A NON-LINEAR ELECTRONIC TEXTBOOK FORMAT TO FACILITATE DEEP LEARNING

by

Eric Spero

Submitted in partial fulfillment of the requirements for the degree of Master of Arts, Human-Computer Interaction at

Carleton University
Ottawa, Canada
May 2018

© Copyright by Eric Spero, 2018
Table of Contents

List of Tables ......................................................... vi
List of Figures ....................................................... vii
Abstract ................................................................. ix
Acknowledgements ...................................................... x

Chapter 1 Introduction .............................................. 1
  1.1 Research Purpose ............................................. 1
  1.2 Thesis Structure ............................................. 2
  1.3 Contributions ............................................... 3
  1.4 Disclaimer on the Use of “We” ................................ 3

Chapter 2 Literature Review ...................................... 4
  2.1 Depth Over Breadth .......................................... 4
  2.2 Project 2061 and the Next Generation Science Standards .... 5
  2.3 Issues with Textbooks ...................................... 6
  2.4 Situation Models and the Construction-Integration Model of Text Comprehension ......................................................... 7
  2.5 Cognitive Load Theory ..................................... 7
  2.6 Bloom’s Taxonomy .......................................... 8
  2.7 Embodied/Grounded Cognition ................................ 9
      2.7.1 Meaning and use ....................................... 9
      2.7.2 The main claim ....................................... 10
      2.7.3 Abstract vs. concrete concepts ................... 10
      2.7.4 Thinking with the environment .................. 11
  2.8 Psychological Essentialism ................................ 11
  2.9 Knowledge is Categorical .................................. 12
  2.10 Knowledge is Hierarchical ............................... 12
Chapter 6 Study ............................................ 49
  6.1 Participants ........................................... 50
  6.2 Materials ............................................. 50
  6.3 Procedure ........................................... 52
  6.4 Third Round of Pilot Testing ......................... 53

Chapter 7 Results ........................................ 55
  7.1 RQ2 Hypothesis Testing ............................. 55
    7.1.1 Post-reading test ................................. 56
    7.1.2 CAP Perceived Learning Scale ................. 60
  7.2 RQ3 Hypothesis Testing ............................. 61
    7.2.1 SUS scores ...................................... 61
    7.2.2 Interface preference ............................ 61
    7.2.3 Themes from the post-experiment interview .. 62

Chapter 8 Discussion ....................................... 65
  8.1 Learning Outcomes ................................ 65
    8.1.1 Differences between the interfaces .......... 65
    8.1.2 Other explanations for lack of findings ...... 68
  8.2 What does this mean for standard print texts? .... 69
  8.3 Is the McNamara et al. really a challenge? ........ 70

Chapter 9 Conclusions ..................................... 71
  9.1 Research Summary ................................ 71
  9.2 Contributions ...................................... 72
  9.3 Limitations and Future Work ......................... 72
    9.3.1 Design .......................................... 72
    9.3.2 Study ........................................... 73

Appendix A Demographics questionnaire .............. 85

Appendix B Cognitive science background test questions 86
Appendix C  Postreading test questions ....................... 90
   C.1  Text-based questions  ........................................ 90
   C.2  Bridging inference questions  .......................... 92
   C.3  Problem solving questions  ............................... 93
   C.4  (One more) bridging inference question .............. 95

Appendix D  System Usability Scale (SUS) questionnaire .... 96

Appendix E  Cognitive, Affective, and Psychomotor (CAP) Perceived Learning Scale questionnaire .......... 97

Appendix F  Research Ethics Protocol Form .................. 98
List of Tables

7.1 Postreading Test Scores ........................................ 57
List of Figures

2.1 ELM-ART .............................................. 15

3.1 Schematic of node structure .................................. 22
3.2 Schematic of node hierarchy .................................. 22
3.3 Early attempt at model design ................................. 24
3.4 Early attempt at mapping content ............................. 25
3.5 Early schematic for content structure ....................... 26
3.6 Early attempt at interface design ............................ 27
3.7 High fidelity prototype ...................................... 28

4.1 Beginning JavaScript development: First iteration .......... 29
4.2 Second JavaScript iteration ................................ 30
4.3 Map view: Visualizing the text structure with D3.js ... 34
4.4 An improved minimap using D3.js .......................... 35
4.5 Final version: Main view ................................... 36
4.6 Final version: Treemap ..................................... 37
4.7 Search functionality ....................................... 38
4.8 Final version: Slider ....................................... 39
4.9 Data structure for text content .............................. 40
4.10 Walkthrough 1 ........................................... 42
4.11 Walkthrough 2 ........................................... 42
4.12 Walkthrough 3 ........................................... 42
4.13 Walkthrough 4 ........................................... 42
4.14 Walkthrough 5 ........................................... 43
4.15 Walkthrough 6 ........................................... 43
4.16 Walkthrough 7 ........................................... 43
4.17 Walkthrough 8

5.1 The control interface

6.1 Setting for the study

6.2 The screen as viewed by participants

6.3 Flow diagram of procedure

7.1 Postreading Test Scores: Six Boxplots

7.2 Learning effects by condition

7.3 CAP scores by condition

7.4 SUS scores by condition

7.5 Interface preference

7.6 Interface preference split by condition
Many educators are calling for an increased emphasis on the interactions between the concepts covered in an educational unit because of its benefits for learning. There is evidence that at present, standard linear textbooks do not adequately support this “deep” approach to curriculum design, and we think this is partially due to the linear format. In this thesis, we propose a non-linear electronic textbook format that organizes content in a way that accords with cognitive science theories of how concepts are represented in the mind. Structuring content in this way makes the relation between ideas more explicit, and reduces extraneous cognitive load, which we believe will aid learning. We implemented a functional prototype of this design, and we designed a study to measure it in terms of pedagogical efficacy (objective and self-reported), usability, and overall likeability, in comparison with a linear control. We found no significant differences between the two interfaces in objective or self-reported learning outcomes, or in usability. We found that a significantly greater number of participants preferred the non-linear interface to the linear interface.
Acknowledgements

This thesis would not have been completed if not for the generosity of Dr. Robert Biddle, and the companionship of Abdul, Mila, Reza, and everyone else at HotSoft. With deepest thanks and appreciation, this thesis is dedicated to you.

Thank you to my supervisors, Dr. Robert Biddle and Dr. Ali Arya, and to Dr. Milica Stojmenović for their guidance and support throughout the research process.

And thank you for everything to my Mom, Dad, and sister, Justine, to my cousin Teri, and to my friends Katie and Mat.
Chapter 1

Introduction

There is a new movement among educators in the fields of science, technology, engineering, and math (STEM) for a fundamental shift in the scope of content coverage from “broad” to “deep”. The dominant “broad” approach emphasizes the total number of domain concepts covered in an educational unit; the “deep” approach, by contrast, covers fewer total domain concepts, but focuses on how these concepts interact with each other, and with concepts from outside the domain [56, 16, 2]. This movement was initiated in response to middling [20, 46] STEM literacy scores by students in the United States.

Textbooks are important tools in the classroom, and will therefore play an important role in any efforts toward education reform [56]. Textbooks in their present form are not well suited for a deep coverage of material [63, 65, 4, 1]. We conjecture that the standard linear format typically used textbooks is partially to blame: a deep approach to teaching emphasizes the many dimensions of a concept, and the standard linear format is inherently multidimensional. We believe that a non-linear, multidimensional format is needed to meet the needs of a depth-based curriculum. In this thesis, we present a design, implementation, and study of such a textbook format.

1.1 Research Purpose

We designed and implemented a non-linear, multidimensional electronic textbook interface that attempts to represent textual content in a way that mimics how concepts are represented in the mind. We believe that presenting information in this way will help learners come to understand how ideas in a domain cohere with each other, and with ideas from other domains.

We designed and ran a study to test the value of our electronic textbook design. Our research questions were: 1. Is it possible to design and implement a non-linear electronic textbook that accords with cognitive science models for how information is
represented in the mind?; 2. Does our non-linear arrangement of information better support the learning process than the standard linear arrangement?; and 3. Do users prefer the proposed non-linear format to the standard linear format?

1.2 Thesis Structure

The structure of the remainder of this thesis is as follows.

Chapter 2: We review some literature relevant to our proposed electronic textbook interface. Key topics include education, embodied cognition, cognitive load theory, situation models and reading comprehension, and a brief review of existing electronic textbook formats available.

Chapter 3: We present the design of our e-textbook interface. Here we discuss key design concepts, leading to our proposed model for how content ought to be organized in a non-linear e-textbook.

Chapter 4: We present our implementation of the model—a high fidelity prototype—and summarize the design process that produced it. We describe the tools we used to implement the design, the data structure we use to store the content of the book, and provide a walkthrough of the interface from a user’s point of view.

Chapter 5: The results of a study by McNamara et al. [44] present a challenge to the model we present. In this chapter, we describe the implications of that study for our interface, and attempt to address the challenge.

Chapter 6: We present the study we designed to test our design. This section includes a description of our research motivations and questions, of the participants, of the materials we used, and of the procedure we asked participants to follow.

Chapter 7: We present the results of the study, which include quantitative and qualitative analyses of the objective and subjective measures we gathered.

Chapter 8: We discuss and attempt to explain the findings of the study.
Chapter 9: Finally, we present concluding remarks, including a summary of the thesis, and possible future directions in which this research could be taken.

1.3 Contributions

The main contributions of this thesis are the principled design and subsequent implementation of a new, non-linear interface for an electronic textbook, and the study and analysis of that design.

1.4 Disclaimer on the Use of “We”

Throughout this thesis, “we” is used to refer to the lead researcher, Eric Spero, under the advisement of his co-supervisors.
Chapter 2

Literature Review

2.1 Depth Over Breadth

When planning the content coverage of an educational unit, there is a trade-off between breadth and depth of coverage of the material. The breadth of a course refers to the total number of domain concepts it covers, whereas the depth of a course refers to the number of interactions between concepts it covers [61]. Depth and breadth are not mutually exclusive, rather they are two ends of a continuum.

The debate over just what is the optimal balance between breadth and depth is long-standing and contentious [61] [5] [34]. A number of prominent 20th century educators have argued for the benefits of a depth-based approach—for example, Bloom et al. [9], Piaget [49], and Vygotsky [70]. The main arguments in favour of a deep approach are: 1. some domain concepts are more important or fundamental than others, so it makes sense to spend more time on them; 2. understanding is context-sensitive, and so learning any concept should happen across a number of contexts; 3. learning a concept requires understanding how that concept interacts with other concepts; and 4. that patterns between data at the micro-level also fit larger theories—in other words, learning at the micro-level “scales up”.

The “deep” approach meets resistance from many educators, parents, and students [5]. The main arguments for maintaining the broad approach are: 1. many find it difficult to accept that covering fewer concepts could have positive effects on learning; 2. teachers are already trained in the methods of the broad approach, and may be under-equipped for the type of teaching required for the deep approach; and 3. it requires a major time investment on behalf of teachers to develop the new skills that the deep approach requires.
2.2 Project 2061 and the Next Generation Science Standards

Project 2061 is a research and development initiative aimed at improving literacy in science, mathematics, and technology through educational reform. The project was created in response to middling scores [14, 20, 46] in science, math, and technology literacy among students in the United States.

Project 2061 believe that the science curricula focus too much on breadth, and call for a “radical” reduction in the total number of concepts students are asked to learn [1]. They believe that real science literacy requires making connections between science, mathematics, and technology, and also with the arts, humanities, and the vocational subjects. They believe that knowing and doing should be learned together and in many different contexts so the concepts students learn can be used outside of the classroom.

Project 2061 regard the design of new and better instructional material such as textbooks as an instrumental component of educational reform needed to improve STEM literacy in the United States. Professional development (i.e. training instructors) is much more expensive in comparison [56].

Partially informed by the work of Project 2061, The Next Generation Science Standards [16] (NGSS) is a framework for new standards in K-12 science education in the United States. It suggests that science education should be built around three dimensions:

- Science and engineering practices;
- Crosscutting concepts that unify the study of science and engineering through their common application across fields; and
- Core ideas in four disciplinary areas: physical sciences; life sciences; earth and space sciences; and engineering, technology, and the applications of science.

We generalize these dimensions as:

- Knowledge of how work in the domain is done;
- Knowledge of how the domain concepts relate to concepts from other domains; and
• Knowledge of the core disciplinary concepts.

The NGSS challenges the designers of curriculum material to present ideas in a \textit{coherent} manner (i.e. ideas should build upon each other over lessons and units): each lessons should link to previous lessons. Coherence also requires that students understand how the new content they are learning relates to prerequisite ideas, or other ideas already present in long-term memory.

\subsection*{2.3 Issues with Textbooks}

In the late 1990s through the early 2000s, Project 2061 carried out a number of studies evaluating textbooks in the domains of math, biology, chemistry. They found that the overwhelming majority of textbooks surveyed were deficient in promoting a depth-based approach to learning. According to George Nelson, a former astronaut and director of Project 2061, “our students are lugging home heavy texts full of disconnected facts that neither educate nor motivate them” \cite{14}.

According to a 2000 report from the American Association for the Advancement of Science \cite{4}, standard textbooks are not well-suited for supporting a depth-based approach to teaching. They say that standard textbooks demonstrate an emphasis on covering a broad range of facts at the cost of an in-depth examination of the central concepts. Information in standard textbooks is presented piecemeal, failing to convey coherence among key ideas within and between scientific disciplines. When these texts include activities, they are often presented without any guidance for interpreting the results in terms of the concepts to be learned.

Shield and Dole \cite{63} evaluated five mathematics textbook series used in the Australian school system. They found the books paid insufficient attention to mathematical structure, and the interrelation between mathematical concepts and concepts from other domains.

McKeown et al. \cite{43} argue that textbooks do not provide enough background information or coherence. They conducted a study where students performed better with the text that provided background knowledge.
2.4 Situation Models and the Construction-Integration Model of Text Comprehension

The successful comprehension of text requires generating representations of knowledge at multiple levels. Readers create linguistic representations of the words and sentences of the text, and also of the state of affairs described by the text. The first of these representations is referred to as the text-base. The second is called the situation model, or the mental model. The situation model is a representation of the referential meaning of the text.

According to the Construction-Integration model of text comprehension, the situation model is created by integrating the text-base representation with prior knowledge. This explains why learners with high background knowledge have a comprehension advantage over learners with low background knowledge. According to McNamara et al., textbooks should “induce students to form appropriate situation models not just text-base representations good only for reproduction”.

A situation model is a multidimensional mental representation, incorporating at least temporal, spatial, causal, motivational, and person- and object-related information. Consider the following description of a situation model of a football game made by Johnson-Laird:

A dynamic model of, say, a football game calls for a temporal sequence of events at various locations, for causal relations between the events, and for the interaction of individuals, interacting physically and socially, governed by physical laws and constrained by the “laws” of the game and social conventions and motivated by various intentions.

A situation model is the integration of these multiple dimensions of meaning.

2.5 Cognitive Load Theory

Cognitive load theory describes the role that working memory limitations play in the learning process: learning is an information processing activity, humans are limited information processors, and so it is important to manage the information workload placed on students.
There are three types of cognitive workload: intrinsic, extraneous, and germane. Some workload is necessary and desirable, and some is not. What makes workload desirable is whether or not it contributes to schema acquisition and automation.

The first type of cognitive workload is intrinsic, which is workload inherent to the material being learned; it is the minimal workload associated with learning a given piece of information. The driving force behind this type of workload is element interactivity, which describes the total number of elements that must be considered at one time in order to understand some piece of information. Each element in a low-element interactivity piece of information can be understood without considering other elements; in high-element interactivity information, a number of elements must be simultaneously held in working memory. This type of cognitive load cannot be reduced without fundamentally altering the nature of the information to be learned.

The second type of cognitive load is extraneous workload. This is cognitive workload that is unnecessary and therefore detrimental to learning. An example of this type of workload is requiring that students search for some piece of information when it could just be provided for them. The crucial management of extraneous workload becomes especially important when element interactivity is high, as the learning task will test the capacity of working memory.

The third type of cognitive load is germane workload, which, like extraneous workload, is extrinsic to the information to be learned, but like intrinsic workload, it enhances learning. Germane workload is contributed by the instructor’s choices regarding how information is presented to learners, and the learning activities students are asked to partake in.

Good instructional texts should keep extraneous cognitive load to a minimum.

2.6 Bloom’s Taxonomy

Bloom’s taxonomy is a set of educational objectives that divides learning into three domains: cognitive, affective, and psychomotor. Cognitive learning is intellectual learning; cognitive learning objectives typically involve recalling or reproducing learning, or successfully using what was learned to complete some intellectual task. Cognitive learning has traditionally been the focus in education. Affective learning
emphasizes students’ feelings or attitudes. Affective learning emphasizes students’ objectives, interests, attitudes, emotions, opinions, attitudes, and values. A study by Rodriguez, Plax, and Kearney [52] suggests there is a strong relationship between affective learning and motivation, and that affective learning influences cognitive learning. Psychomotor learning objectives emphasize applications of muscular motor skills such as speed, dexterity, and the ability to manipulate objects or instruments [57].

2.7 Embodied/Grounded Cognition

2.7.1 Meaning and use

Philosophical pragmatists hold that, phenomenologically speaking, the meaning of objects and categories is “situation-dependent and subjectively determined” [48]. Objects get their meaning from their relation to a goal. For example, a chair is canonically viewed as a tool used for seating, but other situations it can have a different meaning: it can be a tool used for standing-on, or an obstacle that impedes motion. This idea underlies a number of important intellectual works of the last century, including Gibson’s Affordance Theory [28]. Gibson argues that the act of perceiving leads one toward a course of action: when we perceive a door handle, we do not perceive the object in-itself, but rather the ways in which we can interact with it.

Constructivism is ‘meaning is use’ applied to pedagogy. It is a popular and influential alternative theory of learning [17]. Constructivists assert that knowledge is not passively received, but built by learners as a result of experience in situated action. Constructivists believe that formal education should organize itself around this idea. Rorty asserted that “knowledge is not a matter of getting it right, but rather acquiring habits of action for coping with reality” [53]. Dewey thought instead of learning vocations, we should learn domain knowledge (science, math, literature, etc.) through vocations [22]. Bruner felt that educational content must be personally and societaly relevant to keep the learner motivated throughout the learning process [12].
2.7.2 The main claim

The notion that ‘meaning is use’ is a main claim held by proponents of embodied cognition [74]. Embodied cognition is essentially the belief that that the mind is shaped by the body. It is an alternative view to the information processing view of mind, which regards the mind an abstract symbol manipulator which relates to the environment only arbitrarily. Proponents of embodied cognition hold that cognition is for action: the ultimate function of the mind is to guide the body through successful interactions with its environment, and so its various faculties can only be understood by making reference to their relation to situated action.

2.7.3 Abstract vs. concrete concepts

Traditionally, there is thought to be a sharp distinction between concrete concepts and abstract concepts. Concrete concepts, such as bird, car, and sofa, have physical, sensory referents and are easy to visualize mentally [60]; abstract concepts, such as truth, freedom, and invention, have no physical referents. This view is supported by so-called concreteness effects, which are well-established cognitive processing advantages for concrete concepts over abstract concepts. Concrete concepts have an advantage in speed of lexical access and word comprehension, and concrete concepts are linked to stronger memories [8].

However, Schwanenflugel, Shoben, and colleagues [60, 59, 71, 58] showed that this processing advantage for concrete concepts disappears when people are provided with an instantiating situation for abstract concepts.

"...participants might first read about a court trial before studying “truth” for a memory test, or first read about a living room before studying “chair.” When relevant situations were present, abstract words were processed as well as concrete words. Participants accessed and understood both types of words equally quickly, and remembered them just as well.” [8]

An alternative view to the strict abstract/concrete division is that abstract and concrete concepts are not fundamentally different [8]. What distinguishes abstract concepts from concrete concepts is that abstract concepts apply to a much wider range of instances.
Providing concrete examples of abstract concepts is an effective teaching tool, especially for beginners [6]. Examples from real life allow students to anchor new and unusual information in terms of something that they already know. Showing elaborate, realistic examples leads to better knowledge retention [50].

2.7.4 Thinking with the environment

We often make changes to the environment to simplify some mental task [39, 38]. This is one effective way of coping with the fact that we are limited cognitive processors in a highly complex world. Some examples are mathematical notation, written language, and note taking. These actions are beneficial because they alleviate working memory load by off-loading the representation from one’s mind to the environment; it frees up resources which can be allocated to some other aspect of the problem.

Cognitive tools [33] are instruments (including computers) that leverage the representational capacities of the environment to help the learning process.

2.8 Psychological Essentialism

For at least some categories of objects, there are some especially deeply embedded properties that category members must have [27], termed “essences” in the literature. We think of these objects as having a constitutive nature, where its observable features are the result of invisible—and sometimes even unknown—features. An essence is equivalent to an object’s identity.

It is relatively uncontroversial that we think of animals of having essences [23]. When it comes to artifacts, however, we are a little less absolute about category judgments for atypical members (see Bloom [10] for a review). However, researchers such as Bloom [10] have argued that artifacts do indeed have an essence. This would explain why, for example, an original work by a famous artist can sell for an extremely high sum of money, while an exact replica produced by another artist would not. Artifacts of a kind can differ drastically in terms of their physical and functional properties, and yet still maintaining their category membership. For example, a broken chair, which cannot be sat on, is still a chair, a category of objects that seemingly exist to be sat on; some objects, like ottomans, can be sat on, and yet are not chairs; beanbag chairs are chairs, even though they do not have discrete feet, seats,
or backrests. According to Bloom, the defining feature of these objects, and other artifacts, has to do with the intention of their creator: beanbag chairs and broken chairs are created with the intention that they are members of the category ‘chair’. For Bloom, the categorization of artifacts is an abductive process where we observe the object’s current appearance and potential uses, then construct an explanatory hypothesis about the type of object it was created with the intention to be.

There are different classes of essences pertaining to different classes of basic level concepts. The essences of natural objects are equivalent to their fundamental physical structure: DNA in the case of organisms, and molecular structure in the case of inanimate natural substances. For artifacts, essences reside out the object they constitute.

2.9 Knowledge is Categorical

Concepts are a unit of knowledge about the world. Through concepts we divide the complex, continuous world into simple, finite categories, thereby lightening the load of perception [55, 8]. Concepts form the basis for thought and communication [32].

A fundamental property of knowledge is its categorical nature [7]. A conceptual system is not a collection of holistic images like a camera. Rather, it is a collection of category knowledge, where each represented category corresponds to a component of experience, not to an entire holistic experience.

2.10 Knowledge is Hierarchical

There is evidence that concepts are subdivided hierarchically in two ways: taxonomically, and partonomically (or merologically) [68]. A taxonomy organizes things by kind: a pippin is a kind of apple, and an apple is a kind of fruit, but the reverse is not true. A partonomy organizes things by part: a piston is a part of an engine, and also a part of a car, but the reverse is not true. This is true of abstract concepts as well, such as governments: a democracy is a kind of government, and a government consists of the legislative, judicial, and executive branches.

Like objects, events can be organized into hierarchies as well [75, 3]. For example,
getting dressed for work and getting dressed for school are instances of the more general event “getting dressed”. The lower level events are unified by common properties inherited from the higher level parent node, and they have unique properties which distinguish them from each other.

2.11 The State of Electronic Textbooks

Although textbook material presented in an electronic format is widely available through retailers like Apple, Amazon, and Google, electronic textbooks have not enjoyed the same success as ebooks [18]. The typical electronic textbook is structurally very similar to print texts with some additional features added, such as the ability to search the text, and multimedia content. Studies comparing these kinds of electronic texts and standard print texts typically show no differences in learning outcomes, but a preference for using standard print texts [62, 67, 51, 41]. Four out of five students prefer print texts to digitized texts in spite of the fact that print is more expensive and less portable, because they are more familiar, and they better afford highlighting, dog-earring, and annotation [51]. A study by Daniel and Woody [18] showed that students using electronic textbooks spent significantly more time reading. Studies have shown that reading course materials on an LCD screen can induce eye strain. In general, it seems that students feel that the drawbacks of electronic textbooks relative to print texts outweigh the benefits. Authors like Toukonen [67] have argued the current incarnation of electronic textbooks do not properly leverage the advantages of digital media technology such as dynamicity and non-linearity.

More recently, in response to the issues with this first type of electronic textbook, efforts have been made to create ebooks that leverage the interactive nature of computers. Interactive texts contain less text, and add animations of key concepts, more questions, and interactive tools [24]. A new system proposed by Miller and Ranum [45] incorporates video, code editing, execution, and visualization inside the textbook. This provided students the benefit of being able to do all learning activities inside one environment (i.e. they do not need to install Python, a text editor, and so on). zyBooks [25], another example of an interactive textbook, have gained a significant degree of popularity in recent years [77]. zyBooks present text interleaved with interactive examples of the learning material.
While we are certainly intrigued by interactive textbooks and their potential benefits, interactivity in electronic texts is outside the scope of this research project. The model we eventually propose is able to include the key features of interactive texts, and future implementations of our design may do so.

The digital medium is particularly well-suited for supporting hypertext navigation, which represents one potentially significant advantage over print texts as an educational tool. A study by Farag and Shemy [26] showed that some readers perform better using the non-linear navigation style. A non-linear organization of textual material hands a significant measure of control of sequence and pace over to the students [42, 29], which increases motivation [36].

DeStefano and LeFevre [21] review studies on hypertext reading, and discuss the advantages and disadvantages relative to linear texts. Hypertext is more demanding of working memory resources than linear texts because it requires that readers make a decision before selecting a link. The researchers do not specify if this is intrinsic, germane, or extraneous load. They note that in many cases where hypertexts appear to impair learning, it is because these hypertexts exceed the working memory capacity of the reader. In general, structuring the text hierarchically seemed to aid learning, whereas structuring it in some other way (e.g. allowing readers to navigate a semantic network of concepts) sometimes impaired learning. Many hypertexts used visual overviews of the content to reduce the cognitive load burden placed on readers. Visual overviews were helpful only when their structure matched matched the inherent structure of the domain. They note that, while this was not a focus of their research, there may be significant affective benefits to hypertext compared to linear texts. In general, structuring the text hierarchically seemed to aid learning, whereas structuring it in some other way (e.g. allowing readers to navigate a semantic network of concepts) seemed to impair learning.

2.11.1 A similar work: ELM-ART

ELM-ART was a web-based intelligent educational system that sought to combine adaptive hypermedia with intelligent tutoring [72]. The textbook was traditional in that it contained explanations of key concepts, program examples, and problems to solve. It added interactivity by including Q&As, the ability to run software with
different parameters, and solve problems while receiving system feedback. It added intelligence by using AI technologies to support readers – intelligent program analysis, concept-based hyperspace organization with a direct link between page content and domain concepts, intelligent links between examples and problems, and adaptive navigation support (guide readers toward relevant content).

ELM-ART was designed only for programming courses. In 1996 Brusilovsky, Schwarz, and Weber [13] described a generalized version of ELM-ART which was not tied to any particular domain or subject. This system sought a more explicit, direct representation of the concepts students were asked to learn. They felt that the structure of the learning material should match the pedagogical structure of the domain knowledge. This system was built around a domain model and a user model. The domain model consisted of a network of domain concepts. The structure of the domain space could be explored via hyperlinks (hypespace was equal to domain space). The domain structure could also be viewed at a high level with a glossary which took the form of a concept map. The user model tracks the user’s knowledge of the domain knowledge.
The text is divided into two halves: the textbook and a glossary. The textbook window has subsections, or frames. A text window shows the currently active section. For terminal sections, the window shows the title of the section, and the section text. For high-level sections, the window shows a title, a section preface, and a full table of content for the section (list of hierarchically structured titles of its subsections down to the terminal level). To the right of the text section is the concept bar. The concept bar shows concepts related with the current section. In the top-right, there is a nav center/toolbox: the nav center shows position of the current section in the text (predecessors and brothers are provided) providing orientation and navigation functionality; a toolbox provides a set of buttons which can call additional windows (content window, search window, prereq-based help window).

2.12 Summary

There is a movement afoot to switch from a focus on a broad coverage of domain concepts to a deep coverage. Redesigning textbooks will play an important role in this change.

According to the NGSS, education units in STEM domains should teach core disciplinary concepts, show how work in this domain is done, and show how these concepts can be useful in other domains. These latter two requirements seem to tacitly acknowledge that meaning is use. At the very least, they suggest that use is an important component of meaning. It is also important that instructional materials are coherent.

Real understanding requires generating representations of the text that go beyond the text: learners must integrate the content they are reading with their wider world knowledge, and develop a situation model of the text. A curriculum that focuses on depth of coverage in the manner proposed by the NGSS should result in students developing better situation models of the course content.

Effective teaching and learning requires minimizing extraneous cognitive load, which is a type of cognitive load that is unnecessary and therefore detrimental to learning. A “deep” approach to teaching emphasizes the interactions between concepts, which means it has high element interactivity, and therefore higher intrinsic workload than the “broad” approach. Educators and the designers of instructional
material must therefore take extra care to ensure that extraneous workload is kept to a minimum.

Providing people with grounding situations for abstract concepts has been shown to reduce the processing disadvantage normally associated with abstract concepts. Perhaps for the same reason, examples have been shown to be good teaching tools.

We have a tendency to regard invisible properties residing outside of the entity they constitute as an important component of that concept’s meaning because they play a constitutive role in its observable features.

Electronic textbooks ought to be great teaching tools because we can use the environment to help us think, and computers are essentially highly flexible pieces of the environment. The first generation of electronic textbooks do not sufficiently leverage the benefits of digital media. We think that these ebooks do not offer students enough in exchange for giving up the format they are most familiar with which has features that are not offerable in a digital format. We think that interactive textbooks are a step in the right direction as they offer more unique benefits compared to standard print texts. However, these are still linearly organized, and therefore miss out on the potential benefits of a non-linear organization of material.

Knowledge is complex: it is categorical, hierarchical, and multidimensional. If the designers of instructional material can store some of this information relating the various elements of knowledge in the environment, we think it would help make learning new ideas more manageable.

The model we present in this thesis is an attempt at a new way of structuring information in an electronic textbook. In the next chapter, we describe this design, beginning with the key design concepts which were informed by our review of the literature.

Interactivity is not relevant to our research questions, and as such we make no attempts to integrate the defining features of interactive electronic texts (i.e. animations and interactive tools). Our design is fully compatible with interactive texts, however, and future iterations of our implementation could (and perhaps should) include interactive features.
Chapter 3

Design

Based on our review of the literature, we propose a model for an electronic textbook interface that aims to better support a depth-based approach to teaching than the standard linear textbook format.

3.1 Key Design Concepts

Before we present the model, we discuss a few key concepts that inform its design.

3.1.1 Structural harmony between content and concepts

Our model seeks to harmonize the structure of the text with the conceptual structure of its content. We think that unity between the two structures will reduce extraneous workload in at least two ways. The first is by offloading mental work onto the environment. Concepts are complex—they are multidimensional, hierarchical, and categorical—and the interrelations between domain concepts and other concepts are brought to the forefront in the “deep” approach. If the information regarding how concepts relate to each other can be stored in the physical environment, it will free up cognitive resources which can be spent on some other aspect of the learning task at hand, thereby facilitating the learning process.

The second way this design concept could reduce extraneous workload is by minimizing thought about structure that is not also thought about the content. To navigate instructional materials, students must consider its structure. Wherever the structure of the instructional material matches the structure of the content, any effort exerted thinking about the structure of the instructional material is also effort spent learning the content. Conversely, wherever there is disharmony between the two structures, the student will be exerting effort that does not help them learn the content. For this reason, we think that the designers of instructional materials should strive for structural harmony between texts and content.
This design concept is realized in our model in two ways: 1. We divide content into two major categories: core and peripheral; and 2. we allow for the non-linear navigation of content.

### 3.1.2 Dividing content into two types: core and peripheral

We identify two broad categories of content in instructional materials. The first type of content involves domain concepts. Examples of domain concepts from psychology include Operant Conditioning, Behaviourism, Working Memory, and all of their sub-components. We call these concepts core concepts because we believe they are the focus of a learning unit, and that learning these concepts is the primary goal of an educational unit. Domain concepts have a privileged status above all other types of concepts covered in the classroom, and for this reason we believe that core content should be provided a privileged status in the format of instructional materials as well.

The second type of content present in instructional materials is content that enriches the core content. These are ideas that relate the abstract core concept to the real world. Examples of these kinds of content include real-world instances of the concept, the history of the concept, and ways in which the concept can be usefully applied in the world. We identify three categories of enriching content: antecedents, postcedents, and instances.

**Antecedents (inputs):** While each student encounters a domain concept at a particular instant in time, a concept is something that stretches across time, and its past can tell us something about its present. Antecedents are the category of concepts that come prior to (chronologically or causally) the core concept. Examples of antecedent content in scientific disciplines include descriptions of key research, and paradigmatic assumptions underlying the concept.

**Postcedents (outputs):** The second category of enriching content are the ways in which the concept can be usefully applied in the real world. Since applying the concept comes after mastering it, from the reader’s point of view this category of concepts mirror antecedents; we therefore call these postcedents. If a concept’s use is an important element of its meaning (if not meaning itself), this category of concepts will be of great interest to learners. And if the primary goal of
education is to foster the development of skills that are socially useful—which we hold—then this category of concepts is essential.

**Instances (real-world examples):** The concepts students are asked to learn are all abstractions of instances. There is evidence that abstract concepts are easier to understand when they are linked to a grounding situation. An example of an instance of the abstract concept **truth** is a courtroom scenario.

The three types of enriching content provide three different types of grounding for each abstract core concept. We think that linking abstract ideas with ideas relating this knowledge to the real world will help readers develop better situation models than they might without this information.

We consider the proposition that these three types of concepts are important for learning to be uncontroversial, as these concepts already feature prominently in textbooks. For example, in psychology textbooks, a domain concept is typically accompanied by real-world scenarios instantiating the concept; key studies providing empirical substantiation for the concept are often summarized; and information regarding how the concept can be applied, either professionally or not, are often provided. Our proposed design differs from standard texts in its treatment of these enriching concepts by giving them an elevated status in the text’s form.

We believe that this elevated status would result in these concepts being present more consistently than they are now: ideally, for each concept students are asked to learn, they will quickly be able to see three different ways in which it relates back to the material world.

### 3.1.3 Non-linear organization/navigation of content

According to our literature review, concepts are multidimensional, and concepts relate to other concepts hierarchically. That is, concepts have a non-linear structure. However, learning materials are typically presented linearly. Hierarchy is signalled through sentences such as “$x$ is a kind of $y$”), and typographically through headers (e.g. chapter, section, subsection). When the dimensions of are presented, they are presented in-line with the core domain content.

How should the many (interrelated) concepts covered in an educational unit be
arranged relative to one other? The concepts of *hierarchy* and *multidimensionality* are fundamentally at-odds with linearity. Strict linearity forces a single ordering, but it seems to us that there is no single optimal way to arrange this content. What will be optimal for one student will be sub-optimal for others; and what will be optimal for one student at one time under one set of circumstances may for that same student be sub-optimal at some other time under other circumstances. This, we feel, is a fundamental problem for standard texts.

We conjecture that, rather than organizing text linearly and signalling non-linear structure with language and typography, it would be better to organize learning material non-linearly in a way that mimics how it is represented in the mind. A non-linear organization will allow students to choose the order that suits them at the time of reading. Computers make this kind of organization plausible.

We believe we avoid the issues faced by the more problematic hypertexts covered in the review by DeStefano and LeFevre [21] because 1. we structure the material according to the domain hierarchy; 2. we take care to limit navigational freedom so that working memory capacity is not exceeded; and 3. we provide a hierarchical visual overview of the content.

### 3.2 The Model

We propose the following model for an electronic textbook interface, incorporating the key design concepts just discussed.

We elect to position enriching content *around* the core content. From the student’s perspective of time, the antecedents, core content, and postcedents comprise a timeline of past, present, and future. English speakers tend to think of time as passing from left to right, so we propose positioning antecedent content to the *left* of the core, and postcedent content to the *right* of the core. An instance is a taxonomical category, and we have a tendency to think of taxonomies as *vertical* in nature [54], with more abstract elements placed above less abstract elements. For this reason, we position instances *below* the core. Figure 3.1 shows a schematic of the structure of a concept in our model.

There is nothing in principle preventing making core and peripheral status contingent. Peripheral elements are concepts in their own right, with their own antecedents,
postcedents, instances, and hierarchical/partonomical structure. We opted for making core and peripheral status fixed for a number of practical reasons, which are discussed in Section 4.1.2.

Nodes with the structure in Figure 3.1 are the atomic units of the model. Each node corresponds to a single domain concept. Nodes are organized into a tree (i.e. a hierarchy), which acts as a concept map of the domain content students are asked to learn. Figure 3.2 shows a schematic of a hierarchy of nodes. Students can select a node in the tree to navigate to that node.
3.3 Interface Design Process

3.3.1 Inception

Our design process was initiated in response to observations we made in day to day life. The first observation came from listening to other students talk about their experience as students, we noted a recurring theme of students expressing exasperation in the face of a lack of understanding of the practical utility of the information they were being asked to learn. Examples of these types of remarks are “Why am I being asked to learn this?” and “How would this ever be useful to me?”.

The other observation pertained to the structure of learning materials compared to the structure of the underlying concepts. We felt that course material was under-lyingly non-linear, and the linear presentation sometimes made the task of reading the material difficult. We set out to design a non-linear textbook that made informa-tion about how the domain content relates to the material world in a consistent and therefore predictable fashion.

We identified the need for a distinction between core and peripheral early on. A significant challenge was in deciding what were the ‘important’ peripheral categories, and in what manner we should arrange them relative to the core content, and to each other.

3.3.2 Early sketches

The source material for the first attempts at this design came from an introductory programming course. A sketch from this period of development is shown in Figure 3.3. The focus for this course was programming concepts. The course also featured a number of other kinds of concepts pertaining to things relevant to the wider context of computer programming. These include: for each programming concept, it was shown how it was implemented in at least one programming language, and sometimes in multiple languages; information about what is happening in the computer hardware when the code is run; information regarding how the programming concepts were beneficial to either programmers or end users. Figure 3.3 is an attempt to arrange these types of concepts in an optimal way. We tried to arrange the material in such
a way that peripheral information surrounded a core, and that users navigate a flow-chart of the flow of information in the development of software development process. In a design like this, each course would map content according to its own pattern representing the process of doing work in that domain.

The complexity of this design raised a number of potential issues, including: (i) it seemed that it might be difficult to implement; (ii) it seemed like it might be difficult for students to navigate. While we felt it was important for the spatial placement of the different concepts to convey semantic information to the reader, it was not clear to us exactly what value this would have for learners. In future designs, we opted

Figure 3.3: Early attempt at model design
for a simpler arrangement of core and peripheral elements in order to minimize the chances of developing something unusable.

Not long after Figure 3.3 was drawn, after considering many possible categories of peripheral content, and configurations of core and peripheral content, we decided to group nearly all peripheral content into two categories: what comes before the concept, or what we termed the concept’s causal history, and what comes after the concept, or what we termed the causal future of the concepts (i.e. applications of the concept). We placed the causal history elements to the left of the core concept, and the causal future elements to the right of the core concept. Together, the three elements make a timeline of past, present, future.

Around the same time, we felt it prudent to provide taxonomical information related to the core concept: that is, superordinates, and subordinates all the way down to concrete instances. Students could move ‘up’ to see higher-level content, and ‘down’ to see lower-level content.

An attempt to map the content from an introductory psychology textbook according to these design changes is shown in Figure 3.4. Concepts at each level of abstraction have two types of causal history: relevant earlier theories and paradigms, superordinates, and subordinates.

---

1We changed the domain of the content of our implementation from computer programming to psychology because the lead researcher had a stronger background in the latter domain.
and key research and/or studies that helped give rise to the idea. Each concept also has two types of causal future: real-world applications of the theory, and research applications of the theory (i.e. ways in which the concept interacts with other psychological concepts; e.g. learning relates to memory).

Interface sketches made according to the map shown in Figure 3.4 is shown in Figure 3.6. The distinction between internal structure and external structure reflects our desire at this stage to provide information regarding causal history, causal future, and instances both for the concept construed as a psychological theory, and for the
concept construed as a description of human behaviour. For example, the causal history of “Memory” in the first construal might include key studies related to memory, and the causal history in the second construal might include a description of the parts of the brain involved in memory. This distinction was abandoned in future iterations for the same (above-mentioned) reasons we abandoned the design shown in Figure 3.3. These sketches also reflect our desire at this time to make a strict distinction between the partonomy of a concept (the part-whole structure of a concept) and the taxonomy. The partonomical structure would be traversed by selecting a part, which would show the user a new screen looking identical to the one they were just on, except focused on that part.

![Early attempt at interface design: template (left) and more specific example (right)](image)

**Figure 3.6:** Early attempt at interface design: template (left) and more specific example (right)

### 3.3.3 High fidelity prototype

For the next stage of the development process, we implemented a high fidelity prototype of the design using Pencil Evolus, a GUI prototyping program. A screenshot of this prototype is shown in Figure 3.7. The taxonomical structure is simpler here than
in the previous design: only a single superordinate level and a single subordinate level (namely, *instances* of the concept) are provided. Parts (shown above the content) are grouped around a common a superordinate concept. Parts are navigated in the same way as the previous design: selecting a part makes that part the new concept of focus, and the content, peripheral elements, and parts are updated accordingly.

Figure 3.7: High fidelity prototype
Chapter 4

Implementation

4.1 Towards a High Fidelity Functional Prototype Using JavaScript

The final stage of the design process coincided with the first stage of implementation. The very first iteration of what would eventually become the version of the interface used in testing is shown in Figure 4.1. This version showed text in the center area, and featured actionable buttons on the periphery.
In the next iteration, shown in Figure 4.2, buttons on the periphery (the blue-green rectangular areas on the edge of the page) “slid” in a panel showing content related to the button’s label (for an example in the ‘final’ version of the interface, see: the rightmost image in Figure 4.8).

4.1.1 Navigation panel changes

The functionality of the contents of the navigation panel (in Figure 4.2) changed slightly over the course of the design and implementation of the interface. At the stage depicted in Figure 4.2, the bottom-most row of buttons in this group (coloured orange in Figure 4.2) brought the user “down” a level in the domain hierarchy—something which did not change in later versions. The top and middle rows, however, behaved differently than they would in later iterations.
The middle group of buttons (blue-green in colour in Figure 4.2) took the user to associated content, which are concepts related to the current node in some way, but not an immediate ascendant or descendant. The support for linking associated content was abandoned soon after this screenshot was taken due to time constraints—while we feel this feature would be key in a full implementation of this design, it introduced usability and implementation issues, and since we did not believe it was a core element of the interface, it was not urgent that it was a part of our testing.

The top button (coloured blue in Figure 4.2) slid in content about the superordinate concept of the current concept (e.g. for the Memory node, this button would be in information about faculties of mind, because memory is a faculty of mind). At this point the system had no functional navigation system, and we were lacking a way to move back up the hierarchy after moving down by selecting an orange button. We realized that changing the functionality of the top button to move readers “up” a level in the tree would make the hierarchy navigable without losing anything, as information about the superordinate concept of the current node is still communicated to the reader.

In the next iteration, clicking the top button moved the reader “up”, clicking the bottom buttons moved the reader “down”, and clicking the middle group of buttons moved the reader “laterally” to a sibling node. The navigation panel with this functionality is shown in the top image in Figure 4.4.

4.1.2 Limiting the degrees of navigational freedom

Up until the point of implementation, we had plans to include support for associative links between concepts, and also links from peripheral elements to other concepts. We decided to abandon both of these plans for similar reasons: implementing each ‘correctly’—that is, in a way that did not harm the usability of the interface—would be time-consuming, and perhaps not possible. Since they were not core interface features, we left these features out. We briefly discuss what we had in mind for each feature in turn.
Associations

Concepts can be related hierarchically, but also associatively [19, 64, 35]. Allowing for the ability to easily view concepts associated with the concept a student is studying would be helpful for learning. It is also a functionality that is especially well-suited to an electronic non-linear system: in a standard linear book this functionality could be achieved by including associated content in an appendix, but making the student hunt for page numbers in not an efficient management of their cognitive load.

Links in peripheral content

Originally, the ‘peripheral’ category was less fixed. We wanted to link concepts together in input-output relations, but antecedent and postcedent concepts could also be cores, with their own antecedents and postcedents. Moreover, peripheral elements could relate to other elements hierarchically as well: users could move “up” and “down” to taxonomically higher and lower elements. The concept map in 3.4 shows one attempt to hierarchically link peripheral elements in a sensible way.

Apart from the reasons stated above, we abandoned this design feature because it weakened the distinction between core and peripheral concepts, which we felt was important. On top of this, the task of linking all domain concepts in a three-dimensional web seemed far beyond any reasonable scope for the project.

4.1.3 Hard distinction between partonomy and taxonomy

Early on, we planned on making a hard distinction between partonomy and taxonomy in the structure of our design. For each domain concept, we would provide relevant taxonomical information by linking its superordinate concept “above” it, and instances “below” it. We would provide partonomical information by allowing the user to “drill down” into the simpler and simpler sub prats of the concept. A schematic of this design is shown in Figure 3.5.

In later designs, we partially conflate partonomy and taxonomy: in our system, as one moves “down” the hierarchy, the relationship alternates between subordinate (i.e. taxonomical descendant) and part (i.e. partonomical descendant). For example, the descendant of Memory is Sub-processes, and the descendants of Sub-processes are
Encoding, Storage, and Retrieval.

We made the decision to abandon the hard distinction between partonomy and taxonomy for a few reasons. For one, educational units do not teach concepts at a single level of abstraction; core concepts can vary a great deal in their level of abstraction. We would therefore need to let the users move “up” multiple levels, which comes with the same problems mentioned in the “Links in Peripheral Content” section above. Second, it was not clear to us that the added complexity that making a hard distinction would give to the UI would provide pedagogical rewards to make the trade-off worthwhile.

4.1.4 Visualizing the text structure using D3.js

We conducted preliminary round pilot testing of the version of the interface shown in Figure 4.2 and found that users had a difficult time imagining their location in the content structure. “I’m lost” was a frequently-made comment. To help users visualize the structure of the text, we began looking at visualization tools in JavaScript. We made modifications to an example of a collapsible tree diagram using D3.js, populated this structure with our text structure, and made it available by clicking a button in the top peripheral panel. The first version of the treemap is shown in Figure 4.3. When users clicked a node, the panel was retracted, and the main screen was redrawn to show content corresponding to the clicked node.

4.1.5 Improved navigation panel using D3.js

We found the navigation panel pictured in the top image in Figure 4.4 insufficient for a number of reasons. First, it did not clearly convey the structural relations between the various elements. Second, the buttons had the same style as the buttons that brought in slider elements, which we thought might be confusing to users.

After implementing the treemap using D3.js, it occurred to us that if we could show a miniaturized version of the tree in the top area of the main page it would solve the above-mentioned issues with the navigation panel. The consistency between both navigation tools would bring another welcome added benefit.

We implemented the improved navigation panel, which we call the minimap, by
making heavy modifications to the treemap. The minimap shows only three hierarchical levels: the level above the current node, below the current node, and the current node’s siblings. When the user clicks any node in the minimap other than the current node, the tree updates to reflect the shift in context created by the click. When the user clicks the node above the current node, the tree appears to move downward, and a new superordinate node slides into view. When the user clicks a node in the bottom row, the tree appears to move upward, and new child nodes slide up into view. This animation communicates to the user that they are moving “up” and “down” a tree.

4.2 Second Round of Pilot Testing

We concluded the interface development and implementation pilot with a round of pilot testing featuring 7 participants. During these tests, we made a number of
changes to the study methodology, and to the interface. In regards to the study methodology, we found that the post-reading questions were far too difficult, and we eventually switched to a multiple choice format. We also made minor reductions to the amount of content featured in the textbook. In regards to the interface, we uncovered a number of various minor glitches which were subsequently fixed.

4.3 ‘Final’ Version

The design process culminated in the ‘final’ version (i.e. the one used in testing), which is shown in Figures 4.5, 4.6, and 4.8 below. The interface contains content sourced from the Memory chapter in Psychology: Themes and Variations (4th Canadian Edition) by Weiten and McCann [73]—an introductory psychology textbook used in first year Psychology courses at Carleton University.

4.3.1 Main screen

The main screen (Figure 4.5) is divided into a central area encompassing the minimap and the core text area, and a peripheral area containing the four kinds of slider tabs. The minimap is located in the top of the central area. It is a ‘local’ snapshot of the complete concept tree, showing only the nodes directly above, below, and adjacent to
the current node. It is meant to help remind users about their location in the tree, and to provide them with an easy way of moving through the tree.

The current node is indicated with red colouring; nodes with sub-nodes (or “children”) underneath are indicated with blue colouring; visited nodes are indicated with light grey colouring; non-visited nodes with no sub-nodes are white. Labels are truncated after a certain length; on hover, the full label is shown.

### 4.3.2 Treemap

A full map of the text’s content can be seen by clicking the top tab, labelled “Map” (Figure 4.6). We refer to this map as the treemap. It looks and functions the same as
the minimap: the current node is coloured red, and visited nodes are coloured grey; users can visit nodes by clicking on them; truncated labels are shown in full on hover.

![Figure 4.6: Final version: Treemap view](image)

**Search**

Users can enter text in the text box in the top left corner of the treemap to perform a regular expression search through the data structure. If the search string matches the title of a node, the content of a node, or the content or titles of any of its peripheral elements, a link to the node or peripheral element is presented in a window beneath the text box. Users can visit the linked content by clicking the link. A screenshot of the search functionality is shown in Figure 4.7.
4.3.3 Peripheral content

The buttons on the left, right, and bottom periphery show peripheral content (anteecedents, postcedents, and instances, respectively). Figure 4.8 shows the peripheral content view.

Sliders bearing peripheral content can have sub-content, which is accessible by clicking sub-tabs. Two sub-tabs are visible in 4.8.

4.4 Tools

We implemented the textbook in HTML, CSS, and JavaScript (with jQuery). The trees in the treemap and minimap were based on an example of a collapsible tree diagram written in D3.js\footnote{https://bl.ocks.org/d3noob/43a860bc0024792f8803bba8ca0d5ecd}.

The main JavaScript file contained 1515 lines, of which approximately 60% was dedicated to the treemap and minimap.

4.5 Data Structure

The content for the electronic text was generated from a file system directory consisting of sub-directories and text files. A schematic of this structure is shown in Figure

Figure 4.7: Search functionality
4.9

The pre, post, and inst directories hold content for ANTECEDENTS, POSTCEDENTS, and INSTANCES, respectively.

The chunks of text shown on screen are stored in plain text files named “content”. Core domain content is stored in a content file directly under a CONCEPT directory. Peripheral content is stored in a content file under pre, post, or inst directories. The title of each page is read from the directory name.

When a parts directory is under a CONCEPT directory, this directory contains SUB-CONCEPTS (which featured the same structure as CONCEPTS). When a parts directory is under a pre, post, or inst directory, this directory contains content that would be accessible via a button click on a slider (for an example, see: rightmost
Each time the textbook was loaded with a new browser session, an AJAX call is made to a PHP process which recursively traverses the directory containing the data structure, and generates a JSON string mirroring its structure. This string is returned to the JavaScript file and stored in Session Storage (for local reuse), and converted to a JavaScript object for use in the running instance of the textbook.

4.6 Walkthrough

The following describes a simulated scenario wherein a reader browses through a few pages of the non-linear electronic textbook.

The student opens the book for the first time in a few days, and is not quite sure where they left off last time. They open the treemap by clicking the “Map” button (Figure 4.10). By looking at the treemap, they are reminded of the overall structure of the course content. They see that the “Encoding” and “Retrieval” nodes...
are coloured grey, which means they have already been visited. This leaves only one unvisited memory sub-process: “Storage”, which they decide to visit next (Figure 4.11). The reader also takes note that two of the three sub-nodes under “Storage” contain many sub-nodes themselves, whereas one sub-node, “Sensory memory” has no further sub-nodes.

After reading the main content for “Storage”, the reader wishes to see what paradigmatic assumptions underlies psychological models for storage, so they click on the “Paradigms” tab in the left peripheral area (Figure 4.12), which slides in a pane containing “Paradigms” content. There are two sub-tabs of “Paradigms” content. After reading the first, titled “Antiquity”, the reader selects “Information Processing” (Figure 4.13).

After reading the text for “Information Processing”, the reader clicks inside the peripheral border (Figure 4.14) to clear the slider and return to the main view for “Storage”. Looking at the colour of the sub-nodes under “Storage”, the reader is reminded that the “Sensory memory” node contains no descendant nodes beneath it (because the node is coloured white, and not blue). They decide to tackle this ‘lighter’ node before taking on the other ‘heavier’ nodes with many sub-nodes, so they click “Sensory memory” (Figure 4.15).

The reader reads the main content for “Sensory memory”. They understand the words they are reading, but they are having a difficult time visualizing sensory memory in action, so they look for examples by clicking the green tab in the bottom peripheral area titled “Example: Sparkler” (Figure 4.16).

4.7 Answer to Research Question 1

Our first research question was: Is it possible to design and implement a non-linear electronic textbook that adheres to cognitive science models for how information is represented in the mind?. At this point, we believe we found an affirmative answer.
Memory

Memory is the faculty of mind by which information is encoded, stored, and subsequently retrieved. A key function of memory is the retention of information over time for the purpose of influencing future actions. If we could not remember past events, we could not learn to develop language, relationships, or personal identity.

While memory is understood as an informational processing system with explicit and implicit functionality that is wake-up to sensory processing, short-term or working memory, and long-term memory.

Memory can be divided into a number of memory types that each handle a different type of memory function.

Storage

Storage is the more or less passive process of retaining information in the brain, whether in the sensory memory, short-term memory, or the more persistent long-term memory.

Paradigms

In their efforts to understand memory, cognitive psychologists have historically related to this technologies of their age. One of the earliest models used to explain memory storage was the war models, which illustrate the long-term memory systems that differ in their pathways for various activities. According to this model, when an event was no longer meaningful or impressive in nature, the memory would remain intact.
Paradigms

Modern theories of memory reflect the technological advances of the 20th century. For example, many theories formulated at the dawn of the computer age drew analogies between information storage in computers and information storage in human memory. The two most influential of these paradigms, known as memory systems, are the short-term and long-term memory systems. The brain, however, stores and retrieves information in a much more complex fashion. The brain is not divided into distinct memory stores, as these memory systems are, but rather into functionally distinct regions.

Storage

Storage is the process of transferring information from the sensory memory to the long-term memory store. Information that is transferred to the long-term memory store is no longer accessible to the brain unless it is retrieved from memory.

Sensory Memory

The sensory memory preserves information in the original sensory form for a brief time, usually only 5-10 seconds. Sensory memory allows the encoding of new perceptual sounds, sights, or smells. It provides a short-term storage for sensory information in the brain. Sensory memory is not the same as the auditory or visual senses, but rather the information is stored in the brain.

Example: Sparkler

The example demonstrates the existence of afterimages by pouring a red light or green light into the sparkler. If you observe the sparkler long enough, you should see a complete cycle of colors through the light source. The appearance of the sparkler is repeated in the sparkler again because it is a short-term memory.
Chapter 5

Challenge to the Model

The model just presented is a text of higher coherence than the standard linear text format: The present model more explicitly shows the structural relations between entities than standard linear texts. Since the goal of education is to foster learning, and learning involves making connections between ideas, one would assume that a more coherent text would offer an educational advantage over less coherent texts.

However, research by McNamara et al. [44] suggests that high coherence texts might not always be better for learning. In their study, they measured the effects of two types of coherence: local (microstructure) and global (macrostructure). (Local coherence is not relevant to the present study; our interface adds only global coherence.) A text that is high in local coherence explicitly links each sentence to the next; a text that is high in global coherence shows how paragraphs relate to each other and to the larger topic.

McNamara et al. started with a text that was coherent locally, but not globally: sentences were linked together, but there were no headers. They created another text by reducing the local coherence of the initial text (by removing links between sentences); they created another text by adding explicit markers of macrostructure (subtopic headers); and they created a fourth by adding headers to the locally incoherent text. These texts allowed the researchers to study the effects of local and global coherence on reading comprehension independently.

McNamara et al. separated participants into two groups according to their level of background knowledge in the domain, based on their performance in a pre-test. Participants completed a post-test of questions. They asked participants two kinds of questions, one targeting representations of the text base (a linguistic representation, which is relatively shallow), and the other requiring targeting participants’ situation model of the text (the referential meaning of the text, which integrates the text-base with other concepts in long term memory).
For questions targeting the text base, participants scored better the more coherent the text. For questions targeting the situation model, however, only low knowledge participants were helped by textual coherence. The learning of high knowledge participants seems to have been harmed by textual coherence: high knowledge participants performed best when the text was minimally coherent.

McNamara et al. argue that high knowledge participants were harmed because they did less *active processing* when using the high coherence texts. In their words:

“The formation of such situation models requires two things as a basis: active inferencing and adequate prior knowledge. Without adequate prior knowledge, students are limited in their constructive processes. Nevertheless, a well-written, fully coherent, explicit text at least will allow them to form a good text base.” [44]

Low knowledge participants, this argument goes, are helped by the high coherence texts because it provides key knowledge that they do not already have. The high knowledge learners learn more from low coherence texts because they are already in possession of the required background knowledge to form a rich situation model, and a low coherence text forces them to pay attention to the relations between the concepts presented in the text.

5.1 Attempting to Address the Challenge

Our study (described in more detail in the next chapter) is designed to address the challenge to the model presented by McNamara et al. [44]. Sweller, Merrienboer, and Paas [66] offer an alternative explanation for the results in McNamara et al. They argue that high knowledge readers demonstrate signs of shallower learning when using high coherence texts because the extra information contained in high coherence texts is redundant for these readers. That is, the high coherence texts placed a higher *extraneous load* on high knowledge learners, which impairs their learning. When reading the high coherence text, the high knowledge readers were forced to allocate cognitive resources toward generating representations of the text (of information they already knew) which, if they were reading a low-coherence text, would be allocated toward integrating what they are reading with their wider world knowledge.
We prefer the explanation offered by Sweller, Merrienboer, and Paas. We think it is possible that by striving for structural harmony between the text and the concepts it contains, our proposed interface could reduce extraneous workload compared to a standard linear text in three ways:

The first way is by reducing opportunities for thinking about the structure of text that is not also thought about the structure of content. In a standard linear text, there are times when a reader has to think about the structure of the text and not the structure of the content, which is extraneous (i.e. wasted) cognitive load. For example, a student is reading about a concept that has a prerequisite concept that the student has already read about but forgotten, and they would like to revisit this prerequisite concept so they can understand the current concept. When looking for this concept in a standard text, the student might think about how “far back” the concept is, in terms or pages or concepts, which is thought that does not help them learn any of the course material.

In our proposed non-linear interface, on the other hand, to find a concept they are looking for they will have to think about the way the ideas in the text fit together. In the case of domain concepts they will have to think about the concepts’ location in the hierarchy of domain concepts; in the case of peripheral concepts, they will have to think about the internal (and possibly external) structure of the core concept. When students are thinking about the structure of a concept, it should aid understanding of that concept.

The second way is by making it easier for students to skip content they already know. In a standard linear text, core content and peripheral content are often placed alongside each other with no clear separation. This makes it difficult for students to skip content they already know: they have to read the content to know that they know it. For example, if a student will not benefit from seeing any examples of a concept, then making them read examples is extraneous load. Overt markers of structure like headings certainly help, but we feel that these do not go far enough in communicating conceptual structure.

In the proposed non-linear interface, there is more overt categorization of concepts, and identifying concepts that need not be read is therefore easier. If a student does not need to see any examples of a concept, then they know they do not need to look
at any of the content in the bottom of the interface. We feel that this benefit would be especially apparent when re-reading the content, where a student is looking to brush up on a number of key ideas while ignoring many others that they feel more comfortable with. Of course, this benefit is made greater the more information the students have about what content is ‘beyond’ each link (i.e. ‘peekable’ information).

The third way we think our proposed model could reduce extraneous workload compared to highly coherent linear texts is by showing structural relations spatially, not linguistically. The high coherence text in McNamara et al. [44] signals conceptual structure in two ways: with topic headers, and with added sentences, which often state taxonomical information (e.g. “There are many kinds of heart disease, some of which are present at birth and some of which are acquired later”; “Coronary disease is another example of an acquired heart disease”). This way of adding coherence to a text also adds to the amount of reading students have to do.

In the proposed non-linear interface, we also use headers to signify macrostructure. However, we signal taxonomical information spatially via the concepts’ relative position in the domain hierarchy. Readers using our interface will acquire this information when thinking about how they will navigate the content. We think it is possible that signalling textual structure spatially as opposed to linguistically will interfere less with the reading task.

Because we believe that our interface reduces extraneous workload, we hypothesize that our high-coherence text will avoid the learning disadvantages for high knowledge learners observed in McNamara et al. [44].

For this preliminary investigation into the effectiveness of our proposed interface, we decided to examine only the effects of non-linearity, setting aside the effects of dividing content into core and several kinds of periphery. We chose to focus on a single design concept rather than both because of limitations of time and of access to participants: based on the number of participants we expected to register for the study in the amount of time available to us, we could not have more than two conditions in a between-subjects design. We felt that non-linearity was the most distinctive feature of our design, and the one that would have the biggest effect on learning.
5.1.1 Linear interface

To examine the effects of non-linearity on support for learning, we created a second interface (shown in Figure 5.1) that we tried to keep similar to our proposed design in all ways except for the ordering of the content. The new interface is essentially a linearized version of the proposed non-linear interface. We removed the elements of the interface which afforded non-linear navigation: the minitree and the treemap. Instead of navigating up and down a tree representing the content covered in the text, users moved left and right by clicking either the button on the left or right margin of the page. In place of the treemap, we provide a list of pages contained in the text.

The linear interface does not feature headings and sub-headings which signal the underlying structure of the text, which is a standard feature of standard linear texts. In retrospect, the elements of the table of contents view (Figure 5.1) should have shown as a hierarchy of headings and subheadings rather than a simple list.
Chapter 6

Study

The primary goal of a textbook is to support learning. We felt that some core features of the standard linear textbook insufficiently supported the learning process, and so we designed and implemented this electronic textbook format to address some of these perceived weaknesses in the standard format. Our second research question is: Does our non-linear arrangement of information better support the learning process than the standard linear arrangement?

If a format for an electronic textbook is to be a success, it not only has to better afford learning, but people have to like using it. We are therefore also interested in user attitudes regarding the interface. And even if this format offers no learning advantage over the standard format, if participants prefer the non-linear arrangement of content to the linear one, we would regard this as a success. Our third research question is: Do users prefer the proposed non-linear format to the standard linear format?

We designed the following study to address these two questions. To address the second question, we attempt to measure learning outcomes for two interfaces: our proposed non-linear interface, and a linearized version of this interface. At the end of each reading session, participants are asked to complete a 19-item questionnaire featuring three kinds of questions: text-based, bridging inference, and problem solving. With this comparison, we examine the effects of our method of organizing content non-linearly on the learning process. We do not attempt to measure the effects of separating content into core, and three kinds of peripheral elements on the learning processes: both interfaces make this distinction.

In addition to attempting to measure the learning effects of the respective interfaces, we also ask participants to complete a questionnaire assessing perceived learning along the three dimensions of Bloom’s taxonomy: cognitive, affective, and psychomotor.
In regards to the objective learning measures, we hypothesize that participants in the non-linear condition will outperform participants in the linear condition in bridging inference and problem solving questions (i.e. the types of question that measure situation model development) because the non-linear interface helps readers make connections between ideas. In regards to subjective learning measures, we hypothesize that scores for the non-linear condition will be higher than the linear condition for psychomotor and affective learning, because learners in this condition will have a better idea of the significance of what they are learning, that is how it relates to the wider world.

To address the third research question, during debriefing, we present each participant the interface they did not use during their reading session, and ask them which interface they prefer. We also ask participants to complete a Standard Usability Scale (SUS) questionnaire, assessing their feelings regarding the interface’s usability.

In regards to interface preference, we hypothesize that participants will tend to prefer the non-linear interface over the linear interface. Regarding the SUS scores, we hypothesize that participants will rate the non-linear interface as more usable than the linear interface.

This study was cleared by Carleton University Research Ethics Board-B (Clearance #108451).

6.1 Participants

26 participants (13 female; 13 male) were recruited. 12 were students in a first year Cognitive Science seminar course who signed up for the study through Carleton University’s Sona System for Cognitive Science research; 14 came to the study in response to a poster posted on bulletin boards around campus. Participants were alternately assigned to one of the two conditions.

6.2 Materials

Participants sat at a desk featuring a Thinkpad T440p laptop functioning as a desktop computer. The laptop was connected to a mouse and keyboard through USB, and to a Dell UltraSharp monitor (2209WA; 1680x1050 resolution) through DVI. The
laptop ran a Linux-based OS (Debian 9.4) with i3 window manager. The output that appeared on the Dell monitor was duplicated on the laptop’s screen so the researchers could easily and unobtrusively view the participants’ interactions.

![Setting for the study](image1)

Figure 6.1: Setting for the study

![The screen as viewed by participants](image2)

Figure 6.2: The screen as viewed by participants

The electronic textbooks were presented using Mozilla Firefox ESR 52.6.0 (64-bit) featuring add-ons to eliminate all GUI elements except for the main display frame.

The System Usability Scale (SUS) [11] and the Cognitive-Affective-Psychomotor (CAP) Perceived Learning Scale [57] were used to gauge participants’ perceptions.
regarding the interface’s usability, and how much they felt that they learned while using the interface. The time they took to complete the reading task was recorded. The software logged key events such as mouse hovers and mouse clicks, and stored them in a database on our server at mvp.soft.carleton.ca.

Participants completed two tests: one, taken before the reading session, assessing background knowledge in cognitive science; and the other, taken after the reading session, assessing knowledge of the material they just read. Both tests consisted of multiple choice questions (16 and 19 items, respectively) written by us. We decided against having overlapping questions between the two tests (i.e. no questions that were asked in the background assessment test were asked again in the post-reading test) to avoid priming participants. The pre-test of prior cognitive science knowledge can be seen in Appendix B and the postreading test can be seen in Appendix C.

6.3 Procedure

An overview of the procedure is shown in Figure 6.3 below.

Participants were greeted at the front door of the lab, and directed to a smaller side-room where the experiment took place. Participants were given a verbal overview of the study, and of their task, and then were presented with the informed consent form. After signing the consent form, participants completed a demographics questionnaire, and a 16-item questionnaire assessing background knowledge in cognitive science. Performance on this test was later used to divide participants into two groups based on their level of prior knowledge in cognitive science: (i.e. high knowledge and low knowledge).

After the first two questionnaires were completed, participants were presented with a ‘demo’ interface: the interface was identical to the interface they would be using in the main task except for the content. Participants were given a brief verbal description of the content layout, and how to navigate the content, while participants practised using the interface. The demo session typically lasted about 2 minutes.

Once participants were comfortable with the ‘demo’ interface, they were presented with the ‘main’ interface featuring content from an introductory psychology textbook. Participants were asked to read all of the content contained in the interface. The researcher sat next to participants during the reading session and viewed the
participant’s interactions with the interface indirectly through the laptop’s screen.

After the reading session, participants were immediately presented with the post-task questionnaires, starting with the post-test assessing knowledge of the material they just read. After the post-test, participants completed the CAP Perceived Learning Scale, and then the SUS.

When the participant was finished with the questionnaires, participants were debriefed, and asked if they be willing to be interviewed about their experience, their attitudes towards various kinds of books, and about their general approaches to learning. After participants described their experience, they were shown the interface they did not use, and asked which interface they prefer.

Each session lasted a maximum of one hour. At the conclusion of each the session, participants were compensated for their time. Participants who signed up for the study through Sona received 1% course credit; all other participants received $15.

6.4 Third Round of Pilot Testing

Before the study proper began, we conducted a round of pilot testing with 8 participants. They followed the same procedure using the same materials described above. During this round of testing, we discovered that there was too much content contained in our two interfaces for most people to complete in a 20-minute reading session: a number of participants made it less than halfway through the material after the 20 minutes expired. We made major reductions to the total amount of content contained in the books before the final round of testing.
Figure 6.3: Flow diagram of procedure
Chapter 7

Results

Our second research question (RQ2) is: *Does our non-linear organization of content offer a learning advantage compared to the standard linear format?*. We only had access to each participant for up to one hour, which meant that we could only attempt to measure learning over a short interval of time. We hypothesize that the participants using a non-linear text would learn more than participants using a linear text featuring the same content.

Our third research question (RQ3) is: *Do users prefer the proposed non-linear format to the standard linear format?* We hypothesize at least a small (but non-trivial) subset of students would prefer the non-linear format to the standard linear format.

7.1 RQ2 Hypothesis Testing

To test the first hypothesis, we asked participants to complete a 19-item multiple choice post-reading test of the material they read during a 20 minute reading session. The test featured three kinds of questions: text-based questions \(n=11\), bridging-inference questions \(n=4\), and problem-solving questions \(n=4\). Text-based questions assess the reader’s representation of the text (the *text-base*). The latter two types of questions assess the reader’s representation of the situation described by the text (the *situation model*). We are mostly interested in participants’ performances in the situation model questions, because these questions require that the reader go beyond simply reproducing the text: they test a deeper understanding of the material \[44\].

We also asked participants to complete a questionnaire assessing the Cognitive, Affective, Psychomotor (CAP) Perceived Learning Scale: a self-reported measure of how much participants felt that they learned during the reading session.
7.1.1 Post-reading test

A summary of the results of the post-reading scores are shown in Table 7.1. We divide participants into two groups, low and high knowledge, based on their performance in the pre-test assessing background knowledge in cognitive science: participants who scored greater than the median score (Mdn=0.5/1) were placed in the high knowledge group (n = 13), and participants who scored less than the median score were placed in low knowledge group (n = 13). We compare test scores by question type for low knowledge and high knowledge learners, for both interface conditions. For sake of completeness, we include the combined scores of low and high knowledge participants (under “Combined”), and the overall test scores (i.e. regardless of question type) (in the rows labelled “Overall”).

The scores are generally quite similar between the conditions across question types and knowledge levels. The biggest difference between the two conditions is for low knowledge participants in bridging inference questions, where participants in the non-linear condition (n = 7) scored higher (.61) than participants in the linear condition (n = 6) (.33).

Boxplots of the post-reading scores for each condition, knowledge level, and question type are shown in Figure 7.1. For high knowledge participants, for all question types, the post-reading scores in the linear condition are higher than those in the non-linear condition. However, for all three of these boxplot pairs the notches indicating the 95% confidence interval around the median overlap, suggesting that the differences are not significant.

For low knowledge participants, for the text based questions, the scores are quite similar. For bridging inference and problem solving questions, the post-reading scores in the linear condition are lower than the scores in the non-linear condition—the opposite of what we saw with post-reading scores for high knowledge participants. For the problem solving question boxplot pairs, the notches indicating the 95% confidence interval around the means overlap, suggesting that the differences are not significant. For the bridging inference boxplots, however, the notches do not overlap, suggesting

15 participants scored exactly 50%. The chronological first three of these participants were placed in the low knowledge group, and the other two were placed in the high knowledge group. Other ways of addressing this issue include a random assignment of participants into the two groups, or the creation of a middle knowledge group.
Table 7.1: Proportion of Correct Responses in the Postreading Test for the Two Conditions by Knowledge and Question Type

<table>
<thead>
<tr>
<th></th>
<th>Linear</th>
<th>Non-Linear</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High knowledge</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Text based</td>
<td>.68</td>
<td>.62</td>
</tr>
<tr>
<td>Bridging inference</td>
<td>.61</td>
<td>.62</td>
</tr>
<tr>
<td>Problem solving</td>
<td>.50</td>
<td>.42</td>
</tr>
<tr>
<td>Overall</td>
<td>.62</td>
<td>.58</td>
</tr>
<tr>
<td><strong>Low knowledge</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Text based</td>
<td>.56</td>
<td>.56</td>
</tr>
<tr>
<td>Bridging inference</td>
<td>.33</td>
<td>.61</td>
</tr>
<tr>
<td>Problem solving</td>
<td>.42</td>
<td>.39</td>
</tr>
<tr>
<td>Overall</td>
<td>.48</td>
<td>.53</td>
</tr>
<tr>
<td><strong>Combined</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Text based</td>
<td>.62</td>
<td>.59</td>
</tr>
<tr>
<td>Bridging inference</td>
<td>.48</td>
<td>.62</td>
</tr>
<tr>
<td>Problem solving</td>
<td>.46</td>
<td>.40</td>
</tr>
<tr>
<td>Overall</td>
<td>.56</td>
<td>.55</td>
</tr>
</tbody>
</table>

that the bridging inference scores are significantly higher in the non-linear condition than in the linear condition.

A number of one-tailed independent samples t-tests were conducted to compare the performance of high and low knowledge participants in the two conditions for each type of question (text-based, bridging inference, and problem solving), where “performance” means the proportion\(^2\) of correct scores on the post-reading test.

For text-based questions: (i) there was no significant difference between the linear condition (\(M=0.56, SD=0.17\)) and the non-linear condition (\(M=0.56, SD=0.18\)) for low knowledge participants, \(t(10.88)=0.023, p=0.51\); and (ii) there was no significant difference between the linear condition (\(M=0.68, SD=0.16\)) and the non-linear condition (\(M=0.62, SD=0.13\)), \(t(11.00)=0.67, p=0.74\) for high knowledge participants.

For bridging inference questions: (i) the difference between the linear condition

\(^2\)We treat the scores as continuous data.
Figure 7.1: Proportion of Correct Responses in the Postreading Test by condition, knowledge level, and question type
Learning effects by condition


table

(M=0.33, SD=0.26) and the non-linear condition (M=0.61, SD=0.32) for low knowledge participants approached significance, \(t(10.98)=-1.71\) \(p = 0.057\); and (ii) there was no difference between the linear condition (M=0.61, SD=0.38) and the non-linear condition (M=0.62, SD=0.21), \(t(9.56), p=0.45\) for high knowledge participants.

For problem-solving questions: (i) there was no significant difference between the linear condition (M=0.42, SD=0.26), and the non-linear condition (M=0.39, SD=0.24) for low knowledge participants, \(t(10.47)=0.17, p=0.57\); and (ii) there was no difference between the linear condition (M=0.50, SD=0.20) and the non-linear condition (M=0.42, SD=0.20), \(t(10.70)=0.73, p=0.76\) for high knowledge participants.

Finally, we examined the learning effects for the two conditions, where ‘learning effects’ means the post-reading score minus the score for the pre-test assessing background knowledge in cognitive science. Boxplots of the learning effects by condition are shown in Figure 7.2. Learning effects were greater in the non-linear condition than in the linear condition. However, the notches indicating the 95\% confidence interval around the median overlap, suggesting that these results are not significant.
7.1.2 CAP Perceived Learning Scale

Boxplots showing the CAP Perceived Learning Scale for each condition are shown in Figure 7.3. The two boxplots are quite similar: they have the same median, and the notches indicating the 95% confidence interval around the median are nearly identical.

The CAP Perceived Learning Scale is generated from Likert-scale questions, which means we look for differences between the scores for the two conditions using the non-parametric Wilcoxon rank-sum test.

The Wilcoxon rank-sum test indicated that there was no significant difference between the CAP scores for the linear condition ($Mdn=31$) and the non-linear condition ($Mdn=31$), $U=54.5$, $p=0.167$.
7.2 RQ3 Hypothesis Testing

To test the second hypothesis, after each reading session, we show participants the interface they did not use, and ask if they prefer one or the other. We also asked participants to complete a System Usability Scale (SUS) questionnaire for the interface they used.

7.2.1 SUS scores

To see if there was any difference in perceived usability between the two conditions, we asked all participants to complete the System Usability Scale (SUS) questionnaire, and then compared the results for each condition. Boxplots of the results are shown in Figure 7.4. The median is the same for each condition, but the spread of scores is greater for the non-linear condition than the linear condition.

The System Usability Scale is generated from Likert-scale questions, which means we look for differences between the scores for the two conditions using the non-parametric Wilcoxon rank-sum test.

The Wilcoxon rank-sum test indicated that there was no significant difference between the SUS scores for the linear condition \( \text{Mdn}=50.5 \) and the non-linear condition \( \text{Mdn}=50 \), \( U=71.5 \), \( p=0.5 \).

7.2.2 Interface preference

During debriefing, at the end of the study session, participants were shown and given a description of the interface they did not use, and were asked which they preferred. The distribution of interface preferences is shown in Figure 7.5. 18/25 (72%) of participants said that they preferred the non-linear interface to the linear interface.

A chi-square test of independence showed that the difference between the total number of participants who preferred the non-linear interface \( n=18 \) compared to the linear interface \( n=7 \) was significant, \( \chi^2 (1, N=25)=8, p=.005 \).

Figure 7.6 shows the distribution of interface preferences separately for participants in the linear condition (left) and non-linear condition (right). Of the participants who used the linear interface, 7/12 said they would prefer to use the “other”

---

3One participant declined to state their preference.
(i.e. non-linear) interface. Of the participants who used the non-linear interface during the reading task, only 2/13 said they would prefer to use the “other” (i.e. linear) interface.

7.2.3 Themes from the post-experiment interview

Participants who preferred the control interface tended to like its straightforwardness, and found the non-linear interface “confusing”. They generally seemed to find the freedom of choice in terms of navigation overwhelming.

Participants who preferred the treatment interface tended to say it was “well-organized” and “it helps show how the ideas fit together”. Many participants appreciated that the text was categorized into smaller units than one would typically find in a standard textbook. Many participants noted difficulties with the standard
Figure 7.5: Interface preference

Figure 7.6: Interface preference split by condition
prose format, where content is provided in big, relatively undifferentiated blocks of text. Some felt that this format was “intimidating”, and others felt that the process of going through page after page of these big blocks of text “monotonous” and “boring”, and that these texts seem to “go on and on” indefinitely. A number of participants noted that the ideas presented in standard texts tended to get mixed up in their heads, and that they felt that the larger organizing structure provided by the treatment condition would help prevent this from happening. They found that organizing the content in a hierarchical tree made it appear more manageable.

We asked many participants (perhaps half) if they preferred text in textbooks to appear in smaller or larger chunks. Every participant we asked this question to said that they preferred smaller chunks of text. We thought this result was favourable to our non-linear design, which breaks up text into smaller chunks than standard linear texts as a consequence of the division of content into core and three kinds of peripheral, and keeping this content in separate locations in the UI.

A few of the participants who preferred the treatment interface said that they generally did not like reading, and this interface somehow made reading better for them.
Chapter 8

Discussion

8.1 Learning Outcomes

The proposed interface has more coherence at the macrolevel: it more explicitly indicates to the reader the underlying structure holding together paragraphs and larger sections of text. A study by McNamara et al. [44] suggests that high coherence texts may be worse for learning than low coherence texts for high knowledge learners. McNamara et al. argue that low coherence texts force learners to engage in active processing, which results in better-developed situation models. Sweller, Merrienboer, and Paas [66] present an alternative explanation: the high coherence texts contain redundant information (for high knowledge learners), which means they have more extraneous workload than the low coherence texts. The study was designed in part to address the issues raised by McNamara et al. We felt that the proposed non-linear interface would introduce less extraneous workload than a linear text, which we hoped would allow us to avoid the issues faced by the high coherence texts in the study by McNamara et al.

We found, however, no significant differences between the high coherence non-linear text, and the high coherence linear text. In this section, we discuss several explanations and implications for these results.

8.1.1 Differences between the interfaces

In the context of a 20-minute reading session, we suspect that the experience of reading the two texts was quite similar. The important differences between the two interfaces were: (i) the non-linear textbook more explicitly marked categories of peripheral content; (ii) the non-linear text more clearly showed how the ideas comprising the core domain content related to each other; and (iii) the non-linear textbook allowed users to choose their own path through the text. The first difference is a consequence
of the distinction between core and peripheral content, and the latter two differences result from the navigable hierarchy (i.e. the treemap and minitreemap). We discuss these differences in turn.

**Organization and presentation of peripheral content**

Both interfaces divided content into core and peripheral chunks. The only difference was how the chunks were arranged. In the linear text, the enriching content was placed in grey boxes underneath the core content, which always appeared at the top of the screen with a white background. The different kinds of peripheral content were not differentiated from each other in any way, except that the relative position of each element with respect to the other elements always remained the same. By contrast, in the non-linear condition, the categories of peripheral elements were marked more saliently: each category had its own location relative to the core content, and each had its own colour scheme.

In an artificial context such as the one we created for our study, we now suspect that this difference would have much influence on learning outcomes. We feel that the biggest difference this makes to the reading experience is (i) that it reduces the amount of text appearing on the screen at one time, and (ii) in reducing the time it takes to find the type of content one is looking for. The impact of the first difference is diminished in our implementation by the relatively small amount of content for each concept: we felt it necessary to include the number of concepts we did so we could write challenging multiple choice questions, and since we could only ask students to read for 20 minutes, we had to keep the amount of content per-node down. We also feel that the advantage of reducing the amount of text on screen would be largely affective in nature: our participants did not like large chunks of text, and many expressed frustration with the monolithic and seemingly interminable standard texts. In the artificial context where they feel obligated to read everything, and where they only have to read a relatively small amount of text a single time, this affective difference might not reveal itself. But in a real-world context where they have to voluntarily read very large chunks of texts—and much of it more than once—the two

---

1Participants in both conditions were told that there were three categories of enriching content, and what those categories were.
texts may diverge in terms of how well they support the act of reading. We think we found support for this claim in our post-session interview, where we found a universal distaste for large chunks of text compared to small chunks of texts.

We think the impact of the second difference is eliminated totally by the fact that readers were never really looking for any particular content. We asked readers to read all of the content in the text, and they proceeded concept-by-concept in a rather rigid fashion until they had read every chunk of text. They navigated the text non-linearly, but in all but one or two cases, readers never revisited content they had already read. Readers also never used the search function.

The only practical implication this difference would have in our 20-minute session would be that readers can more easily visit the peripheral content in the order they like, which, similar to the above-mentioned difference, seems to us to be a benefit that is mostly affective in nature. Again, we feel that affective benefits are amplified in a real-world context where learners are reading of their own volition. And the increased efficiency resulting from a reduction in extraneous workload would only become manifest when the reader is looking for some specific piece of content—mostly likely during a review (i.e. re-reading) session.

**Hierarchical macrostructure**

The non-linear interface organized content in a partonomical/taxonomical structure (i.e. hierarchy) relating all domain concepts to each other. The entire structure was made visible in the treemap, and a snapshot of the ‘local’ structure was made visible in the minimap.

In the linear condition, the information was laid-out linearly and the relation between concepts was not explicitly stated. Often sub-concepts were presented very close to their parent topics, which might be suggestive of their structural relationship. However, the linear distance between concepts was not a perfectly reliable indicator of structural relationship, as often child concepts would appear many pages away from their parent, and sometimes a deeply-embedded descendant of one node would be placed alongside a node from an entirely different sub-tree.

---

---

2The order is equivalent to a depth-first, left-to-right traversal of the hierarchy used in the non-linear condition.
We expected that this would provide a clear advantage for the non-linear interface, even in the context of a 20-minute reading session. What is particularly perplexing is that several of the bridging inference questions were designed so that knowing the parent concept of a given target concept could help determine the correct answer. For example, for the question “What component of working memory is at work when you mentally rearrange the furniture in your living room?”, half of the potential answers can be ruled out by knowing that they are not components of working memory.

Perhaps the tight time constraints prevented readers from taking advantage of the explicit macrostructure provided to them: participants seemed to spend all of their time reading the text, and none of their time pausing to reflect on the text. Readers in the McNamara et al. [44] study were able to read through the text twice—perhaps a second reading aids the construction of situation models.

**Free order**

The linear interface effectively guided readers through the text in a relatively fixed manner with its *next page / previous page* navigation system. Readers were free to navigate the content in any order they chose through use of the table of contents, but the overwhelming majority opted for the default linear path.

The non-linear text, by contrast, did not suggest any particular path through the content. When readers were presented with more than one option to descend the tree, they had to decide for themselves which way to go.

Making participants think about which node to visit next is an increase in cognitive load, but since this thought doubles as thought about domain concepts and how they fit together, it is not extraneous load. We would expect that giving readers the freedom to choose their own path through the text would give the non-linear interface an advantage in supporting the generation of situation models.

In the context of a 20 minute reading session, perhaps participants made arbitrary choices about which concept to visit next just to ‘get through it’.

**8.1.2 Other explanations for lack of findings**

We have already mentioned the lack of time to reflect on the content caused by the short session time and the fact that they only read the content once. We identify a
few other possible contributing factors to the lack of conclusive differences in learning outcomes between the two interfaces.

Our non-linear interface is a new way of viewing and interacting with text, whereas the linear text was modelled on pre-existing electronic textbooks, which in turn are modelled on physical print media. Learning something new, whether it is text or how to use a new interface, requires the deployment of working memory resources. Our text required more learning, which means there are fewer resources available for learning the content. It is possible that the benefits offered by the non-linear organization of content was overshadowed by the costs associated with learning the new interface.

Our study proper had 26 participants, which meant there were only 6 to 7 participants in each knowledge level-interface condition group. The study by McNamara et al. [44], which we modelled our study after, had 56 participants—more than twice the amount. Perhaps increasing the total number of participants would have resulted in significant results.

Lastly, there could be issues with the post-reading questions themselves. We used multiple choice questions and not short answer questions as in McNamara et al. [44]. Our original study design featured short answer questions, but we found during pilot testing that participants were answering almost no questions correctly (the average score was approximately 25%). We do not know what effect changing the format of the questions had on the scores.

8.2 What does this mean for standard print texts?

This study was not a comparison of a novel text with the types of texts students use in classrooms today. The control interface divided text into smaller chunks than standard textbooks do, and more strictly separated content by category. For example, in the psychology text we used as source material for the content, core content, key studies, examples, and applications are often presented alongside one another, and the order in which they appear varies from topic to topic. Ideally, our study would have included a more traditional text as one of the conditions, but time constraints made this practically impossible.

Our study was a comparison of linearity versus non-linearity. Our study would be
the first time the participants had seen or used an interface like the one we designed, and that this interface was as effective as the more familiar one could be seen as encouraging. We also found that on average, participants exhibited a decided preference for non-linearity. And a number of participants expressed strong enthusiasm about the non-linear interface; for these participants, this design seemed to fulfill a need that linear texts were not providing. Overall, even though we found no conclusive effect of linearity on learning outcomes, we think our results suggest that many students are looking for a non-linear alternative to the standard format.

8.3 Is the McNamara et al. really a challenge?

In hindsight, perhaps we overestimated the severity of the challenge. Their study indeed suggests that high coherence could impair learning for students with a relatively high amount of background knowledge on the topic they are reading about. However, the study also suggests that our design could help readers with a relatively low amount of knowledge build better mental representations.
Chapter 9

Conclusions

We reviewed literature related to learning and found a call for radical reform in curriculum design, namely to focus less on “breadth” of coverage and more on “depth”. We identified a number of propositions relevant for the design of instructional materials that support the “deep” approach.

Based on our review of the literature, we designed a model for a non-linear electronic textbook that better supports the “deep” approach to learning by organizing textual content in a way that attempts to mimic how concepts are organized in the mind. The key features of this e-textbook design are the division of content into core and peripheral elements (a domain concept, and antecedents, postcedents, and instances), and the non-linear navigation of concepts. We summarized the design process leading from inception, to the start of implementation, to the ‘final’ design.

A study by McNamara et al. [44] presented a challenge to our design. The results of their study suggest that high coherence texts are good for low knowledge learners, but bad for high knowledge learners: high knowledge learners performed worse on questions targeting “deep” learning when using a high coherence text compared to a low coherence text. Sweller, Merrienboer, and Paas [66] felt that these results were due to an increase in extraneous load, and we agreed. Since we felt that our design reduced extraneous cognitive load, we hoped our interface would fare better. We designed our study in part to test this hypothesis.

9.1 Research Summary

Our research questions were: 1. Is it possible to design and implement a non-linear electronic textbook that comports with cognitive science models of how concepts are represented in the mind?; 2. Does our non-linear arrangement of information better support the learning process than the standard linear arrangement?, and 3. Do users prefer the proposed non-linear format to the standard linear format?. The answer to
the first question is “yes”: the ideas in the literature review did allow us to design, implement, and iteratively refine an electronic textbook interface.

For the second question, we did not find a conclusive answer: there were no significant differences in objective or subjective learning outcomes between the two interfaces. We think our study design may have precluded the benefits of the non-linear interface, and a modified study might yield more positive results.

We believe we found an affirmative answer to the third question. A majority of participants preferred the non-linear interface to the linear interface, and a non-trivial subset of this group were highly enthusiastic about our proposed interface.

9.2 Contributions

Our contributions are: 1. the principled design and subsequent implementation of a non-linear electronic textbook; and 2. the study and analysis of this design.

9.3 Limitations and Future Work

There were a number of limitations to the design, and to the study of the design, which could be amended in future research.

9.3.1 Design

A more full implementation of the design would have to address some issues related to scaling: our implementation worked reasonably well for the amount of content we covered, but this is equivalent to approximately one-third of a single chapter in an introductory psychology textbook. If the treemap showed the entire contents of even a single chapter, this could be overwhelming to readers (several participants noted that they felt overwhelmed by looking at the treemap featured in our reduced content set). Future would have to find a way to allow users to browse large texts without overwhelming them.

It is unclear how a more contentful implementation would affect usability. For example, perhaps the non-linear design is more taxing on mental resources, which introduces fatigue issues in longer reading sessions.
During our post-study interview, participants often made note of features they would like to see in a full implementation of an electronic textbook, but which were not present in the prototype they used. These include the ability to highlight and annotate text, and support for animations and videos. Future development should include features like this.

A non-linear organization of content affords numerous *paths* through the text. This notion of *paths* will have an effect on the reading and learning experience. For example, some students may feel overwhelmed by the free navigational order, and may benefit from some sort of guidance, perhaps in the form of something like ‘suggested paths’. The notion of *paths* should be given more consideration in future work.

Participants were generally less enthusiastic about the sliding peripheral content than about the other major feature, hierarchical navigation. A common complaint was that it took too much effort to bring the peripheral content into view. More work could be done to consider how to reduce the amount of effort required to bring sliding content in.

Ideally, the book should work for domains other than psychology. We believe it would, since all concepts have a past, a use, are about things in the world, and relate to other concepts hierarchically, but future work would have to verify this.

### 9.3.2 Study

Several of the ways in which the non-linear interface could provide an advantage over a linear text may have been precluded by the experimental design. Going back through a text to review concepts is a common thing students do with textbooks, and the proposed non-linear organization may help make finding known information easier (several participants alluded to this). In a study where participants read through the material only once, this benefit cannot manifest itself. Future studies might test participants over multiple reading sessions.

Our main study featured 26 participants, which is less than the 56 featured in McNamara et al. We divided participants into four groups (2 knowledge levels × 2 interface conditions), which meant that we had group sizes of only 6 or 7. We believe that increasing the number of participants may have resulted in significant results.
We arbitrarily assigned 5 participants to low and high knowledge groups: 5 participants had the median score of .5, and we assigned them to a knowledge level group based on their chronological order of appearance in the study. It is possible that there is an uneven distribution of post-reading test scores in this group of 5, so it is possible that our arbitrary assignment method had a large influence on the learning effects, which of course would not come from either interface’s ability to support learning. More participants could have helped minimize the effect of this arbitrary split, and more questions would have helped reduce the likelihood of having 5 participants with the exact same score.

This last point bring up another issue with the study: we feel that we had too few bridging inference \((n = 4)\) and problem solving \((n = 4)\) questions. More questions would have helped increase the test’s accuracy as a gauge of the development of participants’ situation model of the content: it would help minimize the negative effects of any poorly-formed questions, and correct answers that were random guesses.
Bibliography


[34] Lilian G Katz and James Raths. “Six dilemmas in teacher education”. In: Journal of Teacher Education 43.5 (1992), pp. 376–385.


[43] Margaret G McKeown et al. “The contribution of prior knowledge and coherent text to comprehension”. In: Reading research quarterly (1992), pp. 79–93.


[52] Jose I Rodriguez, Timothy G Plax, and Patricia Kearney. “Clarifying the relationship between teacher nonverbal immediacy and student cognitive learning:
Affective learning as the central causal mediator”. In: *Communication education* 45.4 (1996), pp. 293–305.


Appendix A

Demographics questionnaire

1. Age (in years): ____________

2. Please state your gender:
   
   (a) Male
   (b) Female
   (c) Other

3. Handedness:
   
   (a) Right
   (b) Left
   (c) Ambidextrous

4. Please state your agreement/disagreement with the following questions (sic):
   
   (a) I have a strong background in cognitive science.
   (b) Working with computers makes me nervous.
   (c) Computers make me feel uncomfortable.
   (d) Computers make me feel uneasy.
Appendix B

Cognitive science background test questions

1. Memories of personal experiences are known as:
   (a) Semantic memories
   (b) Episodic memories
   (c) Declarative memories
   (d) Procedural memories

2. Operant conditioning is a type of learning in which:
   (a) a stimulus comes to evoke a response that was originally evoked by another stimulus
   (b) a response is evoked by its consequences
   (c) a response is learned through the observation of others, who are called models

3. B.F. Skinner is best known for his theory of:
   (a) the subconscious
   (b) humanism
   (c) behaviourism
   (d) the collective unconscious

4. The smallest unit of meaning in a language is the:
   (a) phoneme
   (b) morpheme
   (c) word
5. The structure that connects the two hemispheres of the brain is called:
   (a) the myelin sheath
   (b) the corpus callosum
   (c) the axon
   (d) the olympus mons

6. An action potential is:
   (a) a brief shift in a neuron’s electrical charge that travels along an axon
   (b) a neuron’s stable, negative charge when the cell is inactive
   (c) a junction where information is transmitted from one neuron to another
   (d) the amount of energy needed to trigger the release of neurotransmitters

7. The smallest difference in the amount of stimulation that a specific sense can detect is called:
   (a) the minimal distinction difference
   (b) the just noticeable difference
   (c) the salient difference
   (d) the sensory detection difference

8. The system responsible for one’s sense of smell is called:
   (a) the gustatory system
   (b) the kinesthetic system
   (c) the olfactory system
   (d) the vestibular system

9. The technique that measures brain activity by detecting changes associated with blood flow is called:
(a) Computerized tomography (CT)
(b) Electroencephalograph (EEG)
(c) Functional magnetic resonance imaging (fMRI)
(d) Positron emission tomography (PET)

10. Internally produced chemicals that resemble opiates in structure and effects are called:

(a) endorphins
(b) GABA
(c) Acetylcholine
(d) Monoamines

11. The study of how physical stimuli are translated into psychological experience is known as:

(a) psychometrics
(b) psychophysics
(c) physical psychology
(d) applied psychology

12. The neural tissue that receives incoming light is called:

(a) the fovea
(b) the iris
(c) the cornea
(d) the retina

13. Construct validity is:

(a) the measurement consistency of a test
(b) the extent to which there is evidence that a test measures a particular psychological concept
(c) the degree to which the content of a test is representative of the target domain

(d) estimated by correlating subjects’ scores in a test with their scores on an independent criterion of the trait assessed by the test

14. Of the regions of the brain, the _____________ is the largest and most complex.

(a) midbrain
(b) hindbrain
(c) highbrain
(d) forebrain

15. Freud is best known for this theory of:

(a) the subconscious
(b) behaviourism
(c) humanism
(d) the collective unconscious

16. _____________ refers to the tendency of some neural functions to be more dominant in one hemisphere than the other.

(a) split-brain
(b) laterality
(c) plasticity
(d) bisymmetry
Appendix C

Postreading test questions

C.1 Text-based questions

1. The capacity of short-term memory can be increased by grouping items into units called:
   (a) bits
   (b) chunks
   (c) neurons
   (d) nodes

2. When people spontaneously organize information into categories for storage in memory, this is an example of:
   (a) clustering
   (b) a connectionist network
   (c) a semantic network
   (d) schema

3. What is the first step in the process of creating a new memory?
   (a) encoding
   (b) perception
   (c) sensation
   (d) attention

4. Stimuli that bring memories to mind are called:
   (a) context
(b) cues
(c) associations
(d) schemata

5. According to Baddeley (2003), ____________ is a limited capacity storage system that temporarily maintains and stores information by providing an interface between perception, memory, and action.

(a) the episodic buffer
(b) short-term memory
(c) the central executive
(d) working memory

6. Which of the following contributes to the loss of information in short-term memory?

(a) failure to rehearse
(b) shallow encoding
(c) stress
(d) decay

7. The number of items that can simultaneously be held in short-term memory is:

(a) 7 +/− 1
(b) 4 +/− 1
(c) 7 +/− 2
(d) 4 +/− 2

8. A Parallel Distributed Processing system consists of a large network of interconnected computing units, or nodes, that operate much like ____________

(a) the internet
(b) computers
9. When a memory is encoded and stored, it is often accompanied by information from:

(a) other memories retrieved from long-term storage
(b) semantically related memories
(c) one’s expectations based on past experience
(d) the context in which it took place

10. Recall that when viewing a sparkler that is rapidly moving in the dark, it can look to us as if the sparkler is leaving a trail of light. This phenomenon is known as a/an ________ image.

(a) after
(b) fleeting
(c) ghost
(d) transient

11. The phenomenon referred to in the previous question demonstrates the duration of:

(a) working memory
(b) the episodic buffer
(c) sensory memory
(d) the visuospatial sketchpad

C.2 Bridging inference questions

12. What component of working memory is at work when you use recitation to remember a phone number?
13. What component of working memory is at work when you mentally rearrange the furniture in your living room?

(a) conceptual clustering  
(b) episodic buffer  
(c) visuospatial sketchpad  
(d) schema

14. Which type of knowledge representation and organization represents knowledge as patterns of activation in a system?

(a) semantic network  
(b) connectionist network  
(c) conceptual cluster  
(d) associative network

C.3 Problem solving questions

15. After seeing a dog on the street, John is reminded that he has to get home to feed his cat. What type of knowledge structure can help explain this sequence of thoughts?

(a) context cues  
(b) episodic buffer  
(c) semantic network  
(d) schema
16. Mary is thinking back to a visit to a restaurant. She recalls salt and pepper shakers being present on the table, but in actuality these items were not present. What type of knowledge structure can help explain Mary’s mistaken recollection?

(a) episodic buffer
(b) semantic networks
(c) visuospatial sketchpad
(d) schema

17. John witnesses a bank robber fleeing the scene of the crime. He takes mental note of several pieces of identifying information about the robber and their getaway vehicle, and he wants to remember this information so he can later pass it along to the police. What can John do to help ensure this information is maintained in short-term memory?

(a) drawing a mental picture of the event
(b) repeating the information in his head over and over
(c) connecting the information to personally relevant information from long-term memory
(d) connecting the information to the wider environment in which the event took place

18. Mike left his wallet somewhere at the theme park, but he can’t recall where. What should Mike do to help cue the retrieval of this memory?

(a) go back to the park
(b) mentally retrace his steps
(c) think of other times he lost and then found his wallet and look for structural patterns
(d) undergo hypnosis
(e) find the clothes he wore that day and put them back on
C.4  (One more) bridging inference question

19. Knowledge is represented and organized within what type of memory storage?

(a) Short-term memory
(b) The episodic buffer
(c) The central executive
(d) Long-term memory
Appendix D

System Usability Scale (SUS) questionnaire

Participants were asked to state their agreement/disagreement with the following statements on a 1-5 scale where 1 is “strongly disagree”, 5 is “strongly agree”, and 3 is “neutral”.

1. I think that I would like to use this textbook frequently.
2. I found the textbook unnecessarily complex.
3. I thought the textbook was easy to use.
4. I think that I would need the support of a technical person to be able to use this textbook.
5. I found the various functions in this textbook were well integrated.
6. I thought there was too much inconsistency in this textbook.
7. I would imagine that most people would learn to use this textbook very quickly.
8. I found the textbook very cumbersome to use.
9. I felt very confident using the textbook.
10. I needed to learn a lot of things before I could get going with this textbook.
Appendix E

Cognitive, Affective, and Psychomotor (CAP) Perceived Learning Scale questionnaire

Participants were asked to state their agreement/disagreement with the following statements on a 1-5 scale where 1 is “strongly disagree”, 5 is “strongly agree”, and 3 is “neutral”.

1. I could organize the material I read into a logical structure.

2. I could not produce a study guide of this material for future participants.

3. I am able to use the physical skills learned in this reading session outside of this lab.

4. I have changed my attitudes about the subject matter as a result of this reading session.

5. I can intelligently critique the concepts featured in the textbook.

6. I feel more self-reliant as the result of the content learned in this session.

7. I have not expanded my physical skills as a result of this reading session.

8. I can demonstrate to others the physical skills learned in this reading session.

9. I feel that I am a more sophisticated thinker as a result of this reading session.
Appendix F

Research Ethics Protocol Form
# CUREB A&B: Research Ethics Protocol Form

## 1. Project Team

### 1A. Lead Researcher

<table>
<thead>
<tr>
<th>Role</th>
<th>Last name/First name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic Staff</td>
<td>Response: Spero/Eric</td>
</tr>
<tr>
<td>Library or Other Staff</td>
<td>Official university (or other institution) email address: <a href="mailto:ericspero@cmail.carleton.ca">ericspero@cmail.carleton.ca</a></td>
</tr>
<tr>
<td>Post-doctoral Fellow</td>
<td>Response: Human-Computer Interaction, Faculty of Arts &amp; Social Sciences, Carleton University</td>
</tr>
<tr>
<td>Ph.D. Student</td>
<td></td>
</tr>
<tr>
<td>Master's Student</td>
<td></td>
</tr>
<tr>
<td>Undergraduate</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
</tbody>
</table>

### 1B. Academic Supervisor

- Same as lead researcher

<table>
<thead>
<tr>
<th>Academic supervisor(s) Last name/First name. (Note, the supervisor must be copied on all correspondence with CUREB.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response: Arya/Ali and Biddle/Robert</td>
</tr>
<tr>
<td>Official university (or other institution) email address:</td>
</tr>
<tr>
<td>Response: <a href="mailto:AliArya@cunet.carleton.ca">AliArya@cunet.carleton.ca</a>; <a href="mailto:RobertBiddle@cunet.carleton.ca">RobertBiddle@cunet.carleton.ca</a></td>
</tr>
</tbody>
</table>

Indicate your department, faculty and institution

**Response**: School of Computer Science, Faculty of Engineering and Design, Carleton University

### 1C. Project Team Members

- Not applicable/No other team members

List the project team members here. For each team member, provide the following: 1) Last name/First name 2) Email address 3) Role in project 4) Department and institution (E.g. Master’s student in Canadian Studies at Carleton)

**Response**: 1) Stojmenovic/Milica; 2) milica.stojmenovic@carleton.ca; 3) Research Support/Data Analysis; 4) Faculty of Engineering and Design; 4) Post-Doctoral Fellow in Cognitive Science at Carleton University
### Researcher Training

<table>
<thead>
<tr>
<th>1D. Researcher Training</th>
<th>(Detailed instructions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>No training provided/Not applicable</td>
</tr>
<tr>
<td>Response</td>
<td>Describe any additional training the researcher(s) have (or will receive) to work with the participants.</td>
</tr>
<tr>
<td></td>
<td>Researcher will be trained</td>
</tr>
<tr>
<td></td>
<td>Researcher is trained</td>
</tr>
</tbody>
</table>

### Study Overview

#### Project Title

<table>
<thead>
<tr>
<th>2A. Project Title</th>
<th>(Detailed instructions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response</td>
<td>Title of Research Project</td>
</tr>
<tr>
<td></td>
<td>Learning with an electronic textbook</td>
</tr>
</tbody>
</table>

#### Study Goal

<table>
<thead>
<tr>
<th>2B. Study Goal</th>
<th>(Detailed instructions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response</td>
<td>What question will your research answer (1-2 sentences)?</td>
</tr>
<tr>
<td></td>
<td>I have developed a new e-textbook format that attempts to organize content in a way that emulates the structure of conceptual knowledge. I want to know if this method of organization facilitates the learning process.</td>
</tr>
</tbody>
</table>

#### Study Purpose and Benefits

<table>
<thead>
<tr>
<th>2C. Study Purpose and Benefits</th>
<th>(Detailed instructions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response</td>
<td>Study rationale: why should the research be pursued; what are the benefits, and to whom? (Benefits can be to research community, companies, or society in general.)</td>
</tr>
<tr>
<td></td>
<td>The benefit of this research is knowing whether or not organizing information in this way will help students learn. If this method of organizing content is successful in facilitating learning, this research could potentially benefit all students. Regardless of the outcome of this research, the results will be of interest to instructional material designers, educational psychologists, and instructors.</td>
</tr>
</tbody>
</table>

#### Dates of Recruitment/Participant Interaction

<table>
<thead>
<tr>
<th>2D. Dates of Recruitment/Participant Interaction</th>
<th>(Detailed instructions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not applicable/Secondary Data</td>
<td>When will you start recruiting participants? (DD/MM/YYYY)</td>
</tr>
<tr>
<td>Response: 15/01/2018</td>
<td>When will you stop interacting with participants? (DD/MM/YYYY)</td>
</tr>
<tr>
<td>Response: 31/08/2018</td>
<td></td>
</tr>
</tbody>
</table>

#### New or Previously Recorded Data

<table>
<thead>
<tr>
<th>2E. New or Previously Recorded Data</th>
<th>(Detailed instructions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>Primary Analysis (New Data are Collected)</td>
</tr>
<tr>
<td>Response</td>
<td>Does this research collect new data or analyse previously collected data (secondary analysis)? If the research involves secondary analysis, describe the coding of personal identifiers and indirect identifiers within data. Note: Tri-Council has defined anonymized</td>
</tr>
</tbody>
</table>
Directly Identifying Data
Secondary Analysis of Coded or De-identified Data

Additional Reviews
(X) No additional review
Departmental review
Grant council review

Response:

3. Funding and Approvals

Project Funding
Unfunded
(X) Tri-Council Funded
Other Award/Grant
Contract Funded
Personal Consulting or Personal Work
Scholarship

Who is funding this project? If applicable, include the funding source/agency/company, program, award name, and number (from CUResearch)

Response: NSERC Discovery Grant. Robert Biddle "Beyond Usable Security". Fund 315760

Researcher Funding (for research contracts and personal consulting only)

(X) Not applicable/Not contract funded research
No funds are paid directly to the researcher as personal income
The researcher will receive a portion of the funds as personal income
A copy of the contract/agreement has been submitted to the Research Compliance Office

For research contracts and personal consulting only: Is there a real or perceived conflict of interest and how will it be managed? How much funding (dollar amount and the percentage of the total) will the researcher(s) receive as income? Provide the title and date of any contracts. (The REB may review the contract.)

Response:

Minimal Risk Review Request

No
(X) Yes

Would you like to request this protocol be considered for minimal risk review? If so, please briefly justify. If not requesting a minimal risk review, leave this section blank. (The REB will use this information to make a decision as to whether this application will be reviewed at full board or via a delegated process).
### Additional Approvals Required

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is organizational permission required to conduct research (e.g., schools, employers, other universities, correctional services, aboriginal communities, and other data collection locations)? If conducting research in another country, is local permission required?</td>
<td>Indicate if permission has been secured and provide a copy of the permission. Research with biohazards or animals must also secure approval from the appropriate committee at Carleton University.</td>
</tr>
<tr>
<td>Organizational Permission</td>
<td></td>
</tr>
<tr>
<td>Visa/Travel Permits</td>
<td></td>
</tr>
<tr>
<td>Other REBs or Institutional Approvals</td>
<td></td>
</tr>
<tr>
<td>Biohazards</td>
<td></td>
</tr>
<tr>
<td>Animal Care Committee</td>
<td></td>
</tr>
<tr>
<td>Other (please specify)</td>
<td></td>
</tr>
<tr>
<td>Permission letters attached</td>
<td></td>
</tr>
<tr>
<td>Letters to follow</td>
<td></td>
</tr>
</tbody>
</table>

### Methods: Participants

#### Participant Interactions Overview

<table>
<thead>
<tr>
<th>Interaction Type</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directly interacting with participants</td>
<td>Briefly list what will happen to, or will be required of, the participants during the course of the research. (Only a project overview is requested here; methodology details are required in the first question of each section). If the research involves secondary analysis of data that has already been collected, the REB needs information about the original data collection to be confident data were collected ethically.</td>
</tr>
<tr>
<td>Interacting with participants online (e.g. online surveys)</td>
<td></td>
</tr>
<tr>
<td>Observing participants</td>
<td></td>
</tr>
<tr>
<td>Secondary Analysis of Data</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
</tbody>
</table>

Participants will be randomly assigned to one of two conditions determining the interface they will be asked to use during the experiment. The two interfaces are (1) a novel interface with a non-linear organization of content, and (ii) a standard electronic textbook interface, with a linear organization of content (see Appendices document for diagrams of the two interfaces). The content featured in the two interfaces will be identical. This content will be taken from the Human Memory chapter of an introductory psychology textbook. The procedure participants will be asked to follow will be identical across the two conditions. The remainder of this section describes this procedure.
Before the trial, participants will be asked to read and sign the informed consent form, and complete two questionnaires. The first questionnaire asks questions related to demographic information. The second is a sixteen-item multiple choice test assessing their background knowledge in cognitive science.

During the trial, participants will be asked to read the electronic textbook as if they were reading for a university course. Participants will have up to 20 minutes to finish reading, but they do not have to use all of this time. The length of time the participants spend reading will be recorded.

After the reading task, participants will be asked to complete a test measuring their ability to reproduce content they read about during the reading session. This test will feature eight short answer questions: four will be recall questions; three will be bridging-inference questions; and one will be a problem-solving question. After the post-test, participants will also be asked to complete two questionnaires: the Standard Usability Scale (SUS) and the CAP Perceived Learning Scale. The SUS is ten questions in length and gauges the participants’ feelings regarding the usability of the interface. The CAP Perceived Learning Scale is nine questions in length, and will assess participants’ feelings regarding the quality of their learning along three dimensions: psychomotor, affective, and cognitive.

Lastly, participants will be asked for any immediate comments on their experience, then thanked and compensated.

**Description of Participants**

**4B. (Detailed instructions)**

Describe the participants and any inclusion criteria. If applicable, describe any exclusion criteria. If using a separate sample of control participants, describe this group.

**Response:** Participants will be university students. Any student will be eligible, provided they can use a mouse and keyboard and have normal and/or corrected colour vision.

**Number of Participants**

**4C. (Detailed instructions)**

What is the number of participants requested? If multiple groups of participants are involved, breakdown by participant type. Provide a justification including a statistical rationale if appropriate.
### Vulnerable Population

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not Vulnerable Population</td>
<td>Describe any pre-existing vulnerabilities associated with the proposed participant group(s) that may cause additional risks. Describe the associated risks and mitigation strategy.</td>
</tr>
<tr>
<td>Vulnerable Population</td>
<td></td>
</tr>
</tbody>
</table>

**Response:**

### Participant Relationship to Researcher

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>No previous relationship</td>
<td>Describe any relationship that exists between the participants and the research team or any recruiting party or sponsor. Indicate how relationships will be managed so there is no undue pressure put on participants.</td>
</tr>
<tr>
<td>Instructor-Student</td>
<td></td>
</tr>
<tr>
<td>Client</td>
<td></td>
</tr>
<tr>
<td>Employee</td>
<td></td>
</tr>
<tr>
<td>Friends/Family</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
</tbody>
</table>

**Response:**

### Conflict of Interest

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>No conflicts</td>
<td>Describe any real or perceived conflicts of interest for any research team member that could affect participant welfare. If so, describe it here and indicate how it will be managed.</td>
</tr>
<tr>
<td>Financial</td>
<td></td>
</tr>
<tr>
<td>Commercial Entity Benefits</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
</tbody>
</table>

**Response:**

### Recruitment Methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posters</td>
<td>Describe each step of how participants will be recruited. This includes how contact information is obtained, how participants will be made aware of the study, where will recruitment materials be located, and how participants can express their interest. Provide a copy of the recruitment material(s) including any oral scripts, recruitment posters, recruitment emails, social media postings etc.</td>
</tr>
<tr>
<td>Social Media</td>
<td></td>
</tr>
<tr>
<td>Online Panels (e.g. Qualtrics)</td>
<td></td>
</tr>
<tr>
<td>Student Participant Pool (e.g. SONA)</td>
<td></td>
</tr>
<tr>
<td>Emails</td>
<td></td>
</tr>
<tr>
<td>Letters</td>
<td></td>
</tr>
<tr>
<td>Telephone</td>
<td></td>
</tr>
<tr>
<td>Snowballing</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
</tbody>
</table>

**Response:**

We will post Call-for-Participation poster on some of the 20 Official University Bulletin Boards across campus, on the front door of 2110 HCI Building, and also on the Carleton Research Participants Facebook page. Participants will also come from SONA: Cognitive Science Research. For the social media posting, we will use the Call for Research.
<table>
<thead>
<tr>
<th><strong>Location of Recruitment</strong> (Detailed instructions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not applicable</td>
</tr>
<tr>
<td>X Carleton</td>
</tr>
<tr>
<td>Other Canadian School/University</td>
</tr>
<tr>
<td>Canada</td>
</tr>
<tr>
<td>Online</td>
</tr>
<tr>
<td>Other</td>
</tr>
</tbody>
</table>

List all recruitment locations. If some locations require permission prior to recruitment, indicate if permission has been secured.

**Response:**

<table>
<thead>
<tr>
<th><strong>Third Parties in Recruitment</strong> (Detailed instructions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X Not applicable</td>
</tr>
<tr>
<td>Third Parties</td>
</tr>
</tbody>
</table>

If using third parties to recruit, indicate who is doing the recruitment and how it will be accomplished. Does the third party have contact information for the participants? If not, how will it be acquired?

**Response:**

<table>
<thead>
<tr>
<th><strong>Recruitment risks to Participants</strong> (Detailed instructions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X No risks / Not applicable</td>
</tr>
<tr>
<td>Mild risks</td>
</tr>
<tr>
<td>Moderate risks</td>
</tr>
<tr>
<td>High risks</td>
</tr>
</tbody>
</table>

Describe any risks to participants during the recruitment phase.

**Response:**

<table>
<thead>
<tr>
<th><strong>Recruitment risks to Researcher</strong> (Detailed instructions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X No risks / Not applicable</td>
</tr>
<tr>
<td>Mild risks</td>
</tr>
<tr>
<td>Moderate risks</td>
</tr>
<tr>
<td>High risks</td>
</tr>
</tbody>
</table>

Describe any risks to the research team during the recruitment phase.

**Response:**

<table>
<thead>
<tr>
<th><strong>Benefits</strong> (Detailed instructions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Direct Benefits to Participants / Not applicable</td>
</tr>
<tr>
<td>X Direct Benefits to Participants</td>
</tr>
</tbody>
</table>

Describe any direct benefits to the research participants as opposed to society or knowledge.

**Response:** Participants will be learning about the research process, seeing a new interface design (if they are in the novel interface condition), and potentially learning something about human memory. It is my belief that these are all potential direct benefits to participants. However, if the reviewer prefers that we say there are no direct benefits, we will do so.
### Compensation

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Compensation/Not applicable</td>
<td>Describe all compensation/remuneration and indicate when participants will receive the compensation. What is the monetary value of the compensation/remuneration? What happens to the compensation if a participant withdraws?</td>
</tr>
<tr>
<td><strong>X</strong> Money / Gift Card</td>
<td>Response: We will use the Cognitive Science SONA. Participants who come through SONA will receive 1% course credit. Non-SONA will receive $15 compensation. The money will be paid in cash at the conclusion of the study. Participants will sign a form indicating that they have received $15 compensation.</td>
</tr>
<tr>
<td>Reimbursement of Travel Expenses</td>
<td></td>
</tr>
<tr>
<td>Refreshments</td>
<td></td>
</tr>
<tr>
<td><strong>X</strong> Course Credit</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
</tbody>
</table>

### Methods: Informed Consent

#### Obtaining informed consent

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>X</strong> Signed consent</td>
<td>Describe the method for obtaining informed consent from the participants. If signed consent is not used, justify the alternative method chosen. Include a copy of the consent materials.</td>
</tr>
<tr>
<td>Online consent</td>
<td></td>
</tr>
<tr>
<td>Oral consent</td>
<td></td>
</tr>
<tr>
<td>Implied consent</td>
<td></td>
</tr>
<tr>
<td>Parent/Guardian consent</td>
<td></td>
</tr>
<tr>
<td>Assent</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
</tbody>
</table>

#### Deception

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Disclosure (i.e. no deception)</td>
<td>Describe the deception and/or partial disclosure (e.g. what information is withheld). Why must it be used and why not an alternative research method? Describe the magnitude and likelihood of harm. Deception requires debriefing and secondary consent forms. The secondary consent form allows the participant to consent to the use of their data when they have been informed of the true nature of the study. Partial disclosure requires a debriefing form.</td>
</tr>
<tr>
<td><strong>X</strong> Partial Disclosure</td>
<td></td>
</tr>
<tr>
<td>Mild Deception</td>
<td></td>
</tr>
<tr>
<td>More than Mild Deception</td>
<td></td>
</tr>
</tbody>
</table>

#### Debriefing

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not applicable/not required</td>
<td>Will participants be debriefed? If this is the case, describe when and how participants will be debriefed. (Include a copy of any documents that will be provided to participants). Describe any risks during debriefing and how they will be mitigated. According to Tri-Council, debriefing is required in all cases of deception: <a href="http://www.pre.ethics.gc.ca/eng/policy-politique/initiatives/tcp2-CUREB">http://www.pre.ethics.gc.ca/eng/policy-politique/initiatives/tcp2-CUREB</a> A&amp;B: Research Ethics Protocol Form (Version April 2016) 8/14</td>
</tr>
<tr>
<td><strong>X</strong> Participants will be debriefed</td>
<td></td>
</tr>
</tbody>
</table>
**Withdrawal Procedures**

<table>
<thead>
<tr>
<th>X Participants can withdraw</th>
<th>Not applicable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants can only withdraw during the study session</td>
<td>Describe the procedures for a participant to withdraw. What will happen to data from participants who withdraw? Describe any deadlines and limitations on withdrawal.</td>
</tr>
<tr>
<td>Special withdrawal procedures</td>
<td>Response: Participants can withdraw from the study by verbally informing the researchers that they would like to withdraw. If a participant withdraws, their data is destroyed. Withdrawal after the study is not possible because immediately following each study session the participant’s data will be encoded anonymously. This means that, after encoding has taken place, there will be no way to identify data using a name. Participants will still be paid in full should they decide to withdraw.</td>
</tr>
</tbody>
</table>

**7. Methods: Data Collection**

**Data Collection Methods**

<table>
<thead>
<tr>
<th>X Questionnaires / Surveys</th>
<th>Describe the method of data collection being used and provide details of any instruments used. Breakdown by phases, participant groups, or types if required. If data collection is being done online, visit the detailed instructions for full details on what information the REB requires. (CUREB requires a copy of any questionnaires, surveys, or interview guides).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interviews</td>
<td>Response: Participants will complete surveys using LimeSurvey software running on a desktop computer. Please see the procedure and the Appendices document for a more detailed description of the survey procedure and the surveys themselves.</td>
</tr>
<tr>
<td>Focus Groups</td>
<td></td>
</tr>
<tr>
<td>Oral and/or Visual Stimuli</td>
<td></td>
</tr>
<tr>
<td>Equipment and/or software testing</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
</tbody>
</table>

**Location of Data Collection**

<table>
<thead>
<tr>
<th>X Carleton</th>
<th>Where will the participant be during data collection?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada (other than Carleton)</td>
<td>Response: HCI Room 2110</td>
</tr>
<tr>
<td>Workplace</td>
<td></td>
</tr>
<tr>
<td>Public venue</td>
<td></td>
</tr>
<tr>
<td>Online</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
</tbody>
</table>

CUREB A&B: Research Ethics Protocol Form (Version April 2016)
| **Photography or Recordings** | X Not applicable | If the participant will be photographed, video-recorded or audio-recorded, indicate how the data will be acquired and protected (if applicable). |
| **7C. (Detailed instructions)** | Photographs | Response: |
| | Audio Recording | |
| | Video Recording | |

| **Translation or Transcription** | X Not applicable | If you require the services of a translator or transcriber, describe what services you will use and how you will interact with the translator and/or transcriber. If a confidentiality agreement will be used, include a copy. |
| **7D. (Detailed instructions)** | Translation | Response: |
| | Transcription | |
| | Researcher will translate or transcribe | |

| **Online data collection** | Not applicable | Describe the technology platform used to collect online data. |
| **7E. (Detailed instructions)** | Carleton-based server | Describe the security of the data. Will participant IP addresses be recorded? Are there any special limits to privacy? |
| | Commercial server (based in Canada) | Response: We will use the server at mvp.soft.carleton.ca located at 2183 HCI Building. This server is secure: it has daily anti-malware checks, is password-protected and is located secure server room. |
| | Commercial server (outside Canada) | |
| | Other | |

| **Bio-interactions** | X Not applicable | Describe the apparatus and methods to acquire biological specimens or fluids. (e.g., blood, saliva, tissue samples.) How will specimens be safely stored and destroyed? If any will be kept, explain why, how and for how long. |
| **7F. (Detailed instructions)** | Biological specimens/fluids | Response: |

| **Bio-instruments** | X Not applicable | Bio-instruments touch or send energy into the body. (e.g., electrodes, MRI/X-ray.) Describe the apparatus and its use. If applicable, explain any significant risks and compare the dose (e.g., electrical, radiation) to established safety standards. |
| **7G. (Detailed instructions)** | Bio-instruments | Response: |

| **Bio-interventions** | X Not applicable | Describe the apparatus and methods associated with the bio-intervention. (e.g., drug, stress, medical devices.) Explain any risks to the participants and compare it to established safety standards. |
| **7H. (Detailed instructions)** | Bio-interventions | Response: |
### Risk of Psychological Harm

7I. (Detailed instructions)

<table>
<thead>
<tr>
<th>Not applicable/No risks</th>
<th>Explain the rationale for your selection, and, if applicable, explain the nature, magnitude and probability of the risks and how they will be mitigated.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X Mild risks</td>
<td>Response: Participants will be engaged in a learning task using a computer, and then they will be given a test on what they have learned. There is a mild risk of mental fatigue, but the risks will not exceed what students encounter on a regular basis.</td>
</tr>
<tr>
<td>Moderate risks</td>
<td></td>
</tr>
<tr>
<td>High risks</td>
<td></td>
</tr>
</tbody>
</table>

### Risk of Physical Harm

7J. (Detailed instructions)

<table>
<thead>
<tr>
<th>Not applicable/No risks</th>
<th>Explain the rationale for your selection, and, if applicable, explain the nature, magnitude and probability of the risks. Describe how they will be mitigated.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X Mild risks</td>
<td>Response: The potential physical risks here (e.g. eye strain, repetitive stress) are no greater than the physical risks that accompany normal computer use.</td>
</tr>
<tr>
<td>Moderate risks</td>
<td></td>
</tr>
<tr>
<td>High risks</td>
<td></td>
</tr>
</tbody>
</table>

### Risk of Social and/or Economic Harm

7K. (Detailed instructions)

<table>
<thead>
<tr>
<th>Not applicable/No risks</th>
<th>Explain the rationale for your selection, and, if applicable, explain the nature, magnitude and probability of the risks and how they will be mitigated.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X Mild risks</td>
<td></td>
</tr>
<tr>
<td>Moderate risks</td>
<td></td>
</tr>
<tr>
<td>High risks</td>
<td></td>
</tr>
</tbody>
</table>

### Incidental Findings

7L. (Detailed instructions)

<table>
<thead>
<tr>
<th>Not applicable/No incidental findings anticipated</th>
<th>Describe possible incidental findings (unanticipated discoveries that relate to the welfare of participants or others) and how they will be managed. Examples are becoming aware of abuse of a child, or imminent harm to a participant or third party. Your approach to managing any findings should also be described in the informed consent.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low probability</td>
<td>Response:</td>
</tr>
<tr>
<td>High probability</td>
<td></td>
</tr>
</tbody>
</table>

### Methods: Data Storage and Analysis

8. Identifyability of stored data

8A. (Detailed instructions)

<table>
<thead>
<tr>
<th>Anonymous</th>
<th>Describe the identifyability of research data, including how pseudonyms will be assigned, if applicable. If there are different levels of anonymity for different groups, describe each level here.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pseudonyms/Coded</td>
<td>Response: Real names are collected only for the purposes of awarding compensation but survey data will not be linked to identifiers. Data will be encoded by participant number; that is,</td>
</tr>
<tr>
<td>Real participant names with data attributable</td>
<td></td>
</tr>
<tr>
<td>X Real participant names with data non-attributable</td>
<td></td>
</tr>
</tbody>
</table>
### Identifiability of published data

<table>
<thead>
<tr>
<th>X</th>
<th>Anonymous</th>
<th>Aggregate data only</th>
<th>Pseudonyms/Coded</th>
<th>Real participant names with data attributable</th>
<th>Real participant names with data non-attributable</th>
<th>Different levels of anonymity for different groups of participants</th>
</tr>
</thead>
</table>

Describe the identifiability of data that will appear in publications, including how pseudonyms will be assigned, if applicable. If there are different levels of anonymity for different groups, describe each level here.

Response: Participant names will be collected, but the data will be stored without any reference to these names, and there will be no way to recover a link to any names after the fact. That is, the published data will be fully anonymous.

### Data Storage (during the project)

<table>
<thead>
<tr>
<th>X</th>
<th>Encrypted</th>
<th>Password-protected</th>
<th>Anonymous data</th>
<th>Physical documents</th>
<th>Other</th>
</tr>
</thead>
</table>

How are data being stored and kept safe? Provide details for electronic data and hard copies if applicable.

Response: The data will be stored on a secure Carleton server located at 2183 HCI Building (mvp.soft.carleton.ca), and a secure, password-protected laptop in an encrypted, password-protected directory. Both the server and the laptop run Linux-based operating systems (Ubuntu and Debian, respectively) with up-to-date security patches.

### Data Disposition (after the project)

<table>
<thead>
<tr>
<th>X</th>
<th>Retained by the researcher(s)</th>
<th>Stored in a depository</th>
<th>Archived</th>
<th>Shared with research agreement</th>
<th>Shared publicly</th>
<th>Returned</th>
<th>Destroyed</th>
</tr>
</thead>
</table>

After project completion, describe how the data (confidential and non-confidential aspects) will be stored for future use, made publicly available, archived, returned to participants, or destroyed. If shared, with whom? Describe any restrictions on access. If destroyed, how long will data be kept? Will personal identifiers and the actual data be destroyed at different phases? Will participant contact information be kept for future studies?

Response: Withdrawal after the study has completed will not be permitted because participant data will be encoded anonymously, making it impossible to identify which data belongs to which participant.

Data will stored long-term on a secure server at Carleton in 2183 HCI Building, and on a secure laptop. The data will be held until...
### Data Breach Risks

<table>
<thead>
<tr>
<th>Risk Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild risk to participants</td>
<td>Describe how likely a data breach is to occur and how it could affect the participants. If risks are significant, how will they be mitigated?</td>
</tr>
<tr>
<td>Moderate risk to participants</td>
<td>Response: Data is stored anonymously.</td>
</tr>
<tr>
<td>High risk to participants</td>
<td></td>
</tr>
</tbody>
</table>

### 9. Declarations

#### Supervisor Approval

- **9A.** Not applicable
- **X Supervisor Approved**
- Supervisor has not approved

For student projects, please indicate the date that the supervisor approved the application to go forward for REB approval. (All CUREB-A applications must copy the supervisor when submitting an application to ethics@carleton.ca. CUREB-B applications are automatically submitted to the supervisor through the online application system; therefore, the student must inform his/her supervisor that the application has been submitted so the protocol may be approved by the supervisor and received by CUREB-B).

**Response:** Supervisor approval was obtained on Thursday, December 21, 2017.

#### Declaration #1

**9B.** I agree

This ethics application accurately describes the research project or scholarly activity that I plan to conduct.

#### Declaration #2

**9C.** I agree

No recruitment or data collection for this protocol will commence before ethics clearance.

#### Declaration #3

**9D.** I agree

No changes will be made to the research project as described in this protocol without receiving clearance from the Research Ethics Board.

#### Declaration #4

**9E.** I agree

The Research Ethics Board will be notified immediately of any alleged or real ethical breaches or concerns, adverse events, or participant complaints that arise during or after the course of this research project.
### 10. Comments

<table>
<thead>
<tr>
<th>Comments (optional)</th>
<th>Do you have any comments or suggestions on the form</th>
</tr>
</thead>
<tbody>
<tr>
<td>10A.</td>
<td>Response:</td>
</tr>
</tbody>
</table>