third, a multi-touch table, even though increasingly installed in public places (e.g., in museums), is not a common reading device. Thus, our findings should be replicated in more natural reading environments allowing for spatial organization of documents. Fourth, even though participants were asked to write an argumentative essay on the health effects of UV radiation, they still revisited and highlighted the task-irrelevant document (although to a lesser extent than task-relevant documents), possibly because they were provided with all six documents and because they wanted to learn as much of the provided information as possible. Future research using more than one task-irrelevant document is needed to investigate participants' overt reading processes in such scenarios in more detail. Fifth, it should be considered that parts of the reported results are only based on correlational data. Thus, in order to draw conclusions about causality, further experimental work is required, for instance, by explicitly instructing participants to systematically group, sort, or revisit documents in order to compare and link information across documents. Finally, we acknowledge that we measured integrated understanding only based on written products, which may be affected by students' writing skills (List, Du, et al., 2019). Further research is needed to understand how our findings relate to other measures of integrated understanding, such as verification items requiring cross-document integration (Bråten & Strømsø, 2011) or the construction of a graphic representation of multiple documents (e.g., Barzilai et al., 2018; List, 2019).

Notwithstanding these limitations, we believe that our findings provide important novel insights into the interplay between the reading environment, online integrative processes, and offline integrated understanding, supporting theoretical considerations of the RESOLV model (Britt et al., 2018; Rouet et al., 2017), and, in turn, stressing the importance of designing reading environments and educational interventions in a way that fosters comparison and integration of multiple documents (e.g., Barzilai & Ka'adan, 2017).

3.5 Supplementary Material

Supplementary Table 3.5.1

Detailed description of the six documents.

Document type	Argumentative position	Word count	LIX-score[a]	Central information conveyed	Embedded source(s)	Trustworthiness rating[b] (*M*, *SD*)
Journalist-authored article from a reputable liberal-conservative German newspaper	Positive effects of UV radiation	327	52.2	A longitudinal study by a professor and his team revealed a positive effect of vitamin D on the reduction of cancer risk; they claim that sunrays may protect against all types of cancer through the production of vitamin D, and recommend at least 30 min of daily sun exposure.	yes	4.70 (1.13)
Popular science article from a medical science magazine	Positive effects of UV radiation	402	53.7	Interview with a professor who stated that cancer patients living close to the equator (i.e., with high sun exposure) have greater chances of recovery from cancer; he advocates that sunbeds may be a good alternative to natural sun exposure to produce enough vitamin D.	yes	5.53 (1.23)
Popular science article from a university research magazine	Negative effects of UV radiation	373	51.8	Interview with a professor who stated that UV radiation is the most common cause of cancer and that the one-time use of sunbeds can increase cancer risk; claims that rather than exposing oneself to UV radiation, vitamin D should be supplemented.	yes	5.57 (1.56)

Public information text published by the German national cancer association	Negative effects of UV radiation	377	47.6	Describes that UV radiation causes DNA mutations and that using sunbeds before the age of 35 bears a 75% increased risk of cancer; explains how to calculate unharmful sun exposure time as a function of sun protection factor of sun blockers and skin tone.	no	5.80 (1.01)
Excerpt from a high-school science textbook	General information about UV radiation	364	54.1	Describes UV radiation in neutral, academic terms; explains factors affecting the UV-index and the different wave lengths of UV-A, UV-B, and UV-C radiation.	no	5.17 (1.21)
Public information text published by the German federal office for radiation protection	Negative effects of mobile phone radiation	357	48.1	Explains that since the increased use of mobile phones, the risk of cancer of the parotid gland has increased by a factor of 1.5; suggests several possible preventive actions to decrease health risks of mobile phone usage.	no	6.43 (0.84)

[a]*Note*. LIX = readability formula according to Björnsson (1968).

[b]*Note*. Trustworthiness ratings were assessed in an independent pilot study with 27 university students (from different majors) rating the sources' trustworthiness on a scale from 1 (not at all trustworthy) to 7 (highly trustworthy).

Supplementary Table 3.5.2

Examples of different types of intertextual connections.

Type of intertextual connection	Information provided by one document	Information provided by another document	Example for the respective type of intertextual connection
Global contradicting connection	UV radiation is one of the most common causes of skin cancer. (information conveyed by both negative documents)	Sun rays accelerate the body's own production of vitamin D. This can prevent the development of cancer and many other diseases. (information conveyed by both positive documents)	Research is arguing about whether UV radiation can be harmful or even beneficial to health. On the one hand, UV radiation is needed to produce vitamin D; on the other hand, serious diseases can be triggered if one exposes oneself to radiation.
Specific contradicting connection	Studies show that people who have already used a sunbed before the age of 35 are 75% more likely to develop skin cancer.	In the dark winter months, we cannot produce enough vitamin D naturally. Therefore, we have to supplement it or trigger the endogenous production by using a sunbed.	The use of sunbeds is controversially discussed among researchers. Some researchers are in favor of sunbeds (to a certain degree) because it also boosts vitamin D production – others are against it because the artificial UV radiation harms the body.
Consistent positive connection	In a long-term study with 50,000 men, they found that people with high levels of vitamin D in their bodies were less susceptible to cancer.	Sunrays stimulate the body’s vitamin D production. […] The findings show, for example, that people who live closer to the equator have better chances of survival if their internal organs are affected by cancer than people who live further away from the equator.	For example, both a long-term study and a study conducted with people living near the equator show that UV radiation can have a positive effect on cancer risk by stimulating vitamin D production.

Consistent negative connection	In fact, the results of his studies show that even a single use of a sunbed can increase the risk of skin cancer.	Studies show that people who have already used a sunbed before the age of 35 are 75% more likely to develop skin cancer.	Studies show that the use of sunbeds before the age of 35 – or even a one-time use – increases the risk of developing cancer.
Connection of factual information	The lower the wavelength, the higher the energy content of UV radiation.	UV radiation penetrates deeper levels of the skin, the higher its energy content is.	Lower-wave UV radiation penetrates deeper into the skin.
Connection between sources	A team of researchers led by Professor Edward Gillies of Harvard University in Boston conducted a long-term study with 50,000 men. It was found that people with high levels of vitamin D in their bodies were less susceptible to cancer.	"Sunlight accelerates the body's own production of vitamin D. This can prevent the development of cancer and many other diseases," said Professor Moan, researcher at the National Center for Tumor Diseases in Heidelberg.	Researchers from Heidelberg and Boston argue that UV radiation has a positive effect on human health.

Supplementary Table 3.5.3

Control variables, each on a scale from 1 = "very low" to 7 = "very high".

Measure	Item
Self-reported prior topic knowledge	How do you rate your knowledge on the issue of "Health effects of UV radiation"?
Self-reported topic interest	How do you rate your interest in the issue of "Health effects of UV radiation"?
Self-reported positive topic belief	How strong is your agreement with the statement that UV radiation has a positive effect on health?
Self-reported negative topic belief	How strong is your agreement with the statement that UV radiation is harmful to health?

Supplementary Table 3.5.4

Means (and standard deviations) and inferential statistics of control variables, i.e., working memory capacity (a score of 0-15) and self-report items (7-point Likert scales from 1-7) as a function of documents presentation and highlighting.

Documents presentation	Sequential		Simultaneous		
Highlighting	Without	With	Without	With	*F*- and *p*-values
Working memory capacity (0-15)	10.44 (2.64)	10.81 (2.75)	9.77 (2.75)	10.39 (2.69)	DP.: $F(1, 122) = 1.27, p = .262$ HL: $F(1, 122) = 1.04, p = .310$ DP x HL: $F < 1$
Prior topic knowledge	2.78 (1.29)	3.34 (1.26)	3.07 (1.18)	2.97 (1.17)	DP.: $F < 1$ HL: $F(1, 122) = 1.19, p = .278$ DP x HL: $F(1, 122) = 2.28, p = .134$
Prior topic interest	4.50 (1.08)	4.44 (1.32)	4.55 (1.15)	4.61 (1.15)	DP.: $F < 1$ HL: $F < 1$ DP x HL: $F < 1$
Positive topic beliefs (prior to reading)	2.59 (1.54)	2.72 (1.06)	2.58 (1.26)	2.65 (1.36)	DP.: $F < 1$ HL: $F < 1$ DP x HL: $F < 1$
Negative topic beliefs (prior to reading)	5.22 (1.24)	5.38 (1.07)	5.23 (1.15)	5.29 (0.90)	DP.: $F < 1$ HL: $F < 1$ DP x HL: $F < 1$

Note. DP = documents presentation. HL = highlighting.

Supplementary Table 3.5.5

Means (and standard deviations) of participants' self-reported topic beliefs (7-point Likert scales from 1-7) before and after reading as well as repeated-measures ANOVAs for measurement time.

Item	Before reading	After reading	*F*- and *p*-values
Positive topic beliefs	2.64 (1.30)	3.95 (1.37)	DP: $F < 1$ HL: $F < 1$ DP x HL: $F < 1$ Time: $F(1, 122) = 115.32$, $p < .0001$ Time x DP: $F < 1$ Time x HL: $F < 1$ Time x DP x HL: $F < 1$
Negative topic beliefs	5.80 (1.09)	4.84 (1.25)	DP: $F < 1$ HL: $F < 1$ DP x HL: $F < 1$ Time: $F(1, 122) = 15.14$, $p < .001$ Time x DP: $F(1, 122) = 2.45$, $p = .120$ Time x HL: $F < 1$ Time x DP x HL: $F(1, 122) = 1.41$, $p = .237$

Note. Time = measurement time. DP = documents presentation. HL = highlighting.

Supplementary Table 3.5.6

Means (and standard deviations) of the number of revisits to a task-relevant document and to the task-irrelevant document as a function of highlighting. Statistics reflect a repeated-measures ANOVA with document relevance as within-subjects factor.

	Sequential without highlighting	Sequential with highlighting	*F*- and *p*-values
Average # revisits to a relevant document (> 1 sec)	2.30 (1.74)	1.64 (1.65)	Relevance: $F(1, 62) = 16.22$, $p < .001$ HL: $F(1, 62) = 3.32$, $p = .073$ Relevance x HL: $F(1, 62) < 1$, $p = .929$
# Revisits to irrelevant document (> 1 sec)	1.75 (1.72)	1.06 (1.16)	

Note. DP = documents presentation. HL = highlighting. Relevance = document relevance.

Supplementary Table 3.5.7

Means (and standard deviations) of the number of highlights made in a task-relevant document and in the task-irrelevant document as a function of documents presentation. Statistics reflect a repeated-measures ANOVA with document relevance as within-subjects factor.

	Sequential with highlighting	Simultaneous with highlighting	*F*- and *p*-values
Average # highlights in a relevant document	17.18 (13.20)	16.64 (11.36)	Relevance: $F(1, 61) = 16.43$, $p < .001$ DP: $F(1, 61) < 1$, $p = .828$ Relevance x DP: $F < 1$
# Highlights in the irrelevant document	14.50 (13.71)	13.71 (11.29)	

Note. DP = documents presentation. HL = highlighting. Relevance = document relevance.

Supplementary Table 3.5.8

Percentage of intertextual connections between different types of documents as a function of document presentation and highlighting.

Document presentation	Sequential		Simultaneous	
Highlighting	Without	With	Without	With
% Intertextual connections between one positive and one negative document	59.90 (31.74)	53.53 (24.15)	60.85 (28.77)	52.37 (25.54)
% Intertextual connections between the two positive documents	4.17 (18.45)	7.21 (12.41)	5.14 (11.17)	9.89 (13.93)
% Intertextual connections between the two negative documents	8.70 (18.20)	6.96 (14.62)	10.52 (22.79)	14.84 (18.67)
% Intertextual connections between the neutral and one positive document	13.13 (23.40)	14.28 (18.82)	14.09 (21.47)	9.68 (16.82)
% Intertextual connections between the neutral and one negative document	14.11 (24.41)	16.94 (22.14)	8.61 (14.77)	14.84 (18.29)

Supplementary Table 3.5.9

Descriptives and inferential statistics of the percentage of essays containing source citations and essay length as a function of documents presentation and highlighting.

Documents presentation	Sequential		Simultaneous		
Highlighting	Without	With	Without	With	Wald's χ^2 , F-, and p-values
At least one document source cited [% participants]	3.13	9.38	25.81	25.81	DP: Wald's $\chi^2 = 7.42$, $p = .007$ HL: Wald's $\chi^2 = 0.78$, $p = .378$ DP x HL: Wald's $\chi^2 = 0.78$, $p = .378$
At least one embedded source cited [% participants]	12.50	21.88	25.81	25.81	DP: Wald's $\chi^2 = 1.52$, $p = .220$ HL: Wald's $\chi^2 = 0.56$, $p = .455$ DP x HL: Wald's $\chi^2 = 0.56$, $p = .455$
Essay length [#words]	265.56 (50.14)	254.22 (70.24)	256.29 (73.82)	253.13 (55.05)	DP: $F < 1$ HL: $F < 1$ DP x HL: $F < 1$

Note. DP = documents presentation. HL = highlighting.

4. Study II: Investigating the Roles of Document Presentation and Reading Interactions on Different Aspects of Multiple Document Comprehension.

The content of the following chapter has been submitted for publication to the *International Journal of Human-Computer-Interaction.* The proportional contributions of all co-authors to the manuscript are presented in the subsequent table. This article may not exactly replicate the final version going to be published.

Author	Author position	Scientific ideas	Data generation	Analysis interpretation	Paper writing
Caroline Leroy	1st author	50 %	80 %	70 %	60 %
Peter Gerjets	2nd author	10 %	0 %	5 %	5 %
Uwe Oestermeier	3rd author	10 %	0 %	0 %	5 %
Yvonne Kammerer	4th author	30 %	20 %	25 %	30 %

Title of paper:	Investigating the roles of document presentation and reading interactions on different aspects of multiple document comprehension.
Status in publication process:	Submitted for publication: *International Journal of Human-Computer-Interaction*

Leroy, C., Gerjets, P., Oestermeier, U., & Kammerer, Y. (submitted for publication). Investigating the roles of document presentation and reading interactions on different aspects of multiple document comprehension.

Abstract

The present study examined reading interactions and multiple document comprehension of 108 university students who read five partly conflicting documents on a health-related issue. Documents were presented on a large touch interface that either enabled a simultaneous or imposed a sequential presentation of the documents. None of the three multiple document comprehension measures (number of intertextual connections in essays, score in source-content mapping task, and number of source names recalled) was affected by document presentation. However, in the sequential condition, the number of revisits to documents was positively related to memory for source names and source-content integration, but not to intertextual integration. Conversely, in the simultaneous condition, participants who grouped documents during reading showed greater intertextual and source-content integration, but no greater memory for source names than those who had not grouped documents. To conclude, reading interactions played a more important role in readers' multiple document comprehension than document presentation.

Keywords: Multiple document comprehension, document presentation, systematic re-arrangement, revisiting documents, reading interactions

4.1 Introduction

When trying to understand complex and controversial health-related issues such as potential positive or negative health effects of ultraviolet (UV) radiation, individuals often turn to the Internet. There, they are usually faced with multiple documents, which stem from a variety of sources such as online newspapers, magazines, and health institutions, among others. Some documents might provide consistent or complementary information on the issue at hand, while other documents might stand in conflict with each other (Bråten, Anmarkrud, et al., 2014). Thus, building a comprehensive understanding of such a controversial subject matter requires comparing, evaluating, and integrating information across those documents (e.g., Anmarkrud et al., 2014; Britt & Rouet, 2012; List et al., 2021; Mahlow et al., 2020; Primor & Katzir, 2018; Wineburg, 1991).

Complex reading tasks (e.g., on the Internet) as outlined above are usually conducted on conventional computer screens, laptops, or tablets, where screen space is limited, and document reading applications (e.g., PDF Readers, Web Browsers) by default often present only one document at a time (Lombard et al., 2021). However, a sequential presentation of multiple documents may not be ideal for building a comprehensive understanding of a complex issue (Leroy et al., 2021; Lombard et al., 2021; Wiley, 2001). In contrast, as assumed by various researchers from both human-computer interaction research and text-comprehension research, a simultaneous presentation of multiple documents may better support comparing and integrating information across documents (Andrews et al., 2010; Ball & North, 2005; Benshoof et al., 1995; Bi & Balakrishnan, 2009; Jang & Schunn, 2012; Leroy et al., 2021; Wiley, 2001). Yet, empirical research that examines differences in readers' comprehension of multiple documents when documents are read on an interface presenting them simultaneously as compared to sequentially is very limited. The present work aimed to contribute to this line of research by examining the effects of a simultaneous as compared to sequential document presentation on three different aspects of multiple document comprehension. For this purpose, half of the participants read five partly conflicting documents on an interactive touch-interface that enabled a simultaneous presentation of all documents, whereas the other half read the documents on an interactive touch-interface that allowed only for a sequential presentation, that is, with only one document presented at a time.

Furthermore, when presenting texts by means of an interactive reading interface it might be crucial for readers' multiple document comprehension how they use the capabilities of the interface to interact with documents as a tool for document comparison and integration (e.g., Author(s), 2021; Delgado et al., 2020; Leroy et al., 2021; List & Alexander, 2018). Thus, a second major goal of the present research was to investigate how specific interactions with documents (i.e., reading interactions) that might reflect cross-document information comparison are related to different aspects of readers' multiple document comprehension in the respective presentation condition. Specifically, we assessed participants' engagement in revisiting documents (sequential condition) and in grouping the partly conflicting documents on the screen (simultaneous condition) during reading. In the following, we first describe what constitutes an adequate mental representation of multiple documents that skilled readers are assumed to form during reading. Subsequently, we discuss which roles a simultaneous as compared to sequential document presentation and/or specific reading interactions performed might play for constructing this mental representation.

4.2 Theoretical and Empirical Background

4.2.1 Readers' Mental Representation of Multiple Documents

According to the Documents Model framework (Britt et al., 1999; Britt & Rouet, 2012; Perfetti et al., 1999), in addition to the core information of single documents (Kintsch, 2003), an adequate mental representation of multiple documents, a so-called *documents model*, comprises the following components: (a) information on how statements from different documents relate to each other or how documents as a whole relate to each other, respectively (i.e., intertextual connections), (b) information about the origin, that is, the source of the documents (e.g., publication venue or name of the author), and (c) information about which content stems from which source (i.e., connections of source and contents).

Building on the assumptions of the Documents Model Framework, the quality of readers' mental representation of multiple documents (and, thus, their multiple document comprehension) can be assessed in tasks that address the specific components of readers' documents models. For instance, these can be (a) essay tasks that allow to assess the number intertextual connections made across documents (Kobayashi, 2009; Linderholm & Van den Broek, 2002; List, 2019; Salmerón et al., 2018a), (b) source name recall tasks to assess memory of source information (Bråten, Ferguson, et al., 2014; Kammerer, Kalbfell, et al., 2016; Salmerón et al., 2018b), or (c) source-content mapping tasks that address readers' source-content integration (Delgado, Stang, et al., 2020a; Kammerer, Meier, et al., 2016; Stang Lund et al., 2019; Strømsø et al., 2010; for recent reviews also see Barzilai et al., 2018; Primor & Katzir, 2018).

While Britt and colleagues (Britt et al., 1999; Britt & Rouet, 2012) mainly focused on describing and assessing readers' mental representations that result from reading multiple documents, other theoretical work has addressed the processes required to construct such mental representations, such as comparing, evaluating, and integrating information from and across multiple documents (e.g., Afflerbach & Cho, 2009; Braasch & Bråten, 2017; Cho & Afflerbach, 2017; Rouet et al., 2017; Rouet & Britt, 2011; Stadtler & Bromme, 2014). For instance, Cho and Afflerbach (2017, p. 120) provided a taxonomy in which they described various important strategies involved in comprehending multiple documents. These include (a) "comparing and contrasting the content of the text being read with the content of related texts to develop a coherent account of cross-textual contents", (b) "rereading and linking text segments that were previously regarded as unrelated to finalize cross-textual meaning structures", (c) "perceiving that multiple texts related to the same topic can provide diverse and contrasting views about the topic, complementary information about the topic, or both", and (d) "using information about a present source to evaluate and interpret text content". In line with this taxonomy, several empirical studies showed a positive relation between readers' multiple document comprehension and their engagement in cross-document information comparison during reading as reflected by their think-aloud comments (Anmarkrud et al., 2013, 2014; Bråten, Anmarkrud, et al., 2014; Hilbert & Renkl, 2008; Wineburg, 1991), or their re-reading behavior (Goldman, Braasch, et al., 2012; List, Du, et al., 2019).

However, from the above description of the many-faceted strategies involved in comprehending multiple documents it becomes clear that constructing a documents model is a challenging task for most readers (e.g., Kiili & Leu, 2019; List et al., 2019). Therefore, it has become an important research goal to identify characteristics of a reading environment that support readers in performing such strategies (Barzilai et al., 2018; Britt & Rouet, 2012). As we will outline in the following, one such characteristic of a reading environment that has been considered to support cross-document information comparison and integration is whether documents are presented simultaneously or sequentially.

4.2.2 Effects of a Simultaneous Presentation of Documents on Multiple Document Comprehension

Previous findings from observational studies suggest that a simultaneous presentation of multiple documents can support readers in tasks that require cross-document information comparison (Andrews et al., 2010; Ball & North, 2005; Hutchings et al., 2004; Hutchings & Stasko, 2004; O'Hara et al., 2002). Furthermore, experimental research from the field of human-computer interaction revealed positive effects of a reading environment enabling a simultaneous presentation (as compared to imposing a sequential presentation) of documents for analytic problem solving tasks that require information integration across documents (Jang et al., 2011, 2012; Jang & Schunn, 2012, 2014; Takano et al., 2015), and experimental research from the field of multiple document comprehension indicated respective positive effects on cross-document information comparison (Olive et al., 2008), on overall quality of readers' multiple document comprehension (Wiley, 2001), and on readers' intertextual integration (Leroy et al., 2021). However, it remains unclear from previous work whether different aspects of multiple document comprehension are differentially affected by document presentation.

As argued by previous work, the advantage of a simultaneous over a sequential presentation of documents might be due to an easier re-accessing of or switching between multiple documents (Andrews et al., 2010; Bi & Balakrishnan, 2009; Haber et al., 2014; Jang et al., 2012; Jang & Schunn, 2012; O'Hara & Sellen, 1997; Takano et al., 2015), which, in turn, might support cross-document information comparison and integration through increased re-processing of the respective information. Still, results of experimental studies investigating the effect of a simultaneous versus sequential presentation of documents on multiple document comprehension suggest it requires more than simply providing a simultaneous presentation of documents to readers for it to have a beneficial effect. For instance, in a study by Wiley (2001), participants read ten documents either in a split-screen simultaneously presenting two out of ten documents, or in a sequential presentation with only one document visible at a time, and wrote argumentative essays on the topic after reading without the documents available. Results showed that when participants received explanations for why they were given a split-screen interface prior to reading, they wrote essays reflecting better integration compared to participants who had read the documents in the sequential interface. However, the split-screen interface did not yield better integration than the sequential interface when no explanations for the split-screen interface were given. Similarly, Lombard et al. (in press) found that the provision of a tool that enabled a partly simultaneous presentation of (up to three out of five) documents yielded better integration than a sequential presentation of documents *only* for participants who had received explicit instructions to use all features of the tool.

The simultaneous conditions of the aforementioned studies, however, only allowed a partly simultaneous presentation of two or three documents. In contrast, a fully simultaneous presentation of all documents might yield greater support of cross-document information comparison and integration, and thus foster readers' multiple document comprehension even without explicit instructions. To the best of our knowledge, to date, the study by Leroy et al. (2021) is the only one investigating the effect of a simultaneous presentation of *all* documents provided to readers (as compared to a sequential presentation of these documents) on multiple document comprehension. In their study, participants read six partly conflicting documents on a multi-touch table (i.e., a large horizontal display) presented either simultaneously or sequentially, with the possibility to freely re-arrange documents on the screen in both conditions. Participants received instructions on how to use the respective interface but no explicit instructions to use all features of the interface. Results showed that participants in the simultaneous condition *spontaneously* were more likely than those in the sequential condition to spatially organize (i.e., move to similar locations) the partly conflicting documents on the screen. Most interestingly, the likelihood to engage in this reading interaction fully mediated the positive effect of document presentation on intertextual integration. That is, similar to previous findings suggesting that a (partly) simultaneous document presentation might only be beneficial when readers are instructed to use the interface, Leroy et al. (2021) found that the simultaneous presentation of documents is particularly beneficial for readers who use it to group documents during reading. Hence, merely providing readers with a simultaneous presentation of documents per se might not be sufficient to support multiple document comprehension. Rather, how readers interact with an interface during reading might play a crucial role in the effectiveness of the kind of document presentation on their multiple document comprehension.

4.2.3 The Role of Reading Interactions in Forming an Understanding of Multiple Documents

The findings by Leroy et al. (2021) point out the importance of assessing reading interactions reflecting readers' engagement in comparing information across documents during reading (cf. Anmarkrud et al., 2014; Bråten, Ferguson, et al., 2014; Britt et al., 2004; Britt & Aglinskas, 2002; Ferguson et al., 2012; Ferguson & Bråten, 2013; Wineburg, 1991). In addition to having analyzed whether participants had grouped the partly conflicting documents during reading, in the sequential condition, Leroy et al. (2021) also assessed the number of revisits to documents from user interaction logfiles as a second measure of reading interactions. Revisits to documents were assumed to reflect cross-document information comparison. In the following, we will briefly outline previous work explaining the relation between the reading interactions of grouping documents or revisiting documents, respectively, and readers' multiple document comprehension.

4.2.3.1 Grouping Documents

Instructional research has previously suggested that strategical re-organization of information provided by multiple documents is positively related to multiple document comprehension (Barzilai & Ka'adan, 2017; Hilbert & Renkl, 2008; Lombard et al., 2021; Lundstrom et al., 2015). That might be because the respective layout of re-organized information provides visual signals that highlight the relations between pieces of information (Barzilai & Ka'adan, 2017). In their observational studies, O'Hara et al. (O'Hara et al., 2002; O'Hara & Sellen, 1997) furthermore reported that readers also *spontaneously* re-organized printed documents spatially during reading to "gain a sense of overall structure" and to "check on or to relate specific pieces of information across documents" (p. 346; see also Haber et al., 2014). When documents are partly conflicting such that some documents provide arguments in favor of a particular issue, while other documents provide information against that issue, such re-arrangement might take the form of grouping documents according to their overall stance, and the resulting layout of documents might bear contextual information regarding the relation of documents (Andrews et al., 2010b; Bi & Balakrishnan, 2009; Leroy et al., 2021). Of note, in order to group documents according to their overall stance, readers need to be already aware of the fact that some documents provide positive arguments for a subject matter, while others provide negative arguments for it. Yet, the resulting layout of documents (i.e., their arrangement in groups) can make intertextual relations even more salient (cf. Barzilai & Ka'adan, 2017; Jang & Schunn, 2012; Skuballa et al., 2018), and the spatial proximity of documents might further support comparison and integration of information across documents (cf. Ginns, 2006).

To our knowledge, however, the study by Leroy et al. (2021) is the only one that investigated the role of readers' engagement in grouping partly conflicting documents during reading in their multiple document comprehension. They found that participants' intertextual integration (as assessed from essays written after reading) was better when they had grouped documents according to their overall stance during reading than when they had not grouped documents. Yet, Leroy et al. (2021) had assessed participants' intertextual integration as the sole measure of multiple document comprehension, thus leaving the question unanswered whether grouping documents is also related to other measures of multiple document comprehension. The present research aimed to address this question by assessing different measures of multiple document comprehension.

4.2.3.2 Revisiting Documents

Readers' intent to compare or re-evaluate information *across documents* might be reflected in their engagement in re-accessing previously read documents (Goldman, Braasch, et al., 2012; Leroy et al., 2021; List & Alexander, 2018). For example, List and Alexander (2018) found that participants who had revisited at least one document were more likely to include source citations in the essays written after reading. Regarding other measures of multiple document comprehension, Delgado et al. (2020) found a positive relation between accessing source information (provided as pop-up windows in a hypertext) and source-content integration. Similarly, Author(s) (2021) found that the number of revisits to documents was positively related to source-content integration, but was unrelated to intertextual integration – unless participants were able to highlight text (as it was also found by Leroy et al. (2021). Overall, the findings by List and Alexander (2018), Leroy et al. (2021), and Author(s) (2021) thus suggest that (when text-highlighting is not possible) revisits in particular play a role for multiple document comprehension measures related to source information. This is particularly interesting since the documents in these studies were partly conflicting and readers usually are more likely to return to source information when they contain conflicting information (Braasch et al., 2012). Hence, revisits might rather be initiated by readers to re-evaluate information in the light of its source (cf. List & Alexander, 2018; Wineburg, 1991) than to link information between documents.

However, previous work has examined only one or two aspects of multiple document comprehension (e.g., only intertextual connections *or* source-content integration *or* source citations). The present work thus contributes to this line of research by assessing three measures that address different aspects of multiple document comprehension.

4.2.4 Present study

In this study, we built on previous work by Leroy et al. (2021) to examine the effect of a reading interface enabling a simultaneous presentation of multiple documents as compared to one that imposes a sequential presentation of documents on different aspects of multiple document comprehension. Further, we aimed to investigate how reading interactions, namely revisiting previously read documents in the sequential presentation and grouping partly conflicting documents in the simultaneous presentation, are related to readers' multiple document comprehension. To this end, we presented five partly conflicting documents on the topic of health effects of UV radiation on a large multi-touch table (i.e., a large horizontally oriented touch-display). In the simultaneous condition, the documents were initially presented on a stack and could be flexibly re-arranged on the interface by participants to generate a simultaneous presentation (without overlap, if desired) similar to when interacting with printed documents (O'Hara et al., 2002). In the sequential condition, similar to a regular computer screen or tablet interface, the documents were presented statically (i.e., not allowing any re-arrangement of documents on the screen) and only one at a time.

We had the following three main research questions (RQs) in the present study:

RQ1: Does an interface that enables a simultaneous presentation of multiple documents result in an increased use of intertextual connections in argumentative essays as well as in better memory for source names and better source-content integration than an interface that imposes a sequential presentation?

RQ2: Do readers in the simultaneous condition who spontaneously (i.e., without being prompted to) group the partly conflicting documents according to their overall stance show better intertextual integration, better memory for source names, and better source-content integration?

RQ3: Is the extent of revisiting documents in the sequential condition positively related to readers' intertextual integration, memory for source names, and source-content integration?

Regarding RQ1, based on previous findings by Leroy et al. (2021), we expected to find a beneficial effect of the simultaneous compared to the sequential condition for readers' intertextual integration. Furthermore, we explored whether the presentation condition also had an effect on participants' source-name recall or their source-content integration.

Regarding RQ2, based on findings by Leroy et al. (2021), we expected participants' intertextual integration to be positively related to grouping documents. Furthermore, we explored whether participants who had grouped documents during reading in the simultaneous condition also showed better memory for source names or better source-content integration than those who had not grouped documents during reading.

Regarding RQ3, as in previous research we did not expect participants' intertextual integration to be related to the number of revisits made (Author(s), 2021; Leroy et al., 2021). However, based on prior research that found a positive relation between revisiting documents and source citations in essays (List & Alexander, 2018) as well as source-content integration (Author(s), 2021; Delgado et al., 2020), we expected the number of revisits to be positively related to participants' memory for source names as well as to their source-content integration.

4.3 Method

4.3.1 Participants and Experimental Design

One hundred and eight university students from different majors at a large German university participated in this study. However, two participants had to be excluded from the analyses since they did not fulfil the requirements (i.e., they stated their mother tongue was not German in the questionnaire and their essays did not reflect a good level of German). This resulted in a final sample of N = 106 participants (75.47% female; M = 22.53 years, SD = 2.75 years). Participants were compensated with 8€ for their participation. The study was approved by the local ethics committee and participants gave their written consent at the beginning of the study.

Participants were randomly assigned to one of two experimental conditions, which differed in whether the reading environment enabled a simultaneous or imposed a sequential presentation of documents on the multi-touch table. 54 participants were serving in the simultaneous condition and 52 participants in the sequential condition.

4.3.2 Task and Documents

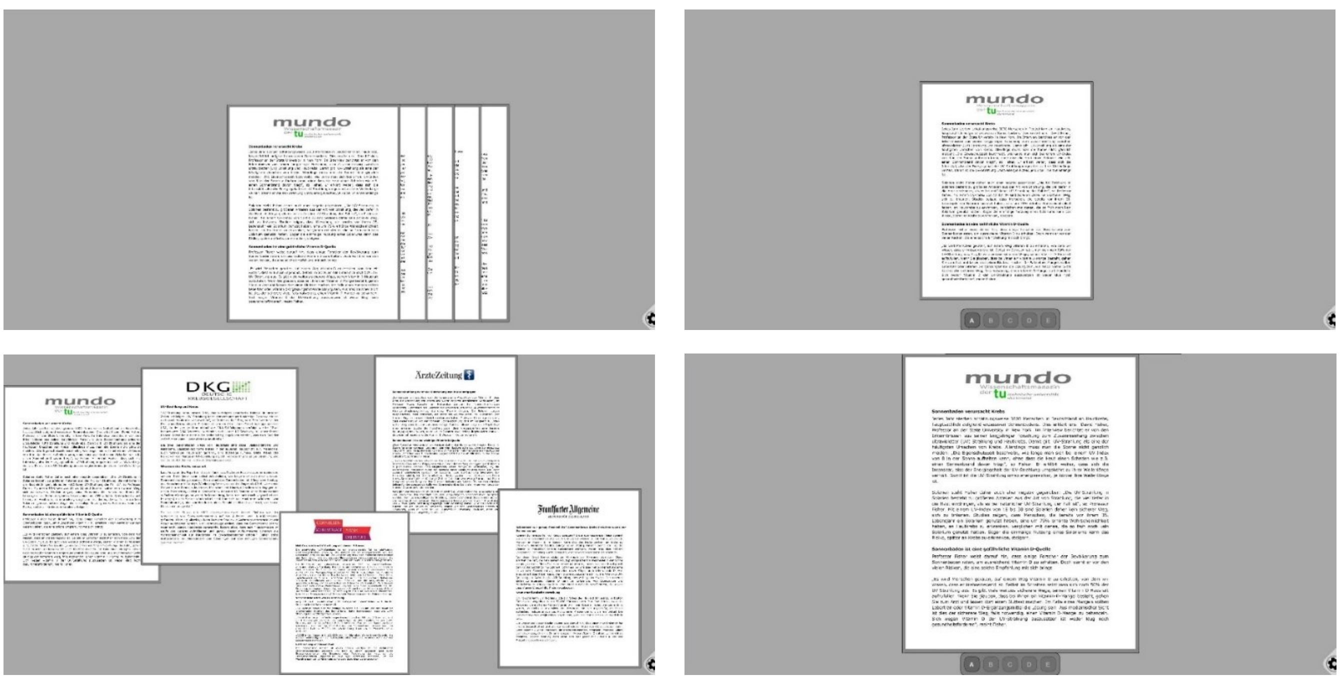
Participants' task was to inform themselves about the health effects of UV radiation by reading five documents that were provided on the multi-touch table in order to, afterwards, write an argumentative essay about the topic as well as answer several questions. The partly conflicting documents were slightly adapted (and translated) versions of documents used in previous studies (Author(s), 2021; Ferguson & Bråten, 2013; Leroy et al., 2021; Lombard et al., 2021; Strømsø et al., 2016). The source of each document was indicated by means of a source logo at the top of each document. The sources of the two documents reporting positive health effects of UV radiation (through the promotion of vitamin D production) were a reputable liberal-conservative German newspaper ('FAZ – Frankfurter Allgemeine Zeitung') and a medical science magazine ('Ärztezeitung'); the sources of the two documents reporting negative health effects of UV radiation (e.g., promoting skin cancer) were a university research magazine ('Mundo – Wissenschaftsmagazin der TU Dortmund') and the German national cancer association ('DKG – Deutsche Krebsgesellschaft'). The source of the neutral document providing general information about UV radiation was a German school book publisher ('Cornelsen'). Each document was structured into two or three sections, which were separated by subheadings. The documents contained an average of $M = 352.40$ words ($SD = 27.32$) and $M = 20.80$ sentences ($SD = 1.64$). The LIX scores (cf. Björnsson, 1968) computed for each text revealed comparable readability and that documents were not too easy to read ($M = 52.06$, $SD = 2.65$), which, however, can be presumed appropriate for university students' reading skills.

4.3.3 Documents Presentation

Participants read the documents on a multi-touch table (ARV EDGE 55'' 4K; resolution: 3840x2160 px, 1.88x1.06 meter). The interface either enabled a simultaneous or imposed a sequential presentation of the five documents. All documents were initially displayed in A4 size but could be freely scaled by participants using a zooming-gesture (i.e., pinching). However, in the sequential condition, each document was displayed in A4 size again upon re-opening. In the simultaneous condition, documents could furthermore be re-arranged freely by using a dragging gesture. While in the sequential condition documents were always displayed in the center of the bottom of the screen, documents were initially stacked in the simultaneous condition with the text of only one document (the topmost document) being visible (see Figure 4.1). Specifically, the document initially presented as the topmost one in the simultaneous condition was presented in the same position as the documents in the sequential condition when they were opened. The purpose of initially presenting documents in a stack in the simultaneous condition rather than distributing them on the screen with all documents being initially visible (cf. Leroy et al., 2021) was to investigate a situation where multiple documents *can* be presented simultaneously, but are not presented simultaneously *per default*. In both presentation conditions, counterbalanced across participants, documents were presented in one of two alternating document orders. In both document orders, the document presented in the third position provided general information about UV radiation without taking a stance. Documents in the first and fifth position took a positive stance in one document order, and a negative stance in the other document order, whereas documents in the second and fourth position took the respective opposing stance. Overall, 27 participants in the simultaneous and 27 participants in the sequential condition were presented a positive document first, and 27 and 25 participants in the respective conditions were presented a negative document first.

Figure 4.1

Simultaneous (left) and sequential (right) experimental user interface. Figures in the top row represent the initial state; figures in the bottom row represent a potential manipulation in the respective condition.



4.3.4 Multiple Document Comprehension Measures

We assessed participants' multiple document comprehension with four dependent variables. First, we assessed participants' intertextual integration through the number of intertextual connections included in their argumentative essays (e.g., (Kobayashi, 2009; Leroy et al., 2021; List, Stephens, et al., 2019; Salmerón et al., 2018a). Second, we assessed participants' spontaneous sourcing in essays as measure of their source memory (e.g., List & Alexander, 2018). Third, since *spontaneous* sourcing might underestimate participants' source memory, participants' memory for document sources was furthermore assessed in a free recall task, in which they were asked to list as many of the five document sources (i.e., the names provided by logos at the top of each document) as they remembered (Kammerer, Meier, et al., 2016). Fourth, to assess participants' source-content integration we used a source-content-mapping task, that is, a task that requires participants to map which information stemmed from what source (e.g., (Delgado et al., 2020; Kammerer, Meier, et al., 2016; Stang Lund et al., 2019; Strømsø et al., 2010).

As in previous work (Author(s), 2021; Leroy et al., 2021), we coded one intertextual connection in essays for each statement that connected information clearly stemming from different documents. Such connections could be made between conflicting (i.e., information stemming from one positive and one negative document) or complementary (i.e., information stemming from two documents taking the same stance or from the neutral and any other document) pieces of information, and could also span several sentences when the connection was semantically apparent. For examples of intertextual connections, see (Author(s), 2021). Two raters independently coded 20 (18.87%) essays for intertextual connections and reached an interrater agreement of 80.92%, which was calculated in a point-wise manner (i.e., agreement was coded when both raters gave a point for the respective statement, and disagreement was coded when only one rater had coded an intertextual connection for the respective statement). Disagreements were discussed and resolved before one rater proceeded with coding of the remaining 86 essays.

Two independent raters judged the correctness of the source names listed by all 106 participants (i.e., they evaluated whether the given name was close enough to the correct, full name of a document's source). Overall, the 106 participants wrote down 309 source names and interrater agreement of their correctness was 93.83%. Disagreements were resolved through discussion.

Finally, for the source-content mapping task, participants were presented with a 2-column table containing source logos of all five documents in alphabetic order in the first column and main statements of each document in the second column (see Figure 4.2). The statements were initially presented in one defined order in which none was in its correct position. Participants' task was to re-arrange (i.e., drag and drop) statements, such that source logos and main statements of each document were positioned in the same row of the table. Because once four statements were correctly positioned, the fifth was automatically correct too, the maximum score in the source-content mapping task was 4.

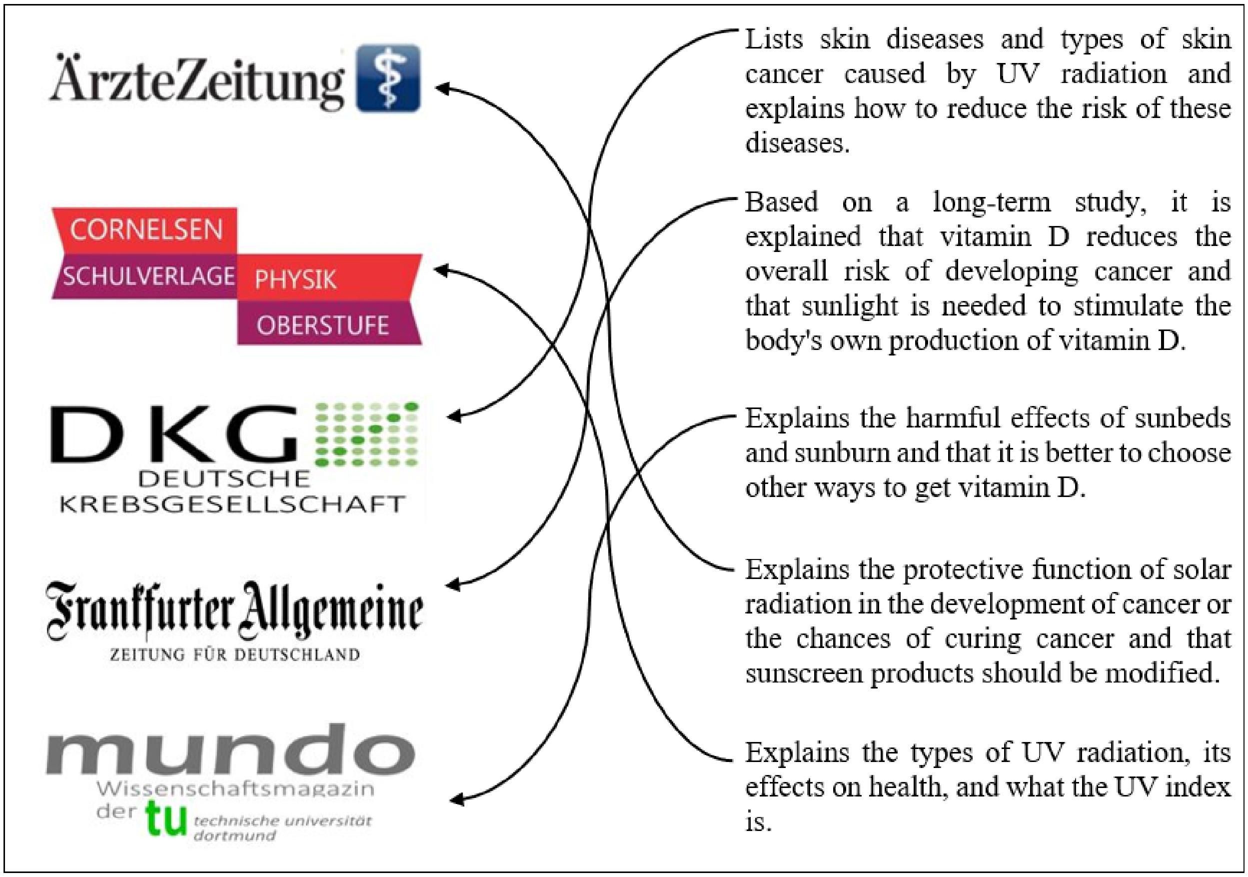
4.3.5 Reading Interactions

Regarding the reading interactions assessed for participants in the simultaneous condition, based on screen recording videos that were recorded for the whole task processing time of each participant, two independent raters evaluated whether the participant had grouped documents according to their overall stance during reading (i.e., whether the two positive documents and/or the two negative documents were grouped). Interrater agreement of all 54 screenrecordings was 92.59 %, and disagreements were resolved through discussion.

Regarding the reading interactions assessed for participants in the sequential condition, we used interaction log files to assess the number of revisits to documents. As in previous research (Leroy et al., 2021), we used a threshold for document openings of 1 sec to exclude short, presumably accidental (re-)openings from our analyses. Consequently, a revisit of a document was counted for each opening that lasted at least 1 sec after a first opening that lasted at least 1 sec. Take, for example, the following partial document opening sequence: A (52.0 sec) – B (0.9 sec) – A (3.2 sec) – B (73.4 sec) – A (11.8 sec). In a first step, we cleared the sequence of all openings with a duration less than 1 sec, resulting in A (52.0 sec) – A (3.2 sec) – B (73.4 sec) – A (11.8 sec). In a second step, we checked for "openings" of the same document in direct succession, which were then accumulated into one opening (i.e., one first reading or one revisit). In the given example, this results in the following final sequence of A (52.0 sec + 3.2 sec) – B (73.4 sec) – A (11.8 sec) with one revisit of document A.

Figure 4.2

Items of the source-content-mapping task (translated from German). The initial order of source logos (left) and statements (right) was as shown here. Participants' task was to drag the statements vertically such that their position matched that of the source logos. Arrows indicate the correct assignments.



4.3.6 Control Variables

To ascertain comparability across experimental conditions, we assessed participants' working memory capacity (15-item computer-based reading span task; cf. Kane et al., 2004), prior topic knowledge (see below), as well as their topic interest (1 item, from 1=very low to 7=very high) and prior topic beliefs (2 items; one addressing positive and another addressing negative effects of UV-radiation on health; from 1=totally disagree to 7=totally agree) on 7-point Likert scales. In order to gain one score for participants' prior topic beliefs, we reverse-coded the negative item and averaged both values (Cronbach's alpha = .76).

Prior topic knowledge was assessed with an essay task prior to reading the documents, in which participants were asked to write down within 5 minutes everything they knew about the potential effects of UV-radiation on human health. Participants were awarded one point for each relevant and correctly reported information (e.g., UV-radiation stems from the sun; sunbeds use UV-radiation; there are 3 types of UV-radiation) or argument concerning health effects of UV-radiation for humans (e.g., UV-radiation causes skin cancer; UV-radiation promotes the production of vitamin D in the body) mentioned in their essay. Two independent raters scored all 106 prior knowledge essays. Interrater agreement for point-wise scoring was 95.56% (i.e., 95.56% of the points were awarded by both raters). Disagreements were resolved through discussion.

4.3.7 Procedure

Participants were tested in single sessions (of approximately 1 h) in the lab. They first completed a demographic questionnaire and the topic knowledge measure using a laptop. Subsequently, participants received a written explanation of the interaction possibilities on the multi-touch table (i.e., how to open and scale documents in the sequential condition, and how to scale and re-arrange documents in the simultaneous condition) and practiced interaction with five blank documents in the respective experimental condition. Once they felt comfortable in operating the interface, they received written task instructions on screen. Specifically, they were told that in the following they would be provided with five documents from the Internet, which they should read carefully, in order to subsequently write an argumentative essay about potential health effects of UV radiation without having the documents available. They were also informed that they had a maximum of 15 minutes to read the documents and that within this time frame they could read the documents as often as they wanted. After reading, participants were asked to write the argumentative essay on a laptop within 15 minutes. They were told to include arguments from the documents in their essay to justify their statements. Afterwards, also on the laptop, they were asked to recall the documents' sources, and to complete the source-content mapping task.

4.4. Results

Overall, we assessed four different measures of multiple document comprehension as dependent measures, namely the number of intertextual connections in essays, the number of source citations in essays, the number of correctly recalled document source names in a free recall task, and the score in a source-content mapping task. However, analyses revealed that only 6 participants (11.11%) in the simultaneous condition and 5 participants (9.62%) in the sequential condition cited at least one document source in their essays. Thus, we refrained from analyzing this measure in the present study. In the following, the results regarding the remaining three measures of multiple document comprehension are reported.

4.4.1 Differences Regarding Control Variables and Multiple Document Comprehension (RQ1) Between Experimental Conditions

Regarding control variables, Welch's t-test with presentation condition as independent variable showed that participants in the simultaneous and sequential condition did not differ with respect to prior topic knowledge, $t(98.65) = 0.76$, $p = .447$, or working memory capacity, t(99.14 = 1.36, $p = .177$. They also did not differ in their topic interest, $t(101.79) = 0.19$, $p = .848$, or in their prior beliefs regarding health effects of UV radiation, $t(102.43) = 0.16$, $p = .874$, which on average were rather negative (i.e., participants, on average, were more aware of negative health effects of UV radiation). Furthermore, neither the overall time participants spent reading the documents, $t(103.71)$, $p = .357$, nor the length of their essays, $t(101.57) = .217$, differed between presentation conditions. Table 4.1 shows means (and standard deviations) for these measures for each presentation condition.

To answer RQ1, we also conducted Welch's t-tests with document presentation as independent variable and the three multiple document comprehension measures as dependent variables (see Table 4.1 for means and standard deviations). However, no significant differences between the simultaneous and sequential condition were shown for any of the three measures of multiple document comprehension (number of intertextual connections: $t(103.55) = 0.89$, $p = .373$; number of source names recalled: $t(103.19) = 0.32$, $p = .752$; source-content mapping score: $t(104.00) = 0.05$, $p = .959$).

Subsequently, to answer RQ2 and RQ3 about the relation between participants' interaction behavior and their multiple document comprehension for the two presentation conditions separate analyses were conducted and are reported in the following.

Table 4.1

Means (and standard deviations) for control variables and multiple document comprehension measures as a function of document presentation.

	Document presentation	
Variables	Simultaneous (n = 54)	Sequential (n = 52)
Prior topic knowledge (from 1=very low to 7=very high)	4.30 (1.79)	4.00 (2.18)
Working memory capacity (0-15)	11.28 (2.33)	10.60 (2.80)
Topic interest (from 1=very low to 7=very high)	4.24 (1.41)	4.19 (1.17)
Prior topic beliefs (negative item reversed; averaged; from 1=very negative to 7=very positive)	2.77 (1.17)	2.73 (1.27)
Time on task [min]	12.09 (2.56)	12.53 (2.33)
Essay length [# words]	262.48 (75.98)	282.00 (85.46)
# Intertextual connections in essays	3.42 (1.66)	3.42 (1.50)
# Source names recalled (0-5)	2.17 (1.44)	2.25 (1.27)
Source-content mapping score (0-4)	2.28 (1.09)	2.29 (1.05)

4.4.2 The Role of Grouping Documents During Reading for Multiple Document Comprehension (RQ2)

Twenty participants (37.04%) in the simultaneous condition had grouped documents according to their overall stance during reading. Eighteen out of those twenty participants had grouped the two positive as well as the two negative documents, one participant had only grouped the positive documents, and one participant had only grouped the negative documents. Table 4.2 provides means and standard deviations of all control variables as well as the three multiple document comprehension measures for participants who had grouped and not grouped documents during reading, and respective results of the mean comparisons (Welch's t-tests).

Regarding the three multiple document comprehension measures, participants who had grouped documents during reading included significantly more intertextual connections in their essays and scored significantly higher in the source-content mapping task. However, there was no significant difference in the number of source names correctly recalled between participants who had and had not grouped documents (see Table 4.2).

Table 4.2

Descriptive and inferential statistics for all measured variables in the simultaneous condition as a function of having or not having grouped documents during reading.

Variables	Having grouped documents		Welch's t-tests
	Yes (n = 20)	No (n = 34)	
Prior topic knowledge	4.95 (1.73)	3.91 (1.73)	$t(39.94) = 2.13$, $p = .040$
Working memory capacity (0-15)	11.90 (2.00)	10.91 (2.45)	$t(46.58) = 1.61$, $p = .114$
Topic interest (from 1=very low to 7=very high)	4.25 (1.33)	4.24 (1.48)	$t(43.37) = 0.04$, $p = .970$
Prior topic beliefs (negative item reversed; averaged; from 1=very negative to 7=very positive)	2.98 (1.33)	2.65 (1.06)	$t(33.20) = 0.94$, $p = .355$
Time on task [min]	12.50 (2.37)	11.86 (2.67)	$t(43.91) = 0.91$, $p = .367$
Essay length [# words]	291.80 (67.66)	245.24 (76.22)	$t(43.88) = 2.33$, $p = .025$
# Intertextual connections in essays	3.90 (1.71)	2.71 (1.49)	$t(35.56) = 2.59$, $p = .014$
# Source names recalled (0-5)	2.55 (1.15)	1.94 (1.56)	$t(49.22) = 1.65$, $p = .106$
Source-content mapping score (0-4)	2.65 (0.93)	2.06 (1.13)	$t(46.05) = 2.08$, $p = .043$

Regarding control variables, participants who had grouped documents during reading had significantly higher prior topic knowledge than those who had not grouped documents during reading. However, neither measure of multiple document comprehension correlated significantly with prior topic knowledge (number of intertextual connections: $r = .14$, $p = .323$, $n = 54$; number of source names recalled: $r = .04$, $p = .779$; source-content mapping score: $r = .18$, $p = .193$, $n = 54$). There was no significant difference between those who had and had not grouped documents in any of the remaining control variables (also see Table 4.2). Also, participants who had grouped and not grouped documents during reading did not differ in their overall time spent reading the documents. Yet, participants who had grouped documents during reading wrote significantly longer essays than those who had not grouped documents during reading (also see Table 4.2).

4.4.3 The Role of Number of Document Revisits for Multiple Document Comprehension (RQ3)

Five participants (9.62%) in the sequential condition did not make any revisits. The remaining 47 participants, on average, made 10.28 revisits ($SD = 8.22$). For the complete sample of 52 participants, Table 4.3 provides correlational analyses between the number of revisits to documents during reading, on the one hand, and control variables and multiple document comprehension measures, respectively, on the other hand.

Regarding multiple document comprehension measures, there was a significant positive correlation between the number of revisits and both the number of source names correctly recalled and the score in the source-content mapping task. In contrast, there was no significant correlation between the number of revisits and the number of intertextual connections in essays (see Table 4.3). The number of revisits was not significantly correlated to any of the control variables. Finally, the number of revisits was positively correlated to the time participants took for reading, but not to essay length (also see Table 4.3).

Table 4.3

Person's correlations between the number of revisits and the respective measured variables in the simultaneous condition.

Variables	Pearson's correlations with # revisits
Prior topic knowledge	$r = -.19, p = .175$
Working memory capacity (0-15)	$r = .12, p = .402$
Topic interest (from 1=very low to 7=very high)	$r = .17, p = .219$
Prior topic beliefs (negative item reversed; averaged; from 1=very negative to 7=very positive)	$r = -.04, p = .757$
Time on task [min]	$r = .41, p = .003$
Essay length [# words]	$r = -.00, p = .995$
# Intertextual connections in essays	$r = .09, p = .533$
# Source names recalled (0-5)	$r = .37, p = .007$
Source-content mapping score (0-4)	$r = .32, p = .020$

4.5 Discussion

With the present study we aimed to extend previous work examining whether a simultaneous as compared to a sequential presentation of multiple documents fosters multiple document comprehension (Andrews et al., 2010; Ball & North, 2005; Bi & Balakrishnan, 2009; Czerwinski et al., 2003; Hutchings & Stasko, 2004; Leroy et al., 2021; Olive et al., 2008; Wiley, 2001). We assessed different aspects of multiple document comprehension, that is, intertextual integration, source memory, and source-content integration, and also investigated the role that reading interactions played in these aspects. To our knowledge only one previous study (Leroy et al., 2021) had investigated the effect of document presentation on multiple document comprehension by also taking reading interactions into account. However, readers' intertextual integration, that is, the number of statements in their essays that combined information from two documents, was assessed as the sole measure of multiple document comprehension in that study. Thus, multiple document comprehension measures reflecting readers' mental representation of source information or of source-content integration were not considered in Leroy et al.'s (2021) study.

4.5.1 Multiple Document Comprehension When Reading Multiple Documents in a Reading Environment (not) Enabling Simultaneous Presentation

Other than expected, the simultaneous condition did not foster participants' intertextual integration. Furthermore, exploratory analyses revealed that also participants' memory for source names and their source-content integration did not differ across presentation conditions. Regarding intertextual integration, these findings are in contrast to those by Leroy et al. (2021), who found that reading documents in the simultaneous condition yielded a significantly higher number of intertextual connections in essays written after reading (i.e., the same measure of intertextual integration as was used in the present work) than reading documents in the sequential condition.

One potential post hoc explanation for the inconsistent findings might be that other than in Leroy et al.'s (2021) study, in the simultaneous condition documents were initially presented on a stack in the present study rather than spread out on the screen from the start. That is, the simultaneous condition of the present study *enabled* a simultaneous presentation of documents, yet readers had to actively "create" a simultaneous presentation of documents by spreading out the initially stacked documents on the screen. Thus, only for those readers who actually made use of the possibility of creating a simultaneous document presentation, potential beneficial effects of a simultaneous presentation could have been brought to bear.

Second, while documents were presented in random order for each participant in the study by Leroy et al. (2021), they were presented in a specific alternating order in the present study, with conflicting documents being presented in direct succession. While the document order was only initially given in the simultaneous condition (because documents could be re-arranged, which might have changed their order), it was rather static in the sequential condition, in which documents had to be opened from a menu. Hence, the given order might have played a more important role in the sequential condition than in the simultaneous condition. As previous work showed, the order of partly conflicting documents can affect readers' mental representation of multiple documents and their relations (Braasch et al., 2021; Maier & Richter, 2013). Thus, reading documents that take an opposite stance in direct succession may have enhanced readers' conflict awareness, hence supporting their intertextual integration. In consequence, this might have diminished any detrimental effect of a sequential as compared to simultaneous presentation of multiple documents. In the same vein, an increased conflict awareness in the sequential condition due to the alternating document order might also have fostered participants' attention to sources (cf. Braasch & Bråten, 2017). This, in turn, might have diminished any potential detrimental effects of a sequential as compared to simultaneous condition on readers' memory for sources or their source-content integration. Of course, these assumptions are speculative in nature and will need to be examined in future research.

4.5.2 The Role of Grouping Documents for Multiple Document Comprehension in a Reading Environment Enabling a Simultaneous Presentation of Documents

As in previous work (Leroy et al., 2021), participants' strategic re-organization, that is, their grouping of documents in the simultaneous condition, was positively related to their intertextual integration. This may be due to the spatial contiguity resulting from grouping documents which "can make the relations between multiple [documents] more salient" (Barzilai & Ka'adan, 2017, p. 199; see also Ginns, 2006) and thus support readers in cross-textual information comparison and integration (cf. Leroy et al., 2021). Regarding readers' memory for source information as a second measure of multiple document comprehension, the present study revealed that regardless of whether participants had grouped documents during reading, they only recalled about two document sources. This is in line with previous research showing that readers typically only pay little attention to source information (e.g., Braasch & Bråten, 2017; Britt et al., 2004; Le Bigot & Rouet, 2007; Rouet et al., 1996). However, participants who had grouped documents according to their overall stance were better able to map main statements to the respective source names than were participants who had not grouped documents. In conclusion, readers' spontaneous engagement in grouping documents might not be related to their overall attention to sources, but rather support integration processes within and across documents (cf. Ginns, 2006), yielding better intertextual integration as well as source-content integration (i.e., having mentally represented which source said what). However, further research is needed to understand the differences in integration processes between readers who do and do not (spontaneously) group documents during reading.

4.5.3 The Role of Document Revisits for Multiple Document Comprehension in a Reading Environment Imposing a Sequential Presentation of Documents

A different pattern of results was shown for the relation between multiple document comprehension and the number of revisits to documents in the sequential condition. As in previous work (Author(s), 2021; Leroy et al., 2021), the number of revisits to documents was not related to participants' intertextual integration. In contrast, the number of revisits to documents was positively related to both multiple document comprehension measures related to source information, that is, to participants' memory for source names, and to participants' source-content integration, that is, their performance in the source-content mapping task. That is, the more revisits readers made during reading, the more document sources they recalled after reading and the better they performed in the source-content mapping task. These findings corroborate previous work by List and Alexander (2018) regarding source citations and by Author(s) (2021) and Delgado et al. (2020) regarding source-content integration, and indicate that during revisits readers (also) pay additional attention to source information. Since the document set in the present study was partly conflicting, this is in line with previous work showing that conflicting information fosters readers' attention to source information (e.g., Braasch et al., 2012; Braasch & Bråten, 2017). That is, when encountering information that stands in conflict to previously read information, readers might revisit the previously read document, possibly in an attempt to restore coherence in their overall mental representation of the subject matter (Braasch et al., 2012; Braasch & Bråten, 2017). This reasoning is also in line with previous assumptions that revisits might be initiated in an attempt to re-evaluate information in the light of information found in another document (cf. Anmarkrud et al., 2014; Bråten et al., 2013; Cerdán & Vidal-Abarca, 2008; List & Alexander, 2018; Wineburg, 1991).

4.5.4 Conclusion

In conclusion, while document presentation did not affect multiple document comprehension, the two reading interactions assessed in the present study (i.e., grouping documents in the simultaneous condition and the number of revisits to documents in the sequential condition) were differentially related to different measures of multiple document comprehension. Overall, the findings of the present study suggest that re-organizing partly conflicting documents into groups according to their overall stance is positively related to integrative processes within and across documents rather than to readers' attention to source information per se. Conversely, regarding the number of revisits to documents (which might be initiated in order to restore coherence after encountering conflicting information) findings of the present study rather suggest a positive relation to readers' mental representation of source information, but not to integrative processes.

4.4.5 Limitations and Future Work

We acknowledge that the present research comes with certain limitations. First, we examined a homogenous sample of university students, which can be regarded as proficient readers who might engage more strategically in reading interactions than non-proficient readers (cf. (Britt & Rouet, 2012; Goldman, Braasch, et al., 2012; Rouet et al., 2017). Hence, whether our findings extend to other samples, such as readers with lower reading skills, remains an open question that we would like to see addressed in future research – especially when keeping in mind that, for example, school students could benefit from trainings in how to engage with reading materials in effective ways.

Second, each reading interaction could only be assessed in one presentation condition. In the simultaneous condition it was not possible to assess the number of revisits to documents (i.e., from logfiles), since a re-inspection of information could have also occurred without overt interaction with the document. Furthermore, participants' engagement in grouping documents could not be assessed in the sequential condition, since the re-arrangement of documents was not possible in this condition in the present study, just like in most sequential reading environments in which documents usually cannot be spatially re-arranged on the screen either.

Third, it remains unclear from the present study which intentions participants had in engaging in the respective reading interactions. For instance, eye-tracking data or (concurrent or retrospective) think-aloud protocols could help shed light on whether readers who group documents during reading do so directly after realizing that documents support or contradict each other. Furthermore, eye-tracking data might also help shed light on how the grouping of documents affects readers' further reading process with regard to information integration within and across documents. The use of eye-tracking methodology would also allow to investigate revisits to documents that are presented simultaneously on the screen. In that case, both in a simultaneous and in a sequential document presentation, think-aloud methodology could also help shed light on readers' intentions in revisiting documents (e.g., Anmarkrud et al., 2014; Bråten et al., 2014; Goldman et al., 2012; Wineburg, 1991).

Finally, we acknowledge that our findings regarding the relation between reading interactions and multiple document comprehension measures are only correlational in nature. Thus, while the positive relations between reading interactions and specific aspects of multiple document comprehension suggest that training readers in strategically performing such reading interactions might support them in building a comprehensive mental representation of multiple documents, further experimental work is needed to test this assumption. Notwithstanding these limitations, the findings of the present study indicate that reading interactions do play an important role in multiple document comprehension. Thus, we encourage future research to take into account reading interactions when investigating the effect of different reading interfaces or tools.

5. *Study III*: Reading Multiple Documents on a Health-Related Issue: The Roles of a Text-Highlighting Tool and Re-reading Behavior on Integrated Understanding

The content of the following chapter has been submitted for publication to *Behaviour and Information Technology*. The proportional contributions of all co-authors to the manuscript are presented in the subsequent table. This article may not exactly replicate the final version published in the journal. It is not the copy of record.

Author	Author position	Scientific ideas	Data generation	Analysis interpretation	Paper writing
Caroline Leroy	1st author	70 %	90 %	80 %	65 %
Yvonne Kammerer	2nd author	30 %	10 %	20 %	35 %

Title of paper:	Reading Multiple Documents on a Health-Related Issue: The Roles of a Text-Highlighting Tool and Re-Reading Behaviour on Integrated Understanding
Status in publication process:	Submitted for publication: *Behaviour and Information Technology*

Leroy, C. & Kammerer, Y. (submitted for publication). Reading multiple documents on a health-related issue: The roles of a text-highlighting tool and re-reading behaviour on integrated understanding

Abstract

This study aimed to investigate the roles of a text-highlighting tool and readers' re-reading behaviour in their integrated understanding of multiple documents. University students (N = 95) read five partly conflicting documents on a health-related issue on a touch display with or without a text-highlighting tool. Integrated understanding of documents was assessed by the number of intertextual connections in essays written after reading and by a source-content mapping task. The provision of the text-highlighting tool resulted in longer initial reading times even when subtracting the time taken for highlighting, but shorter re-reading times, particularly for participants with a high number of re-readings. Further, only for participants with a high number of re-readings, the provision of the text-highlighting tool resulted in more intertextual connections than when no text-highlighting tool was provided. Participants' source-content integration was positively related to the number of re-readings, regardless of whether the text-highlighting tool was provided. Finally, additional exploratory eye-tracking analyses revealed that for two out of the five documents, participants in the with-highlighting condition focused on significantly smaller parts of the documents during re-reading than controls.

Keywords: text-highlighting tool, multiple documents, multiple document integration, re-reading behaviour, eye-tracking

5.1 Introduction

In today's information society, due to the ease of accessing a multitude of documents on almost any topic, reading multiple documents on a particular issue has become a common reading task. For example, when turning to the Internet to inform oneself about the health effects of ultraviolet (UV) radiation, multiple documents stemming from various sources, such as different publishers of newspapers, magazines, or books, can provide consistent, complementary, or even conflicting information on the issue at hand (Bråten, Anmarkrud, et al., 2014). Thus, to gain an overarching understanding of the issue addressed in multiple documents, readers do not only need to understand information *within* individual documents, but also to compare, evaluate, and integrate information *across* multiple potentially conflicting documents (e.g., Britt & Rouet, 2012; Mahlow et al., 2020; Primor & Katzir, 2018; Wineburg, 1991). This is a challenging task for most readers (e.g., Kiili & Leu, 2019; List et al., 2019), and failure to integrate information across multiple documents can lead to an incomplete or even one-sided understanding of complex subject matter (e.g., Anmarkrud et al., 2014; List & Du, 2021). Therefore, during the past few years, research into characteristics of the reading environment that can foster multiple document integration has increased considerably (e.g., Britt et al., 2004; Haber et al., 2014; Lombard et al., 2018; Margolin et al., 2013; Olive et al., 2008; Salmerón et al., 2009, 2010, 2018; Salmerón & Llorens, 2019). Reading environments, for instance, can differ with respect to whether they provide computer-based tools that make it possible to highlight or annotate text (for recent overviews of support tools investigated, see, e.g., Barzilai et al., 2018; Britt & Rouet, 2012). However, the beneficial effects of such tools, in general, might depend on how readers interact with these tools or how they interact with the reading environment as a whole, that is, which reading actions readers perform while reading (e.g., Bråten, Ferguson, et al., 2014; Du & List, 2020; Ferguson & Bråten, 2013; Freund et al., 2016; Goodwin et al., 2020; List & Alexander, 2017, 2020; Rouet et al., 2017).

In line with this reasoning, a recent study by Leroy et al. (2021) has shown that the provision of a text-highlighting tool supported readers' integrated understanding of multiple partly conflicting documents; however, this was true only for readers with a higher number of re-accesses to previously read documents. In contrast, for participants who infrequently re-accessed previously read documents, no difference was found between those who were provided with a text-highlighting tool and those who were not. The study presented here builds on that previous work by Leroy et al. (2021), extending their experimental approach in two main ways: First, by means of eye-tracking methodology, we aimed to define and examine readers' re-reading behaviour in greater detail. Second, we assessed the same measure of integrated understanding as Leroy et al. (2021), as well as one additional integrated understanding measure (i.e., one measure addressing intertextual integration and a second measure addressing source-content integration). In so doing, we aimed to obtain a more complete picture of the interplay between the provision of a text-highlighting tool and readers' re-reading behaviour on their integrated understanding of multiple documents.

In what follows, we first outline theoretical considerations of how readers integrate multiple documents, and of what role re-reading previously read documents might play in gaining an integrated understanding. Then, we review literature on theoretical considerations of why text-highlighting might benefit text comprehension and discuss why it might especially benefit readers who engage in re-reading. We also discuss how the re-processing of previously read documents might be different when text can be highlighted during reading versus when it cannot.

5.2 Theoretical and Empirical Background

5.2.1 Readers' Integrated Understanding of Multiple Documents

According to the Documents Model framework (DMF; Britt et al., 1999; Britt & Rouet, 2012; Perfetti et al., 1999), an adequate mental representation of multiple documents, a so-called *documents model*, comprises information on how statements from different documents relate to each other or how documents as a whole relate to each other, respectively (i.e., intertextual integration) and information about which content stems from which source (i.e., source-content integration). Thus, to build an integrated understanding of multiple documents, readers need to compare, contrast, connect, and evaluate information from different documents and sources (e.g., Afflerbach & Cho, 2009; Barzilai et al., 2018; List & Alexander, 2018), which might be achieved by engaging in reading actions, such as re-reading (parts of) previously read documents (e.g., Anmarkrud et al., 2014; Goldman et al., 2012; Latini et al., 2020; List & Alexander, 2017; Wineburg, 1991).

Furthermore, as an extension of the DMF, the RESOLV (Reading as problem SOLVing) model (Britt et al., 2018; Rouet et al., 2017) suggests that the effects of characteristics of the reading environment (e.g., whether or not particular computer-based support tools are provided) on readers' integrated understanding of multiple documents depend on the reading actions they perform and vice versa. In support of this notion, in the abovementioned study by Leroy et al. (2021), the provision of a text-highlighting tool yielded better intertextual integration for readers with a higher number of re-accesses to previously read documents, but not of readers with a low number of re-accesses. Likewise, only when the text-highlighting tool was provided, the extent of re-accessing previously read documents was positively related to readers' intertextual integration, but not when no text-highlighting tool was provided.

Regarding readers' source-content integration, to our knowledge, no previous study has examined the interplay between text-highlighting and re-reading behaviour. Yet, also supporting the assumption made by the RESOLV model that reading behaviours play an important role for readers' integrated understanding, two previous studies are indicative that re-reading may also be positively related to source-content integration. In a study by List and Alexander (2018), participants who had re-accessed at least one previously read document included more source citations in essays written after reading six documents on a political issue than participants who had not re-accessed any document. Hence, during re-reading, participants may have paid additional attention to source information, which is an integral part of source-content integration. Furthermore, Delgado et al. (2020) asked participants to read four documents on the effects of UV radiation on health in a hypertext environment. Results showed that the more often participants opened hyperlinked pop-up windows that provided hints about conflicting information in another document together with the source of that document, the better their source-content integration was after reading. However, whether source-content integration is also positively related to the extent of re-reading previously read documents when no hints about conflicting information across documents are provided remains an open question which is addressed in the present study.

5.2.2 Mechanisms of Text-Highlighting

Since the seminal work of Fowler and Barker (1974), which provided first indications that active text-highlighting supported text comprehension, two distinct mechanisms have been discussed as to why (see e.g., Dunlosky et al., 2013; Winchell et al., 2020): First, active text-highlighting might serve an encoding function, because readers have to make active decisions about which parts of the text to highlight (Ponce & Mayer, 2014; Yue et al., 2014). Second, highlighted text might serve a retrieval function, because highlighted text stands out (or "pops out"; e.g., Leroy et al., 2021; Strobelt et al., 2016; Winchell et al., 2020). It can thus be relocated and reviewed more easily than non-highlighted text (e.g., Chi et al., 2007; Yeari et al., 2017). Furthermore, Yeari et al. (2017, p. 2) assume that during re-reading readers "presumably focus mostly on the highlighted parts and thus save time and effort in reprocessing the textual content". Hence, text-highlighting should be particularly beneficial for comprehension when readers engage in *re-reading*. Overall, however, it should be noted that empirical evidence for the beneficial effects of text-highlighting for comprehension is rather inconclusive (cf. Dunlosky et al., 2013; Miyatsu et al., 2018). This might be due to the existence of factors that moderate the effectiveness of text-highlighting (e.g., Ben-Yehudah & Eshet-Alkalai, 2018; Goodwin et al., 2020; Kobayashi, 2009; Li et al., 2016; for recent overviews, see Dunlosky et al, 2013; Winchell et al., 2020). One such moderating factor might be the extent to which readers engage in re-reading (Leroy et al., 2021).

5.2.3 The Effects of Text-Highlighting on Integrated Understanding

In the context of multiple document comprehension, to our knowledge only three previous studies examined the effects of text-highlighting (Kobayashi, 2009; Leroy et al., 2021; Li et al., 2016). In the study by Kobayashi (2009)[2], participants were asked to read in a self-paced manner six documents on the introduction of English education into elementary schools in Japan, and to find relations between texts. Afterwards, they had to write an argumentative essay on the topic as well as to recall as many arguments from the documents as possible. While they were reading, participants were either allowed or not allowed to use external reading strategies such as text-highlighting, underlining, or annotating which make information "pop out". Among the participants who were allowed to use external reading strategies, text-highlighting was the most frequently used strategy. Regarding comprehension outcomes, results showed that the groups did not differ in the number of arguments they recalled from individual texts. Yet, participants who were allowed to use external reading strategies included more intertextual connections in their essays than participants who were not. In contrast, in a study by Li et al. (2016) the provision of a text-highlighting tool neither had an effect on participants' recall of keywords from individual texts, nor on their intertextual integration when reading a set of hypertext documents on the topic of cancer. While these inconclusive findings regarding readers' intertextual integration might be due to methodological differences across these two studies (i.e., for example hypertext in the study by Li et al. versus non-hypertext in the study by Kobayashi, see Freund et al., 2016), the possibility that highlights might be especially beneficial *during re-reading* (see Section 2.2) was not addressed

[2] Please note that participants received one of two reading instructions, namely, to form an opinion or to find relations between texts. However, Kobayashi (2009) only found differences in the relation-finding group. These results are reported here.

in either of these studies. Kobayashi (2009) did not report at all on participants' re-reading behaviour. Li et al. (2016) only reported that the overall number of accesses to documents (i.e., the length of the navigation path) did not differ between groups. It thus remains unclear from these studies whether there was a moderating role of participants' re-reading behaviour on the effectiveness of text-highlighting, as Leroy et al. (2021) found.

Specifically, Leroy et al. (2021) asked participants to read five partly contradicting documents on the effects of UV radiation on human health in order to write an argumentative essay about the issue. The number of re-accesses to documents was assessed from logfiles as an indication of how much participants had engaged in re-reading. As for the overall number of document-accesses in Li et al. (2016), Leroy et al. (2021) also reported a comparable number of document re-accesses in the with-highlighting and without-highlighting conditions. However, with respect to readers' intertextual integration (as measured by the number of intertextual connections in essays that participants wrote after reading), findings by Leroy et al. (2021) revealed a significant interaction between the possibility to highlight text and the number of re-accesses to documents. The provision of the text-highlighting tool supported intertextual integration only for readers who had re-accessed previously read documents to a considerable extent, but not for participants with only few document re-accesses. Furthermore, intertextual integration and the number of document re-accesses were positively related only for participants who were provided with the text-highlighting tool.

The finding that text-highlighting supported intertextual integration only for participants with a high number of document re-accesses might be due to the fact that highlighted information is easier to re-locate during re-accesses. This is because it *pops out* and thus might save readers time in reprocessing textual content (cf. Yeari et al., 2017). However, to our knowledge, no previous study has investigated a potential difference in the distribution of readers' visual attention during re-reading or in their re-reading time when text-highlighting is possible versus when it is not. With regard to reading times, previous research has only investigated effects of text-highlighting on the overall time devoted to the task, indicating that with the possibility to highlight text, overall reading times were longer than without the possibility to highlight text (e.g., Ben-Yehudah & Eshet-Alkalai, 2018; Johnson, 1988; Leroy et al., 2021; Li et al., 2016). This could, however, be explained by the fact that highlighting itself takes time (i.e., for deciding what to highlight and for executing the highlighting). As highlights seem to be made predominantly during initial readings of documents (Leroy et al., 2021; Peterson, 1991), these longer overall reading times might stem from longer initial reading times rather than from longer re-reading times. Thus, in the present study, we used eye-tracking methodology to distinguish between initial readings and re-readings of documents, in order to examine whether the possibility to highlight text results in shorter re-reading times and more focused (i.e., less distributed) re-readings.

5.2.4 Present Study

With the present study, first, we aimed to replicate the finding by Leroy et al. (2021) that the beneficial effect of a text-highlighting tool on intertextual integration was moderated by the extent to which readers engaged in re-reading of documents. Similar to the study by Leroy et al. (2021), participants were asked to read five one-paged, partly conflicting documents on the effects of UV radiation on human health in order to write an argumentative essay on the issue afterwards. Furthermore, we sought to extend the findings by Leroy et al. (2021) in several ways, one being the assessment of a second measure of integrated understanding, namely readers' source-content integration (i.e., their mental representation of which information stems from which source; Britt & Rouet, 2012). For this purpose, after the essay-writing task participants had to complete a source-content mapping task (also see Delgado et al., 2020; Kammerer et al., 2016; Stang Lund et al., 2017, 2019; Strømsø et al., 2010). Second, using eye-tracking methodology, we sought to examine differences in participants' re-reading behaviour when highlighting was possible versus not. That is, we aimed to further extend the study by Leroy et al. (2021) by investigating whether the time taken to re-read as well as the visual distribution on text during re-reading differed between highlighting conditions. Third, we aimed to extend Leroy et al.'s (2021) previous work by further methodological analyses. That is, Leroy et al. (2021) had defined each re-access to a previously opened document as re-reading when the duration of the document access exceeded 1 sec. Yet, since the choice of threshold directly affects the number of re-readings, it might bias the moderation effect of re-reading in the effect of text-highlighting on multiple document comprehension. Using a combination of eye-tracking and logfile data, we therefore differentiated more precisely than Leroy et al. (2021) between initial *full* readings (rather than first accesses) and re-readings of documents. In order to explore whether it was adequate to use a threshold of 1 sec to define re-readings, then, (1) we examined how many fixations were directed towards a document during short re-readings between below 1 sec and up to 5 sec (i.e., we used 5 different thresholds) and (2) tested whether the interactive effect between text-highlighting and the number of re-readings on readers' integrated understanding depended on the choice of the threshold for re-readings. Overall, we had the following research questions (RQ), hypotheses, and exploratory goals:

(RQ1) Does the number of re-readings moderate the effect of text-highlighting on readers' multiple document comprehension?

Specifically, in line with the results by Leroy et al. (2021), we expected an interaction effect between the possibility to highlight text and the number of re-readings on *intertextual integration*. That is, we expected that for participants with a higher number of re-readings, the possibility to highlight text would result in more intertextual connections than without the possibility for text-highlighting; in contrast, for participants with a lower number of re-readings, we expected no differences between the with-highlighting and without-highlighting conditions regarding the number of intertextual connections (H1a). Furthermore, we expected the number of intertextual connections in essays to be positively related to the number of re-readings for participants in the with-highlighting condition, but not for those in the without-highlighting condition (H1b). Likewise, regarding participants' source-content integration, we explored whether the provision of the highlighting-tool or the number of re-readings (cf. Delgado et al., 2020) were related to participants' score in the source-content mapping task and whether there was any similar interaction effect as was expected for participants' intertextual integration.

(RQ2) How does the possibility to highlight text affect readers' re-reading behaviour?

Assuming that participants would make most of their highlights during initial readings (Kobayashi, 2007; Leroy et al., 2021), we expected initial reading times to be longer when text-highlighting was possible than when it was not, even when excluding the time taken for highlighting (H2a). In contrast, the time taken to re-read documents was expected to be shorter for participants who had the possibility to highlight text (H2b), because they might use highlights to navigate their attention, thus potentially saving time in re-locating information (Yeari et al., 2017). Finally, we explored whether participants in the with-highlighting condition would also re-read a smaller proportion of the texts than participants in the without-highlighting condition, because the highlights might guide their visual attention. For this explorative analysis, we used eye-tracking methodology to generate heatmaps, in which fixations made during re-readings were accumulated, for each document. This allowed us to assess the percentage of the document's area that had been fixated during re-readings.

5.3. Method

5.3.1 Participants and Experimental Design

One hundred university students (75.00 % female; *M* = 22.68 years, *SD* = 3.27 years) from different majors of a large German university participated in this experiment. They were compensated with 10€ for their participation. The study was approved by the local ethics committee and participants gave their written consent at the beginning of the study.

Participants were randomly assigned to one of two conditions, which differed in whether a text-highlighting tool was provided (with-highlighting condition) or not (without-highlighting condition). One participant did not read all documents (one document was only opened for around 4 seconds) and was thus excluded from all analyses. Furthermore, due to problems with calibration (e.g., due to astigmatism), we were not able to record eye-tracking data for four participants. Since most analyses required a differentiation between initial readings and re-readings based on a combination of eye-tracking and logfile data, these four participants were also excluded from all analyses. This resulted in a final sample of *N* = 95 participants, with 47 and 48 participants serving in the with- and without-highlighting condition, respectively.

5.3.2 Task and Documents

Participants' task was to read five documents dealing with the topic of the effects of UV radiation on human health in order to write an argumentative essay about the topic afterwards without having the documents available. The document set comprised five partly conflicting, one-paged documents. The documents were slightly adapted (and translated) versions of documents used in previous studies (e.g., Ferguson & Bråten, 2013; Leroy et al., 2021; Strømsø et al., 2016). As in Leroy et al. (2021), the sources of the five documents were a reputable liberal-conservative German newspaper ('FAZ – Frankfurter Allgemeine Zeitung'), a medical science magazine ('Ärztezeitung'), a university research magazine ('Mundo – Wissenschaftsmagazin der TU Dortmund'), the German national cancer association ('DKG – Deutsche Krebsgesellschaft'), and a German school book publisher ('Cornelsen'). Overall, two documents took a positive stance on UV radiation by stating that it had positive effects on health by promoting the production of vitamin D ('Ärztezeitung' and 'FAZ – Frankfurter Allgemeine Zeitung'). Two documents took a negative stance on UV radiation by stating that it had negative effects on health by promoting skin cancer ('DKG – Deutsche Krebsgesellschaft' and 'Mundo – Wissenschaftsmagazin der TU Dortmund').

One document gave general and neutral information about UV radiation without taking a stance (‘Cornelsen’). Complementary information was provided by each pair of documents (e.g., one positive document stated that people with Mediterranean skin type can expose themselves to sunlight at a UV index of 8 for up to 40 min without risk, while one negative document stated that people with skin type IV should not expose themselves for more than 40 min to sunlight at a UV index of 8). In addition, the two positive and two negative documents, respectively, provided conflicting claims (e.g., “UV radiation damages genetic material, which can lead to cell death or even cancer” vs. “UV radiation promotes Vitamin D production, which, in turn, counteracts cell degeneration”). The documents with the same stance provided consistent claims (e.g., each positive document reported positive effects of UV radiation). Each of the four documents taking a stance reported on an interview with a professor who worked in the respective research area. For a more detailed description of the documents’ contents, see Table 5.6.1 of the Supplementary Material. All documents were comparable in length ($M = 24.00$ sentences, $SD = 1.92$; $M = 352.40$ words, $SD = 27.32$) and readability (with an average LIX score of $M = 52.06$, $SD = 2.65$; Björnsson, 1968). All documents contained two to three subheadings and a source logo at the top of the page displaying the name of the source.

The documents were presented in one of two defined document orders. In one document order, the documents in the first and fifth positions took a positive stance, whereas in the other document order they took a negative stance. Moreover, the documents in the second and fourth positions took an opposing stance to those in the first and fifth positions. The document in the third position was always the one giving general information about UV radiation without taking a stance. This resulted in two sets of partly conflicting documents arranged in direct succession (i.e., those in the first and second as well as those in the fourth and fifth positions took opposing stances, whereas documents in the first and fifth as well as those in the second and fourth positions took the same stance). Overall, 25 participants in the with-highlighting and 24 participants in the without-highlighting condition were presented a positive document first, and, accordingly, 22 and 24 participants in the respective conditions were presented a negative document first.

5.3.3 Documents Presentation

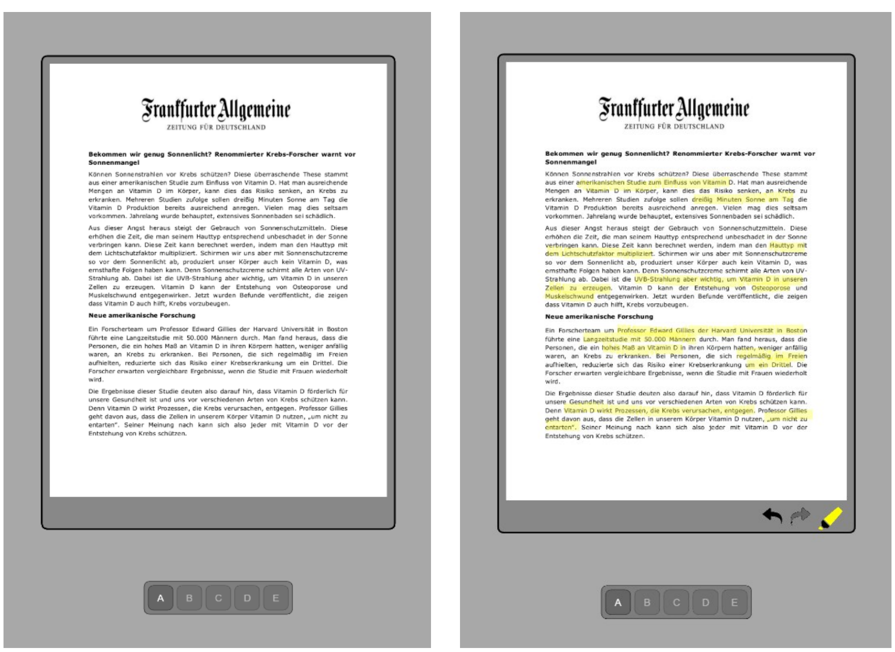
Because we sought to replicate previous findings of a study where participants had read documents on a multi-touch display where highlighting was done with the finger (Leroy et al., 2021), in the present study we also presented documents on a touch display. Specifically, documents were presented sequentially (i.e., one at a time) on an upended 27″ touch display (3M Multi Touch M2767PW, resolution: 1920 x 1080 px) which was diagonally positioned on a height-adjustable table. Thus, the upper end of the screen was higher than the lower end, and participants could interact with it more easily than had it been either lying flat or standing vertically on the table (see Figure 5.1). For each participant, the table was adjusted in height until they felt that interacting with the user interface was comfortable. Each document could be opened from a menu at the bottom of the display, which then was displayed in the centre of the screen in a size slightly bigger than A4 (33.9 cm x 23.9 cm). In the with-highlighting condition, participants could highlight text (in yellow) with their finger by using a swiping gesture over the text. Highlighting could be (de-)activated by tapping a button in the bottom-right corner of the document's frame, the most recently generated highlight could be undone by tapping a left-arrow button, and the last undone highlight could be retrieved by tapping a right-arrow button. These buttons were also positioned at the bottom of the document's frame (see Figure 5.2).

Figure 5.1

Re-enactment of a participant standing in front of the height-adjustable table with the inclined display on which the documents were presented.

Figure 5.2

Experimental user interface without highlighting-tool (left) and with highlighting-tool (right). Highlighting could be activated by tapping the highlighter-button which, in activated mode, would appear yellow. Text could then be highlighted with the finger.



5.3.4 Measures

5.3.4.1 Integrated Understanding

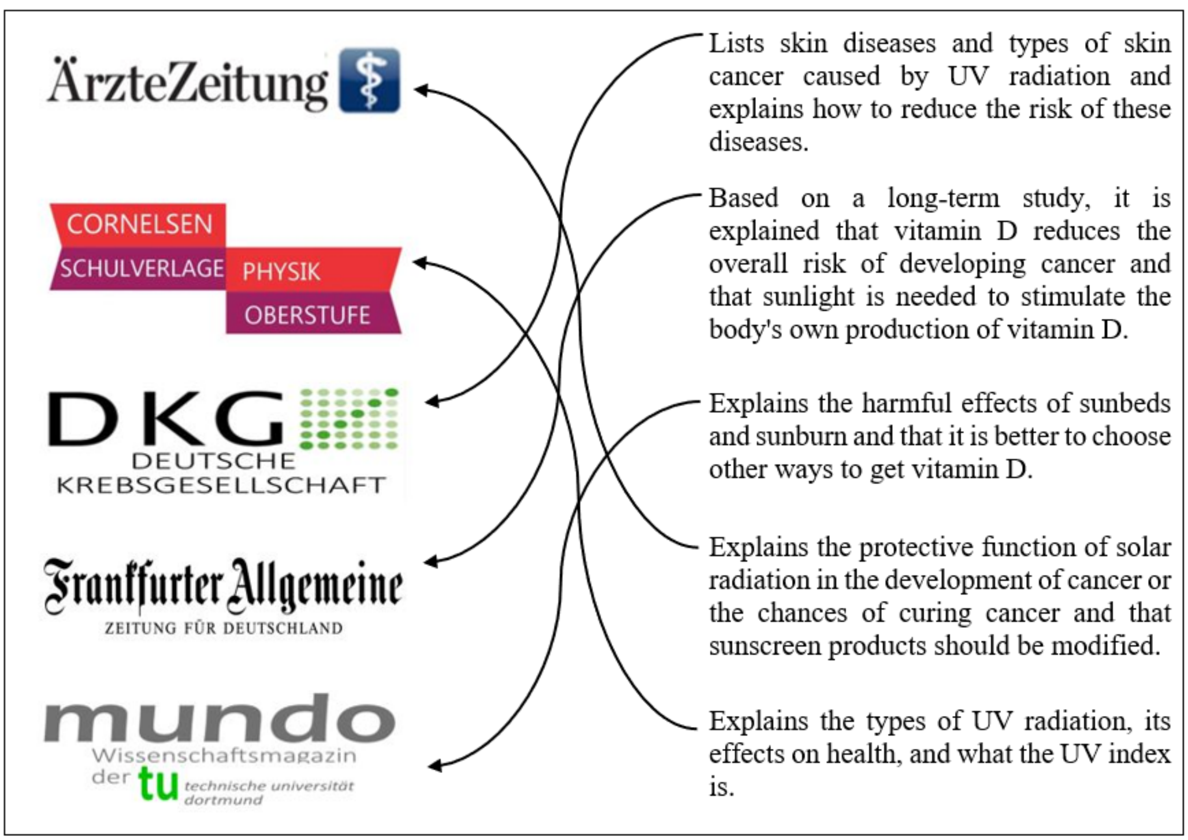
Participants' integrated understanding was assessed with two dependent measures. First, we assessed their *intertextual integration* through the number of intertextual connections included in their argumentative essays (e.g., Kobayashi, 2009; Leroy et al., 2021; List et al., 2019; Salmerón et al., 2018). Second, to assess participants' *source-content integration* we used a source-content-mapping task, that is, a task that requires participants to map which information stemmed from what source (e.g., Delgado et al., 2020; Kammerer et al., 2016; Stang Lund et al., 2019; Strømsø et al., 2010).

Statements in participants' argumentative essays that combined information from two documents were coded as intertextual connections. Both connections between document-specific information and connections between the global stances taken by the documents were coded as intertextual connections (i.e., connections of conflicting or complementary information; cf. Primor & Katzir, 2018; see Table 5.6.2 of the Supplementary Material for examples). An intertextual connection could span several sentences (e.g., being connected through an adverb, such as "furthermore", "also", "however", or "on the contrary"), or be contained in a single sentence. More than one intertextual connection could be contained in a single sentence (e.g., when information from three documents was connected within one sentence, that sentence contained two intertextual connections). In general, only correct statements and connections were considered. Interrater agreement was coded for each intertextual connection coded by both of two raters, or for each sentence that had not been coded as containing any intertextual connections by either rater, respectively. Disagreement was coded for each intertextual connection that had only been coded by one rater. Two raters (the first author and a trained research assistant) independently coded 20 (20.20%) essays for intertextual connections and reached an interrater agreement of 86.93%. Disagreements were discussed and resolved, before one rater (the first author) proceeded with the coding of the remaining 79 essays.

In the source-content mapping task (see Figure 5.3), participants were presented with the documents' source logos in one column, and one-sentenced main statements, each distinctively describing the content of one document, in a second column. The source logos were presented in alphabetic order and the one-sentenced main statements were presented in the same scrambled order for each participant, in which none was initially placed in its correct position (see Figure 5.3). Participants were asked to move the statements next to the source logo of the document that provided the respective information. Participants received one point for each correctly mapped statement. However, because once four statements were correctly mapped, the fifth was automatically correct as well, we defined a maximum score of 4 for this task.

Figure 5.3

Items of the source-content-mapping task (translated from German). The initial order of source logos (left) and statements (right) was as shown here. Participants' task was to drag the statements vertically such that their position matched that of the corresponding source logos. Arrows indicate the correct place assignments.



5.3.4.2 Task Processing Measures

We assessed the number of re-readings, reading times for initial readings and re-readings, as well as the fixation coverage of documents during re-readings as task processing measures. To differentiate between initial readings and re-readings, eye-tracking data was used in combination with logfiles. In the with-highlighting condition we also assessed how many highlights were made during initial readings and re-readings as well as the respective time it took to perform the highlights (highlighting-time).

Eye-tracking data was recorded by means of the mobile eye-tracking system Tobii Pro Glasses 2 Live View Wireless 50, with a temporal resolution of 50 Hz, and the software Tobii Pro Glasses Controller. All eye-tracking measures were generated using the Tobii Pro Lab Software (V.1.145.28180).

Distinguishing between initial readings and re-readings. In Leroy et al. (2021), each re-accessing of a previously opened document was defined as a revisit when the document was opened for at least 1 sec. This approach, however, did not guarantee that the document had been fully read during its first access. In contrast, in the present study we aimed to differentiate between participants' initial readings of a document until they had fully read it, and their subsequent re-readings of that document. By using eye-tracking data, we were able to assess when participants had fully read a document (this could span several document accesses), and, accordingly, how many times they subsequently re-accessed a document to re-read it. For this purpose, in the eye-tracking videos of all 95 participants, first we manually tagged all document accesses using the Tobii Pro Lab software. Subsequently, we used Tobii Pro Lab's automatic mapping functionality with the gaze filter Tobii I-VT (Fixation), which uses a default threshold of 60 ms for fixations, and its visualization tool to create heatmaps (based on the number of fixations) for each document access of each participant. For each participant and each document, we then used the heatmap created for the first document access to determine whether the document had been fully read (i.e., whether the resulting heatmap covered the entire text of the document), or whether further document accesses needed to be taken into account to make up the initial reading of the respective document (see Figure 5.4). In case the heatmap of the first document access revealed only partial reading of the document, we accumulated heatmaps of subsequent document accesses until the resulting heatmap revealed full coverage of the text. Overall, for 22 documents (from 17 participants) another document had been (re-)accessed before the respective document had been

Figure 5.4

Heatmaps of a participant's fixations on the document 'Ärztezeitung' during its first access (left; 107 sec duration), second access (middle; 28 sec duration), and re-readings (accumulated; right; 137 sec duration). Please note that the first and second accesses together account for the initial reading.

fully read. That is, for these 22 cases, initial document readings comprised more than just the first document access. The total initial reading time (in sec) for each participant was calculated by summing up the time taken for the initial full readings of the five documents.

All accesses of a document after it had been fully read were defined as re-readings, given that they lasted for a particular minimum duration. Specifically, we examined five different thresholds for the duration of re-readings, namely 1, 2, 3, 4, and 5 sec. Please note that Leroy et al. (2021) had instead used a fixed threshold of 1 sec. However, re-accesses that last for 2 or 3 sec, for example, are still very short and allow only little information processing, and thus might not necessarily imply intentional re-accessing. By defining thresholds between 1 sec and up to 5 sec, we thus aimed to explore how many fixations were directed towards the respective document (i.e., approximately how much processing occurred) during short re-accesses of up to 5 sec (for details see Section 3.4.2.3), and whether (and if so, how) the choice of the threshold for the duration of re-readings made any difference for the investigated effects. However, for the sake of conciseness, in the Results Section (i.e., Sections 4.2, 4.3, and 4.4), we will only report results of analyses using the 1 and 3 sec thresholds. Further analyses using the 2, 4, and 5 sec thresholds are provided in the Supplementary Material.

Determining the number of re-readings. In order to assess the number of re-readings for a given duration-threshold, similar to Leroy et al. (2021), we used logfile data to determine the sequence of all document accesses for each participant and how long each document access lasted. For the 1 sec threshold, for example, we erased all document accesses that lasted less than 1 sec from this sequence. When the resulting sequence contained two accesses of the same document in direct succession, they were counted as one document access which lasted for the sum of both document accesses. Take, for example, the following document accessing sequence, with the bold print reflecting initial readings that were completed after the first document access and the underlined print reflecting initial readings spanning across two document accesses: "**A (130 sec)**, **B (111 sec)**, **C (132 sec)**, D (107 sec), C (13 sec), A (0.7 sec), C (18 sec), D (28 sec), **E (109 sec)**, D(137 sec)". After erasing all document accesses that lasted for less than 1 sec (i.e., for applying a 1 sec threshold), this results in the sequence "**A (130 sec)**, **B (111 sec)**, **C (132 sec)**, D (107 sec)**,** C (13 + 18 sec), D (28 sec), **E (109 sec)**, D(137 sec)". This reveals one re-reading of document C (for 13 + 18 sec) and document D (for 137 sec), and no re-readings of documents A, B, and E. The initial reading of document D spans across the first two accesses (see Figure 5.4 for example heatmaps generated for document D from one participant's real data).

Number of fixations on documents during short re-accesses. In addition to determining the number of re-readings, eye-tracking data was also used for more fine-grained fixation-based analyses. For this purpose, the data of 6 participants had to be excluded because of imprecise data quality. For the remaining 89 participants, we took the following approach using the Tobii Pro Lab Software to assess the number of fixations on documents during short re-accesses below 5 sec: First, we categorized re-accesses according to the following five categories of time thresholds: below 1 sec; at least 1 but less than 2 sec; at least 2 but less than 3 sec; at least 3 but less than 4 sec; or at least 4 but less than 5 sec. Using Tobii Pro Lab's automatic mapping functionality with the gaze filter Tobii I-VT (Fixation, see above), we then determined the number of fixations on the respective document for each of these short re-accesses. Finally, we calculated the average number of fixations per re-access separately for each of the five categories. Please note that only the data of participants who made at least one re-access with a duration within the respective time span was considered in these analyses.

Fixation coverage during re-readings. Goldberg and Kotval (1999) argued that the spatial distribution of fixations on an interface, that is, the fixation coverage (or spatial density) indicates the extent of search and processing, with a fixation coverage of "a small area reflect[ing] direct and efficient search" (p. 640). Accordingly, we used the fixation coverage during re-readings as a measure of participants' focused re-reading. Only the data of the 89 participants with valid eye tracking was used for these analyses. For each participant and re-accessed document, we created one heatmap comprising all fixations on the respective document during re-readings, by using a blank background instead of the document. We defined a radius of 35px for each fixation's contribution to the heatmap, and used Tobii Pro Lab's default scale value to define colouring (note that colour type played no role in our fixation coverage analyses). The resulting heatmap was non-white in all areas where a participant had fixated during re-readings. That is, 100% coverage would mean that the whole document (including white space) was covered with fixations. Using the R function *as.raster*, we created bitmaps of each fixation coverage heatmap and calculated the percentage of the respective document's non-white area (Sharafi et al., 2015). For example, the re-reading fixation coverage in the example shown in Figure 5.4 (right), notably with a white background instead of the document's text, is 47.13%.

5.3.4.3 Control Variables

To ascertain comparability across the two conditions, we assessed participants' working memory capacity (15-item computer-based reading span task; cf. Kane et al., 2004), topic interest (1 item) and prior topic beliefs (2 items; one addressing positive and another addressing negative effects of UV radiation on health) on 7-point Likert scales, as well as prior topic knowledge (see below). The two items on prior topic beliefs were averaged after reverse-scoring the negative item (Cronbach's alpha = .79). Thus, the higher the score, the more positive a participant's topic beliefs were.

Prior topic knowledge was assessed with an essay task prior to reading the documents, in which participants were asked to write down in 5 minutes everything they knew about potential effects of UV radiation on human health. Participants were awarded one point for each relevant and correctly reported general fact (e.g., UV radiation comes from the sun; sunbeds use UV radiation; there are 3 types of UV radiation) or specific argument concerning negative or positive health effects of UV radiation (e.g., UV radiation causes skin cancer; UV radiation promotes the production of vitamin D in the body). Two independent raters (the first author and a trained research assistant) scored a random selection of 20 (20.20%) prior knowledge essays using a pre-defined list of 35 facts and arguments related to the topic. Interrater agreement was assessed for point-wise scoring, that is, agreement was coded when both raters gave a point for a statement, and disagreement was coded when only one rater gave a point for a statement. Interrater agreement for the 20 essays was 87.06%. Disagreements were resolved through discussion, and subsequently one rater (the trained research assistant) scored the remaining 79 prior topic knowledge essays.

5.3.5 Procedure

Participants were tested in single sessions (of approximately 75 min) in the lab. First, they were asked to rate their topic beliefs, write the prior-knowledge essay, provide demographic data, and complete the working memory task, using a laptop. Subsequently, participants received a written explanation of the interaction possibilities on the touch display (i.e., how to open documents in the without-highlighting condition, and how to open and highlight documents in the with-highlighting condition). Each participant practiced interaction with the interface by using five blank documents in the respective highlighting-condition. Once they felt comfortable operating the interface, they received a written task instruction on the screen. Specifically, they were told that they would be provided with five documents from the internet, which they should read carefully, in order to subsequently write an argumentative essay about potential health effects of UV radiation without having the documents available. They were also informed that they had a maximum of 15 minutes to read the documents and that within this time frame they could read the documents as often as they wanted. After reading, participants were asked to write the argumentative essay in 15 minutes on a laptop. They were told to include in their essay arguments and findings stated in the documents to justify their claims. After having completed their essay, they worked on the source-content-mapping task on the laptop.

Table 5.1

Means (and standard deviations) for control variables as a function of text-highlighting.

	With highlighting	Without highlighting
Age	22.77 (3.25)	22.56 (3.41)
Prior topic knowledge	3.43 (1.86)	4.00 (1.81)
Working memory capacity (0-15)[a]	10.07 (3.03)	10.35 (2.43)
Topic interest (from 1=very low – 7=very high)	4.28 (1.38)	4.44 (1.05)
Prior topic beliefs (from 1=very negative to 7=very positive)	2.69 (1.05)	2.49 (0.84)

[a] *Note*. Due to technical issues, one participant's data is missing.

5.4. Results

5.4.1 Comparability of Conditions

Participants in the with-highlighting and without-highlighting conditions did not differ with respect to age, $t(92.94) = 0.30$, $p = .767$, prior knowledge as assessed from essays written before reading, $t(92.77) = -1.52$, $p = .131$, working memory capacity, $t(86.19) = -0.51$, $p = .612$, topic interest, $t(85.96) = -0.64$, $p = .525$, or prior beliefs about potential health effects of UV radiation, $t(87.53) = 0.65$, $p = .516$ (see Table 5.1 for means and standard deviations).

Furthermore, the number of re-readings of documents did not differ between conditions (for any of the five re-reading thresholds, see Table 5.2 for the 1 sec 3 sec thresholds and Table 5.6.3 of the Supplementary Materials for the other thresholds). Table 5.2 also provides the number of re-readings of each of the five documents separately. Descriptively, the neutral document and one negative document were re-read most often by participants, whereas the other negative document was re-read least often in both conditions. Furthermore, the essays composed after reading also did not significantly differ in length between conditions, $t(93) = 1.06$, $p = .294$ (with highlighting: $M = 273.55$ words, $SD = 70.69$; without highlighting: $M = 258.10$ words, $SD = 72.04$).

Table 5.2

Means (and standard deviations) of the number of re-readings with the respective duration-threshold as a function of text-highlighting, as well as inferential statistics.

Threshold	Document(s)	Condition		Inferential statistics
		With highlighting	Without highlighting	
# re-readings ≥ 1 sec	All	9.75 (7.70)	10.92 (7.28)	$t(92.45) = -0.76, p = .448$
	Ärztezeitung (positive)	1.94 (1.79)	2.15 (1.77)	$t(92.93) = -0.57, p = .568$
	FAZ (positive)	1.92 (1.80)	2.06 (1.45)	$t(88.11) = -0.44, p = .662$
	Cornelsen (neutral)	2.15 (1.60)	2.52 (1.60)	$t(90.82) = -1.03, p = .307$
	Mundo (negative)	2.32 (2.06)	2.39 (1.77)	$t(89.43) = -0.03, p = .973$
	DKG (negative)	1.40 (1.53)	1.83 (1.42)	$t(92.18) = -1.42, p = .160$
# re-readings ≥ 3 sec	All	7.72 (5.96)	8.85 (5.94)	$t(92.94) = -0.93, p = .357$
	Ärztezeitung (positive)	1.62 (1.55)	1.81 (1.65)	$t(92.88) = -0.59, p = .553$
	FAZ (positive)	1.32 (1.18)	1.71 (1.25)	$t(92.86) = -1.56, p = .123$
	Cornelsen (neutral)	1.87 (1.45)	2.06 (1.59)	$t(92.57) = -0.61, p = .544$
	Mundo (negative)	1.85 (1.53)	1.83 (1.29)	$t(89.81) = 0.06, p = .952$
	DKG (negative)	1.06 (1.24)	1.44 (1.09)	$t(90.95) = -1.56, p = .123$

Table 5.3

Means (and standard deviations) of the number of fixations on documents per re-access for short re-accesses that lasted up to 5 sec as a function of the re-access duration and text-highlighting. The number of participants who made at least one such short re-access as well as the total number of re-accesses across participants for each time span are also provided.

Measure	Condition	Duration of document re-access				
		[0, 1 sec[	[1, 2 sec[	[2, 3 sec[	[3, 4 sec[	[4, 5 sec[
# Fixations	With highlighting	1.66 (1.38), $n = 20$ (31 re-accesses)	2.78 (1.49), $n = 23$ (42 re-accesses)	4.70 (2.82), $n = 23$ (41 re-accesses)	8.29 (3.66), $n = 15$ (27 re-accesses)	11.36 (4.77), $n = 17$ (24 re-accesses)
	Without highlighting	1.77 (0.95), $n = 10$ (21 re-accesses)	3.69 (1.51), $n = 24$ (53 re-accesses)	4.90 (2.07), $n = 20$ (33 re-accesses)	7.85 (2.67), $n = 19$ (13 re-accesses)	11.23 (5.81), $n = 13$ (19 re-accesses)

5.4.2 Number of Fixations on Documents During Short Re-Accesses

As a first step in our methodological examination of the re-reading threshold, we explored how much information, on average, participants processed during short re-accesses that lasted up to 5 sec. To this end, we calculated the average number of fixations per document re-access according to the five duration thresholds examined. Means (and standard deviations) of the average number of fixations on documents during short re-accesses are shown in Table 5.3 for the with-highlighting and the without-highlighting conditions. The longer a re-access lasted, the more fixations participants made in the respective document. Furthermore, in both conditions, on average, during re-accesses that lasted less than 1 sec, participants made less than two fixations, and in re-accesses that lasted more than 2 but less than 3 sec, participants made less than five fixations. Hence, simply put, the application of the 1 sec threshold excluded re-accesses in which participants had an average of less than two fixations. Likewise, the 3 sec threshold excluded re-accesses in which participants had an average of less than five fixations.

5.4.3 Interaction Effects Between the Possibility to Highlight Text and the Number of Re-Readings on Integrated Understanding Measures (RQ1)

On average, participants in the with-highlighting and without-highlighting conditions included $M = 4.09$ ($SD = 1.82$) and $M = 3.79$ ($SD = 1.65$) intertextual connections in their essays. Hypothesis H1a stated that only for participants with a higher number of re-readings, the possibility to highlight text would result in an increased number of intertextual connections (H1a). Further, hypothesis H1b stated that only in the with-highlighting condition the number of re-readings would be positively related to the number of intertextual connections (H1b). In order to test our hypotheses and also how the choice of re-reading threshold might affect the expected moderation of the number of re-readings on the effect of text-highlighting, we conducted five moderated linear regression analyses, one for each of the five re-reading thresholds. For the sake of conciseness, in the following, we report the results of these analyses for the 1 and 3 sec re-reading thresholds. Analyses using the 2, 4, and 5 sec thresholds are provided in Table 5.6.4 of the Supplementary Material. In each regression analysis we used highlighting (with vs. without) as a dichotomous predictor (coded as -1 and +1, respectively) and the respective number of re-readings (z-standardized) as a continuous predictor, as well as the interaction term between the two predictors, and number of intertextual connections as dependent variable. In all analyses, we used two-sided tests of significance.

For the 1 sec threshold, neither text-highlighting, nor the number of re-readings, nor the interaction between text-highlighting and the number of re-readings were significant predictors for the number of intertextual connections. In contrast, for the 3 sec threshold the interaction was significant[3] (for inferential statistics, see Table 5.4).

Figure 5.5 illustrates this interaction between text-highlighting and the number of re-readings for the 3 sec threshold. Simple comparison analyses[4] according to the procedure outlined by Aiken and West (1991) showed that – in line with H1a – for participants with a high number of re-readings (1 *SD* above the mean; i.e., 14.24 re-readings) the number of intertextual connections was higher in the with-highlighting condition than in the without-highlighting condition, β = .52, $t(91)$ = 2.09, p = .039, but not for participants with a low number of re-readings (1 *SD* below the mean; i.e., 2.35 re-readings), β = -.18, $t(91)$ = -0.73, p = .466. Likewise, in line with H1b, for the 3 sec threshold for participants in the with-highlighting condition, the number of re-readings was significantly positively correlated to participants' intertextual connections, β = .58, $t(91)$ = 2.34, p = .021, but not for participants in the without-highlighting condition, β = -.12, $t(91)$ = -0.49, p = .627.

With regard to the second integrated understanding measure, that is, participants' score in the source-content mapping task, participants in the with-highlighting and without-highlighting conditions achieved an average score of M = 2.32 (*SD* = 1.24) and M = 2.40 (*SD* = 1.07), respectively. We aimed to explore whether an interaction effect similar to that with the number of intertextual connections in participants' essays would be shown for participants' source-content integration, and whether this interaction would be affected by the choice of re-reading threshold. Thus, we conducted the same moderated regression analyses as described above for the number of intertextual connections. However, for the source-content mapping score as the dependent variable, the interaction between text-highlighting and the number of re-readings was not significant; neither for the 1 sec threshold nor for the 3 sec threshold (see Table 5.4) nor for the other thresholds (see Table 5.6.4 of the Supplementary Material). We therefore reran the regression analyses without the interaction term. Results showed that text-highlighting wasn't a significant predictor in any of the five regression analyses. However, the number of re-readings was a significant positive predictor of participants' source-content mapping score for the re-reading threshold of 1 sec, and a marginally positive predictor for the 3 sec threshold[5] (see Table 5.4).

---

[3] This interaction was marginally significant for the 2 sec threshold, and significant for the 4 and 5 sec thresholds (see Table 5.6.4 of the Supplementary Material for all analyses).
[4] Note that these analyses are reported for the 3 sec threshold here. Yet, this pattern of results was also the same for the 2, 4, and 5 sec thresholds (see Table 5.6.5 of the Supplementary Material for all analyses).
[5] For the 2, 4, and 5 sec thresholds the effect was also marginally significant (see Table 5.6.4 of the Supplementary Material).

Table 5.4

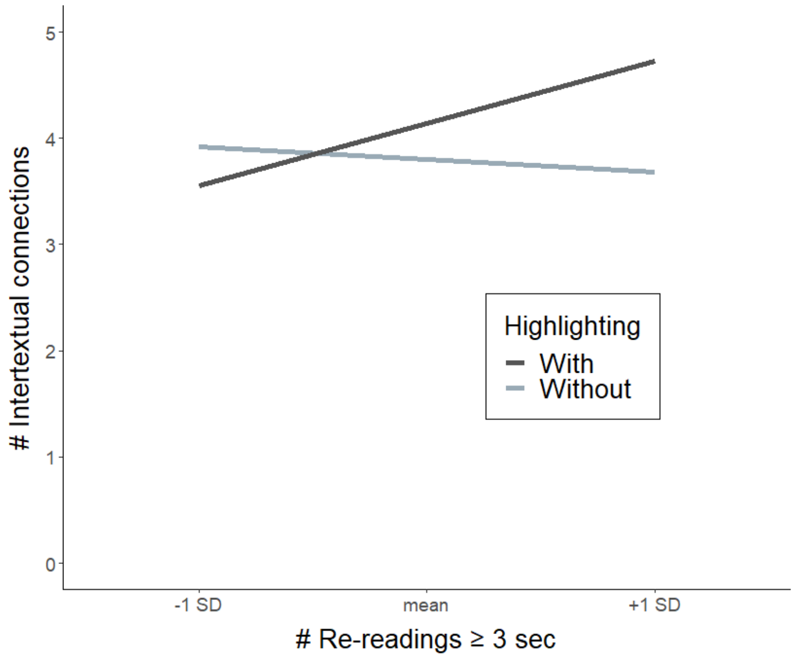
Inferential statistics for main effects of and interactions between text-highlighting and the number of re-readings (for the 1 and 3 sec re-reading thresholds) on the integrated understanding measures (number of intertextual connections and source-content mapping score).

Measure	Re-reading threshold	Inferential statistics
# intertextual connections	1 sec	text-highlighting: β = .10, $t(91)$ = 0.95, p = .358 #re-readings ≥ 1 sec: β = .12, $t(91)$ = 1.17, p = .244 interaction: β = .16, $t(91)$ = 1.51, p = .135
	3 sec	text-highlighting: β = .10, $t(91)$ = 0.97, p = .336 # re-readings ≥ 3 sec: β = .14, $t(91)$ = 1.33, p = .186 **interaction**: β = .20, $t(91)$ = 1.99, p = .049
Source-content-mapping score	1 sec	interaction: β = -.01, $t(91)$ = -0.07, p = .941 regression without interaction term: text-highlighting: β = -.02, $t(92)$ = -0.15, p = .879 **# re-readings**: β = .23, $t(92)$ = 2.25, p = .027
	3 sec	interaction: β = .03, $t(91)$ = 0.31, p = .759 regression without interaction term: text-highlighting: β = -.02, $t(92)$ = -0.15, p = .879 **# re-readings**: β = .19, $t(92)$ = 1.82, p = .072

Note. (Marginally) significant results are displayed in bold print. All p-values are two-tailed p-values.

Figure 5.5

Interaction between text-highlighting and the number of re-readings ≥ 3 seconds (z-standardized) with regard to the number of intertextual connections.



5.4.4 Effect of the Possibility to Highlight Text on Initial Reading and Re-Reading Behaviour (RQ2)

In the following, results are reported with regard to our hypotheses H2a and H2b, which assumed that participants in the with-highlighting condition would have longer initial reading times across the five documents (H2a) but shorter re-reading times (H2b) than participants in the without-highlighting condition. All but two participants in the with-highlighting condition made use of the highlighting-tool. These 45 participants, on average, made *M* = 90.67 highlights (*SD* = 41.87) with a total length of all highlights of *M* = 21,126.28 px (*SD* = 10,373.05). Specifically, they made *M* = 86.48% (*SD* = 28.77) of their highlights during initial readings.

Table 5.5 shows means (and standard deviations) of overall reading times, initial reading times, and re-reading times (for the 1 sec and 3 sec thresholds; for all thresholds see Table 5.6.6 of the Supplementary Material) for the with-highlighting and without-highlighting conditions as well as the respective inferential statistics. Two-sided Welch's t-tests showed that participants in the with-highlighting condition spent significantly more time overall for reading – but only when this time included the time participants in the with-highlighting condition had taken for highlighting. When the time participants in the with-highlighting condition had taken for highlighting was subtracted from the overall time on task, there was no difference in overall reading time between conditions. Yet, more detailed analyses of the reading times revealed that – even when the time participants in the with-highlighting time was subtracted from their initial reading time – participants in the with-highlighting condition took significantly more time for *initial* readings than participants in the without-highlighting condition, confirming H2a. In contrast, participants in the with-highlighting condition took less time for re-reading than those in the without-highlighting condition, thus also confirming H2b.

In order to examine the time-saving effect of text-highlights for re-reading, we further explored whether the time difference between the with-highlighting and the without-highlighting condition was moderated by the number of re-readings. We conducted a moderated linear regression analysis using text-highlighting (with vs. without) as a dichotomous predictor (coded as -1 and +1, respectively) and the number of re-readings (≥ 1 sec and ≥ 3 sec, respectively; z-standardized) as a continuous predictor, as well as the interaction term between the two predictors, and the re-reading time as the dependent variable. Results revealed a significant interaction[6] between predictors for both the 1 sec threshold, β = -.17, $t(91)$ = -2.51, p = .014 (text-highlighting: β = -.46, $t(91)$ = -2.37, p = .020; # re-readings: β = .73, $t(91)$ = 10.84, p < .001), and the 3 sec threshold, β = -.17, $t(91)$ = -2.60, p = .011 (text-highlighting: β = -.15, $t(91)$ = -2.35, p = .021; # re-readings: β = .75, $t(91)$ = 11.63, p < .001). Figure 5.6 illustrates this interaction for the 3 sec threshold. Simple comparison analyses[7] revealed that for participants with a high number of re-readings (1 *SD* above the mean), the re-reading time was significantly shorter in the with-highlighting condition than in the without-highlighting condition, β = -.90, $t(91)$ = -3.49, p < .001, but not for participants with only a few re-readings (1 *SD* below the mean), β = .05, $t(91)$ = 0.26, p = .849. This indicates that particularly for participants with a higher number of re-readings, the provision of the highlighting-tool yielded a time-saving effect during re-readings.

Finally, regarding the distribution of visual attention during re-readings, we assessed potential differences for each document individually. Two-sided Welch's t-tests indicated that during re-reading, participants in the with-highlighting condition fixated on significantly smaller parts of the two positive documents (i.e., 'Ärztezeitung' and 'FAZ – Frankfurter Allgemeine Zeitung'; see Figure 5.7) than participants in the without-highlighting condition

---

[6] This interaction was also significant for the 1, 2, and 4 sec thresholds, and marginally significant for the 5 sec threshold (see Table 5.6.7 of the Supplementary Material).

[7] Note that these analyses are reported for the 3 sec threshold here. Yet, this pattern of results was also the same for the 1, 2, 4, and 5 sec thresholds (see Table 5.6.7 of the Supplementary Material for all analyses).

(Ärztezeitung, $n = 64$: $t(61.99) = -2.20$, $p = .031$; FAZ, $n = 68$: $t(64.80) = -2.14$, $p = .036$). In contrast, the fixation coverage for the two negative documents and the neutral document did not differ between conditions (DKG, $n = 57$: $t(55.00) = -1.35$; $p = .182$; Mundo, $n = 72$: $t(70.00) = 0.05$, $p = .959$; Cornelsen, $n = 68$: $t(62.61) = -1.10$, $p = .276$).

Table 5.5

Means (and standard deviations) for overall reading times, initial reading times (both including and excluding the time taken for highlighting during initial readings), and re-reading times as a function of text-highlighting.

Reading-time measure	Specifics	Condition		Inferential statistics
		With highlighting	Without highlighting	
Overall reading time [min]	Including highlighting-time	15.56 (2.13)	13.96 (3.07)	$t(84.27) = 2.98$, $p = .004$
	Excluding highlighting time	14.32 (2.04)	13.96 (3.07)	$t(82.01) = 0.68$, $p = .501$
Initial reading time [min]	Including highlighting-time	12.20 (2.98)	9.36 (2.17)	$t(83.98) = 5.30$, $p < .001$
	Excluding highlighting time	11.07 (2.75)	9.36 (2.17)	$t(87.40) = 3.56$, $p = .001$
Re-reading time [min] [a]	Re-accesses ≥ 1 sec	3.35 (2.38)	4.58 (3.20)	$t(86.71) = -2.14$, $p = .035$
	Re-accesses ≥ 3 sec	3.27 (2.32)	4.52 (3.18)	$t(86.05) = -2.20$, $p = .031$

Note. Analyses are based on all $N = 95$ participants.
[a] Since the re-reading time was already shorter in the with-highlighting condition, we refrained from additional analyses of re-reading time excluding highlighting time. All p-values are two-tailed p-values.

Figure 5.6

Interaction between text-highlighting and the number of re-readings ≥ 3 sec (z-standardized) with regard to re-reading time (only re-accesses ≥ 3 sec taken into account).

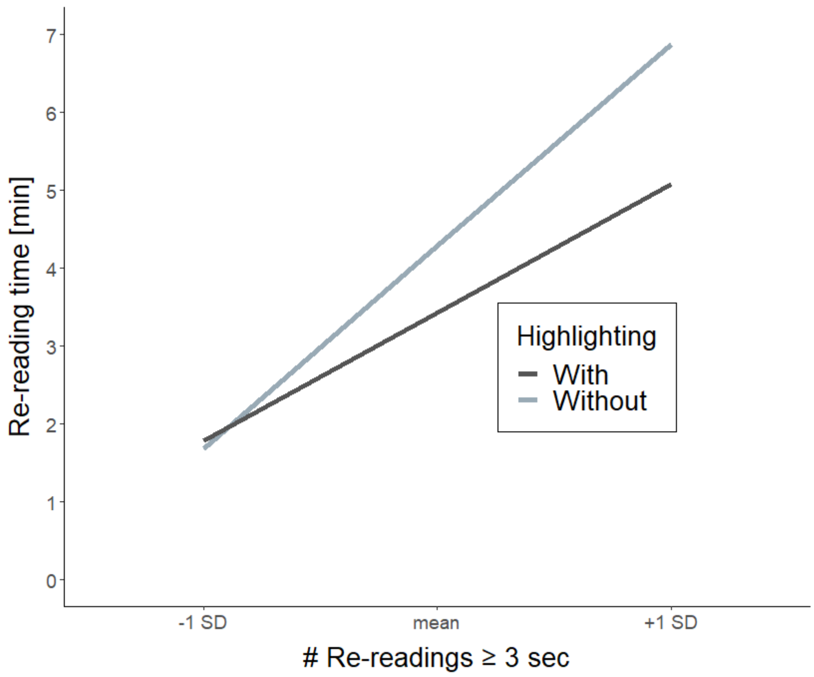
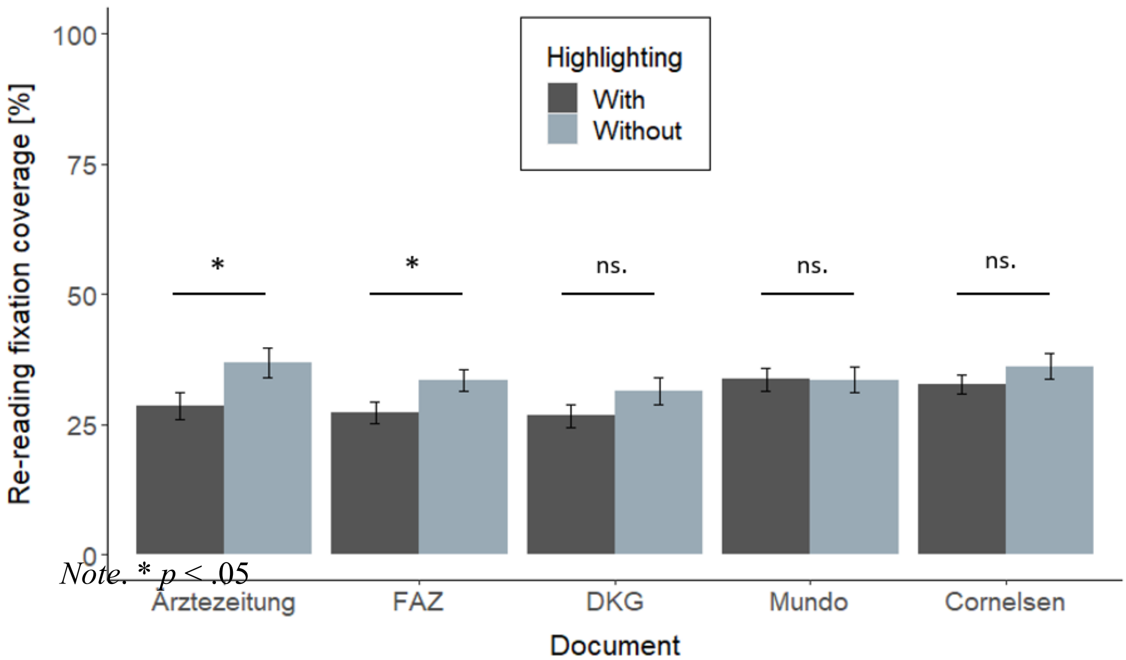


Figure 5.7

Mean percentage of participants' re-reading fixation coverage for each document as a function of text-highlighting. For these analyses, only the data of participants who re-accessed the respective document are considered.



Note. * $p < .05$

5.5 Discussion

In the present study we sought to replicate and extend previous findings that especially for participants with a higher number of re-readings, the provision of a text-highlighting tool has a positive effect on readers' integrated understanding of multiple, partly conflicting documents about a health-related issue (Leroy et al., 2021). Specifically, we aimed to extend that work by assessing not only readers' intertextual integration, but also their source-content integration, and by investigating how initial and re-reading times as well as re-reading fixation coverage might differ when a text-highlighting tool is provided versus when it is not. Moreover, we aimed to expand Leroy et al.'s (2021) work by some methodological examinations. First, they had defined initial readings simply as the first access, and re-readings as each re-accessing of a document (that lasted longer than 1 sec). In contrast, we used eye-tracking methodology to differentiate more precisely between initial *full* readings of a document (which could comprise several accesses) and re-readings of a document (i.e., re-accesses of a document after it had been fully read). Our eye-tracking analyses revealed that in 22 instances (of 17 participants; out of 95 participants * 5 documents, i.e., 475 instances) the initial reading of a document comprised more than just its first access. Hence, our analyses indicate that defining initial readings as the first access of a document can be a reasonable approximation of the number of *re-readings* for reasons of simplicity. Second, in order to examine how the re-reading threshold (i.e., the minimum duration of a document access) chosen would affect the potential interaction effect between the number of re-readings and text-highlighting on readers' integrated understanding, we analysed this interaction by using five different re-reading thresholds between 1 and 5 sec. As we will discuss in detail below, indeed statistical results differed to some extent between the 1 sec threshold and the other thresholds. Related to this issue of the re-reading threshold chosen, as a third methodological examination aimed to expand the findings by Leroy et al. (2021) we investigated how many fixations were directed towards documents during short re-accesses of up to 5 sec. Our analyses revealed that in re-accesses that lasted longer than 1 but shorter than 2 sec, participants, on average, had fewer than four fixations on the document, indicating that re-processing of information was very limited during such short re-accesses. In re-accesses that lasted between 3 and 4 sec, participants, on average, had about eight fixations, which is still quite few. However, assuming that during silent reading, each word receives between one and two fixations (e.g., Hautala et al., 2011; Huestegge & Bocianski, 2010; Krieber et al., 2016) participants can be estimated to have read an average of at least four to eight words in re-accesses that lasted between 3 and 4 sec.

In the following section, we discuss our findings regarding the role of re-reading in readers' integrated understanding when a text-highlighting tool was provided versus when it was not (RQ1). Related to these analyses, we critically discuss our findings regarding the choice of re-reading thresholds, which, based on the present findings, should be made under careful consideration of the respective integrated understanding measure assessed. Finally, we discuss how the provision of the highlighting-tool affected readers' initial reading and re-reading behaviour (RQ2).

5.5.1 The Moderating Role of Re-Reading in the Effect of the Text-Highlighting Tool on Intertextual Integration and the Role of Re-Reading in Source-Content Integration (RQ1)

In line with our first hypothesis (specifically H1a), the provision of the text-highlighting tool increased the number of intertextual connections only for those readers with a high number of re-readings of previously read documents, but not for those with a low number of re-readings. Further, as predicted in H1b, only for readers who were provided with the text-highlighting tool, the number of intertextual connections in essays was positively related to the number of re-readings, but not for those who were not. That is, the more readers who were provided with the text-highlighting tool engaged in re-reading previously read documents, the more intertextual connections they included in their essays. These findings replicate the results by Leroy et al. (2021). However, it should be noted that these results were only significant for the 3 sec threshold (as well as for the 4 and 5 sec thresholds, and marginally significant for the 2 sec threshold), but not for the 1 sec threshold.

In addition, exploratory analyses of our second integrated understanding measure, that is, of readers' performance in the source-content mapping task, showed no effect of the text-highlighting tool. However, regardless of whether the text-highlighting tool was provided, a positive relation with the number of re-readings was shown. That is, the more participants engaged in re-reading previously read documents, the higher they scored in the source-content mapping task. Here it should be noted that these results were significant for the 1 sec threshold (and for the 2 sec threshold), but only marginally significant for the 3 sec threshold (as well as for the 4 and 5 sec thresholds). The positive relationship between engagement in re-reading and source-content integration is in line with the finding by Delgado et al. (2020) who have reported a positive relation between the number of times participants had opened pop-up windows that provided hints about conflicting information in other documents and participants' source-content integration. However, since these pop-up windows explicitly provided participants with the conflicting information, the present study is the first reporting a positive relation between the number of re-readings and readers' source-content integration when no hints about conflicting information across documents are provided.

Several conclusions can be drawn from our findings. First, the interaction between text-highlighting and the number of re-readings regarding the number of intertextual connections in essays suggests that there is no general beneficial effect of providing a text-highlighting tool. Further, similar as in the study by Leroy et al. (2021) participants made more than 85% of their highlights during initial readings – hence, most highlights were already in the documents at the time of re-reading. Thus, our finding that the provision of the text-highlighting tool only benefitted intertextual integration of readers with a high number of re-readings supports the assumption that highlighted text facilitates the comparison and corroboration of information during re-reading. That is, the pop-out effect of highlighted text may yield more effective re-processing of contents (cf. Yeari et al., 2017) compared to when text cannot be highlighted. This, in turn, may support intertextual integration, especially during re-reading. Second, the positive relation between participants' score in the source-content mapping task and the number of re-readings that we found in both conditions, suggests that source-content integration benefits from re-reading, regardless of whether a text-highlighting tool is provided or not. Since the reading material included two documents taking a positive and a negative overall stance, respectively, the increased reprocessing of information across documents through re-reading might have supported readers in differentiating among the documents, as well as in differentiating among specific arguments.

Taken together, when readers were provided with the text-highlighting tool, both measures of integrated understanding were positively related to the number of re-readings. For readers who were not provided with the text-highlighting tool, only their score in the source-content mapping task was positively related to the number of re-readings. This suggests differential effects of reading actions and support tools for different aspects of readers' integrated understanding. In conclusion, these findings support the assumption made by the RESOLV model (Britt et al., 2018; Rouet et al., 2017), which posits that it depends on the reading environment (or the support tools provided by it) whether reading actions taken are more or less beneficial for readers' integrated understanding of multiple documents. Finally, from a methodological perspective regarding the role of the chosen re-reading threshold, for *intertextual integration*, analyses yielded significant results only for re-reading thresholds of at least 3 sec, but not when also very short re-accesses of less than 3 sec were included. In contrast, with regard to *source-content integration*, analyses showed significant correlations with the number of re-readings especially when re-accesses that lasted less than 3 sec were also included.

The generation of an elaborate mental representation of source-content links requires readers to connect the gist of the document (as well as unique information provided by that document) to its source. As documents' sources were provided saliently at the top of the document in the form of large wordmark logos, this might also be possible during short re-accesses of up to 3 sec. In contrast, comparing, contrasting, or connecting a piece of information to information provided by another document in order to generate intertextual connections requires more in-depth processing or re-processing of the text, which is unlikely to be possible during very short re-accesses. In conclusion, re-accesses that last less than 3 sec might be more suitable for the generation of source-content links than for the generation of intertextual connections. Nonetheless, it should be mentioned that the overall results pattern both for the number of intertextual connections and for source-content mapping performance were similar for all five re-reading thresholds investigated.

5.5.2 Differences in (Re-)Reading Behaviour With Versus Without Text-Highlighting (RQ2)

As in previous work (e.g. Ben-Yehudah & Eshet-Alkalai, 2018; Leroy et al., 2021; Li et al., 2016), participants of the present study who were provided with the text-highlighting tool took longer overall for the reading task than those without the text-highlighting tool. However, when the time participants in the with-highlighting condition took for highlighting was subtracted from the overall time on task, there was no difference between conditions. Hence, the longer reading time for participants who are able to highlight text might be due to the act of highlighting itself. Yet, our more fine-grained differentiation into initial readings and re-readings revealed that, in line with our second hypothesis (specifically, H2a and H2b) regardless of whether the time taken for highlighting was subtracted or not, participants in the with-highlighting condition had significantly longer initial reading times (H2a), but shorter re-reading times (H2b) than participants in the without-highlighting condition. These findings are particularly interesting because participants in the with-highlighting condition made most of their highlights during initial readings. Hence, the longer initial reading times may have originated from the decision process of which information to highlight. Regarding the difference in re-reading times, additional exploratory analyses revealed that the difference was especially pronounced for participants with a high number of re-readings. That is, particularly readers who were provided with the text-highlighting tool and made many re-readings had shorter re-reading times than readers who were not provided with the text-highlighting tool and made many re-readings. Because overall the number of re-readings did not differ between the with-highlighting and without-highlighting conditions, this latter finding supports the notion of a time-saving re-processing of information during re-reading of previously highlighted text compared to text without highlighting (Yeari et al., 2017). This is particularly interesting, since the provision of the text-highlighting tool was found to benefit intertextual integration only of participants with a high number of re-readings. Hence, for readers with a high number of re-readings, those who had the possibility to highlight text showed *better* intertextual integration and also took *less* time for re-reading documents than those who did not have the possibility to highlight text. This indicates that participants who were not provided with the text-highlighting tool may have had to take more time to re-locate information during re-readings (cf. Yeari et al., 2017), which, however, did not benefit their intertextual integration.

The assumption that previously made highlights serve readers to focus their visual attention during re-readings was also at least partly supported by our exploratory analyses regarding fixation coverage (Goldberg & Kotval, 1999; Sharafi et al., 2015) during re-readings (that lasted at least 5 sec). Results showed significantly smaller fixation coverage in the with-highlighting than in the without-highlighting condition, however, only for two out of the five documents. Yet, it should be noted that our sample for the analyses regarding the fixation coverage during re-readings was quite small, because not all participants had re-read each document. Therefore, statistical power might have been too low to detect differences in the other documents.

5.5.4 Limitations and Future Work

This study does not come without limitations. First, we examined a homogenous sample of university students, which can be considered advanced readers. Yet, the fact that our findings showed an important role of re-reading (when text-highlighting was possible) even in a sample of presumably advanced readers, stresses the importance of reading actions even more. In this regard, future research should investigate whether trainings in re-reading might benefit readers who usually engage in such reading actions less.

Second, in the present study, participants read the documents on a touch display. This methodological approach was taken because we wanted to replicate and extend previous findings of a study in which documents had also been read on a touch display (Leroy et al., 2021), and because previous research has found indications that the effect of text-highlighting on readers' resulting text comprehension might be affected by the reading medium and, thus, the highlighting methodology (i.e., highlighting with the finger or highlighter-pen might be more effective than highlighting with a computer-mouse; e.g., Ben-Yehudah & Eshet-Alkalai, 2014; Goodwin et al., 2020). Hence, even though numerous applications exist that provide highlighting-tools particularly for touch devices, such as e-readers and tablets, future research should also investigate whether our findings regarding the moderating role of readers' engagement in re-reading in the effect of text-highlighting on intertextual integration can be replicated when documents are read on a computer screen and text is highlighted with a mouse.

Third, we assessed the number of re-readings as an indicator of the degree to which participants had compared and corroborated information across documents (cf. Anmarkrud et al., 2014; Goldman et al., 2012; List & Alexander, 2017). Yet, the fact that the extent of engaging in re-reading was also related to source-content integration suggests that re-readings may (also) be initiated for other reasons than for comparing and corroborating information *across* documents. Future research could use think-aloud methodology (e.g., Anmarkrud et al., 2014; Bråten, Anmarkrud, et al., 2014; Goldman et al., 2012; Wineburg, 1991) to investigate the reasons behind individuals' re-reading strategies when using a text-highlighting tool or not.

5.5.5 Conclusion

The present study replicated the findings by Leroy et al. (2021) of a moderating role of the number of re-readings on the effect of text-highlighting on readers' intertextual integration when reading multiple, partly conflicting documents about a health-related issue on a touch display. We extended these findings in several ways. First, additional analyses revealed that particularly amongst participants with a high number of re-readings, those who were provided with the text-highlighting tool had shorter re-reading times than those who read the documents without a text-highlighting tool. Hence, for participants with a high number of re-readings, the provision of a text-highlighting tool resulted in both time-saving re-processing and better integrated understanding. Second, an examination of a second measure of readers' integrated understanding, their source-content integration, revealed that participants' score in the source-content mapping task was positively related to their engagement in re-reading, regardless of whether the text-highlighting tool was provided or not.

Together, these findings suggest that both the characteristics of the reading environment and readers' engagement in reading actions play a role in readers' integrated understanding of multiple documents (cf. Britt et al., 2018; Rouet et al., 2017). This stresses the importance of assessing reading actions – such as readers' engagement in re-reading – when investigating the potential effects of characteristics of the reading environment on readers' integrated understanding. Furthermore, our findings suggest that the relation between the characteristics of the reading environment and readers' engagement in reading actions may differ depending on the concrete aspect of integrated understanding assessed. Related to this latter issue, also the optimal threshold to define re-readings might depend on the concrete measure of readers' integrated understanding, and, thus, should be chosen deliberately.

5. 6 Supplementary Material

Supplementary Table 5.6.1

Detailed description of the five documents.

Document type	Document's source name	Argumentative stance	Word count	LIX-score[a]	Central information conveyed	Embedded source
Journalist-authored article from a reputable liberal-conservative German newspaper	FAZ – Frankfurter Allgemeine Zeitung	Positive effects of UV radiation	320	49.6	Interview with a professor who describes a longitudinal study that revealed a positive effect of vitamin D on the reduction of cancer risk; he claims that sunrays may protect against all types of cancer through the production of vitamin D because vitamin D may prevent DNA-mutations, and recommends at least 30 min of daily sun exposure; explains how to calculate unharmful sun exposure time (self-protection time) as a function of sun protection factor of sun blockers and skin tone.	Professor Gillies from Harvard University Boston
Popular science article from a medical science magazine	Ärztezeitung	Positive effects of UV radiation	377	55.6	Interview with a professor who stated that cancer patients living close to the equator (i.e., with high sun exposure) have greater chances of recovery from cancer; he advocates that sunbeds may be a good alternative to natural sun exposure to produce enough vitamin D in winter; states that melanoma are mainly caused by long-wave sunray and vitamin D production is mainly triggered by short-wave sunrays, and suggests modifying sun blockers such that only short-wave sunrays can pass into the skin cells.	Professor Moan from the National CanceResearch Center

Popular science article from a university research magazine	Mundo – Wissenschafts-magazin der TU Dortmund	Negative effects of UV radiation	365	51.3	Interview with a professor who stated UV radiation causes DNA mutations and that using sunbeds before the age of 35 bears a 75% increased risk of cancer; claims that rather than exposing oneself to UV radiation, vitamin D should be supplemented instead of being promoted by the use of sunbeds.	Professor Fisher from State University New York
Public information text published by the German national cancer association	DKG – Deutsche Krebs-gesellschaft	Negative effects of UV radiation	326	54.0	Interview with a professor who describes that UV-radiation causes DNA damage which ultimately can lead to cancer; describes a study with 1800 kindergartners that suggests that, contrary to wearing clothes outside, sun blocker is effective in preventing skin cancer.	Professor Brog, dermatologist at the German Cancer Society
Excerpt from a high-school science textbook	Cornelsen Schulverlag Physik Oberstufe	General information about UV radiation	356	49.8	Describes UV radiation in neutral, academic terms; explains factors affecting the UV-index and the different wave lengths of UV-A, UV-B, and UV-C radiation.	none

[a]*Note.* LIX = readability formula according to Björnsson (1968).

Supplementary Table 5.6.2

Examples of different types of intertextual connections (second column) as well as original statements (translated from German) from participants, which, when integrated, yield the exemplary intertextual connection. The respective sources of the statements provided in the third column are indicated in brackets by their original German name, as well as their stance taken (-/+).

Type of intertextual connection	Example of the respective type of intertextual connection	Specific pieces of information each stemming from one distinct document
Global conflicting connection	Scientific findings about health effects of UV radiation are controversial. On the one hand, UV-radiation can cause skin cancer, on the other hand, UV-radiation promotes the production of vitamin D, which in turn is said to be able to prevent cancer.	• UV radiation is one of the most common causes of skin cancer. (*information conveyed by both negative documents,* -) • Sun rays accelerate the body's own production of vitamin D. This can prevent the development of cancer and many other diseases. *(information conveyed by both positive documents,* +)
Specific conflicting connection	One professor suggests using sunbeds in winter months to trigger the production of vitamin D, *whereas* another professor reports on a study that found an increased likelihood to develop skin cancer amongst sunbed users.	• Studies show that people who have already used a sunbed before the age of 35 are 75% more likely to develop skin cancer. (*Mundo,* -) • In the dark winter months, we cannot produce enough vitamin D naturally. One possible way to trigger the production of vitamin D in those months are sunbeds. (*Ärztezeitung,* +)
Complementary connection (positive stance)	Findings of a longitudinal study with 50,000 men and a study with participants who lived closer to the equator suggest that vitamin D can decrease the likelihood to develop cancer.	• In a longitudinal study with 50,000 men, they found that people with high levels of vitamin D in their bodies were less susceptible to cancer. (*FAZ,* +) • Sunrays stimulate the body's vitamin D production. […] The findings show, for example, that people who live closer to the equator have better chances of survival if their internal organs are affected by cancer than people who live further away from the equator. (*Ärztezeitung,* +)

Complementary connection (negative stance)	Studies with sunbed users and kindergartners have shown that UV-radiation can cause skin cancer; and that wearing clothes seems to be the only way to prevent it.	• Studies show that people who have already used a sunbed before the age of 35 are 75% more likely to develop skin cancer. (*Mundo,* -) • A study with 1800 kindergartners found no effect of using sunscreen on the likelihood of developing skin cancer. Wearing clothes, however, reduced the likelihood of skin cancer in participants. (*DKG,* -)
Complementary connection of factual information	Lower-wave UV radiation penetrates deeper into the skin.	• The lower the wavelength, the higher the energy content of UV radiation. (*Mundo,* -) • UV radiation penetrates deeper levels of the skin, the higher its energy content is. (*Ärztezeitung,* +)
Intertextual connection across sentences	Depending on a person's skin type, one should not stay in the sun for longer than 40 minutes without protection. However, this again depends on the season, weather conditions and geographical location.	• People with skin type IV are recommended not to stay in the sun for more than 40 minutes without protection. In contrast, people with skin type I should stay without sun protection for a maximum of 10 minutes on a summer day, even in Germany. (*DKG,* -) • Besides the season, other factors such as geographical location, altitude, and weather conditions also influence the strength of UV radiation. (*Cornelsen)*
Connection between embedded sources[a] or <u>document sources</u>[b]	Researchers from **Heidelberg and Boston** argue that UV radiation has a positive effect on human health. Both, the <u>Frankfurter Allgemeine and the Ärztezeitung</u> report positive effects of UV radiation on human health.	• A team of researchers led by **Professor Edward Gillies of Harvard University in Boston** conducted a long-term study with 50,000 men. It was found that people with high levels of vitamin D in their bodies were less susceptible to cancer. (<u>*FAZ*</u>*,* +) • “Sunlight accelerates the body's own production of vitamin D. This can prevent the development of cancer and many other diseases," said **Professor Moan, researcher at the National Center for Tumor Diseases in Heidelberg**. (<u>*Ärztezeitung*</u>*,* +)

Complex utterance containing two intertextual connections in one sentence	To prevent **DNA from clotting** as a consequence of UV radiation, **sun blockers should be modified** in a way that long-wave rays (UVA) are blocked completely, but shorter wave rays (UVB) can pass, because the latter ones promote the production of vitamin D in our bodies.	• UV-radiation penetrates the skin and causes the **DNA in skin cells to clot**, which can ultimately trigger skin cancer. (*DKG*) • UVA-rays have a wavelength between 280 and 315 nm and are, thus, the rays with the longest wavelength. UVB rays have a wavelength between 280 and 315 nm. UVC rays have the shortest wavelength (100 280 nm). They are completely absorbed by the atmosphere and can't be found on Earth. (*Cornelsen*) • Melanoma are mainly caused by long-wave sunrays, whereas the production of vitamin D is mainly triggered by short-wave sunrays. Thus, **sun blockers should be modified such that they block the long-wave rays**, but let the short-wave rays pass. (*Ärztezeitung*)

[a] These intertextual connections were only valid if the embedded sources mentioned were uniquely attributable to different documents.

[b] These intertextual connections were only valid if the document source names were correct.

Note. In the last two examples, text printed in bold and/or underlined in the second column represents one intertextual connection resulting from an integration of text printed in bold and/or underlined in the third column.

Supplementary Table 5.6.3

Means (and standard deviations) of the number of re-readings with the respective duration-threshold as a function of text-highlighting, as well as inferential statistics.

Threshold	Document(s)	Condition		Inferential statistics
		With highlighting	Without highlighting	
# re-readings ≥ 1 sec	All	9.75 (7.70)	10.92 (7.28)	$t(92.45) = -0.76, p = .448$
# re-readings ≥ 2 sec	All	8.66 (6.69)	9.63 (6.52)	$t(92.79) = -0.71, p = .478$
# re-readings ≥ 3 sec	All	7.72 (5.96)	8.85 (5.94)	$t(92.94) = -0.93, p = .357$
# re-readings ≥ 4 sec	All	7.02 (5.25)	8.27 (5.53)	$t(92.91) = -1.13, p = .262$
# re-readings ≥ 5 sec	All	6.45 (4.81)	7.88 (5.37)	$t(92.29) = -1.37, p = .175$

Table 5.6.4

Inferential statistics for main effects of and interactions between text- highlighting and the number of re-readings (for the 5 different re-reading thresholds) on the integrated understanding measures (number of intertextual connections and source-content mapping score).

Measure	Re-reading threshold	Inferential statistics
# intertextual connections	1 sec	text-highlighting: β = .10, $t(91)$ = 0.95, p = .358 #re-readings ≥ 1 sec: β = .12, $t(91)$ = 1.17, p = .244 interaction: β = .16, $t(91)$ = 1.51, p = .135
	2 sec	text-highlighting: β = .10, $t(91)$ = 0.94, p = .351 # re-readings ≥ 2 sec: β = .14, $t(91)$ = 1.34, p = .183 **interaction**: β = .19, $t(91)$ = 1.88, p = .063
	3 sec	text-highlighting: β = .10, $t(91)$ = 0.97, p = .336 # re-readings ≥ 3 sec: β = .14, $t(91)$ = 1.33, p = .186 **interaction**: β = .20, $t(91)$ = 1.99, p = .049
	4 sec	text-highlighting: β = .10, $t(91)$ = 1.02, p = .312 # re-readings ≥ 4 sec: β = .15, $t(91)$ = 1.45, p = .150 **interaction**: β = .25, $t(91)$ = 2.47, p = .016
	5 sec	text-highlighting: β = .11, $t(91)$ = 1.04, p = .302 # re-readings ≥ 5 sec: β = .14, $t(91)$ = 1.37, p = .173 **interaction**: β = .29, $t(91)$ = 2.56, p = .012
Source-content-mapping score	1 sec	interaction: β = -.01, $t(91)$ = -0.07, p = .941 regression without interaction term: text-highlighting: β = -.02, $t(92)$ = -0.15, p = .879 **# re-readings**: β = .23, $t(92)$ = 2.25, p = .027
	2 sec	interaction: β = .02, $t(91)$ = 0.23, p = .818 regression without interaction term: text-highlighting: β = -.02, $t(92)$ = -0.17, p = .862 **# re-readings**: β = .21, $t(92)$ = 2.10, p = .039
	3 sec	interaction: β = .03, $t(91)$ = 0.31, p = .759 regression without interaction term: text-highlighting: β = -.02, $t(92)$ = -0.15, p = .879 **# re-readings**: β = .19, $t(92)$ = 1.82, p = .072
	4 sec	interaction: β = .04, $t(91)$ = 0.36, p = .721 regression without interaction term: text-highlighting: β = -.01, $t(92)$ = -0.13, p = .901 **# re-readings**: β = .18, $t(92)$ = 1.73, p = .088
	5 sec	interaction: β = .10, $t(91)$ = 0.73, p = .466 regression without interaction term: text-highlighting: β = -.01, $t(92)$ = -0.08, p = .940 **# re-readings**: β = .18, $t(92)$ = 1.78, p = .079

Note. (Marginally) significant results are displayed in bold print. All p-values are two-tailed p-values.

Supplementary Table 5.6.5

Results of the simple slopes analyses according to Aiken and West (1991) regarding the interaction effect between text-highlighting (with vs. without) and the number of re-readings (z-standardized continuous predictor) for a given re-reading threshold, on the number of intertextual connections *in participants' essays.*

Re-reading threshold	Simple slopes	
	# re-readings	Condition
1 sec[a]	**high (+1 *SD*)**: $\beta = .43$, $t(91) = 1.72$, $p = .088$	**with highlighting**: $\beta = .48$, $t(91) = 1.94$, $p = .056$
	low (-1 *SD*): $\beta = -.11$, $t(91) = -0.42$, $p = .677$	without highlighting: $\beta = -.06$, $t(91) = -0.25$, $p = .806$
2 sec	**high (+1 *SD*)**: $\beta = .50$, $t(91) = 2.00$, $p = .049$	**with highlighting**: $\beta = .57$, $t(91) = 2.30$, $p = .024$
	low (-1 *SD*): $\beta = -.17$, $t(91) = -0.67$, $p = .503$	without highlighting: $\beta = -.10$, $t(91) = -0.40$, $p = .693$
3 sec	**high (+1 *SD*)**: $\beta = .52$, $t(91) = 2.09$, $p = .039$	**with highlighting**: $\beta = .58$, $t(91) = 2.34$, $p = .021$
	low (-1 *SD*): $\beta = -.18$, $t(91) = -0.73$, $p = .466$	without highlighting: $\beta = -.12$, $t(91) = -0.49$, $p = .627$
4 sec	**high (+1 *SD*)**: $\beta = .61$, $t(91) = 2.46$, $p = .016$	**with highlighting**: $\beta = .68$, $t(91) = 2.69$, $p = .009$
	low (-1 *SD*): $\beta = -.26$, $t(91) = -1.04$, $p = .302$	without highlighting: $\beta = -.18$, $t(91) = -0.76$, $p = .447$
5 sec	**high (+1 *SD*)**: $\beta = .63$, $t(91) = 2.52$, $p = .013$	**with highlighting**: $\beta = .69$, $t(91) = 2.62$, $p = .010$
	low (-1 *SD*): $\beta = -.27$, $t(91) = -1.10$, $p = .275$	without highlighting: $\beta = -.22$, $t(91) = -0.91$, $p = .360$

[a] Please note that the interaction effect between text-highlighting and the number of re-readings was not significant for the 1 sec threshold. All p-values are two-tailed p-values.

Supplementary Table 5.6.6

Means (and standard deviations) for re-reading times as a function of text-highlighting.

Reading-time measures	Condition		Inferential statistics
	With highlighting	Without highlighting	
Re-reading time (re-accesses ≥ 1 sec) [min]	3.35 (2.38)	4.58 (3.20)	$t(86.71) = -2.14$, $p = .035$
Re-reading time (re-accesses ≥ 2 sec) [min]	3.32 (2.36)	4.56 (3.19)	$t(86.62) = -2.14$, $p = .035$
Re-reading time (re-accesses ≥ 3 sec) [min]	3.27 (2.32)	4.52 (3.18)	$t(86.05) = -2.20$, $p = .031$
Re-reading time (re-accesses ≥ 4 sec) [min]	3.23 (2.30)	4.42 (3.16)	$t(85.98) = -2.11$, $p = .038$
Re-reading time (re-accesses ≥ 5 sec) [min]	3.19 (2.30)	4.39 (3.15)	$t(85.96) = -2.13$, $p = .036$

Note. All p-values are two-tailed p-values.

Supplementary Table 5.6.7

Results of the moderated linear regression analyses and respective simple slopes analyses according to Aiken and West (1991) regarding the interaction effect between text-highlighting (with vs. without) and the number of re-readings (z-standardized continuous predictor) for a given re-reading threshold, on the re-reading time.

Re-reading threshold	Inferential statistics	
	Linear regression	Simple slopes
1 sec	**highlighting**: $\beta = -.16$, $t(91) = -2.37$, $p = .020$ **# re-readings**: $\beta = .73$, $t(91) = 10.84$, $p < .001$ **interaction**: $\beta = -.17$, $t(91) = -2.51$, $p = .014$	**with highlighting**: $\beta = 1.62$, $t(91) = 6.05$, $p < .001$ **without highlighting**: $\beta = 2.59$, $t(91) = 9.19$, $p < .001$ **# re-readings high (+1 *SD*)**: $\beta = -.94$, $t(91) = -3.45$, $p < .001$ # re-readings low (-1 *SD*): $\beta = .03$, $t(91) = 0.11$, $p = .910$
2 sec	**highlighting**: $\beta = -.16$, $t(91) = -2.47$, $p = .016$ **# re-readings**: $\beta = .74$, $t(91) = 11.17$, $p < .001$ **interaction**: $\beta = -.17$, $t(91) = -2.57$, $p = .012$	**with highlighting**: $\beta = 1.62$, $t(91) = 6.15$, $p < .001$ **without highlighting**: $\beta = 2.60$, $t(91) = 9.59$, $p < .001$ **# re-readings high (+1 *SD*)**: $\beta = -.95$, $t(91) = -3.56$, $p < .001$ # re-readings low (-1 *SD*): $\beta = .02$, $t(91) = 0.08$, $p = .934$
3 sec	**highlighting**: $\beta = -.15$, $t(91) = -2.35$, $p = .021$ **# re-readings**: $\beta = .75$, $t(91) = 11.63$, $p < .001$ **interaction**: $\beta = -.17$, $t(91) = -2.60$, $p = .011$	**with highlighting**: $\beta = 1.65$, $t(91) = 6.39$, $p < .001$ **without highlighting**: $\beta = 2.60$, $t(91) = 10.04$, $p < .001$ **# re-readings high (+1 *SD*)**: $\beta = -.90$, $t(91) = -3.49$, $p < .001$ # re-readings low (-1 *SD*): $\beta = .05$, $t(91) = 0.19$, $p = .849$
4 sec	**highlighting**: $\beta = -.13$, $t(91) = -2.00$, $p = .049$ **# re-readings**: $\beta = .76$, $t(91) = 12.17$, $p < .001$ **interaction**: $\beta = -.15$, $t(91) = -2.38$, $p = .020$	**with highlighting**: $\beta = 1.72$, $t(91) = 6.74$, $p < .001$ **without highlighting**: $\beta = 2.57$, $t(91) = 10.57$, $p < .001$ **# re-readings high (+1 *SD*)**: $\beta = -.77$, $t(91) = -3.08$, $p = .003$ # re-readings low (-1 *SD*): $\beta =.07$, $t(91) = 0.28$, $p = .779$
5 sec	**highlighting**: $\beta = -.11$, $t(91) = -1.75$, $p = .084$ **# re-readings**: $\beta = .77$, $t(91) = 12.51$, $p < .001$ **interaction**: $\beta = -.12$, $t(91) = -1.88$, $p = .064$	**with highlighting**: $\beta = 1.86$, $t(91) = 7.13$, $p < .001$ **without highlighting**: $\beta = 2.50$, $t(91) = 10.78$, $p < .001$ **# re-readings high (+1 *SD*)**: $\beta = -.63$, $t(91) = -2.54$, $p = .013$ # re-readings low (-1 *SD*): $\beta =.03$, $t(91) = 0.11$, $p = .914$

Note. (Marginally) significant results (for two-tailed tests) are displayed in bold print.

6. General Discussion

The three studies of the present dissertation investigated the potential fostering effects of a reading environment enabling a simultaneous rather than sequential presentation of multiple documents and the possibility to highlight text on different aspects of multiple document comprehension as well as the roles that specific reading interactions (i.e., interactions with documents during reading) might play in these effects. As such, the present work contributes to previous research in the field of multiple document comprehension in several ways. First, previous work investigating the effect of a reading environment enabling a simultaneous rather than imposing a sequential presentation of documents on multiple document comprehension has used only partly simultaneous presentation formats (i.e., split-screens or a simultaneous presentation of up to three documents; (Lombard et al., 2021; Olive et al., 2008; Wiley, 2001), or used only two short documents as reading material (Kobayashi, 2009). In contrast, Studies 1 and 2 of the present dissertation enabled a fully simultaneous presentation of all documents provided (i.e., six documents in Study 1 and five documents in Study 2). Second, previous work investigating the effect of text-highlighting on multiple document comprehension is rather inconclusive (Dunlosky et al., 2013; Miyatsu et al., 2018), and implemented digital text-highlighting with a computer mouse (e.g., Ben-Yehudah & Eshet-Alkalai, 2018), which, however, has been suggested to be more cumbersome than highlighting text with a pen on paper (Goodwin et al., 2020). While the present work did not investigate the effect of text-highlighting that is initiated via a computer mouse versus via touch, the present work (i.e., Studies 1 and 3) contributes to this line of research in that it was the first to investigate the effect of digital text-highlighting on multiple document comprehension with touch-based highlighting, which can be assumed to be more comparable (regarding the effort required) to highlighting printed text with a pen. Third, the retention hypothesis of highlighted text suggests that being able to highlight text might especially be beneficial during re-reading (rather than during initial reading, i.e., when initiating the highlighting), and previous findings suggest that highlights might "save time and effort in reprocessing the textual content" (Yeari et al., 2017, p.2). However, no previous work had investigated the time-saving re-processing of text (i.e., during re-reading) when text-highlighting was possible versus not. By means of eye-tracking technology, this was addressed in Study 3 by also investigating whether readers who are able to highlight text would re-read smaller proportions of texts than those not able to highlight text. Fourth, previous work investigating the effects of document presentation and text-highlighting on multiple document comprehension only investigated single aspects of readers' multiple document comprehension, or only their overall quality of multiple document comprehension rather than specific aspects of multiple document comprehension. The present work thus extends previous findings by assessing several aspects of readers' multiple document comprehension. Fifth, the present work was the first to investigate the role of reading interactions in the effects that the document presentation and text-highlighting have on multiple document comprehension.

In the following subsections, first, the results of the present work regarding the effects of document presentation and text-highlighting on multiple document comprehension and the roles of specific reading interactions therein are discussed in the light of previous work and, specifically, in the light of the reading processes described by the RESOLV model (Britt et al., 2018; Rouet et al., 2017). For these purposes, Table 6.1 comparatively summarizes the findings of the three studies of the present dissertation. Subsequently, the fact that the main effects of document presentation and text-highlighting on intertextual integration that was found in Study 1 could not be replicated in Studies 2 and 3, respectively, will be discussed. Finally, suggestions for future research will be made regarding further elaborations on the findings of the present work, and conclusions will be drawn from the present findings on how readers might best be supported in building a comprehensive understanding of multiple, partly conflicting documents.

6.1 The Effect of Document Presentation on Multiple Document Comprehension (RQ1), and the Role of Readers' Spontaneous Engagement in Grouping Documents Therein (RQ2)

Regarding research question 1 (RQ1), which asked the question whether a simultaneous as compared to a sequential presentation of multiple documents would foster multiple document comprehension, Study 1 (N = 126) revealed that participants in the simultaneous condition included more intertextual connections in their essays than those in the sequential condition (i.e., confirming Hypothesis 1). However, this main effect of document presentation on intertextual integration could not be replicated in Study 2 (N = 95; i.e., contradicting Hypothesis 1). Furthermore, regarding the exploratory investigations regarding research question 1, participants in the simultaneous condition of Study 1 were more likely to cite at least one source in their essays than those in the sequential condition, whereas neither the number of correctly recalled source names, nor the score in the source-content mapping task differed between presentation conditions in Study 2. Hence, the findings of Studies 1 and 2 yield conflicting results regarding the effect of a simultaneous rather than sequential presentation of multiple documents on multiple document comprehension, and might overall not be convincing.

However, the fact that the main effects found in Study 1 could not be replicated in Study 2 might be due to some methodological differences between the studies, as will be discussed in Section 6.3.

Regarding research question 2 (RQ2), which asked the question of the role of readers' spontaneous engagement in grouping documents in the effect of document presentation on multiple document comprehension, Study 1 revealed a full mediation of grouping in the effect of document presentation on intertextual integration (i.e., confirming Hypotheses 2.1, 2.2, and 2.3). That is, in Study 1, presenting documents simultaneously rather than sequentially had a positive effect on intertextual integration due to the fact that the simultaneous presentation bore a greater affordance to group documents during reading (cf. Haber et al., 2014; O'Hara et al., 2002; Shibata et al., 2013; Takano et al., 2015). Study 2, in which documents could only be re-arranged in the simultaneous condition, replicated the finding of the positive relation between grouping documents and intertextual integration (i.e., further confirming Hypothesis 2.2). However, while in Study 1 participants who had grouped documents during reading were also more likely to include source citations in their essays, participants in Study 2 who had grouped documents in the simultaneous condition did not recall more source names correctly than those in the simultaneous condition who had not grouped documents during reading. Yet, Study 2 revealed that participants in the simultaneous condition who had grouped documents during reading scored higher in the source-content mapping task (which was not assessed in Study 1) than those in the simultaneous condition who had not grouped documents during reading.

Together, the full mediation of grouping documents in the effect of document presentation on intertextual integration in Study 1 as well as the findings of Study 2 that participants who engaged in grouping documents during reading showed better intertextual and source-content integration than those who did not group documents imply that a simultaneous presentation of documents bears the potential to support multiple document comprehension *when* readers make use of the reading environment (i.e., use the screen space to group documents). These findings are thus somewhat in line with previous work that found a positive effect of a partly simultaneous presentation of multiple documents on multiple document comprehension (i.e., the overall quality or level of integration reflected in essays) only when participants had received instructions to use the respective interface provided (Lombard et al., 2021; Wiley, 2001).

Table 6.1

*Comparative presentation of the results across the three studies of the present dissertation concerning the effects of the reading environment on specific aspects of multiple document comprehension and the role of reading interactions therein. Where an analysis of the respective effects and relations was possible (i.e., where the respective reading environment was manipulated or reading interactions assessed), significance of the respective statistical analyses is denoted by *** for* $p < .001$*, ** for* $p < .01$*, * for* $p < .05$*, and † for* $p < .10$*. '✓' denotes agreement, and '✗' denotes disagreement with the respective hypothesis, where no such symbols are provided, effects and relations were exploratively investigated.*

Reading Environment	Measure	Role of reading interaction	Study 1 (Sim vs. Seq + HL vs. noHL)	Study 2 (Sim vs. Seq; noHL)	Study 3 (HL vs. noHL; Seq)
Sim vs. Seq	Intertextual integration		* Sim > Seq ✓	ns. ✗	
		Grouping[2]	Full mediation ✓	Sim: grouping > no grouping ✓	
	Source names[1]		** Sim > Seq ✓	ns. ✗	
		Grouping[2]	*Not applicable*[3]	Sim: grouping = no grouping	
	Source-content integration			ns.	
		Grouping[2]		Sim: grouping > no grouping	
HL vs. noHL	Intertextual integration		** HL > noHL ✓	*Only noHL condition*	ns. ✗
		# revisits / # re-readings	Seq: HL-effect only for medium / high #revisits & pos. correlation only in HL ✓	Seq (noHL): no correlation ✓	HL-effect only for high #revisits & pos. correlation only in HL ✓
	Source names[1]		ns.	*Only noHL condition*	
		# revisits / # re-readings	*Not applicable*[3]	Seq (noHL): pos. correlation	
	Source-content integration			*Only noHL condition*	ns.
		# revisits / # re-readings		Seq (noHL): pos. correlation	Pos. correlation (for re-reading thresholds of < 3 sec)

Note: 'HL' abbreviates 'highlighting', 'Sim' abbreviates 'simultaneous condition', and 'Seq' abbreviates 'sequential condition'.

[1] In Study 1, participants' memory for source names was assessed in whether they had spontaneously cited at least one source in their essays, whereas in Study 2, participants' memory for source names was assessed in a source name recall task.

[2] Please note that whether participants had grouped documents according to their overall stance was only assessed in the simultaneous condition in Study 2, but in both the sequential and simultaneous condition in Study 1, and that re-reading was only assessed in the respective sequential condition of Studies 1, 2 and 3 – with different definitions (i.e., in Studies 1 and 2, *revisits* were defined as each re-accessing of a document that lasted at least 1 sec after an initial access to that document that lasted at least 1 sec, whereas in Study 3, *re-readings* were defined for each re-accessing of a document that lasted at least 1 sec (i.e., for each revisit) after the document had been fully read for the first time, which was assessed by means of eye-tracking methodology).

[3] Since the number of spontaneous source citations in essays was rather low, this analysis was not applicable.

The fact that across Studies 1 and 2, the relation between grouping and multiple document comprehension was positive for measures requiring information integration (i.e., intertextual integration and source-content integration), but not significant for participants' memory for sources furthermore implies that this reading interaction is related to linking and comparing information (across documents) during reading. Yet, the present findings do not allow to draw conclusions regarding the nature of this positive relation between grouping and information integration. That is, whether participants who grouped documents did so *as a result* of their information integration or *in order to* support their information integration remains unclear from the present work. In this regard, however, previous work suggests that readers might engage in such re-organization to reduce their mental load (Skuballa et al., 2018). That is, by grouping documents taking the same stance, readers create a layout of documents that bears contextual information about which document takes a positive versus negative stance. Furthermore, irrespective of participants' reason for grouping documents, the resulting spatial proximity of documents taking the same stance might have further supported information integration across documents by the easier re-accessing of documents or specific pieces of information (Andrews et al., 2010; Bi & Balakrishnan, 2009; Haber et al., 2014; Jang et al., 2012; Jang & Schunn, 2012; O'Hara et al., 2002; O'Hara & Sellen, 1997; Takano et al., 2015), and thus easier cross-document information comparison and integration (Jang et al., 2011; Jang & Schunn, 2012; Olive et al., 2008) compared to when the document is further away, or needs to be re-opened in the sequential condition.

As such, the findings of Studies 1 and 2 regarding the effect of document presentation on multiple document comprehension, and the role of grouping documents therein, can be interpreted in line with the benefit-cost assumption made by the RESOLV model (Britt et al., 2018; Rouet et al., 2017): First, a reading environment enabling a simultaneous rather than imposing a sequential presentation of multiple documents might bear a greater affordance to strategically re-arrange documents during reading for two reasons. On the one hand, the *cost* of re-arranging documents according to their overall stance is lower in a simultaneous presentation since the already re-arranged documents are always visible, and readers do not need to keep in memory which location reflects what overall stance or which document they have previously re-arranged to what location. On the other hand, readers might be more likely to re-arrange documents in a simultaneous presentation because they evaluate the *benefits* as higher compared to a sequential condition, since the resulting spatial proximity of documents can reduce readers' cognitive load (Skuballa et al., 2018) as well as support their information integration more in a simultaneous than in a sequential condition, in which documents are displayed in the same or similar location, but only one of them is visible at a time. Second, and in relation to the latter suggestion, participants who had grouped documents might have benefited from the resulting spatial proximity of documents in that it reduced the cost or increased the benefit of cross-document information comparison, thus yielding better intertextual and source-content integration.

6.2 The Effect of the Possibility to Highlight Text on Multiple Document Comprehension (RQ3), and the Role of Readers' Engagement in Re-Reading Documents Therein (RQ4)

Regarding research question 3 (RQ3), which asked the question whether the provision of a digital text-highlighting tool would support multiple document comprehension, in line with previous findings (Kobayashi, 2009), Study 1 revealed a beneficial effect of text-highlighting on the number of intertextual connections in essays. However, in Study 1, there was no effect of text-highlighting on source citations in essays, and the main effect of text-highlighting on intertextual integration could not be replicated in Study 3 in which all participants read documents in a sequential condition. Furthermore, in Study 3 (N = 95), participants' source-content integration was unaffected by whether they were able to highlight text or not. Hence, regarding the hypothesized effect of a digital text-highlighting tool on intertextual integration (Hypothesis 3), Studies 1 and 3 are inconclusive – as is previous work on the effect of text-highlighting on retention and text comprehension overall (Dunlosky et al., 2013; Miyatsu et al., 2018). However, this may also be due to the same methodological differences between the studies that might have caused the discrepant findings regarding the effect of document presentation on multiple document comprehension between Studies 1 and 2 (see Section 6.3).

Regarding research question 4 (RQ4), which posed the question of the role of re-reading in the effect of text-highlighting on multiple document comprehension, Study 1 revealed an interaction between text-highlighting and the number of revisits on the number of intertextual connections in essays. That is, text-highlighting was only beneficial regarding intertextual integration for participants with many revisits (i.e., confirming Hypothesis 4.1), and there was a positive relation between the number of intertextual connections in essays and the number of revisits only in the with-, but not in the without-highlighting condition (i.e., confirming Hypothesis 4.2). These findings could be replicated across all three studies of the present dissertation. That is, in Study 2, in which participants could not highlight text, there was also no relation between the number of revisits and the number of intertextual connections in essays. In Study 3, like in Study 1, the positive correlation between the number of re-readings and the number of intertextual connections in essays was only present for participants in the with-, but not for those in the without-highlighting condition, and text-highlighting was only beneficial for those with a high number of re-readings.

Moreover, Study 3 investigated potential differences in task processing between participants who could and could not highlight text during reading that might explain the moderating role of the number of revisits (or re-readings) in the effect of text-highlighting on intertextual integration. Eye-tracking data revealed that participants who were able to highlight text re-read smaller portions of text in two out of the five documents compared to those who were not able to highlight text (i.e., partly confirming Hypothesis 4.3). The fact that this was only found for two out of the five documents might be somewhat little convincing regarding the assumption that readers of highlighted text would re-read smaller portions (i.e., presumably the highlighted parts) of text than readers of non-highlighted text - however, building on a sample of 95 participants in that study, the number of participants who actually re-read the respective document at least once (i.e., the dataset used for these analyses) ranged between 57 and 72 participants (*SD* = 5.68) per document. This resulted in data of between 25 and 32 participants for one document in the with-highlighting condition, and data of between 32 and 40 participants in the without-highlighting condition, which still vary greatly in the time spent re-reading. Overall, these results therefore ought to be interpreted with caution since the rather small dataset might point towards a power issue in this analysis (Ärztezeitung: 30 and 34 participants; FAZ: 31 and 37 participants; DKG: 25 and 32 participants; Mundo: 32 and 40 participants; Cornelsen: 33 and 35 participants in the with- and without-highlighting condition, respectively).

Nonetheless, the findings of the present work regarding the interaction effect between the number of re-readings and the number of intertextual connections in essays together with the findings of the lower re-reading fixation coverage (at least in two documents) are especially interesting when also considering the additional analysis in Study 3 regarding the time participants in the with- and without-highlighting condition took for re-reading. This analysis revealed a somewhat reversed picture compared to that regarding the role of re-reading in the effect of text-highlighting on intertextual integration (see Figures 5.5 and 5.6). Namely, while the re-reading time was overall positively correlated with the number of re-readings, there was a difference in re-reading time across highlighting conditions only amongst participants with a high number of re-readings. In contrast to the interaction effect regarding the number of intertextual connections in essays, however, amongst those participants with a high number of re-readings, participants in the without-highlighting condition took longer for re-reading (whilst including less intertextual connections in essays) compared to participants in the with-highlighting condition (who included more intertextual connections in essays). Hence, the additional time participants with a high number of re-readings in the without-highlighting condition took for re-reading compared to those in the with-highlighting condition, did not benefit their intertextual integration. Together with the finding of the smaller re-reading fixation coverage (in two out of five documents) for participants in the with- compared to those in the without-highlighting condition, these findings suggests that there is a time saving effect of text-highlighting on information re-processing (Yeari et al., 2017) *without* any negative influence on intertextual integration. That is, highlights might navigate and focus readers' visual attention during re-reading, whereas re-reading without highlights might require searching the respective information visually (cf. O'Hara et al., 2002), thus "scanning" larger portions of the text and taking longer to re-read information. Therefore, the findings of Studies 1 and 3 regarding the role of readers' re-reading behavior in the effect of text-highlighting on intertextual integration support the assumption that highlights serve a retrieval function (rather than an encoding function).

In contrast to those findings regarding the relation between re-reading and intertextual integration, Study 2 revealed a positive relation between revisiting documents and memory for source names (List & Alexander, 2018), and Studies 2 and 3 revealed a positive relation between revisits (in Study 2) or re-readings (in Study 3) and source-content integration – irrespective of whether text-highlighting was possible, respectively. This is in line with previous work showing that participants who revisited documents were more likely to include source citations in their essays (List, Stephens, et al., 2019; List & Alexander, 2018) and scored higher in a source-content mapping task (Delgado et al., 2020) than those who did not revisit documents. Please note that the present work thereby extends and further qualifies this previous work in the fact that revisiting was assessed as a *continuous* rather than dichotomous factor, showing *positive correlations* with the number of source names recalled and the number of correctly mapped statements to source logos. This suggests that readers re-access documents not only to compare or link information (i.e., in the text) across documents (Goldman, Braasch, et al., 2012; List & Alexander, 2018; Wiley et al., 2009; Wineburg, 1991), but also to lay their attention on source information – which is likely to be a result of cross-document conflict detection (Braasch & Bråten, 2017). For the sake of conciseness, in the manuscript of Study 3 (see Section 5), no further distinction was made between the extent to which participants had re-processed source information versus text during re-readings. Yet, processing source information is as essential as processing the statements of a document for building source-content links. Therefore, Section 6.2.1 provides additional analyses of the eye-tracking data of

Study 3 regarding readers' attention to source information versus text during re-reading in relation to multiple document comprehension.

Overall, the findings regarding the role of re-reading in the effect of text-highlighting on multiple document comprehension can also be interpreted in line with the benefit-cost assumption of the RESOLV model (Britt et al., 2018; Rouet et al., 2017): First, the fact that the number of revisits (or re-readings) did not differ across highlighting conditions suggests that participants evaluated the costs and benefits of re-accessing a document (i.e., in the sequential condition) equally, irrespective of whether they were able to highlight text or not. Second, at re-access, readers of *highlighted* text might end up with a negative benefit-cost assessment of further re-reading sooner than readers of non-highlighted text, thus taking less time for re-reading overall – potentially because highlights make them evaluate the re-reading process as more effective since highlights navigate their attention to presumably important parts of the text. Third, regarding intertextual integration, the benefit of re-accessing documents is higher with compared to without highlighting. Of note, Study 3 revealed that this beneficial effect of text-highlighting during re-reading is greater than a potential beneficial effect of taking additional time for re-reading without text-highlighting.

On a last note, these findings specifically contribute to research on the effectiveness of text-highlighting on multiple document comprehension (or text comprehension in general) in that they reveal an important role of re-reading in the effect. Specifically, the rather inconclusive findings regarding the effectiveness of text-highlighting in previous work might be explained by a large variety in participants' re-reading behavior across studies. Since no previous work in the area has examined the role of re-reading, or reported on participants' re-reading behavior, however, this assumption can not be tested, and further investigations regarding the role of re-reading behavior in the effect of text-highlighting on, for instance, retention, are necessary.

6.2.1 Readers' Visual Attention to Source Information Versus Text During Re-Reading. Additional Analyses of Eye-Tracking Data in Study 3

For the sake of conciseness, in the manuscript of Study 3, it was decided to only report analyses regarding the role of the number of re-readings in the effect that text-highlighting has on readers' multiple document comprehension, rather than to further distinguish between their visual attention to texts and source information during re-readings. Yet, a differentiation between readers' visual attention on source information versus on text during re-readings is particularly interesting especially for the investigation of the role of re-reading behavior for source-content integration (e.g., Delgado et al., 2020; Kammerer, Meier, et al., 2016). Eye-tracking data of Study 3 were thus re-analyzed by defining these specific two AOIs for each document. Table 6.2 presents the fixation durations on text versus source information (i.e., source logos) during initial readings and re-readings, whereas, for correspondence to the reported analyses in the manuscript for Study 3, all analyses regarding re-readings are reported using the 1 and 3 sec re-reading thresholds only. Regarding the relations between fixation durations (i.e., on texts and source-logos) and multiple document comprehension, correlational analyses were calculated where there was a difference in fixation duration between conditions, and moderated linear regression analyses were calculated (with fixation duration as z-standardized predictor) where there was no difference in fixation duration between conditions.

Please note that analyses regarding potential relations to participants' intertextual integration were only conducted with the respective fixation durations *on text*, since there is no indication for a relation between readers' attention to source information and information integration (i.e., pieces of text) across documents.

Regarding participants' *fixation duration on texts*, analyses revealed that there was no difference across highlighting conditions for overall reading time. However, in line with the difference in initial reading and re-reading times across conditions, fixation durations on text were significantly longer in the with- than in the without-highlighting condition during initial readings, but significantly shorter during re-reading. Regarding participants' intertextual integration, the relation between the number of intertextual connections in essays showed no relation to participants' fixation duration on texts – neither for reading time overall, nor for initial readings, or for re-readings. Together with the findings of the present work that in the highlighting-condition, but not in the without-highlighting condition, intertextual integration was positively related to the number of re-readings (or revisits in Studies 1 and 2), this finding further suggests that the *quality* rather than the quantity of re-reading is essential for intertextual integration, that is, switching between documents to compare or re-evaluate information rather than reading text longer (and thus fixating it longer) is required for building intertextual connections. Moreover, participants' score in the source-content mapping task was positively related to fixation duration on text during re-readings *only* for participants in the with-highlighting condition. This is somewhat in parallel with the findings of the present work regarding the positive relation between the number of re-readings (or revisits) and intertextual integration *only* when text-highlighting was possible. That is, these findings furthermore indicate that highlights focus readers' attention during re-reading, thus making the re-reading process more beneficial for information integration (i.e., also for source-content integration) compared to when highlighting is not possible.

Table 6.2

Means (standard deviations) of fixation durations on text and source logos during initial readings and re-readings as a function of text-highlighting, as well as inferential statistics regarding differences across conditions and relations to multiple document comprehension measures assessed. (Marginally) significant values are indicated by bold print.

Fixation location	Reading Interval	Condition		Inferential statistics		
		With HL	Without HL	Difference across conditions	Relation to #intertextual connections in essays	Relation to the score in the source-content mapping task
Texts	Overall [min]	10.99 (1.88)	10.40 (2.42)	$t(84.20) = 1.29$, $p = .202$	HL: $\beta = .08$, $t(85) = 0.70$, $p = .486$ fix. dur.: $\beta = .004$, $t(85) = 0.03$, $p = .977$ interact.: $\beta = .01$, $t(85) = 0.07$, $p = .945$	HL: $\beta = -.05$, $t(85) = -0.48$, $p = .632$ fix. dur.: $\beta = .13$, $t(85) = 0.98$, $p = .331$ interact. $\beta = -.03$, $t(85) = -0.25$, $p = .805$
	Initial readings [min]	8.67 (2.16)	7.25 (1.92)	$t(84.20) = 3.27$, **$p = .002$**	With HL: $r = -.07$, $p = .637$ noHL: $r = .02$, $p = .896$	With HL: $r = -.16$, $p = .293$ noHL: $r = .15$, $p = .318$
	Re-readings ≥ 1 sec [min]	2.31 (1.67)	3.14 (2.26)	$t(82.71) = -1.99$, **$p = .050$**	With HL: $r = .11$, $p = .466$ noHL: $r = -.01$, $p = .937$	**With HL: $r = .28$, $p = .068$** noHL: $r = .04$, $p = .800$
	Re-readings ≥ 3 sec [min]	2.29 (1.66)	3.13 (2.25)	$t(82.66) = -2.00$, **$p = .049$**	With HL: $r = .12$, $p = .457$ noHL: $r = -.01$, $p = .940$	**With HL: $r = .28$, $p = .066$** noHL: $r = .04$, $p = .808$
Source logos	Overall [sec]	5.26 (4.38)	7.88 (7.28)	$t(74.64) = -2.07$, **$p = .042$**	*Not senseful*	With HL: $r = .11$, $p = .498$ **noHL: $r = .32$, $p = .028$**
	Initial reads [sec]	3.13 (2.50)	4.57 (4.10)	$t(75.26) = -2.02$, **$p = .047$**	*Not senseful*	With HL: $r = .07$, $p = .640$ noHL: $r = .17$, $p = .268$
	Re-readings ≥ 1 sec [sec]	2.11 (3.37)	3.29 (4.56)	$t(82.67) = -1.40$, $p = .166$	*Not senseful*	HL: $\beta = -.01$, $t(85) = -0.08$, $p = .93$ **fix. dur.: $\beta = .31$, $t(85) = 2.34$, $p = .022$** interact. $\beta = -.12$, $t(85) = -0.91$, $p = .365$
	Re-readings ≥ 3 sec [sec]	1.84 (2.92)	2.88 (3.97)	$t(82.49) = -1.42$, $p = .161$	*Not senseful*	HL: $\beta = -.01$, $t(85) = -0.12$, $p = .905$ **fix. dur.: $\beta = .25$, $t(85) = 1.93$, $p = .057$** interact. $\beta = -.09$, $t(85) = -0.71$, $p = .478$

Note: ‘HL’ abbreviates ‘highlighting’, ‘fix. dur.’ abbreviates ‘fixation duration’, and ‘interact.’ abbreviates ‘interaction’.

Regarding participants' *fixation duration on source information*, most interestingly, the additional analyses revealed longer durations for participants in the without- than in the with-highlighting condition overall as well as during initial readings, whereas there was no difference in fixation duration on source information during re-readings across highlighting conditions. This might be caused by readers focusing primarily on (highlighting) text during initial readings, and, thus, less on source information when highlighting is possible compared to when no highlighting is possible. This assumption, however, needs further investigation by future research which could use a combination of thinking-aloud and eye-tracking methodology to examine readers' intentions in paying attention to source information during initial readings.

Of particular interest in these additional analyses was the relation between participants' fixation duration on source information and their score in the source-content mapping task. These analyses revealed a (marginally significant) positive relation between participants' score in the source-content mapping task and (a) their fixation duration on source information for the *overall reading time* only for participants in the without-highlighting condition (which might be due to the longer fixation duration on source information during initial readings without compared to with highlighting itself), and (b) their fixation duration on source information during *re-readings* irrespective of highlighting-condition. Together with the finding that there was no relation between the score in the source-content mapping task and fixation duration on source logos during initial readings, this indicates that attention to source information is especially beneficial for source-content integration during re-processing, irrespective of text-highlighting. This is in line with the Discrepancy-Induced Source Comprehension (D-ISC) model (Braasch et al., 2012; Braasch & Bråten, 2017) which states that readers will *strategically* process source information when they encountered conflict across documents in order to re-establish coherence by associating discrepant pieces of information to their sources. Yet, the findings regarding the positive relation between source-content integration and re-reading fixation duration on source information is somewhat in conflict with previous work which found no relation between the time participants had spent re-processing source information and source citations in essays, or source names recall (Salmerón et al., 2018b). Since source-content integration naturally also requires readers to represent source information, one would assume that source-content integration *and* source names recall would equally be related to readers' re-processing of source information. Yet, documents and especially source information was structured differently in that study than it was in the documents of the present work. Hence, future research is needed to investigate under which circumstances readers' attention to source information (during re-reading) is and is not related to their source-content integration or memory of source names.

6.2.2 Is the Definition of Revisits in Studies 1 and 2 Valid? Insights from Eye-Tracking in Study 3

One main contribution of Study 3 to this dissertation lays in the additional insights gained from eye-tracking methodology. While in Studies 1 and 2, the number of revisits (in the sequential condition) was assessed from logfiles as each document re-access that lasted at least 1 sec, eye-tracking methodology in Study 3 allowed to define re-readings for each revisit *after* the respective document had been fully read (i.e., when all the text of the respective document had received fixations). This methodology revealed that only 22 instances of the initial readings of a total of 17 participants (i.e., out of 95 participants reading 5 documents each, thus yielding a total of 475 instances) comprised more than the respective first access until the document had been fully read. Hence, most participants had first fully read a document for the first time before accessing another one. Thus, the number of revisits that was derived from logfiles in Studies 1 and 2 seem to be a fair approximation of the number of re-readings.

Furthermore, using eye-tracking methodology in Study 3 enabled an investigation of the number of fixations on text during short re-accesses. That is, while Studies 1 and 2 used a revisit threshold of 1 sec in order to exclude accidental accesses (in which presumably no meaningful information processing occurred), it remained unclear from these studies, how much text participants had actually fixated (i.e., re-read) during short re-accesses. In Study 3, thus, five different re-reading thresholds (i.e., 1 sec, 2 sec, 3 sec, 4 sec, and 5 sec) were defined in order to investigate the number of fixations on text during short re-accesses. Descriptive analyses revealed that during re-accesses that lasted less than 1 sec, participants in the with- and without-highlighting condition, on average, made 1.66 and 1.77 fixations on text, and during re-accesses that lasted between 4 and 5 sec, participants, on average, made 11.36 and 11.23 fixations on text (see Table 5.3). Using an average of between 1 and 2 (i.e., 1.5) fixations per word during reading (Hautala et al., 2011; Huestegge & Bocianski, 2010; Krieber et al., 2016), these values translate to no more than one and eight words being re-read during short re-accesses of below 1 sec and between 4 and 5 sec, respectively. Regarding the revisit threshold of 1 sec used in Studies 1 and 2, the analysis of eye-tracking data in Study 3 thus suggests that a revisit (or re-reading) threshold of 1 sec can be a valid measure to exclude (accidental) document accesses in which very limited information processing occurred.

Moreover, analyses of the role of re-reading in the effect of text-highlighting on multiple document comprehension in Study 3, which were done for each of the five revisit thresholds, revealed the same pattern of results regarding intertextual integration (i.e., an interaction between the number of re-readings and highlighting condition) for all thresholds greater than 1 sec (i.e., for the 2 – 5 sec thresholds). Regarding source-content integration, the positive correlation with the number of re-readings was significant for the 1 sec and 2 sec thresholds, and marginally significant for the 3, 4, and 5 sec thresholds. Hence, different re-reading thresholds might be suited for different measures of multiple document comprehension. This is because intertextual integration requires readers to (re-)read pieces of information for cross-textual comparison and linking, which presumably requires more information processing in a document than can possibly be done in, for instance, less than one or two seconds. Conversely, source-content integration requires not only processing of the information provided by the respective document (which takes more time), but also processing of source information, which, in all three studies of the present dissertation was provided as colored source logos comprising between one and seven words. Since processing of such relatively short source information might have also occurred during short re-accesses, lower thresholds seem to be better suited to exclude 'less meaningful' re-accesses for source-content integration than for intertextual integration.

In conclusion, analyses of eye-tracking data in Study 3 imply that the definition of revisits in Studies 1 and 2 can be assumed valid, since very little is re-read during short re-accesses that last below 1 sec, and because most participants seem to fully read the respective opened document during its first opening.

6.3 Discussion of Discrepant Results Across Studies

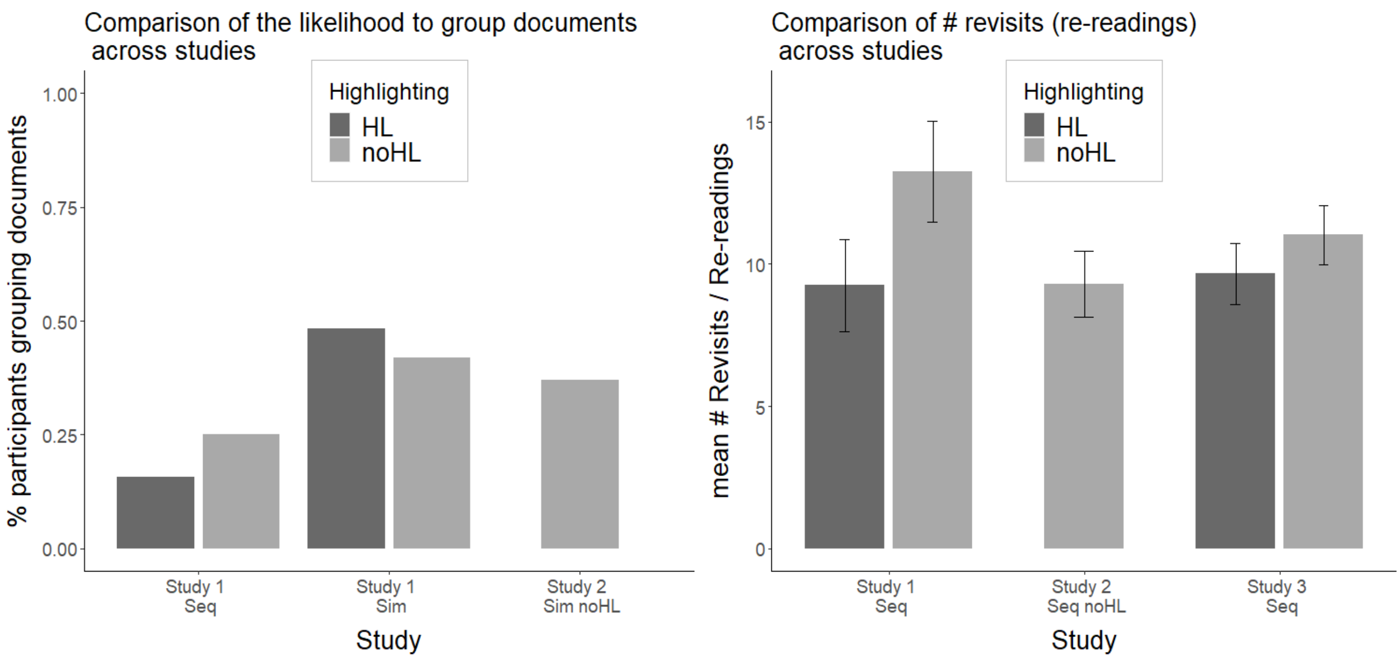
The most striking difference regarding research findings across studies of the present dissertation is the fact that neither the main effect of document presentation, nor the main effect of text highlighting on intertextual integration that was found in Study 1 could be replicated in Studies 2 and 3. As mentioned in Sections 6.1 and 6.2, this might be due to some methodological differences across studies.

First, one possible post-hoc explanation for the fact that the effect of document presentation on intertextual integration could not be replicated in Study 2 is the different implementations of the sequential condition across Studies 1 and 2. Specifically, in the simultaneous condition of Study 1, documents were initially presented in a grid with three documents in each of the two rows, that is, documents were spread out in a simultaneous presentation initially. In contrast, in the simultaneous condition of Study 2, documents were initially presented on a stack such that participants had to actively re-arrange documents in order to create a simultaneous presentation of documents. Hence, participants in Study 1 might have benefited from the initial layout of documents without re-arranging them actively, thus somewhat biasing the effect of document presentation in Study 1 compared to Study 2 by providing a greater *initial* advantage of the simultaneous condition in Study 1 than in Study 2. However, as Figure 6.1 indicates, the number of intertextual connections in essays was, descriptively, not different across simultaneous conditions of Studies 1 and 2.

Second, theoretically, the non-replicated effect of document presentation on intertextual integration in Study 2 could also be due to participants in the simultaneous condition of Study 2 engaging in grouping the partly conflicting documents less (i.e., not using the simultaneous presentation of documents in this strategical way that is positively related to intertextual integration) than participants in Study 1. Likewise, in theory, the non-replicated effect of text-highlighting on intertextual integration in Study 3 (at least regarding the sequential condition of Study 1) could also be due to participants of Study 3 engaging in re-reading documents less often than participants in Study 1. However, as Figure 6.1 descriptively shows, neither the percentage of participants who grouped documents during reading differs across the simultaneous without-highlighting conditions of Studies 1 and 2, nor does the number of re-readings/revisits differ (i.e., descriptively) across studies for the respective highlighting conditions. these measures did not vary much across studies. Hence, the differences in findings

Figure 6.1

Comparative presentation of the likelihood to group documents during reading (left) across Studies 1 and 2, and of the number of revisits / re-readings (right) across Studies 1, 2, and 3.



regarding the main effects of document presentation and text-highlighting on intertextual connections are unlikely to be rooted in behavioral differences across studies.

Third, another possible post-hoc explanation for the fact that neither the main effect of document presentation nor the effect of text-highlighting on intertextual integration could be replicated is the chosen document orders that differed across studies. That is, in order to exclude any order effects, the document order in Study 1 was randomized for each participant. However, this might have resulted in a blocked and alternating order of documents (i.e., with documents taking the same or opposing stance in direct succession) for some participants, which can affect conflict detection (Braasch et al., 2021; Maier & Richter, 2013) and, in turn, intertextual integration. To control for such effects, in Studies 2 and 3, it was decided to present documents in one of two alternating orders, that is, with the first and fifth document taking a positive or negative stance, the second and fourth document taking a negative or positive stance, and the third document always being the neutral one, which resulted in the first and second as well as fourth and fifth document being of opposing positions. This, however, might have resulted in enhanced conflict detection (and, thus, intertextual integration) in Studies 2 and 3 compared to Study 1. Of note, this was presumably even more the case in the sequential conditions (i.e., in the sequential rather than the simultaneous condition of Study 2 and in both highlighting-conditions of Study 3), since documents in the simultaneous condition could be re-arranged, which might have changed their order. In consequence, an increased awareness of conflicts across documents in the sequential conditions in Studies 2 and 3 as compared to Study 1 might have diminished the potential negative effect of the sequential as compared to the simultaneous condition in Study 2, and might have counteracted the positive effect of text-highlighting in Study 3. As can be seen in Figures 6.2 and 6.3, indeed, the mean number of intertextual connections in essays of participants in the sequential conditions differed across Study 1 and Studies 2 and 3. That is, the mean number of intertextual connections in essays of participants in the respective sequential conditions was, descriptively, higher in Studies 2 and 3 than in Study 1.

In conclusion, the descriptive analyses of the mean number of intertextual connections across the studies of the present dissertation suggests that the missing main effects of document presentation and text-highlighting on intertextual connections in Studies 2 and 3 might be due to the alternating document order. However, this assumption requires further research that, for instance, investigated the effect of document order (i.e., alternating versus blocked document order of partly conflicting documents) in a reading environment that enables a simultaneous rather than imposes a sequential presentation of documents, as well as in a reading environment that allows for text-highlighting versus not.

Figure 6.2

Comparative presentation of the mean number of intertextual connections across Studies 1 (left and middle pair of bars) and 2 (right pair of bars) as a function of document presentation. Since text-highlighting was not possible in Study 2, for better comparability of Studies, the middle pair of bars presents the mean number of intertextual connections of the subset of Study 1 of only those participants in the without-highlighting conditions.

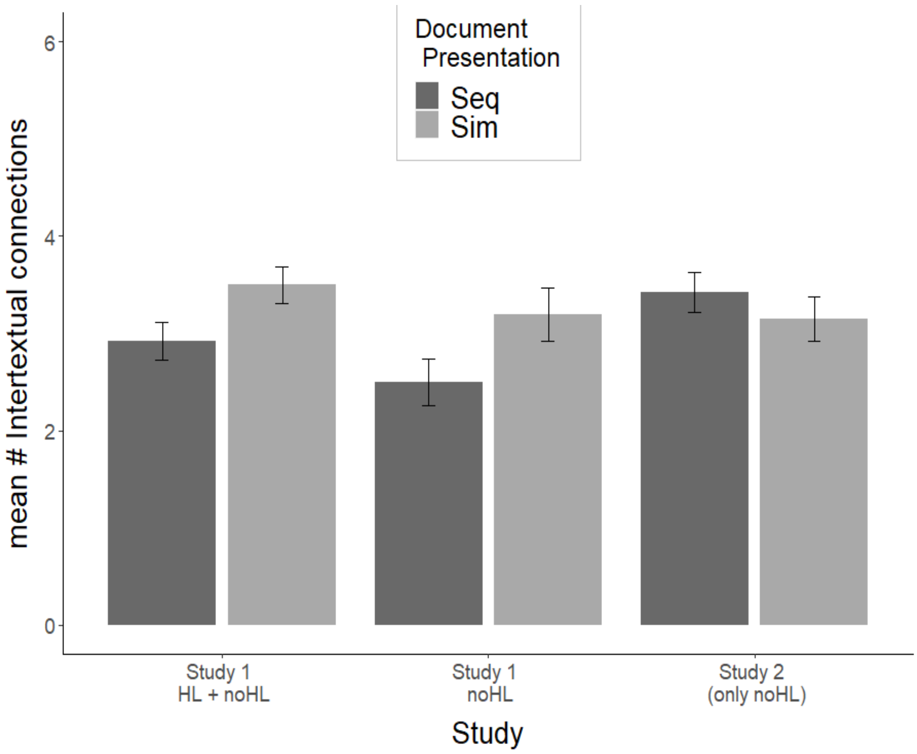
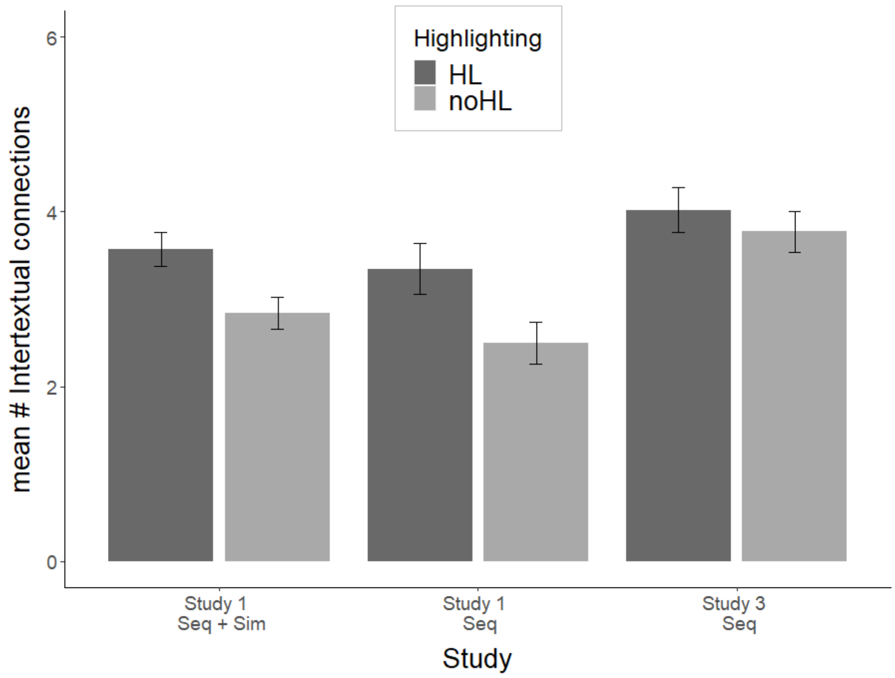


Figure 6.3

Comparative presentation of the mean number of intertextual connections across Studies 1 (left and middle pair of bars) and 3 (right pair of bars) as a function of text-highlighting. Since documents were only presented sequentially in Study 3, for better comparability of Studies, the left pair of bars presents the mean number of intertextual connections of the subset of Study 1 of only those participants in the sequential condition.



6.5 Theoretical and Practical Implications

The findings of the present work allow some theoretical and practical implications. First, the findings of the present work overall support the suggestion made by the RESOLV model (Britt et al., 2018; Rouet et al., 2017) regarding the complex interplay between reading environment, reading (inter-)actions, and multiple document comprehension. That is, the role of grouping documents during reading in the effect of document presentation (i.e., the full mediation of grouping documents in the effect of document presentation on intertextual integration that was found in Study 1) as well as the role of re-reading in the effect of text-highlighting (i.e., the moderation of the number of re-readings or revisits of the effect of text-highlighting on intertextual integration that was found in Studies 1 and 3) on multiple document comprehension stress the importance of reading (inter-)actions for multiple document comprehension. Thereby, as an important theoretical implication of the present work, the fact that the reading interactions assessed were differentially related to different measures of multiple document comprehension (see Table 6.1) furthermore suggests that the aspect of multiple document comprehension assessed is an additional, potentially crucial factor in the interplay between reading environment, reading (inter-)actions, and multiple document comprehension. Hence, future theoretical frameworks of multiple document comprehension, potentially building on the RESOLV model, should also consider the importance of the aspect of multiple document comprehension assessed.

Related to this complex interplay between the reading environment, reading (inter-)actions, and (measure of) multiple document comprehension that the present work suggests, some practical implications can be derived: First, future research investigating the effects of characteristics of a reading environment on multiple document comprehension should take into account reading (inter-)actions. That is, for instance, the moderation of re-reading in the effect of text-highlighting that was found in two studies of the present dissertation suggests that the rather inconclusive findings in previous work on the effectiveness of text-highlighting (Dunlosky et al., 2013; Miyatsu et al., 2018) might be due to differences participants' re-reading behavior across studies. Future research should further address this assumption. Second, previous work has put great effort into understanding the cognitive processes (Anmarkrud et al., 2013, 2014; Bråten & Strømsø, 2011; Ferguson et al., 2013; Ferguson & Bråten, 2013; Goldman, Lawless, et al., 2012; Hagen et al., 2014; Kobayashi, 2016; Wineburg, 1991; Wolfe & Goldman, 2005) and (meta-)cognitive strategies (Ackerman & Goldsmith, 2011; Barzilai & Ka'adan, 2017; Cho et al., 2018; Lee & Baylor, 2006; Leopold & Leutner, 2015; Stadtler & Bromme, 2004, 2007; Yue et al., 2014) positively related to multiple document comprehension. In turn, it has been suggested that readers might benefit from being trained or instructed to use these cognitive processes or strategies (Lachner et al., 2017; List & Alexander, 2020; Nückles et al., 2009; Renkl & Scheiter, 2017), or from tools or scaffolds provided by a reading environment that might foster these cognitive processes (Barzilai & Ka'adan, 2017; Lombard et al., 2021), for a review see (Barzilai et al., 2018). Yet, the findings of the studies of the present dissertation suggest that such trainings or interventions should not only cover these cognitive processes, and that readers should not simply be provided with tools or scaffolds in a reading environment. Rather, in order to best benefit from trainings or the tools or scaffolds provided, readers should *also* be taught (or instructed; cf. Lombard et al., 2021; Wiley, 2001) *how* to interact with them (i.e., potentially in relation to what cognitive processes or (meta-)cognitive strategies the respective tools, scaffolds, or reading (inter-)actions aim to support).

6.6 Limitations, Strengths, and Future Work

Of course, the present work does not come without limitations. First, the fact that university students were used as participants in all three studies of the present dissertation raises the question of generalizability of the respective (and replicated) findings. That is, university students can be regarded as advanced readers, who might engage in reading (inter-)actions such as re-reading more than less advanced readers (Catrysse, Gijbels, & Donche, 2018; Catrysse, Gijbels, Donche, et al., 2018). However, reading interactions were taken into account in all three studies – which is also the greatest strength of the present work. Specifically, even though participants could be regarded as advanced readers, there was a great variability in the number of re-readings (or revisits), and even in the simultaneous conditions (i.e., in Studies 1 and 2) which can be assumed to bear a greater affordance to group partly conflicting documents according to their overall stance (see Study 1), the percentage of participants who had grouped documents was below 50%. The fact that analyses revealed moderating or even mediating effects of the respective reading interaction on the effectiveness of the respective characteristic of the reading environment on multiple document comprehension stresses the importance of such reading interactions (or strategical reading behavior in general) even amongst presumably advanced readers. In turn, the findings of the present work stress the importance of training less advanced readers in such overt reading strategies. Of course, however, whether the findings of the present work can be replicated with a sample of less advanced readers with lower working memory capacity or metacognitive abilities (cf. Barzilai & Strømsø, 2018; Britt & Rouet, 2012; List & Alexander, 2017a; Maier & Richter, 2013; Perfetti, 1997; Rouet & Britt, 2011), and whether less advanced readers, or readers in general, can benefit from the mentioned trainings or interventions remains an open question from the present work.

Second, the fact that the same documents were used in all three studies of the present dissertation raises the question of generalizability of the respective (and replicated) findings to other subject matters than that of the health effects of UV radiation. Yet, the present work addressed particularly new research questions by investigating the role of reading interactions in the effect of specific characteristics of the reading environment on multiple document comprehension. Since the reading material itself might affect multiple document comprehension (e.g., Braasch et al., 2012, 2021; Ferguson et al., 2013; List et al., 2021), for the sake of comparability of the present findings to those of previous work in the field of multiple document comprehension, it was considered important to address these new research questions with (translated and slightly adapted versions of) documents used in previous work (e.g., Bråten, Anmarkrud, et al., 2014; Ferguson & Bråten, 2013; Lombard et al., 2021; Strømsø et al., 2016). Furthermore, for similar reasons (i.e., for comparability across studies), the same documents were used in all three studies, since Studies 2 and 3 aimed to extend findings of Study 1 by using additional measures of multiple document comprehension. However, future research is needed to understand whether the roles of reading interactions that the present work suggests in the effects of document presentation and the possibility to highlight text on multiple document comprehension also come to bear when readers are faced with other topics, and especially when they are faced with topics they have more or less prior knowledge about (Goldman, Braasch, et al., 2012; Kaakinen et al., 2003; Shapiro, 2004), or stronger or weaker attitudes towards (Kobayashi, 2016; Maier & Richter, 2013; Van Strien et al., 2014).

Third, as mentioned in Subsections 6.1 and 6.2, Studies 2 and 3 could not replicate the main effects of document presentation or the possibility to highlight text on readers' intertextual integration. As discussed in Subsection 6.3, this might be due to the specific order of the partly conflicting documents in Studies 2 and 3, which might have fostered conflict detection and thus intertextual integration in the respective (presumed) disadvantageous condition (i.e., the sequential and the without-highlighting condition, respectively) compared to Study 1. Yet, this assumption requires further investigation. Please note, however, that the relation between the respective reading interactions and measures of multiple document comprehension were consistently found in all three studies, which is another strength of the present work.

Fourth, based on previous work, in the studies of the present dissertation, it was assumed that the reading interactions of re-arranging partly conflicting documents into groups (cf. O'Hara et al., 2002; O'Hara & Sellen, 1997) according to their overall stance during reading, as well as re-reading (or revisiting) documents (cf. Goldman, Lawless, et al., 2012; List & Alexander, 2018a; Wiley et al., 2009; Wineburg, 1991) are positively related to multiple document comprehension since these reading actions might reflect readers' engagement in cross-document information comparison and re-evaluation. However, readers' intentions in engaging in these reading interactions remains unclear from the present work and should be investigated in future research, for instance by means of think-aloud methodology. Specifically, since the respective findings of the present work are correlational in nature, it remains an open question whether participants grouped documents *because* they had already integrated information across documents (and had thus also realized the partly conflicting nature of documents) or whether the resulting groups of documents *further* supported their cross-document information corroboration and integration – or both. Likewise, the positive relation between re-reading (or revisiting) and intertextual integration (Studies 1 and 3; i.e., with text-highlighting) as well as source-content integration (Study 3) supports the assumption that readers engage in re-reading (or revisiting) in order to compare, contrast, and integrate information across documents. Yet, the fact that with as well as without highlighting, the number of revisits was also positively related to the number of source names correctly recalled in Study 2 suggests that re-readings (or revisits) might (also) be initiated by readers in order to re-process source information. Please note that the latter is especially indicated by the additional analyses of eye-tracking data presented in Section 6.2.1, which showed a positive relation between the fixation duration on source logos during re-readings and participants' score in the source content mapping task. Hence, future research is needed to shed light on readers' intentions in their engagement in such reading (inter-)actions as well as to set out other reading (inter-)actions that might reflect cognitive processes required for a comprehensive understanding of multiple documents, such as comparison, re-evaluation, and integration of information across documents, or sourcing.

Fifth, in the present work, the assessment of re-readings (or revisits) was only possible in the respective sequential conditions since in the simultaneous conditions, participants could also have re-inspected a document without interacting with it (i.e., by turning their head). To investigate the role of re-reading (or revisiting) in multiple document comprehension when the reading environment enables a simultaneous presentation of documents, or to investigate potential differences in readers' engagement in re-reading (or revisiting) in a reading environment enabling a simultaneous rather than imposing a sequential presentation of documents, future work could use eye-tracking methodology.

Finally, as a last limitation of the present work, it should be noted that the simultaneous conditions in Studies 1 and 2 were implemented on a multi-touch table. While multi-touch tables are increasingly installed in public places such as museums, they are not typically used for reading tasks. Yet, interaction with the interface was intuitive for participants. Also, the simultaneous document presentation, which allowed flexible re-arrangement of documents, resembles reading printed documents on a desk (yet in a digital reading environment), and the sequential document presentation resembles reading digital documents on a tablet (yet on a larger screen). Furthermore, by using a multi-touch table instead of printed documents, the assessment of participants' engagement in grouping and re-reading (or revisiting) documents was possible via logfiles, which enabled a more accurate assessment of, for instance the initial versus re-reading times compared to an assessment from videotaped interactions with printed documents. As such, the present work does not claim to have investigated the roles of reading interactions in the effects of document presentation and text-highlighting on multiple document comprehension in a common reading environment. Rather, the present work aimed to provide more detailed insights into the relation between the reading environment, reading (inter-)actions, and (specific aspects of) multiple document comprehension, and to contribute to the growing body of literature investigating potential ways on how to support readers in building a comprehensive understanding of complex subject matters discussed across multiple documents. Yet, the present findings thereby indicate a great potential in large information spaces such as presented in the Productivity Future Vision of Microsoft© in 2016 (as shown in the video between 2:20 and 2:50 min; Microsoft©, 2016).

Notwithstanding these limitations, the present work provides important new insights for the field of multiple document comprehension by (a) supporting the assumption of a complex interplay between the reading environment, reading (inter-)actions, and multiple document comprehension suggested by the RESOLV model (Britt et al., 2018; Rouet et al., 2017), (b) suggesting that one further important factor in this interplay might be the specific aspect of multiple document comprehension assessed, (c) showing a positive relation between readers' spontaneous engagement in reading interactions and their multiple document comprehension, (d) suggesting that one potential reason for the rather inconclusive findings in previous work investigating the effectiveness of text-highlighting for (multiple) document comprehension might be readers' varying engagement in re-reading, and (e) providing first empirical data suggesting a time-saving reprocessing of information when text-highlighting is possible compared to when text-highlighting is not possible.

6.7 Conclusion

In starting the conclusion of the present work rather critically, across the studies of the present dissertation, the main effects of document presentation and text-highlighting on multiple document comprehension that were found in Study 1 could not be replicated in Studies 2 and 3. Yet, this might be due to the methodological differences across studies (see Section 6.3). Furthermore, and most importantly, findings of all three studies support the assumption of a complex interplay between the reading environment, reading (inter-)actions, and readers' multiple document comprehension after reading as suggested by the RESOLV model (Britt et al., 2018; Rouet et al., 2017). This is because throughout the studies of the present dissertation, (a) the reading environment enabling a simultaneous rather than imposing a sequential presentation of documents was most beneficial regarding multiple document comprehension (i.e., intertextual integration in Studies 1 and 2, and source-content integration in Study 2) for participants who used the (potential) simultaneous presentation for grouping the partly conflicting documents, and (b) the number of revisits (or re-readings in Study 3), on the one hand, moderated the effect of text-highlighting on intertextual integration, and, on the other hand, was positively related to source-content integration (i.e., in Studies 2 and 3). Thereby, the fact that Studies 2 and 3 revealed differential roles of reading interactions in the effect of the respective reading environment on the different measures of multiple document comprehension assessed furthermore indicate that future theoretical models of multiple document comprehension that might potentially build on the RESOLV model (Britt et al., 2018; Rouet et al., 2017) should consider the aspect of multiple document comprehension assessed as additional factor. Furthermore, the present findings regarding the complex interplay between reading environment, reading (inter-)actions, and multiple document comprehension suggest that trainings or interventions aimed to teach readers in how to best build a comprehensive understanding of multiple documents should consider not only the provision of specific support tools in a reading environment, but also readers' interactions with the documents or the reading environment as whole.

On a last note, overall, the moderating and mediating roles of reading interactions found across the studies of the present work raise the question whether reading (inter-)actions are more important for readers' multiple document comprehension than the characteristics of a reading environment, or the support tools it provides. This question can not be fully answered on the basis of the present work. Yet, findings of the present work suggests that *because* readers are more likely to engage in grouping partly conflicting documents during reading when the reading environment enables a simultaneous presentation of documents, a reading environment enabling a simultaneous rather than imposing a sequential presentation of documents, document presentation might be a crucial prerequisite for readers' interactions with documents that are positively related to their multiple document comprehension. Likewise, findings of the present work suggest that *because* re-reading might be more beneficial for readers' multiple document comprehension when text-highlighting is possible, providing readers with the possibility to highlight text might be a crucial prerequisite for their multiple document comprehension.

As such, hopefully, the findings of the present work as well as the further research questions raised based on the respective findings will be fruitful for future research examining the roles of reading interaction in the effect of characteristics of the reading environment on the several aspects of multiple documents comprehension.

References

Ackerman, R., & Goldsmith, M. (2011). Metacognitive regulation of text learning: On screen versus on paper. *Journal of Experimental Psychology: Applied*, *17*(1), 18–32. https://doi.org/10.1037/a0022086

Afflerbach, P., & Cho, B.-Y. (2009). Determining and describing reading strategies: Internet and traditional forms of reading. In H. S. Waters & W. Schneider (Eds.), *Metacognition, strategy use, and instruction* (pp. 69–90). Routledge.

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Sage.

AlGhamdi, K. M., & Moussa, N. A. (2012). Internet use by the public to search for health-related information. *International Journal of Medical Informatics*, *81*(6), 363–373. https://doi.org/10.1016/j.ijmedinf.2011.12.004

Amer, A. A. (1994). The effect of knowledge-map and underlining training on the reading comprehension of scientific texts. *English for Specific Purposes*, *13*(1), 35–45. https://doi.org/10.1016/0889-4906(94)90023-X

Andrews, C., Endert, A., & North, C. (2010). Space to think: Large, high-resolution displays for sensemaking. *Proceedings of the 28th International Conference on Human Factors in Computing Systems - CHI '10*, 55–64. https://doi.org/10.1145/1753326.1753336

Anmarkrud, Ø., Bråten, I., & Strømsø, H. I. (2014). Multiple-documents literacy: Strategic processing, source awareness, and argumentation when reading multiple conflicting documents. *Learning and Individual Differences*, *30*, 64–76. https://doi.org/10.1016/j.lindif.2013.01.007

Anmarkrud, Ø., McCrudden, M. T., Bråten, I., & Strømsø, H. I. (2013). Task-oriented reading of multiple documents: Online comprehension processes and offline products. *Instructional Science*, *41*(5), 873–894. https://doi.org/10.1007/s11251-013-9263-8

Annis, L., & Davis, J. K. (1978). Study techniques and cognitive style: Their effect on recall and recognition. *Journal of Educational Research*, *71*(3), 175–178.

Ariasi, N., & Mason, L. (2011). Uncovering the effect of text structure in learning from a science text: An eye-tracking study. *Instructional Science*, *39*(5), 581–601. https://doi.org/10.1007/s11251-010-9142-5

Author(s). (2021). *Reading multiple documents on a health-related issue: The roles of a text-highlighting tool and re-reading behavior on integrated understanding*.

Ball, R., & North, C. (2005). Analysis of user behavior on high-resolution tiled displays. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *3585 LNCS*, 350–363. https://doi.org/10.1007/11555261_30

Barzilai, S., & Ka'adan, I. (2017). Learning to integrate divergent information sources: The interplay of epistemic cognition and epistemic metacognition. *Metacognition and Learning*, *12*(2), 193–232. https://doi.org/10.1007/s11409-016-9165-7

Barzilai, S., & Strømsø, H. I. (2018). Individual differences in multiple document comprehension. In J. L. G. Braasch, I. Bråten, & M. Mccrudden (Eds.), *Handbook of Multiple Source Use* (pp. 99–116). Routledge. https://doi.org/10.4324/9781315627496

Barzilai, S., Zohar, A. R., & Mor-Hagani, S. (2018). Promoting integration of multiple texts: A review of instructional approaches and practices. *Educational Psychology Review*, *30*(3), 973–999. https://doi.org/10.1007/s10648-018-9436-8

Bell, K. E., & Limber, J. E. (2010). Reading skill, textbook marking, and course performance. *Literacy Research and Instruction*, *49*(1), 56–67. https://doi.org/10.1080/19388070802695879

Ben-Yehudah, G., & Eshet-Alkalai, Y. (2018). The contribution of text-highlighting to comprehension: A comparison of print and digital reading. *Journal of Educational Multimedia and Hypermedia*, *27*(2), 153–178.

Ben-Yehudah, G., & Eshet-Alkalai, Y. (2020). Print versus digital reading comprehension tests: Does the congruency of study and test medium matter? *British Journal of Educational Technology*, *52*(1), 426–440. https://doi.org/10.1111/bjet.13014

Ben-Yehudah, G., & Eshet-Alkalai, Y. (2014). The influence of text annotation tools on print and digital reading comprehension. *Proceedings of the 9th Chais Conference for Innovation in Learning Technologies*, *November 2014*, 28–35.

Benshoof, L. A., Graves, M., & Hooper, S. (1995). The effects of single and multiple window presentations on achievement, instructional time, window use, and attitudes during computer-based instruction. *Computers in Human Behavior*, *11*(2), 261–272.

Bi, X., & Balakrishnan, R. (2009). Comparing usage of a large high-resolution display to single or dual desktop displays for daily work. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1005–1014.

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

Braasch, J. L. G., & Bråten, I. (2017). The discrepancy-induced source comprehension (D-ISC) model: Basic assumptions and preliminary evidence. *Educational Psychologist*, *52*(3), 167–181. https://doi.org/https://doi.org/10.1080/00461520.2017.1323219

Braasch, J. L. G., Killion, S. C., & Bråten, I. (2021). Contextual factors that affect adolescents' detection of and memory for conflicts across multiple texts. *Journal of Research in Reading*. https://doi.org/10.1111/1467-9817.12348

Braasch, J. L. G., Rouet, J. F., Vibert, N., & Britt, M. A. (2012). Readers' use of source information in text comprehension. *Memory and Cognition*, *40*(3), 450–465. https://doi.org/10.3758/s13421-011-0160-6

Bråten, I., Anmarkrud, Ø., Brandmo, C., & Strømsø, H. I. (2014). Developing and testing a model of direct and indirect relationships between individual differences, processing, and multiple-text comprehension. *Learning and Instruction*, *30*, 9–24. https://doi.org/10.1016/j.learninstruc.2013.11.002

Bråten, I., Brante, E. W., & Strømsø, H. I. (2018). What really matters: The role of behavioural engagement in multiple document literacy tasks. *Journal of Research in Reading*, *41*(4), 680–699. https://doi.org/10.1111/1467-9817.12247

Bråten, I., Britt, M. A., Strømsø, H. I., & Rouet, J.-F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Toward an integrated model. *Educational Psychologist*, *46*(1), 48–70. https://doi.org/10.1080/00461520.2011.538647

Bråten, I., Ferguson, L. E., Anmarkrud, Ø., Strømsø, H. I., Br, I., Anmarkrud, Ø., & Strømsø, H. I. (2013). Prediction of learning and comprehension when adolescents read multiple texts: The roles of word-level processing, strategic approach, and reading motivation. *Reading and Writing*, *26*(3), 321–348. https://doi.org/10.1007/s11145-012-9371-x

Bråten, I., Ferguson, L. E., Strømsø, H. I., & Anmarkrud, Ø. (2013). Justification beliefs and multiple-documents comprehension. *European Journal of Psychology of Education*, *28*(3), 879–902. https://doi.org/10.1007/s10212-012-0145-2

Bråten, I., Ferguson, L. E., Strømsø, H. I., & Anmarkrud, Ø. (2014). Students working with multiple conflicting documents on a scientific issue: Relations between epistemic cognition while reading and sourcing and argumentation in essays. *British Journal of Educational Psychology*, *84*(1), 58–85. https://doi.org/10.1111/bjep.12005

Bråten, I., & Strømsø, H. I. (2011). Measuring strategic processing when students read multiple texts. *Metacognition and Learning*, *6*(2), 111–130. https://doi.org/10.1007/s11409-011-9075-7

Bråten, I., Strømsø, H. I., & Andreassen, R. (2016). Sourcing in professional education: Do text factors make any difference? *Reading and Writing*, *29*(8), 1599–1628. https://doi.org/10.1007/s11145-015-9611-y

Britt, M. A., & Aglinskas, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction*, *20*(4), 485–522. https://doi.org/10.1207/S1532690XCI2004_2

Britt, M. A., Perfetti, C. A., Sandak, R., & Rouet, J.-F. (1999). Content integration and source separation in learning from multiple texts. In S. R. Goldman, A. C. Graesser, & P. Van den Broek (Eds.), *Narrative comprehension, causality, and coherence*. Lawrence Erlbaum Associates.

Britt, M. A., & Rouet, J.-F. (2012). Learning with multiple documents. In J. R. Kirby & M. J. Lawson (Eds.), *Enhancing the quality of learning: dispositions, instructions, and learning processes* (pp. 276–314). Cambridge University Press. https://doi.org/10.1017/cbo9781139048224.017

Britt, M. A., Rouet, J.-F., & Durik, A. (2018). Representations and processes in multiple source use. In J. L. G. Braasch, I. Bråten, & M. T. McCrudden (Eds.), *Handbook of multiple source use* (pp. 17–33). https://doi.org/10.4324/9781315627496

Britt, M. A., Wiemer-Hastings, P., Larson, Aaron, A., & Perfetti, C. A. (2004). Using intelligent feedback to improve sourcing and integration in students' essays. *International Journal of Artificial Intelligence in Education*, *14*(3), 359–374.

Bromage, B. K., & Mayer, R. E. (1986). Quantitative and qualitative effects of repetition on learning from technical text. *Journal of Educational Psychology*, *78*(4), 271–278. https://doi.org/10.1037/0022-0663.78.4.271

Callender, A. A., & McDaniel, M. A. (2009). The limited benefits of rereading educational texts. *Contemporary Educational Psychology*, *34*(1), 30–41. https://doi.org/10.1016/j.cedpsych.2008.07.001

Carrier, L. M. (2003). College students' choices of study strategies. *Perceptual and Motor Skills*, *96*, 54–56.

Cashen, V. M., & Leicht, K. L. (1970). Role of the isolation effect in a formal educational setting. *Journal of Educational Psychology*, *81*(8), 484–496. https://doi.org/10.1037/h0030286

Catrysse, L., Gijbels, D., & Donche, V. (2018). It is not only about the depth of processing: What if eye am not interested in the text? *Learning and Instruction*, *58*, 284–294. https://doi.org/10.1016/j.learninstruc.2018.07.009

Catrysse, L., Gijbels, D., Donche, V., De Maeyer, S., Lesterhuis, M., & Van den Bossche, P. (2018). How are learning strategies reflected in the eyes? Combining results from self-reports and eye-tracking. *British Journal of Educational Psychology*, *88*(1), 118–137. https://doi.org/10.1111/bjep.12181

Cerdán, R., & Vidal-Abarca, E. (2008). The effects of tasks on integrating information from multiple documents. *Journal of Educational Psychology*, *100*(1), 209–222. https://doi.org/10.1037/0022-0663.100.1.209

Chi, E., Gumbrecht, M., & Hong, L. (2007). Visual foraging of highlighted text: An eye-tracking study. In J. A. Jacko (Ed.), *Human-Computer Interaction* (pp. 589–598). Springer. https://doi.org/10.1007/978-3-540-73110-8_64

Cho, B.-Y., & Afflerbach, P. (2017). An evolving perspective of constructively responsive reading comprehension strategies in multilayered digital text environments. In S. E. Israel (Ed.), *Handbook of Research on Reading Comprehension* (2nd ed., pp. 109–134). Guilford Publications.

Cho, B.-Y., Afflerbach, P., & Han, H. (2018). Strategic processing in accessing, comprehending, and using multiple sources online. *Handbook of Multiple Source Use*, *May*, 133–150. https://doi.org/10.4324/9781315627496

Coles, P., & Foster, J. (1975). Typographic cueing as an aid to learning from typewritten text. *Programmed Learning and Educational Technology*, *12*(2), 102–108. https://doi.org/10.1080/1355800750120206

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal Of Verbal Learning And Verbal Behavior*, *11*(6), 671–684.

Crouse, J. H., & Idstein, P. (1972). Effects of encoding cues on prose learning. *Journal of Educational Psychology*, *63*(4), 309–313. https://doi.org/10.1037/h0033574

Czerwinski, M., Smith, G., Regan, T., Meyers, B., Starkweather, G., & Way, O. M. (2003). Toward characterizing the productivity benefits of very large displays. In M. Rauterberg, M. Menozzi, & J. Wesson (Eds.), *International Conference on Human-Computer Interaction INTERACT*. IOS Press.

Delgado, P., Stang Lund, E., Salmerón, L., & Bråten, I. (2020). To click or not to click: Investigating conflict detection and sourcing in a multiple document hypertext environment. *Reading and Writing*, *33*, 2049–2072. https://doi.org/10.1007/s11145-020-10030-8

Du, H., & List, A. (2020). Researching and writing based on multiple texts. *Learning and Instruction*, *66*. https://doi.org/10.1016/j.learninstruc.2019.101297

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Fass, W., & Schumacher, G. M. (1978). Effects of motivation, subject activity, and readability on the retention of prose materials. *Journal of Educational Psychology*, *70*(5), 803–807. https://doi.org/10.1037/0022-0663.70.5.803

Ferguson, L. E., & Bråten, I. (2013). Student profiles of knowledge and epistemic beliefs: Changes and relations to multiple-text comprehension. *Learning and Instruction*, *25*, 49–61. https://doi.org/10.1016/j.learninstruc.2012.11.003

Ferguson, L. E., Bråten, I., & Strømsø, H. I. (2012). Epistemic cognition when students read multiple documents containing conflicting scientific evidence: A think-aloud study. *Learning and Instruction*, *22*(2), 103–120. https://doi.org/10.1016/j.learninstruc.2011.08.002

Ferguson, L. E., Bråten, I., Strømsø, H. I., & Anmarkrud, Ø. (2013). Epistemic beliefs and comprehension in the context of reading multiple documents: Examining the role of conflict. *International Journal of Educational Research*, *62*, 100–114.

https://doi.org/10.1016/j.ijer.2013.07.001

Fowler, R. L., & Barker, A. S. (1974). Effectiveness of highlighting for retention of text material. *Journal of Applied Psychology*, *59*(3), 358–364. https://doi.org/10.1037/h0036750

Freund, L., Kopak, R., & O'Brien, H. (2016). The effects of textual environment on reading comprehension: Implications for searching as learning. *Journal of Information Science*, *42*(1), 79–93. https://doi.org/https://doi.org/10.1177/0165551515614472

Gaddy, M. L., van den Broek, P. W., & Sung, Y.-C. (2001). The influence of text cues on the allocation of attention during reading. In T. Sanders, J. Schileroord, & W. Spooren (Eds.), *Text representation. Linguistic and psycholinguistic aspects* (pp. 89–110). https://doi.org/10.1075/hcp.8.05gad

Gier, V. S., Herring, D., Hudnell, J., Montoya, J., & Kreiner, D. S. (2010). Active reading procedures for moderating the effects of poor highlighting. *Reading Psychology*, *31*(1), 69–81. https://doi.org/10.1080/02702710903397082

Gier, V. S., Kreiner, D. S., Hudnell, J., Montoya, J., & Herring, D. (2011). Using an electronic highlighter to eliminate the negative effects of pre-existing, inappropriate highlighting. *Journal of College Reading and Learning*, *41*(2), 37–52. https://doi.org/10.1080/10790195.2011.10850341

Gier, V. S., Kreiner, D. S., & Natz-Gonzalez, A. (2009). Harmful effects of preexisting inappropriate highlighting on reading comprehension and metacognitive accuracy. *The Journal of General Psychology*, *136*(3), 287–300.

Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction*, *16*(6), 511–525. https://doi.org/10.1016/j.learninstruc.2006.10.001

Goldberg, J. H., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: Methods and constructs. *International Journal of Industrial Ergonomics*, *24*(6), 631–645. https://doi.org/10.1016/S0169-8141(98)00068-7

Goldhammer, F., Naumann, J., Stelter, A., Tóth, K., Rölke, H., & Klieme, E. (2014). The time on task effect in reading and problem solving is moderated by task difficulty and skill: Insights from a computer-based large-scale assessment. *Journal of Educational Psychology*, *106*(3), 608–626. https://doi.org/10.1037/a0034716

Goldman, S. R., Braasch, J. L. G., Wiley, J., Graesser, A. C., & Brodowinska, K. (2012). Comprehending and learning from internet sources: Processing patterns of better and poorer learners. *Reading Research Quarterly*, *47*(4), 356–381. https://doi.org/https://doi.org/10.1002/RRQ.027

Goldman, S. R., Lawless, K., & Manning, F. (2012). Research and development of multiple source comprehension assessment. In M. A. Britt, S. R. Goldman, & J.-F. Rouet (Eds.), *Reading-from words to multiple texts* (pp. 194–213). Routledge.

Goodwin, A. P., Cho, S. J., Reynolds, D., Brady, K., & Salas, J. (2020). Digital versus paper reading processes and links to comprehension for middle school students. *American Educational Research Journal*, *57*(4), 1837–1867. https://doi.org/10.3102/0002831219890300

Griffin, T. D., Wiley, J., & Thiede, K. W. (2008). Individual differences, rereading, and self-explanation: Concurrent processing and cue validity as constraints on metacomprehension accuracy. *Memory and Cognition*, *36*(1), 93–103. https://doi.org/10.3758/MC.36.1.93

Gurung, R. A. R., Weidert J., & Jeske A. (2010). Focusing on how students study. *Journal of the Scholarship of Teaching and Learning*, *10*(1), 28–35.

Haber, J., Nacenta, M. A., & Carpendale, S. (2014). Paper vs. tablets: The effect of document media in co-located collaborative work. In F. Garzotto, A. de A. Paolini, G. Jacucci, A. Malizia, M. Matera, & R. Lanzilotti (Eds.), *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces* (pp. 89–96). ACM. https://doi.org/10.1145/2598153.2598170

Hagen, Å. M., Braasch, J. L. G., & Bråten, I. (2014). Relationships between spontaneous note-taking, self-reported strategies and comprehension when reading multiple texts in different task conditions. *Journal of Research in Reading*, *37*(S1), S141–S157. https://doi.org/10.1111/j.1467-9817.2012.01536.x

Hartley, J., Bartlett, S., & Branthwaite, A. (1980). Underlining can make a difference-sometimes. *Journal of Educational Research*. https://doi.org/10.1080/00220671.1980.10885239

Hartwig, M. K., & Dunlosky, J. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement? *Psychonomic Bulletin and Review*, *19*(1), 126–134. https://doi.org/10.3758/s13423-011-0181-y

Hautala, J., Hyönä, J., & Aro, M. (2011). Dissociating spatial and letter-based word length effects observed in readers' eye movement patterns. *Vision Research*, *51*(15), 1719–1727. https://doi.org/10.1016/j.visres.2011.05.015

Hayes, A. F. (2014). Introduction to Mediation, Moderation, and Conditional Process Analysis: *Journal of Educational Measurement*, *51*(3), 335–337. https://doi.org/10.1111/jedm.12050

Hershberger, W. (1964). Self-evaluational responding and typographical cueing: Techniques for programing self-instructional reading materials. *Journal of Educational Psychology*, *55*(5), 288–296.

Hilbert, T. S., & Renkl, A. (2008). Concept mapping as a follow-up strategy to learning from texts: What characterizes good and poor mappers? *Instructional Science*, *36*(1), 53–73. https://doi.org/10.1007/s11251-007-9022-9

Huestegge, L., & Bocianski, D. (2010). Effects of syntactic context on eye movements during reading. *Advances in Cognitive Psychology*, *6*(6), 79–87. https://doi.org/10.2478/v10053-008-0078-0

Hutchings, D. R., Smith, G., Meyers, B., Czerwinski, M., & Robertson, G. (2004). Display space usage and window management operation comparisons between single monitor and multiple monitor users. *Proceedings of the Working Conference on Advanced Visual Interfaces - AVI '04*, 32–39.

Hutchings, D. R., & Stasko, J. (2004). Revisiting display space management: Understanding current practice to inform next-generation design. *Proceedings of the Conference on Graphics, Visualization & HCI (GI'04)*, *62*, 127–134. http://portal.acm.org/citation.cfm?id=1006074

Idstein, P., & Jenkins, J. R. (1972). Underlining versus repetitive reading. *Journal of Educational Research*, *65*(7), 321–323. https://doi.org/10.1080/00220671.1972.10884331

Jang, J., Bell Trickett, S., Schunn, C. D., & Gregory Trafton, J. (2012). Unpacking the temporal advantage of distributing complex visual displays. *International Journal of Human Computer Studies*, *70*(11), 812–827. https://doi.org/10.1016/j.ijhcs.2012.07.003

Jang, J., & Schunn, C. D. (2012). Performance benefits of spatially distributed versus stacked information on integration tasks. *Applied Cognitive Psychology*, *26*(2), 207–214. https://doi.org/10.1002/acp.1811

Jang, J., & Schunn, C. D. (2014). A framework for unpacking cognitive benefits of distributed complex visual displays. *Journal of Experimental Psychology: Applied*, *20*(3), 260–269. https://doi.org/10.1037/xap0000022

Jang, J., Schunn, C. D., & Nokes, T. J. (2011). Spatially distributed instructions improve learning outcomes and efficiency. *Journal of Educational Psychology*, *103*(1), 60–72. https://doi.org/10.1037/a0021994

Johnson, L. L. (1988). Effects of underlining textbook sentences on passage and sentence retention. *Reading Research and Instruction*, *28*(1), 18–32. https://doi.org/10.1080/19388078809557955

Kaakinen, J. K., Hyömä, J., & Keenan, J. M. (2003). How prior knowledge, WMC, and relevance of information affect eye fixations in expository text. *Journal of Experimental Psychology: Learning Memory and Cognition*, *29*(3), 447–457. https://doi.org/10.1037/0278-7393.29.3.447

Kaakinen, J. K., & Hyönä, J. (2007). Perspective effects in repeated reading: An eye movement study. *Memory and Cognition*, *35*(6), 1323–1336. https://doi.org/10.3758/BF03193604

Kammerer, Y., Kalbfell, E., & Gerjets, P. (2016). Is this information source commercially biased? How contradictions between web pages stimulate the consideration of source information. *Discourse Processes*, *53*, 430–456. https://doi.org/10.1080/0163853X.2016.1169968

Kammerer, Y., Meier, N., & Stahl, E. (2016). Fostering secondary-school students' intertext model formation when reading a set of websites: The effectiveness of source prompts. *Computers and Education*, *102*, 52–64. https://doi.org/10.1016/j.compedu.2016.07.001

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, *133*(2), 189–217. https://doi.org/10.1037/0096-3445.133.2.189

Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, *331*(6018), 772–775. https://doi.org/10.1126/science.1199327

Kiewra, K. A., Mayer, R. E., Christensen, M., Kim, S., & Risch, N. (1991). Effects of repetition on recall and note-taking: Strategies for learning from lectures. *Journal of Educational*

Psychology, *83*(1), 120–123.

Kiili, C., & Leu, D. J. (2019). Exploring the collaborative synthesis of information during online reading. *Computers in Human Behavior*, *95*, 146–157. https://doi.org/10.1016/j.chb.2019.01.033

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *98*, 163–182. https://doi.org/10.1037/0033-295X.95.2.163

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, MA: Cambridge University Press.

Kintsch, W. (2004). The construction–integration model of text comprehension and its implications comprehension. In R. B. Ruddell & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 1270–1328). International Reading Association.

Kobayashi, K. (2007). The influence of critical reading orientation on external strategy use during expository text reading. *Educational Psychology*, *27*(3), 363–375. https://doi.org/10.1080/01443410601104171

Kobayashi, K. (2009). Comprehension of relations among controversial texts: Effects of external strategy use. *Instructional Science*, *37*(4), 311–324. https://doi.org/https://doi.org/10.1007/s11251-007-9041-6

Kobayashi, K. (2016). Relational Processing of Conflicting Arguments. *Comprehensive Psychology*, *5*, 1–13. https://doi.org/10.1177/2165222816657801

Kornell, N., & Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review*, *14*(2), 219–224. http://www.ncbi.nlm.nih.gov/pubmed/17694904

Krieber, M., Bartl-Pokorny, K. D., Pokorny, F. B., Einspieler, C., Langmann, A., Körner, C., Falck-Ytter, T., & Marschik, P. B. (2016). The relation between reading skills and eye movement patterns in adolescent readers: Evidence from a regular orthography. *PLoS ONE*, *11*(1), 1–13. https://doi.org/10.1371/journal.pone.0145934

Lachner, A., Burkhart, C., & Nückles, M. (2017). Mind the gap! automated concept map feedback supports students in writing cohesive explanations. *Journal of Experimental Psychology: Applied*, *23*(1), 29–46. https://doi.org/10.1037/xap0000111

Latini, N., Bråten, I., Anmarkrud, Ø., & Salmerón, L. (2019). Investigating effects of reading medium and reading purpose on behavioral engagement and textual integration in a multiple text context. *Contemporary Educational Psychology*, *59*, 101797. https://doi.org/10.1016/j.cedpsych.2019.101797

Latini, N., Bråten, I., & Salmerón, L. (2020). Does reading medium affect processing and integration of textual and pictorial information? A multimedia eye-tracking study. *Contemporary Educational Psychology*, *62*. https://doi.org/10.1016/j.cedpsych.2020.101870

Le Bigot, L., & Rouet, J.-F. (2007). The impact of presentation format, task assignment, and prior knowledge on students' comprehension of multiple online documents. *Journal of Literacy Research*, *39*(4), 445–470. http://journals.sagepub.com.ezproxy.uky.edu/doi/pdf/10.1080/10862960701675317

Lee, M., & Baylor, A. L. (2006). Designing Metacognitive Maps for Web-Based Learning . Designing Metacognitive Maps for Web-Based Learning Disorientation and Metacognition in Web-Based Learning Environments. *Educational Technology and Society*, *9*(1), 344–348.

Leopold, C., & Leutner, D. (2015). Improving students' science text comprehension through metacognitive self-regulation when applying learning strategies. *Metacognition and Learning*, *10*(3), 313–346. https://doi.org/10.1007/s11409-014-9130-2

Leroy, C., Gerjets, P., Oestermeier, U., & Kammerer, Y. (2021). Simultaneous presentation of multiple documents and text-highlighting: Online integrative processes and offline integrated understanding. *Scientific Studies of Reading*, *25*(2), 179–192. https://doi.org/https://doi.org/10.1080/10888438.2020.1784903

Li, L.-Y., Tseng, S. T., & Chen, G. D. (2016). Effect of hypertext highlighting on browsing, reading, and navigational performance. *Computers in Human Behavior*, *54*, 318–325. https://doi.org/10.1016/j.chb.2015.08.012

Linderholm, T., & Van den Broek, P. (2002). The effects of reading purpose and working memory capacity on the processing of expository text. *Journal of Educational Psychology*, *94*(4), 778–784. https://doi.org/10.1037/0022-0663.94.4.778

List, A. (2019). Drawing is integrating: An examination of students' graphic representations of multiple texts. *Reading Psychology*, *40*(6), 491–524. https://doi.org/10.1080/02702711.2019.1629517

List, A., & Alexander, P. A. (2017a). Cognitive affective engagement model of multiple source use. *Educational Psychologist*, *52*(3), 182–199. https://doi.org/10.1080/00461520.2017.1329014

List, A., & Alexander, P. A. (2017b). Text navigation in multiple source use. *Computers in Human Behavior*, *75*, 364–375. https://doi.org/10.1016/j.chb.2017.05.024

List, A., & Alexander, P. A. (2018). Corroborating students' self-reports of source evaluation. *Behaviour and Information Technology*, *37*(3), 198–216. https://doi.org/10.1080/0144929X.2018.1430849

List, A., & Alexander, P. A. (2020). Strategy use in learning from multiple texts: An investigation of the integrative framework of learning from multiple texts. *Frontiers in Education*, *5*. https://doi.org/10.3389/feduc.2020.578062

List, A., & Du, H. (2021). Reasoning beyond history: Examining students' strategy use when completing a multiple text task addressing a controversial topic in education. *Reading and Writing*, *34*(4), 1003–1048. https://doi.org/10.1007/s11145-020-10095-5

List, A., Du, H., & Lee, H. Y. (in press). How do students integrate multiple texts? An investigation of top-down processing. *European Journal of Psychology of Education*.

List, A., Du, H., & Lee, H. Y. (2021). Examining relation formation across consistent and conflicting texts. *Discourse Processes*, *58*(2), 134–154. https://doi.org/10.1080/0163853X.2020.1834328

List, A., Du, H., Wang, Y., & Lee, H. Y. (2019). Toward a typology of integration: Examining the documents model framework. *Contemporary Educational Psychology*, *58*, 228–242. https://doi.org/10.1016/j.cedpsych.2019.03.003

List, A., Stephens, L. A., & Alexander, P. A. (2019). Examining interest throughout multiple text use. *Reading and Writing*, *32*(2), 307–333.

Lombard, J., Amadieu, F., Bråten, I., & Van de Leemput, C. (2018). Reading multiple documents on tablet: Effects of applications and strategic guidance on performance and acceptance. In P. Zaphiris & A. Ioannou (Eds.), *Learning and Collaboration Technologies. Design, Development and Technological Innovation.* (Vol. 10924, pp. 157–169). Springer, Cham. https://doi.org/10.1007/978-3-319-91743-6_12

Lombard, J., Bråten, I., van de Leemput, C., & Amadieu, F. (2021). Performance and acceptance when using tablets as a multiple document learning tool: Do application and guidance matter? *Instructional Science*, *49*, 197–221. https://doi.org/10.1007/s11251-021-09537-6

Lorch, R. F. (1989). Text-signaling devices and their effects on reading and memory processes. *Educational Psychology Review*, *1*(3), 209–234. https://doi.org/10.1007/BF01320135

Lorch, R. F., Lorch, E. P., & Kluseqitz, M. A. (1995). Effects of typographical cues on reading and recall of text. *Contemporary Educational Psychology*, *20*, 51–64.

Lundstrom, K., Diekema, A. R., Leary, H., Haderlie, S., & Holliday, W. (2015). Teaching and learning information synthesis: An intervention and rubric based assessment. *Communications in Information Literacy*, *9*(1), 60–82. https://doi.org/10.15760/comminfolit.2015.9.1.176

Mahlow, N., Hahnel, C., Kroehne, U., Artelt, C., Goldhammer, F., & Schoor, C. (2020). More than (single) text comprehension? – On university students' understanding of multiple documents. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.562450

Maier, J., & Richter, T. (2013). Text-belief consistency effects in the comprehension of multiple texts with conflicting information. *Cognition and Instruction*, *31*(2), 151–175. https://doi.org/10.1080/07370008.2013.769997

Margolin, S. J., Driscoll, C., Toland, M. J., & Kegler, J. L. (2013). E-readers, computer screens, or paper: Does reading comprehension change across media platforms? *Applied Cognitive Psychology*, *27*(4), 512–519. https://doi.org/10.1002/acp.2930

Mayer, R. E. (1983). Can you repeat that? Qualitative effects of repetition and advance organizers on learning from science prose. *Journal of Educational Psychology*, *75*(1), 40–49. https://doi.org/10.1037/0022-0663.75.1.40

Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge Handbook of Multimedia Learning, Second Edition* (pp. 279–315). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.015

McCrudden, Matthew T., Magliano, J. P., & Schraw, G. (2010). Exploring how relevance instructions affect personal reading intentions, reading goals and text processing: A mixed methods study. *Contemporary Educational Psychology*, *35*(4), 229–241.

https://doi.org/10.1016/j.cedpsych.2009.12.001

McCrudden, Matthew T., & Schraw, G. (2007). Relevance and goal-focusing in text processing. *Educational Psychology Review*, *19*(2), 113–139. https://doi.org/10.1007/s10648-006-9010-7

Microsoft©. (2016). *Microsoft: Productivity Future Vision*. https://www.microsoft.com/en-us/research/video/productivity-future-vision/

Millis, K. K., & King, A. (2001). Rereading strategically: The influences of comprehension ability and a prior reading on the memory for expository text. *Reading Psychology*, *22*(1), 41–65. https://doi.org/10.1080/02702710117227

Millis, K. K., Simon, S., & TenBroek, N. S. (1998). Resource allocation during the rereading of scientific texts. *Memory and Cognition*, *26*(2), 232–246. https://doi.org/10.3758/BF03201136

Miyatsu, T., Nguyen, K., & McDaniel, M. A. (2018). Five popular study strategies: Their pitfalls and optimal implementations. *Perspectives on Psychological Science*, *13*(3), 390–407. https://doi.org/10.1177/1745691617710510

Nist, S. L., & Hogrebe, M. C. (1987). The role of underlining and annotating in remembering textual information. *Reading Research and Instruction*, *27*(1), 12–25. https://doi.org/10.1080/19388078709557922

Nückles, M., Hübner, S., & Renkl, A. (2009). Enhancing self-regulated learning by writing learning protocols. *Learning and Instruction*, *19*(3), 259–271. https://doi.org/10.1016/j.learninstruc.2008.05.002

O'Hara, K. P., & Sellen, A. (1997). A comparison of reading paper and on-line documents. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, 335–342. https://doi.org/10.1145/258549.258787

O'Hara, K. P., Taylor, A., Newman, W., & Sellen, A. J. (2002). Understanding the materiality of writing from multiple sources. *International Journal of Human Computer Studies*, *56*(3), 269–305. https://doi.org/10.1006/ijhc.2001.0525

Olive, T., Rouet, J. F., François, E., & Zampa, V. (2008). Summarizing digital documents: Effects of alternate or simultaneous window display. *Applied Cognitive Psychology*, *22*(4), 541–558. https://doi.org/10.1002/acp.1380

Perfetti, C. A. (1997). Sentences, individual differences, and multiple texts: Three issues in text comprehension. *Discourse Processes*, *23*(3), 337–355. https://doi.org/10.1080/01638539709544996

Perfetti, C. A., Rouet, J.-F., & Britt, M. A. (1999). Toward a theory of documents representation. In H. Von Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 88–108). Lawrence Erlbaum Associates.

Peterson, P., Baker, E., & McGaw, B. (2010). International Encyclopedia of Education. In *International Encyclopedia of Education*. https://doi.org/10.1016/0738-0593(94)90060-4

Peterson, S. E. (1991). The cognitive functions of underlining as a study technique. *Reading Research and Instruction*, *31*(2), 49–56. https://doi.org/10.1080/19388079209558078

Ponce, H. R., & Mayer, R. E. (2014). An eye movement analysis of highlighting and graphic organizer study aids for learning from expository text. *Computers in Human Behavior*, *41*, 21–32. https://doi.org/10.1016/j.chb.2014.09.010

Ponce, H. R., Mayer, R. E., Loyola, M. S., López, M. J., & Méndez, E. E. (2018). When two computer-supported learning strategies are better than one: An eye-tracking study. *Computers and Education*, *125*, 376–388. https://doi.org/10.1016/j.compedu.2018.06.024

Praveen, S. D., & Rajan, P. (2013). Using graphic organizers to improve reading comprehension skills for the middle school ESL students. *English Language Teaching*, *6*(2), 155–170. https://doi.org/10.5539/elt.v6n2p155

Primor, L., & Katzir, T. (2018). Measuring multiple text integration: A review. *Frontiers in Psychology*, *9*, 1–16. https://doi.org/10.3389/fpsyg.2018.02294

Rawson, K. A., & Kintsch, W. (2005). Rereading effects depend on time of test. *Journal of Educational Psychology*, *97*(1), 70–80. https://doi.org/10.1037/0022-0663.97.1.70

Reader, W. R., & Payne, S. J. (2007). Allocating time across multiple texts: Sampling and satisficing. *Human-Computer Interaction*, *22*(3), 263–298. https://doi.org/10.1080/07370020701493376

Renkl, A., & Scheiter, K. (2017). Studying visual displays: How to instructionally support learning. *Educational Psychology Review*, *29*(3), 599–621. https://doi.org/10.1007/s10648-015-9340-4

Reznitskaya, A., Kuo, L. jen, Glina, M., & Anderson, R. C. (2009). Measuring argumentative reasoning: What's behind the numbers? *Learning and Individual Differences*, *19*(2), 219–224. https://doi.org/10.1016/j.lindif.2008.11.001

Richter, T., & Maier, J. (2017). Comprehension of multiple documents with conflicting information: A two-step model of validation. *Educational Psychologist*, *52*(3), 148–166. https://doi.org/10.1080/00461520.2017.1322968

Rickards, J. P., & August, G. J. (1975). Generative underlining strategies in prose recall. *Journal of Educational Psychology*, *67*(6), 860–865. https://doi.org/10.1037/0022-0663.67.6.860

Rinck, M., Gámez, E., Díaz, J. M., & De Vega, M. (2003). Processing of temporal information: Evidence from eye movements. *Memory and Cognition*, *31*(1), 77–86.

Robertson, G., Czerwinski, M., Larson, K., Robbins, D. C., Thiel, D., & Dantzich, M. Van. (1998). Data mountain: Using spatial memory for document management. *Proceedings of the ACM Symposium on User Interface Software and Technology*, 153–162. https://doi.org/http://doi.acm.org/10.1145/288392.288596

Rothkopf, E. Z. (1968). Textual constraint as function of repeated inspection. *Journal of Educational Psychology*, *59*, 20–25. https://doi.org/10.1037/h0025378

Rouet, J.-F., & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M.T. McCrudden, J. P. Maglilano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 19–52). Greenwich, CT: Information Age Publishing.

Rouet, J.-F., Britt, M. A., & Durik, A. M. (2017). RESOLV: Readers' representation of reading contexts and tasks. *Educational Psychologist*, *52*(3), 200–215. https://doi.org/10.1080/00461520.2017.1329015

Rouet, J.-F., Le Bigot, L., Pereyra, G. De, & Britt, M. A. (2016). Whose story is this? Discrepancy triggers readers' attention to source information in short narratives. *Reading and Writing*, *29*, 1549–1570. https://doi.org/10.1007/s11145-016-9625-0

Rouet, J.-F., Mason, R. A., Perfetti, C. A., & Britt, M. A. (1996). Using multiple sources of evidence to reason about history. *Journal of Educational Psychology*, *88*(3), 478–493. https://doi.org/10.1037/0022-0663.88.3.478

Rouet, J.-F., Vidal-Abarca, E., & Millogo, V. (2001). Effects of information search tasks on the comprehension of instructional text. *Discourse Processes*, *31*(2), 163–186.

https://doi.org/10.1207/S15326950DP3102

Salmerón, L., Baccino, T., Cañas, J. J., Madrid, R. I., Fajardo, I., Canas, J. J., Madrid, R. I., & Fajardo, I. (2009). Do graphical overviews facilitate or hinder comprehension in hypertext? *Computers & Education*, *53*(4), 1308–1319. https://doi.org/10.1016/j.compedu.2009.06.013

Salmerón, L., & Garcia, V. (2012). Children's reading of printed text and hypertext with navigation overviews: The role of comprehension, sustained attention, and visuo-spatial abilities. *Journal of Educational Computing Research*, *47*(1), 33–50.

Salmerón, L., & García, V. (2011). Reading skills and children's navigation strategies in hypertext. *Computers in Human Behavior*, *27*(3), 1143–1151. https://doi.org/10.1016/j.chb.2010.12.008

Salmerón, L., Gil, L., & Bråten, I. (2018a). Effects of reading real versus print-out versions of multiple documents on students' sourcing and integrated understanding. *Contemporary Educational Psychology*, *52*, 25–35. https://doi.org/https://doi.org/10.1016/j.cedpsych.2017.12.002

Salmerón, L., Gil, L., & Bråten, I. (2018b). Using eye-tracking to assess sourcing during multiple document reading: A critical analysis. *Frontline Learning Research*, *6*(3), 105–122. https://doi.org/10.14786/flr.v6i3.368

Salmerón, L., Gil, L., Bråten, I., & Strømsø, H. (2010). Comprehension effects of signalling relationships between documents in search engines. *Computers in Human Behavior*, *26*(3), 419–426. https://doi.org/10.1016/j.chb.2009.11.013

Salmerón, L., & Llorens, A. C. (2019). Instruction of digital reading strategies based on eye-movements modeling examples. *Journal of Educational Computing Research*, *57*(2), 343–359. https://doi.org/https://doi.org/10.1177/0735633117751605

Schnell, T. R., & Rocchio, D. J. (1978). A comparison of underlying strategies for improving reading comprehension and retention. *Reading Horizons*, *18*(2), 4. http://scholarworks.wmich.edu/reading_horizons/vol18/iss2/4

Schugar, J. T., Schugar, H., & Penny, C. (2011). A Nook or a book? Comparing college students' reading comprehension levels, critical reading, and study skills. *International Journal of Technology in Teaching & Learning*, *7*(2), 174–192.

Shapiro, A. M. (2004). How including prior knowledge as a subject variable may change outcomes of learning research. *American Educational Research Journal*, *41*(1), 159–189. https://doi.org/10.3102/00028312041001159

Sharafi, Z., Shaffer, T., & Sharif, B. (2015). Eye-tracking metrics in software engineering. *Asia-Pacific Software Engineering Conference (APSEC)*, 96–103. https://doi.org/10.1109/APSEC.2015.53

Shi, Z., Tang, T., & Yin, L. (2020). Construction of cognitive maps to improve reading performance by text signaling: Reading text on paper compared to on screen. *Frontiers in Psychology*, *11:571957*. https://doi.org/10.3389/fpsyg.2020.571957

Shibata, H., Takano, K., & Omura, K. (2013). Why is paper superior to computer displays in cross-reference reading for multiple documents? *Proceedings of the International Display Workshop (IDW)*, 5–8.

Silvers, V. L., & Kreiner, D. S. (1997). The effects of pre-existing inappropriate highlighting onreading comprehension. *Reading Research and Instruction*, *36*(3), 217–223. https://doi.org/10.1080/19388079709558240

Skuballa, I. T., Dammert, A., & Renkl, A. (2018). Two kinds of meaningful multimedia learning: Is cognitive activity alone as good as combined behavioral and cognitive activity? *Learning and Instruction*, *54*, 35–46. https://doi.org/10.1016/j.learninstruc.2018.02.001

So, J. C. Y., & Chan, A. H. S. (2009). Influence of highlighting validity on dynamic text comprehension performance. In S. I. Ao, O. Castillo, C. Douglas, D. D. Feng, & J.-A. Lee (Eds.), *International MultiConference of Engineers and Computer Scientists: Vol. II* (pp. 1970–1973). Newswood Limited.

Stadtler, M., & Bromme, R. (2004). Laypersons searching for medical information on the web: The role of metacognition. *Proceedings of the 26th Annual Conference of the Cognitive Science Society*, 1638. http://www.cogsci.northwestern.edu/cogsci2004/ma/ma293.pdf

Stadtler, M., & Bromme, R. (2007). Dealing with multiple documents on the WWW: The role of metacognition in the formation of documents models. *International Journal of Computer-Supported Collaborative Learning*, *2*, 191–210. https://doi.org/10.1007/s11412-007-9015-3

Stadtler, M., & Bromme, R. (2014). The content-source integration model: A taxonomic description of how readers comprehend conflicting scientific information. In D. N. Rapp & J. L. G. Braasch (Eds.), *Processing inaccurate information - Theoretical and applied perspectives from cognitive science and the educational sciences* (pp. 379–402). MIT Press.

Stang Lund, E., Bråten, I., Brandmo, C., Brante, E. W., & Strømsø, H. I. (2019). Direct and indirect effects of textual and individual factors on source-content integration when reading about a socio-scientific issue. *Reading and Writing*, *32*, 335–356. https://doi.org/10.1007/s11145-018-9868-z

Stang Lund, E., Bråten, I., Brante, E. W., & Strømsø, H. I. (2017). Memory for textual conflicts predicts sourcing when adolescents read multiple expository texts. *Reading Psychology*, *38*(4), 417–437. https://doi.org/10.1080/02702711.2016.1278417

Stordahl, K. E., & Christensen, C. M. (1956). The effect of study techniques on comprehension and retention. *Journal of Educational Research*, *49*(8), 561–570. https://doi.org/10.1080/00220671.1956.10882321

Strobelt, H., Oelke, D., Kwon, B. C., Schreck, T., & Pfister, H. (2016). Guidelines for effective usage of text highlighting techniques. *IEEE Transactions on Visualization and Computer Graphics*, *22*(1), 489–498. https://doi.org/10.1109/TVCG.2015.2467759

Strømsø, H. I., Bråten, I., Anmarkrud, Ø., & Ferguson, L. E. (2016). Beliefs about justification for knowing when ethnic majority and ethnic minority students read multiple conflicting documents. *Educational Psychology*, *36*(4), 638–657. https://doi.org/10.1080/01443410.2014.920080

Strømsø, H. I., Bråten, I., & Britt, M. A. (2010). Reading multiple texts about climate change: The relationship between memory for sources and text comprehension. *Learning and Instruction*, *20*(3), 192–204. https://doi.org/10.1016/j.learninstruc.2009.02.001

Strømsø, H. I., Bråten, I., & Samuelstuen, M. S. (2003). Students' strategic use of multiple sources during expository text reading: A longitudinal think-aloud study. *Cognition and Instruction*, *21*, 113–147. https://doi.org/10.1207/S1532690XCI2102

Takano, K., Shibata, H., & Omura, K. (2015). Effects of paper on cross-reference reading for multiple documents: Comparison of reading performances and processes between paper and computer displays. *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction (OzCHI '15)*, 497–505.

https://doi.org/10.1145/2838739.2838745

Van Den Broek, P., & Kendeou, P. (2015). Building coherence in web-based and other non-traditional reading environments: Cognitive opportunities and challenges. *Reading at a Crossroads?: Disjunctures and Continuities in Current Conceptions and Practices*, *January 2015*, 104–114.

Van Gog, T., Paas, F., Van Merriënboer, J. J. G., & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied*, *11*(4), 237–244. https://doi.org/10.1037/1076-898X.11.4.237

Van Strien, J. L. H., Brand-Gruwel, S., & Boshuizen, H. P. A. (2014). Dealing with conflicting information from multiple nonlinear texts: Effects of prior attitudes. *Computers in Human Behavior*, *32*, 101–111. https://doi.org/10.1016/j.chb.2013.11.021

von Restorff, H. (1933). Über die Wirkung von Bereichsbildungen im Spurenfeld. In *Psychologische Forschung* (Vol. 18, pp. 299–342). https://doi.org/10.1007/BF02441202

Wertheimer, M. (1925). Über Gestalttheorie. *Philosophische Zeitschrift Für Forschung Und Aussprache*, *1*.

Wiley, J. (2001). Supporting understanding through task and browser design. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *23*(23), 375–389.

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. https://doi.org/10.3102/0002831209333183

Winchell, A., Lan, A., & Mozer, M. (2020). Highlights as an early predictor of student comprehension and interests. *Cognitive Science*, *44*(11). https://doi.org/10.1111/cogs.12901

Wineburg, S. S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, *83*(1), 73–87. https://doi.org/https://doi.org/10.1037/0022-0663.83.1.73

Wolfe, M. B. W., & Goldman, S. R. (2005). Relations between adolescents' text processing and reasoning. *Cognition and Instruction*, *23*(4), 467–502.

Yan, V. X., Thai, K.-P., & Bjork, R. A. (2014). Habits and beliefs that guide self-regulated learning: Do they vary with mindset? *Journal of Applied Research in Memory and Cognition*, *3*(3), 140–152. https://doi.org/10.1016/j.jarmac.2014.04.003

Yang, A., Chen, I. Y. L., Flanagan, B., & Ogata, H. (2020). How does the quality of students' highlights affect their learning performance in e-Book reading? In H. J. So (Ed.), *28th International Conference on Computers in Education* (pp. 289–294). Asia-Pacific Society for Computers in Education (APSCE).

Yeari, M., Oudega, M., & van den Broek, P. (2017). The effect of highlighting on processing and memory of central and peripheral text information: Evidence from eye movements. *Journal of Research in Reading*, *40*(4), 365–383. https://doi.org/10.1111/1467-9817.12072

Yik, N. C., Yi, L. X., Somadam, S. T. N., Amirudin, A. E. Bin, & Ananthan, S. (2018). Effect of highlighting text on concentration, memory and attention among undergraduate medical students: A Randomized controlled trial. *American Journal of Educational Science*, *4*(4), 149–158. http://www.aiscience.org/journal/ajeshttp://creativecommons.org/licenses/by/4.0/

Yue, C. L., Storm, B. C., Kornell, N., & Bjork, E. L. (2014). Highlighting and its relation to distributed study and students' metacognitive beliefs. *Educational Psychology Review*, *27*(1), 69–78. https://doi.org/10.1007/s10648-014-9277-z

List of Tables

Table 1.1 Examples of tasks and measures used to assess multiple document comprehension, as well as the component of the documents model that the respective task addresses, and examples of previous research that used these tasks……….………………………………………………..…………...…7

Table 1.2 Non-comprehensive summary of previous work investigating the effect of experimenter-generated (Exp.) or reader-generated (Self) text-highlighting by measure category assessed……………………………………………..…19

Table 2.1 Summary of the three studies conducted as part of the present dissertation with respect to the characteristics of the reading environment manipulated as well as reading interactions, multiple document comprehension measures, and additional behavioral measures assessed…………………..………...…………………36

Table 3.1 Means (and standard deviations) of reading time, number of highlights, and dependent measures as a function of documents presentation and highlighting…………………………………………………………………….57

Table 3.5.1 Detailed description of the six documents. …………………………………….63

Table 3.5.2 Examples of different types of intertextual connections……………………….65

Table 3.5.3 Control variables, each on a scale from 1 = "very low" to 7 = "very high"…..….67

Table 3.5.4 Means (and standard deviations) and inferential statistics of control variables, i.e., working memory capacity (a score of 0-15) and self-report items (7-point Likert scales from 1-7) as a function of documents presentation and highlighting. ….68

Table 3.5.5 Means (and standard deviations) of participants' self-reported topic beliefs (7-point Likert scales from 1-7) before and after reading as well as repeated-measures ANOVAs for measurement time…………………………………..………..69

Table 3.5.6 Means (and standard deviations) of the number of revisits to a task-relevant document and to the task-irrelevant document as a function of highlighting…..……………………………………………….…………...70

Table 3.5.7 Means (and standard deviations) of the number of highlights made in a task-relevant document and in the task-irrelevant document as a function of documents presentation.…………………………………………….……………….…….70

Table 3.5.8 Percentage of intertextual connections between different types of documents as a function of document presentation and highlighting…………………...…..…71

Table 3.5.9 Descriptives and inferential statistics of the percentage of essays containing source citations and essay length as a function of documents presentation and highlighting…………………………………………………………………..72

Table 4.1 Means (and standard deviations) for control variables and multiple document comprehension measures as a function of document presentation….………….. 91

Table 4.2 Descriptive and inferential statistics for all measured variables in the simultaneous condition as a function of having or not having grouped documents during reading……………………………………………………………………92

Table 4.3 Descriptive and inferential statistics for all measured variables in the sequential condition as a function of having or not having grouped documents during reading………..…………………………………………………………..…94

Table 5.1 Means (and standard deviations) for control variables as a function of text-highlighting…………………………………………………………………...120

Table 5.2 Means (and standard deviations) of the number of re-readings with the respective duration-threshold as a function of text-highlighting, as well as inferential statistics……………………………………………………………………….121

Table 5.3 Means (and standard deviations) of the number of fixations on documents per re-access for short re-accesses that lasted up to 5 sec as a function of the re-access duration and text-highlighting. ……………………………………………122

Table 5.4 Inferential statistics for main effects of and interactions between text-highlighting and the number of re-readings (for the 1 and 3 sec re-reading thresholds) on the integrated understanding measures (number of intertextual connections and source-content mapping score)……………………………..………..……125

Table 5.5 Means (and standard deviations) for overall reading times, initial reading times (both including and excluding the time taken for highlighting during initial readings), and re-reading times as a function of text-highlighting……..……...128

Table 5.6.1 Detailed description of the five documents…………………………..…..……137

Table 5.6.2 Examples of different types of intertextual connections (second column) as well as original statements (translated from German) from participants, which, when integrated, yield the exemplary intertextual connection. ....…………....139

Table 5.6.3 Means (and standard deviations) of the number of re-readings with the respective duration-threshold as a function of text-highlighting, as well as inferential statistics…………………………………………………………………….142

Table 5.6.4 Inferential statistics for main effects of and interactions between text- highlighting and the number of re-readings (for the 5 different re-reading thresholds) on the integrated understanding measures (number of intertextual connections and source-content mapping score)…………………….………………….……143

Table 5.6.5 Results of the simple slopes analyses according to Aiken and West (1991) regarding the interaction effect between text-highlighting (with vs. without) and the number of re-readings (z-standardized continuous predictor) for a given re-reading threshold, on the number of intertextual connections in participants' essays………………………………………………………………….……...144

Table 5.6.6 Means (and standard deviations) for re-reading times as a function of text-highlighting…………………………………………………………………...145

Table 5.6.7 Results of the moderated linear regression analyses and respective simple slopes analyses according to Aiken and West (1991) regarding the interaction effect between text-highlighting (with vs. without) and the number of re-readings (z-standardized continuous predictor) for a given re-reading threshold, on the re-reading time…………………………………………………………….…….146

Table 6.1 Comparative presentation of the results across the three studies of the present dissertation concerning the effects of the reading environment on specific aspects of multiple document comprehension and the role of reading interactions therein……………………………………………………………………..….150

Table 6.2 Means (standard deviations) of fixation durations on text and source logos during initial readings and re-readings as a function of text-highlighting, as well as inferential statistics regarding differences across conditions and relations to multiple document comprehension measures assessed…….……………………158

List of Figures

Figure 1.1 Schematic illustration of the documents model (oval with grey background on the top; Perfetti et al., 1999) a reader of two documents might construct, and the iterative reading process suggested by the MD-TRACE model (bottom; Britt & Rouet, 2012)……………………………………………………………….……6

Figure 1.2 Schematic representation of the RESOLV model (Britt et al., 2018; Rouet et al., 2017). Rhombuses represent decision steps which, in comparison to the MD-TRACE model, include one regarding whether the respective reading action should be taken based on the benefit-cost assessment………..………………..9

Figure 3.1 Reenactment of a participant conducting the reading task on the experimental multi-touch-table user interface in the simultaneous with-highlighting condition, currently highlighting text……………………………………………………51

Figure 3.2 Multi-touch table interface: Initial positions of documents in the sequential (left column) and simultaneous (right column) conditions, without (top row) and with highlighting (middle row). The bottom row shows an exemplary manipulation of documents in the respective documents presentation condition with highlighting………………………………………………………………..54

Figure 3.3 Interaction of highlighting condition and number of revisits to relevant documents (z-standardized) with regard to the number of intertextual connections in the sequential conditions………………………………………………………59

Figure 4.1 Simultaneous (left) and sequential (right) experimental user interface. Figures in the top row represent the initial state; figures in the bottom row represent a potential manipulation in the respective condition…………………...……....85

Figure 4.2 Items of the source-content-mapping task (translated from German). ………....88

Figure 5.1 Re-enactment of a participant standing in front of the height-adjustable table with the inclined display on which the documents were presented……..………..112

Figure 5.2 Experimental user interface without highlighting-tool (left) and with highlighting-tool (right). ………………………………………………...…………....112

Figure 5.3 Items of the source-content-mapping task (translated from German)……..….114

Figure 5.4 Heatmaps of a participant's fixations on the document 'Ärztezeitung' during its first access (left; 107 sec duration), second access (middle; 28 sec duration), and re-readings (accumulated; right; 137 sec duration). ....………………………..…116

Figure 5.5 Interaction between text-highlighting and the number of re-readings ≥ 3 seconds (z-standardized) with regard to the number of intertextual connections……….126

Figure 5.6 Interaction between text-highlighting and the number of re-readings ≥ 3 sec (z-standardized) with regard to re-reading time (only re-accesses ≥ 3 sec taken into account)........……………………………………………………………...…..129

Figure 5.7 Mean percentage of participants' re-reading fixation coverage for each document as a function of text-highlighting. ………..…………………………………129

Figure 6.1 Comparative presentation of the likelihood to group documents during reading (left) across Studies 1 and 2, and of the number of revisits / re-readings (right) across Studies 1, 2, and 3………………..………………………………….…162

Figure 6.2 Comparative presentation of the mean number of intertextual connections across Studies 1 (left and middle pair of bars) and 2 (right pair of bars) as a function of document presentation……………………………………………...………164

Figure 6.3 Comparative presentation of the mean number of intertextual connections across Studies 1 (left and middle pair of bars) and 3 (right pair of bars) as a function of text-highlighting…………………………………..…………………..…….165