Traversal Patterns for Content Designed Web Environment

PERWAIZ B. ISMAILI
RICHARD M. GOLDEN
School of Behavioral and Brain Sciences
University of Texas at Dallas
UNITED STATES OF AMERICA
Pbi021000@utdallas.edu

Abstract: - This paper describe new ways of observing effects of content presentation and domain knowledge upon navigation behaviors by designing web (hypertext) presentation format that adheres to content design inspired by discourse and text comprehension literature. More specifically logical connections between web pages at macro level for all web sites were constructed meticulously and kept consistent across all three knowledge domains. Twenty undergraduate Psychology students participated in this preliminary study in investigating domain knowledge and content presentation influence on Hypertext (web) site traversal behavior. Classical data analysis (ANOVA) were used to explore these qualitative phenomena. Contrary to our belief expertise difference were not significant for total number of web pages (nodes) visited or overall time spent on each knowledge domain web sites. However, these differences were significantly strong for super-ordinate nodes, nodes with more semantic (logical) connections and irrelevant nodes.

Key-Words: - Hypertext, Navigation, User Behavior, Content Design, Expertise, Scientific Text, Web Design

1 Introduction

When reading, not all the information that is needed to comprehend a passage is presented in the text. Moreover, the inclusion of all such details would greatly obstruct the comprehension and writing processes. Hence, it is beneficial on the behalf of the author and the reader to omit these superfluous details. Readers depend on structural patterns at the local and global level in order to recognize the type of text and integrating relevant parts for better comprehension [1, 2], and construct a coherent understanding of a text [3, 4]. Inferences make assumptions about the reader's internal representation of the text. Such structural patterns may be modeled as semantic networks where ideas are represented as nodes and relationships between ideas are represented as connections [5]. Furthermore, super-ordinate nodes [6] and nodes with more semantic connections [7] have been found to be more memorable in number of studies.

Most research in text comprehension, however, has focused upon texts which are intended to be read in a sequential sentence by sentence and paragraph by paragraph manner. Such texts which are in print-like format will be referred to as linear texts. More recently, important alternative versions of texts such as hypertext have become widely used as a result of the wide-spread use of the World Wide Web as a means of communicating complex information. The "web" may be regarded as a "new space" where authors can "guide" readers through wealth of diverse information sources and types [8] that were never seen before in human history. However, among other things, there are associated costs associated with being ''lost in hyperspace" which have been technically referred to as costs associated with 'cognitive overload', [9] defines cognitive load as the amount of effort needed to identify and, coherently integrate information from the text. Not only do issues of disorientation and lack of experience with the technology constitute barriers but also differences in the way textual elements are displayed and connected play an influential role in aiding or hindering navigation performance of the hypertext reader. Indeed, navigation through websites is almost never sequential and thus many reading comprehension situations involving the worldwide web are most effectively characterized as texts which are non-sequential and which will be referred to as nonlinear text or hypertext.

Hypertext has been attributed to reading comprehension difficulties encountered by web navigators. Furthermore, past research has demonstrated that non-linear, hypertext Web presentations may lead to decreased free recall, and learning of factual information compared to traditional, print-like linear Web designs [10, 11]. On the other hand, a minimal amount of nonlinearity might be expected to help readers in understanding the underlying linguistic message of a concept [12]. Indeed, some authors have argued that the nonlinear nature of hypertexts might be effectively exploited for the purposes of improving reading comprehension. Recent evidence suggests that nonlinear designs may facilitate learning of the interconnections (structure) of the presented information [13, 14]. This latter goal might be achieved by designing hypertexts which more effectively exploit the semantic organization (interconnections) which the author wishes to communicate to the reader.

For example, Eveland and Dunwoody found that disorientation can be reduced by providing cues and therefore minimizing nonlinearity helps readers with text coherence. Studies of different structural organization (presentation) and found that recall was better for only certain types of tasks while using certain type presentations. The role of structural organization in the hypertext learning environment was also evident in a study [13]. Furthermore, different navigation patterns have been noticed with varying levels of reader knowledge [15]. In addition, domain knowledge plays a significant role in predicting the recall and improved comprehension for certain type of readers [16]. However, it must be noted that research investigating the role of domain knowledge in supporting hypertext comprehension performance is not conclusive and sometimes contradictory. For example, some research show positive effects of prior knowledge [17], whereas other research has shown no effects of prior knowledge in assisting comprehension for low knowledge readers in hypertext environments. For low knowledge readers, the hypertext environment may sometimes be perceived as "confusing" since there may be a large cognitive load on the reader in decoding and incorporating the textual information into existing knowledge base for learning. Therefore, a linear text environment may be more appropriate for supporting learning and comprehension.

Saremi & Montazer [18] rightly pointed out that from an electronic commerce (e-commerce) perspective the necessity of more effective web sites has become a major concern in the industry since the web content and its structure plays a critical role customers' perception of products and decision making. Furthermore, educational institutions are soon realizing the potentials of learning on the web and are trying to create "Scaffolding Learning paths (SLP)" [19], knowledge maps for fragmentary facts [20], frameworks and curriculum designs that enables learning such as in the technical area of computer programming [21].

2 Specific Aims

Following the recent work on comprehension in hypertext environment, this study investigated whether traversal behavior is influenced differently when reading takes place in two different albeit qualitatively similar types of hypertext (web) environments (semantically [logically] organized versus fully connected), relative to a linear text environment. Meticulous content design process was implemented in creating logically organized web sites. Furthermore, the proposed research investigated the influence of domain knowledge expertise by considering twenty undergraduate Psychology students for this preliminary study. These students navigated web sites which are in the knowledge domains of Psychology, Neuroscience, and (as a control condition) Archeoastronomy. Traditional (analysis-of-variance) data analysis methodologies was used to establish the presence or absence of qualitative phenomena.

2.1 Content

Readers depend on structural patterns at for better comprehension [1], and coherent understanding of a text [3]. Inferences make assumptions about the reader's internal representation of structure for the read text which may be modeled as semantic networks where ideas are represented as nodes and relationships between ideas are represented as connections or links.

Content analysis has been used in communication, media sciences, and by psychologists to decipher underlying messages, cues for visuals, marketing advertisement and thought processes of individuals such as authors and politicians. It has been shown that people respond to a complex content or stimuli quite differently than they do to a simpler stimulus [22]. Watt & Welch [23] attempted to define 'simple' has having an ordered, predictable and structured characteristics, while complex being disordered, unpredictable and random. Therefore in order to create a structured, predictable web site meticulous content design, inspired by text and discourse findings, were used in creating meaningful relationships (links) between the web pages (nodes) of the semantic-based web site for all knowledge domains.

Towards this end, paragraphs of expository (scientific) texts were designed each paragraph containing two sentences. Each paragraph represents a web-page of a particular knowledge domain web-site. The first sentence of the paragraph serves as the title sentence while the following sentence is an elaboration of the preceding topic sentence. There were three knowledge domain areas: Psychology, Neuroscience and (as a neutral) Archeoastronomy. The semantic associative relationships among the topic sentences of each of these paragraphs had qualitatively same patterns across all three knowledge domains. Please see Table 1 for Psychology web page content sample, Table 2 for Archaeoastronomy (Neutral) content, and Figure 1 for the content design template. Furthermore, meticulous attention was given in order to ensure the balanced content design such as: word count, level of difficulty etc., so that Stimuli characteristics across all three knowledge areas (Psychology, Neuroscience and Archeoastronomy) are qualitatively equivalent with each domain consisting of five facts, two intermediate conclusions, two irrelevant fact nodes and a final conclusion. For the ease of viewing, the five 'fact' nodes are shown in black font, the two 'intermediate' and 'final' conclusion nodes in blue, and, 'irrelevant' fact nodes in red. The two Irrelevant Fact nodes were mutually associated however not with the main content semantic relationships (see Table 3). For example, the psychology 'main' content chain is about context –dependent processing models whereas, the Irrelevant-fact nodes chain explores the memory storage details.

Table 1: Psychology text read by participants. Like other knowledge domain (Neuroscience & Archeoastronomy) web sites, Psychology web site consist of ten web-pages: Five web pages for Facts 1-5, Two web-pages for the two Irrelevant Fact nodes, and, three web-pages presenting two Intermediate conclusions and a Final conclusion.

Psychology Content
F1 (Fact 1)- Craik and Lockhart proposed that stimuli subjected to semantically processed stimuli ("deeply processed stimuli") will be more memorable than perceptually processed stimuli ("shallow processed stimuli"). In other words, memory performance is dependent on the "levels of processing" (LOP): Stimuli processed at deeper levels are more effectively recalled.
F2 (Fact 2)- Tulving proposed that memory performance increases as the similarity between encoding and retrieval contexts increases. This principle has been referred to as Tulving's Encoding Specificity Principle (ESP) because it specifies how encoding conditions interact with retrieval conditions to influence memory performance.
F3 (Fact 3)- Although many scientific studies have reported experimental results supporting both LOP and ESP, some research has identified situations where LOP theory fails while ESP theory is successful. For example, "shallow" rhyming encoding strategies are highly effective learning strategies when retrieval cues are also rhyme-oriented."
F4 (Fact 4)- Although these early context-independent memory models were highly influential, later experimental findings showed these early models could not account for experimental findings as effectively as context-dependent models such as LOP and ESP, since they ignored encoding and retrieval factors. Early context-independent memory models assumed items stored in short-term memory will be remembered differently from items stored in long-term memory.
F5 (Fact 5)- The psychologist Abernathy (1940) studied context effects on test performance and showed memory recall performance improved when students were tested in the same physical environment as the environment in which they received test instructions. The finding of this experiment provides another example consistent with predictions of context-dependent memory models.
IF1 (Irrelevant Fact 1)- One traditional theory of long-term memory storage assumes that items that are retained in short-term memory for extended periods of time will eventually be transferred to long-term memory. A classical example is "rehearsal" where the verbal repetition of an item multiple times aids in transferring that item into long-term memory.
IF2 (Irrelevant Fact 2)- Associative theories of long-term memory storage postulate that items with more semantic connections are more effectively retained. For example, in order to improve memory performance for the irrelevant

word pair "dog-bicycle", it is helpful to construct a rich mental image of a dog riding a bicycle.
IC1 (Intermediate Conclusion 1)- In the current scientific literature, many cognitive scientists tend to agree that the Encoding Specificity Principle (ESP) provides a more comprehensive explanation of experimental findings as well as a more theoretically satisfying model of memory phenomena relative to Levels of Processing (LOP) theory. Accordingly, ESP models are considered preferable over LOP models.
IC2(Intermediate Conclusion 2)- In the current scientific literature, many cognitive scientists now agree that it is important to include context-dependent factors in memory models in order to provide adequate explanations of empirical phenomena. It is now widely recognized that memory models need to incorporate context-dependent factors.
FC(Final Conclusion)- In summary, ESP models are preferable to LOP models because they incorporate interactions between encoding and retrieval factors. Therefore, ESP models are context-dependent memory models which emphasize that memory performance is specific to the environment in which it is embedded.

Figure 1: All three knowledge domain (Psychology, Neuroscience & Archeoastronomy) area content were kept consistent for semantic (logical) associations between nodes and in consisting five 'fact' nodes (F1 – F5), two 'irrelevant fact' (IF-1 & IF-2) nodes, and, three nodes presenting two 'intermediate conclusions' (IC-1 & IC-2) and a 'final conclusion' (FC).

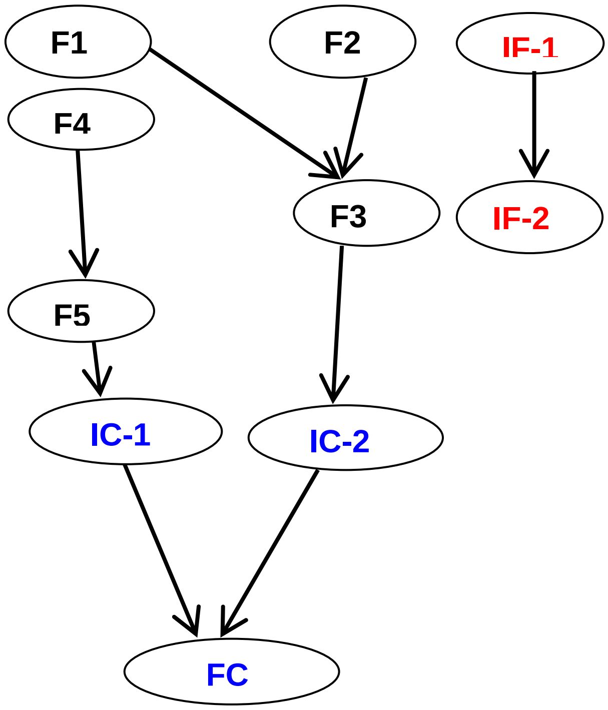


Table 2: Archaeastronomy (Neutral) text read by participants. Like other knowledge domain (Neuroscience & Psychology) web sites, Neutral content web site consist of ten web-pages: Five web pages for Facts 1-5, Two web-pages for the two Irrelevant Fact nodes, and, three web-pages presenting two Intermediate conclusions and a Final conclusion. It must be noted that being a control text it, the idea behind having such a text was to make sure that it does not give any unfair advantage to either one of the psychology or neuroscience group without compromising the content design structure and stimuli characteristics.

Archaeoastronomy (Neutral) Content
F1 (Fact 1)- Archaeologists and ethnographers look at various historical sources such as written records and architectural structures to study ancient beliefs and cultures. For example, archaeologists discovered the "Rosetta stone" which possessed inscriptions of the same passage in multiple languages leading to a scientific breakthrough in deciphering hieroglyphic writings.
F2 (Fact 2)- Astronomy is the study of celestial bodies and phenomena through the use of observational methodologies. For example, astronomers study the distribution of matter in outer space by examining the motion, location, composition, energy-emissions, and velocity of planets, stars, meteor showers, galaxies, supernovas, and other extra-terrestrial matter.

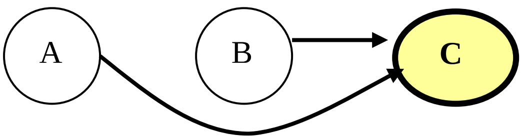
F3 (Fact 3)- Archaeoastronomy is a field of scientific inquiry which incorporates scientific methodologies from the fields of Archaeology and Astronomy. This new area of study investigates ancient people, their observations, interpretations and achievements with regards to the distribution and motion of celestial objects within our solar system and beyond.
F4 (Fact 4)- Archaeoastronomers have unearthed ancient structures which revealed that ancient civilizations were actively interested in observing and studying the stars, planets and other celestial bodies. Such structures have been identified by scientists throughout the world including locations in the USA, United Kingdom, Ireland, and Egypt.
F5 (Fact 5)- In addition to incorporating scientific methodologies from the fields of Archaeology and Astronomy, Archaeoastronomists sometimes use additional heuristic methodologies. For example, comparing structures at different sites through the use of discovered charts of stars and constellations at those sites in order to establish when those structures were utilized.
IF1 (Irrelevant Fact 1- Thermoluminescence dating of structures works because radioactive elements in the soil (possibly generated by cosmic rays) become less radioactive over time. Structures built by ancient civilizations can be dated using advanced thermoluminescence dating technology which involves using a measurement of the amount of radiant energy emitted from a structure to estimate the age of the structure.
IF2 (Irrelevant Fact 2)- Thermoluminescence dating has been used successfully by archaeologists to identify approximate construction dates of a number of ancient structures. For example, this method was used to date the Helleniko pyramid to 2730 B.C. with error factor of ± 720 years, and also for the dating of the Ligourio pyramid to 2260 B.C. with error factor of ± 710 years.

IC1(Intermediate Conclusion 1) - Archaeoastronomy is a field of scientific inquiry which incorporates scientific methodologies from the fields of Archaeology and Astronomy for the purpose of understanding the beliefs and culture of ancient civilizations. Researchers with training both the fields of Archaeology and Astronomy are well-positioned for pursuing research in Archaeoastronomy.
IC2(Intermediate Conclusion 2) - Archaeoastronomy is a field of scientific inquiry which not only incorporates scientific methodologies from established scientific fields of inquiry, but additionally uses heuristics for the purpose of understanding the nature of ancient civilizations. Thus, some Archaeoastrononomical findings may not be as reliable as findings from other scientific domains.
FC(Final Conclusion)- In summary, Archaeoastronomy investigates the beliefs and culture of ancient civilizations through a variety of scientific methodologies including those utilized by Archaeologists and Astronomers. However, some of these alternative methodologies which are fundamentally heuristic in nature will tend to limit the precision of the conclusions.

In reference to content design structure as shown in Figure 1, all three knowledge areas (Psychology, Neuroscience & Archeoastronomy) area content were kept consistent for semantic (logical) associations between nodes and in consisting five 'fact' nodes (F1 – F5), two 'irrelevant fact' (IF-1 & IF-2) nodes, and, three nodes presenting two 'intermediate conclusions' (IC-1 & IC-2) and a 'final conclusion' (FC). These logical associations provide coherence to text structure which is proposed to be a possible mental presentation of a reader. Additionally, 'F3' and 'FC' are considered as 'link nodes' since more than one associative links terminates there, while, 'IC-1', 'IC-2' and 'FC' were considered as 'super-ordinate' nodes in this study.

The first step in the comprehension process is achieved by a reader striving to discover the super-ordinate concept or thought in the text that encompasses all events in the text, in other words, incorporates the higher level concise summary of presented information. Therefore, text read in IC-1, IC-2 condensed two lines thoughts respectively. For example, from psychology content domain, IC-1 summarized the thoughts on levels of processing while IC-2 focused on condensing texts on context dependence. Final node (FC) further aggregated the thoughts presented in both IC-1 and IC-2. On the other hand, for this study, link nodes are defined as nodes that provide logical relationships with the preceding web pages by binding, validating or regenerating information presented on those previously read web pages. For example, consider the following link node depiction and node content:

Figure 2: Sample depiction of a link node. In this example node 'C' is serving as a binding node for both nodes (web pages) A and B. Semantic (logical) connections are between A → C and B → C.



A → *'Archaeology is a scientific study of past cultures and the way they lived.'*

B → *'Astronomy is a scientific study of celestial bodies and phenomenon.'*

C → *'Archaeoastronomy combines archaeology and astronomy in studying past cultures.'*

Here node '**C**' is considered a link node since it provides binding function and serves as a conjecture for nodes A and B.

Table 3: Stimuli characteristic of all three presented scientific knowledge domains (Psychology, Neuroscience & Archeoastronomy [Neutral])

Characteristics	*Psychology*	*Neuroscience*	*Neutral*
Avg Word Count/website	47.9	47.9	48.7
Avg Flesch-Kincaid/website	18.01	18.35	18.7
Avg Words per sentence per website	23.95	23	23.5
No. of Sentences per webpage	2	2	2
No. of Webpage per Website	12	12	12

In order to have balanced stimuli content across all three domains, along with other measures, Flesch-Kincaid scores were also used which rates text on a U.S. school grade level by using the following formula:

$$(.39 \times ASL) + (11.8 \times ASW) - 15.59$$

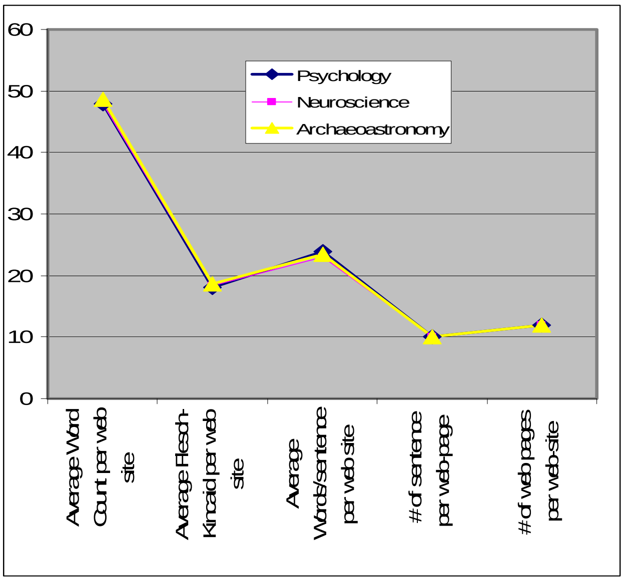
Where:

ASL = average sentence length (the number of words divided by the number of sentences)

ASW = average number of syllables per word (the number of syllables divided by the number of words)

A score of 8.0, for example, means that an eighth grader can understand the document. As it is shown in Table 3 that content read by all undergraduate students across all three domain knowledge was equated to be at level 18, which equates for advance graduate level class course content.

Figure 3: Stimuli characteristic graph was used to visualize any difference between the content of all three knowledge domains. As it is seen there were no apparent differences found between psychology, neuroscience and Archeoastronomy (neutral) experimental texts.



2.2 Hypertext (Web) Sites

Towards this end, paragraphs of expository (scientific) texts were designed each paragraph page. Each web site was organized into five fact topic sentences, two irrelevant fact nodes and two intermediate conclusions, and a final conclusion, which were represented as separate web pages. Each content presentation (Linear, Meshed-hypertext, Semantic-based) had a navigation bar on the left hand side of the web pages with buttons to go directly to the introduction page, the end page, and all the interconnected nodes, with some variations.

Although outside the scope of this paper, the linear navigation presentation format (Figure 4) was designed to encourage subjects to traverse sequentially while in the Meshed-hypertext environment (Figure 5), subjects were free to move around between pages using navigation bar on the left hand side which provided a meaningful image of the organization of the web site's connectivity. Although, the semantic-based web site (Figure 6) had a similar layout as the meshed-hypertext, subjects were allowed to move among all pages, but the navigation scheme included suggested traversals as solid lines between nodes depicting recommended traversal paths (since those paths link semantically related nodes). The semantic-based hypertext presentation was created using the outcome of the content design where semantic association, in our example 'Support', provides the link between these nodes. All subjects were exposed to three contents (Psychology, Neuroscience & Archeoastronomy) however, presentation viewed were counterbalanced to reflect all three presentation formats were seen by each subject.

It should be noted that Figure 4 through Figure 6 shown below, were presented to participants but is considered outside the scope of this particular research article.

Figure 4: Linear

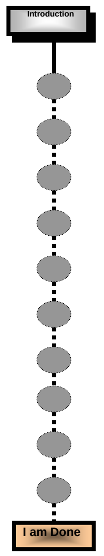


Figure 5: Meshed

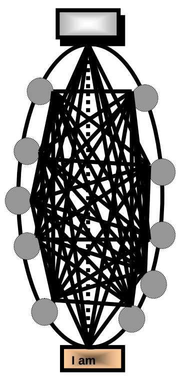
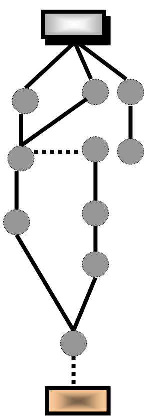


Figure 6: Semantic-Based



4 Methods

Twenty Psychology undergraduate student participated in this exploratory study. Each student acted as an Expert (reading Psychology text), as a Novice (reading Neuroscience text) and, as a control group participant by reading neutral Araeoastronomy text.

Each subject read all three knowledge domains i.e., Psychology, Neuroscience and Araeoastronomy as well as were exposed to all three Hypertext (web) presentation (Linear, Meshed and Semantic-based) formats. The order of web presentation format were counterbalanced across all the participants.

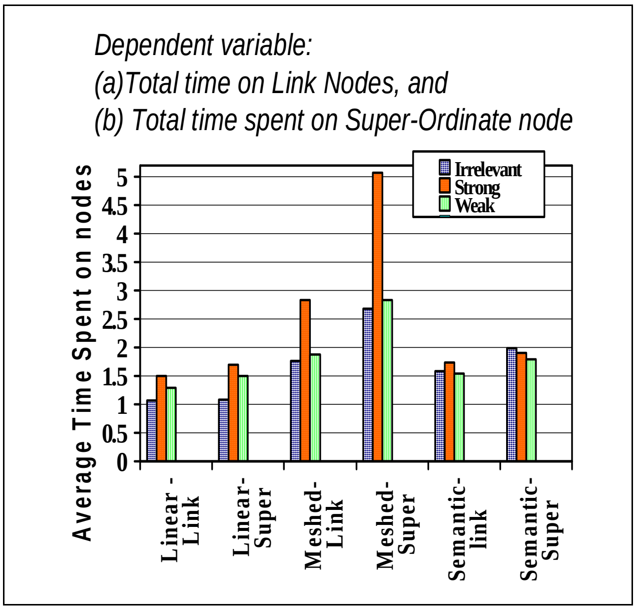
Each subject started with a Meshed Hypertext filler text – Astrophysics, which was not analyzed followed by the three counterbalanced experimental texts. After reading each website within allotted time participants were asked to summarize their understanding of the presented text before moving on to the next website. Each participants started with the 'Introduction' page and ended the website traversal by selecting 'I am Done' button in the bottom. Only participant traversal behavior analysis has been presented in this paper.

5. Results and Discussion

Two different types of performance data was submitted to an ANOVA with Expertise (Expert and novice) as Between-Group, and Format (Linear, Meshed-Hypertext and Semantic-based Hypertext) as Between-Group variables. Only two specialized texts (Psychology & Neuroscience) and the neutral (Archaeoastronomy) text was analyzed for this ppilot study.

Contrary to our original belief 'total number of nodes' visited which includes revisiting nodes more than once, although reaching significance at $p<0.5$ for presentation formats (Linear, Mesh & Semantic) and expertise (Strong, Weak, Irrelevant) levels, did not prove significant for format-expertise interactions. In addition 'total time' spent traversing web sites did not reach significance either except for formats.

The total time spent on nodes with most semantic connections (link nodes, i.e., F3 & FC) reached significance for presentation formats $F(2, 4) = 10.27$, $MSe = 4.044$, $p= 0.0002$, as well as for expertise-format interactions, $F(2, 4) = 2.90$, $MSe= 1.142$, $p= 0.0307$. In general experts are spending more time reading nodes with more semantic connections than other nodes especially in the Mesh-Hypertext environment. Although number of times link nodes (including revisits) visited were more or less compared with novice, however, experts were found to be spending more time on link nodes (nodes with more connections) in all three (Linear, Meshed, Semantic-based) however proportionately more in Meshed-hypertext presentation format. See Figure 7.



The total time spent on super-ordinate nodes (i.e., IC-1, IC-2 & FC) reached significance for expertise match F (2, 4) = 9.92, MSe = 5.83, p= 0.0002, as well as for presentation format, F (2, 4) = 39.20, MSe= 23.04, p< 0.001. In addition, interactions were quite strong as well, F (2, 4) = 6.78, MSe = 3.98, p= 0.0002. In general experts are spending more time reading super-ordinate nodes than other nodes especially in the Mesh-Hypertext environment. Although number of times a Super-ordinate node visited was only significant p<.001 for presentation formats and not for interactions or expertise. Again, experts were found to be spending more time on super-ordinate nodes as compared with novices in all three (Linear, Meshed, Semantic-based) however proportionately more in Meshed-hypertext presentation format. Please see Figure 7.

We saw a strong effect for the number of times irrelevant nodes were visited since expertise, format and, interactions all reached p<.001. Furthermore time spent on irrelevant nodes shows similar strong effects of p<.001 for expertise, format and interactions. Novice visited Irrelevant nodes more frequently then experts in all three presentation formats and especially for Meshed Hypertext . In addition, compared with experts, novice spend more time on Irrelevant nodes in all three presentation formats especially in Meshed Hypertext. visited more number of times The results from this exploratory study seems very promising for further elaborate studies.

Figure 7: Time spent on Link (web pages with most semantic connections) nodes and, time spent on Super-ordinate nodes are compared against expertise and different presentation formats. In general experts tend to spend more time on both the Link and Super-ordinate nodes, in particular when traversing Meshed web site environments.

Several studies have proposed that expertise is hard to quantify but some qualities associated with this notion can be defined. Some researchers have viewed it in terms of having prior exposure in the domain area [15, 24] and formal training [25]. Furthermore, it has been argued that complexity of structures plays a critical role [22, 23] where complexity between an expert versus novice lies within the conceptual organization [26] of semantic structures (referred as situational models or mental models) that are formed during an activity such as reading. Watt & Welch [23] attempts in defining 'simple' (ordered, predictable and structured) characteristics, while 'complex' (disordered, unpredictable and random) seems well suited to shed some insights into these comparative conceptual organization between experts and novice.

Although preliminary however this study seems to manifest encouraging trends of expertise and web presentation format effects which not only address the qualitative characteristics associated with expertise but also provides venues into quantifying these characteristics. Furthermore, the suggested ideas on 'simple' versus 'complex' were easily tested in designing and presenting participants with different web presentation formats. Furthermore, this study was useful in discerning expertise as a function of traversal behaviors by an expert when presented with web presentation formats that are inspired by and designed according to text comprehension underpinnings.

In general, strong expertise requires more time reading but only for the nodes with most semantic connections and super-ordinate nodes. Novice and weak expertise tend not to spend proportionately more time on these nodes. It seems that experts are able to identify which nodes or information are most important. Furthermore, given the random nature of Mesh-Hypertext environment, it seems that having expertise in domain knowledge tends to assist in finding and in reading the most critical nodes.

Furthermore, in line with Lawless and Kulikowich who identified that there are navigational paths differences between focused and less focus readers [24]. We found that novice is more likely to visit greater number of Irrelevant nodes and spend more time on it as oppose to experts.

In summary, we found experts are not only to identify web pages that are critical in understanding of a topic, but are also able quickly recognize the ones that seems less important or irrelevant. Furthermore, experts tend spend more time on most important web pages are oppose to novice.

Much of the hypertext research are focused on computer science and interface issues rather than the cognitive aspects associated with hypertext learning. It seems that further explorations in this domain such as distance and web learning may prove useful and perhaps, it may lie at the nexus of discourse and multimedia domains.

6. Acknowledgment

We are grateful to National Science Foundation (NSF) for their generous support. This project is fully funded by NSF Grant 0624983 from the Methodology, Measurement, and Statistics (MMS) Program in the Division of Social and Economic Sciences. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

References:

[1] Halliday, M., Hasan, R. (1976). Causal coherence and memory for events in narratives. *Cohesion in English*. London: Longman.

[2] Fahnestock, J., Secor, M. (1983). Teaching argument: A theory of types. College Composition and Communication, 34, 20-30.

[3] Long, D. L.; Oppy, B. J.; Seely, M. R. (1997). A "global-coherence" view of event comprehension: Inferential processing as question answering.

[4] Lorch, R. F. Jr. & O'Brien, Edward J. (1995). *Sources of coherence in reading*.

[5] Trabasso, T., and van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of memory and language, 24,* 612-630.

[6] Rumelhart, D. (1977). Understanding and summarizing brief stories. In D. Laberge & J. Samuels (Eds.), *Computational models of discourse.* Cambridge, MA: MIT Press.

[7] Trabasso, T., and van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of memory and language, 24,* 612-630.

[8] Trench, B. (2000). Science journalism on the Web. Paper presented at the Seminaire sur les Nouveaux Territories de la Communication Scientifique de la Cite' des Sciences, Paris, France, November.

[9] Sweller, J., Chandler, P. Tierney, P., Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General* 119 (2) 176-192.

[10] Gorden, S., Gustavel, J., Moore, J., and Hankey, J. (1988). The effect of hypertext on reader knowledge representation. In *Proceedings of the 32nd Annual Meeting of the Human Factors Society,* 296-300, Santa Monica, CA: Human Factors Society.

[11] Macedo-Rouet, M., Rouet, J.-F., Epstein, I. and Fayard, P. (2003). Effects of Online Reading on Popular Science Comprehension. *Science Communication 25* (2), 99-128.

[12] Foltz, P. W. (1996). Comprehension, coherence, and strategies in hypertext and linear text. In J.-F. Rouet, J. J. Levonen, A. Dillon, & R. J. Spiro (Eds.), *Hypertext and cognition* (pp.109-136).

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

[14] Dee-Lucas, D. & Larkin, J.H. (1995). Learning from electronic texts: Effects of interactive overviews for information access. *Cognition & Instruction* 13 (3), 431-468.

[15] Lawless, K. A., Brown, S. W., Mills, R. & Mayall, H. J. (2003). Knowledge, Interest, Recall and Navigation:A Look at Hypertext Processing. *Journal of Literacy*, 35 (3), 911-934.

[16] McDonald, S., Stevenson, R. J., (1998). Navigation in hyperspace: An evaluation of the effects of navigational tools and subject matter expertise on browsing and information retrieval in hypertext. *Interacting with Computers 10*, 129–142.

[17] Puntambekar, S., Stylianou, A., & Hübscher, R. (2003). Improving navigation and learning in hypertext environments with navigable concept maps. *Human-Computer Interaction,18*, 395-428.

[18] Saremi, H., Q. & Montazer, G, M., (2007). An application of type-2 Fuzzy notions in website structures selection: Utilizing extended TOPSIS method. *WSEAS Transactions on Computers*, 7, 8-15.

[19] Lee, C-H., Lee, G-G. & Leu, Y., (2007). Analysis on the adaptive Scaffolding Learning Path

and learning. *WSEAS Transactions on Information Science & Applications*, 4, 320-330.

[20] Chen, M., F. & Chuang, L. P., (1999). Effects of Knowledge Map construction on e-Learning performance in a web-based learning. *The 8th. Internation Conference on Computer-Assisted Instruction* (ICCAI 1999), Taichung City, Feng Chia University.

[21] Yousoof, M., Sapian, M. & Ramasamy, K. (2008). Performance evaluation of the software visualization tools and a new framework to mange cognitive load in computer program learning. *WSEAS Transactions on Information Science & Application*, 5, 655-663.

[22] Berylne, D. E. (1971). *Aesthetics and psychobiology*. New York: Appleton-Century-Craft.

[23] Watt, J., H., Jr., & Welch, A. (1983). Effects of static and dynamic complexity on children's attention and recall of televised instruction. In Jennings Bryant & Daniel R. Anderson (Eds), *Children's understanding of television: Research on attention and comprehension* (pg. 69-102). New York: Academic Press.

[24] Lawless, K. A., & Kulikowich, J. M. (1998). Domain knowledge, interest, and hypertext navigation: a study of individual differences. *Journal of Educational Multimedia and Hypermedia*, *7*, 51-70.

[25] Simmons, P. E., & Lunetta, V.N. (1993). Problem-solving behaviors during a genetics computer simulation: Beyond the expert/novice dichotomy. *Journal of Research in Science Teaching*, 30 (2), 153-173.

[26] Spires, H. A., & Donley, J., (1998). Prior knowledge activation: Inducing engagement with informational texts. *Journal of Educational Psychology*, 90 (2), 249-260.