Cognitive load in hypermedia reading comprehension: Influence of text type and linearity

Joerg Zumbach a,*, Maryam Mohraz b

a University of Salzburg, Department of Science Education and Teacher Training, Hellbrunnerstr. 34, Salzburg, Austria
b University of Heidelberg, Department of Educational Psychology, Germany

Available online 30 March 2007

Abstract

In this paper the assumption of cognitive overhead in hypermedia learning is specified by cognitive load theory. This analysis is based on different types of cognitive load, the dimension of linearity/non-linearity as well as text characteristics. We propose a model stating that extraneous cognitive load in hypermedia learning is basically determined by the interaction of text presentation format (linear/non-linear) with text type (text with and without narrative structures). This assumption was tested by means of a $2 \times 2$ experimental design. Sixty participants completed a computer-based learning program that contained a narrative text or an encyclopaedia text in either linear or non-linear presentation format. Results confirm the suggested interaction hypothesis postulating that non-linear information presentation of narrative text structure increases cognitive load and decreases knowledge acquisition. However, for encyclopaedia text participants’ knowledge acquisition was not affected by linear or non-linear presentation format. Furthermore, results suggest a cross-validation of cognitive load measures and propositional analysis.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Hypertext; Hypermedia; Cognitive load; Reading comprehension
1. Introduction

Much of past research in information retrieval from hypertext or hypermedia has addressed the “lost-in-hyperspace”-phenomenon or cognitive overhead (cf. Conklin, 1987). Especially interventions for overcoming lost-in-hyperspace has emerged many different approaches (e.g., Chiu & Wang, 2000; Dias & Sousa, 1997; McDonald & Stevenson, 1998; van Nimwegen, Pouw, & van Oostendorp, 1999). Nevertheless, approaches to investigate the cognitive overhead phenomenon are rare. In this paper we specify this phenomenon by basic cognitive mechanisms of text comprehension and cognitive load theory.

2. Cognitive overhead and cognitive load theory

Cognitive overhead addresses the limitation of the human working memory and cognitive processes in hypertext/hypermedia learning (Conklin, 1987): Learners have to process information represented in hypertext nodes and plan their (further) navigation simultaneously which demands higher resources of working memory (cf. Bannert, 2004; Gerdes, 1996). Niederhauser, Reynolds, Salmen, and Skolmoski (2000) compared reading and navigation strategies in hypermedia learning on their impact of learning outcomes. Results suggest that learners using a sequential and almost linear information retrieval show higher learning success than participants using a non-linear browsing strategy.

Despite this basic evidence for occurrence of the cognitive overhead phenomenon, it is a general assumption rather than a specific theoretically grounded concept that would allow predicting learning success or failure. A more specific and appropriate approach for examining learning processes with linear and non-linear knowledge media is provided by cognitive load theory (CLT; Paas, Renkl, & Sweller, 2004; Sweller, 1994). CLT differentiates between intrinsic cognitive load (ICL), extraneous cognitive load (ECL) and germane cognitive load (GCL; Sweller, van Merriënboer, & Paas, 1998). All three proposed types of cognitive load are assumed to be additive. In case of exceeding working memory resources information processing will be decreased which inhibits knowledge acquisition (Paas, Tuovinen, Tabbers, & Van Gerven, 2003).

Especially the ECL seems to be what is meant by “cognitive overhead”. If the assumption of cognitive overhead would be adequate, learners should always profit more from linear text than from non-linear text because learning with non-linear (hyper-)text requires always additional mental effort in navigation planning and monitoring. However, cognitive load theory postulates that additional mental effort does not always inhibit learning, for example, if this additional effort leads to deeper elaboration processes.

Our main argument in this paper is that inappropriate instructional format causing a high ECL could also affect a higher GCL. In hypertext learning, students are enforced to reflect upon their prior knowledge in order to make decisions about navigation. This additional cognitive effort could lead to deeper elaboration and, thus, contribute to learning success (limited resources due to ECL help to activate GCL). Learners have here to activate prior knowledge and possible schemata in order to overcome obstacles of the ECL.

Studies examining differences in knowledge acquisition from text and hypertext are often difficult to interpret because dimensions of linearity/non-linearity are vague and influence of text type is neglected.
3. Models of text comprehension and learning with (non-)linear text

Contemporary models of text comprehension describe the process of understanding on multiple levels of representation: surface code, propositional text base, and situation model (e.g., Graesser, Olde, & Klettke, 2002; Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983). Text comprehension takes place within a limited working memory. Thus, activation of prior knowledge and the drawing of inferences in order to build local and global coherence is crucial (cf. Graesser & Britton, 1996). Coherence building is difficult if hypertext navigation leads to accessing text that is “discontinuous”.

Difficulties in building a mental model from continuous and discontinuous text have been examined by Schnotz (1988) showing that learners reading a continuous text could remember content better and were able to draw more conclusions than learners with a discontinuous text which made the construction of a mental model more difficult. Also other research has shown that topic continuity plays an important role in mental model construction from text (cf. Garnham, Oakhill, & Johnson-Laird, 1982; Schnotz, 1993).

Discontinuous text is not necessarily problematic. Following Johnson-Laird (1990), the construction of mental models is possible, if references of text base and existing mental models are available. Only if learners cannot activate a mental model or schema, they will face difficulties in coherence building processes. Based on these considerations it is possible to explain empirical findings showing that disadvantages of hypertext learning resources disappear when learners have a certain degree of prior knowledge (e.g., Gerdes, 1996; Lawless & Brown, 1997). A steady interconnectivity of activated mental models and the text base allows learners with prior knowledge to close gaps in the text and to assimilate new information into their existing knowledge structures.

Nevertheless, the analysis of differences in hypertext and text comprehension is still a major problem due to shortcomings of established theories in differencing dimensions of linearity and non-linearity. Following Kintsch and van Dijk (1978), differences between text and hypertext comprehension can only be explained on a higher level of processing (level of macrostructures and level of superstructure). Traditional linear text types (on the level of a text super structure) like script-based episodes or fairy tales seem to be inappropriate to be represented in a non-linear manner (cf. Kintsch & Kozminska, 1977).

The transition between both extremes of linear and non-linear text and especially factors influencing processes of understanding needs further research. In order to bridge the gap between theories of text comprehension and media specific attributes addressing dimensions of linearity and non-linearity, text design, text type and cognitive load, Zumbach (2006) has proposed a model involving all these factors (see Fig. 1).

The basic assumption in this model is that variations in cognitive load in text or hypertext learning environments are caused by an orientation problem. We postulate that disorientation is possible to occur in linear as well as non-linear texts. If there is no disorientation, there will be no need of planning navigation and hence no corresponding ECL in learners with no or little prior knowledge. In case of learners with expertise in the learning domain, information search and retrieval processes might rather affect ICL than ECL. Basically, disorientation could be a consequence of unstructured texts (e.g., missing paragraphs, headings, or text that contains complex sentence structures). It can also occur if the complexity of the content or the learning environment exceeds the capacity of an individual. Disorientation is also likely to occur when (a) the narration format (e.g., like
in a fairy tale) and the presentation format (linear vs. non-linear) are not corresponding. This should lead to an increased ECL.

Furthermore, the complexity of the content and/or task should play a major role in evoking ECL. In this case, the suggested Extraneous Load will evoke GCL. Faced with a highly demanding task, hypertext learners should benefit at least to the same amount from their learning media as traditional text learners do. First empirical research on this model related to influence of text design and text complexity has shown that the interaction hypothesis of text complexity and (non-)linearity is adequate and the assumption of a global cognitive overhead is inadequate (Zumbach, 2006; Zumbach & Pixner, 2006): A non-linear presentation of a complex content does not imply cognitive load per se, but rather triggers learning by initiating schema activation and, thus, GCL as a consequence from overcoming orientation problems and, thus, ECL.

In addition to complexity of content, another crucial element of the proposed model of CL within text and hypertext comprehension is the text type and its underlying narrative structures which have been neglected in recent hypertext research.

We assume that linear text types with narrative structures following conventional time logic (e.g., a script; cf. Schank & Abelson, 1977) trigger on basis of existing processing schemata an automated navigation strategy and, thus, Extraneous Load is not likely to occur. Chronological aspects during construction of the situation model are determined by sequence of propositions in the text base. Most learners will implicitly assume that a new episode will timely occur after the preceding (Givon, 1993). Within a non-linear presentation all information has to be mentally restructured in order to build global coherence. This should require additional cognitive effort.

4. Hypotheses

Based on these considerations we hypothesize that a linear text type presented in a non-linear manner will increase ECL. Non-linear text type can be processed in a linear as well as non-linear manner without evoking different ECL. The level of ECL resulting from the interaction of text type and (non-)linear presentation format should influence reading comprehension and, thus, knowledge acquisition. We assume that the increase of knowl-
edge depending on a linear text type will be higher in linear text presentation than in non-linear hypertext. With a non-linear text type, we do not expect differences in knowledge acquisition depending on the presentation format (linear vs. non-linear; see Fig. 2).

5. Method

5.1. Design and sample

In order to test these hypotheses an experiment was conducted. We used a $2 \times 2$ factor design including the two between group factors text type (text with narrative structures, text without narrative structures) and presentation format (linearity/non-linearity). Sixty university students (40 female and 20 male) majoring in different fields participated in the experiment (mean age of students was 24.0, SD = 6.09). Subjects were randomly assigned to one of four experimental groups. Each student received €10 for participation.

5.2. Dependant variables

5.2.1. Learning outcomes

In order to measure individual’s learning success a knowledge pre- and post-test was conducted. It included a 12-item multiple-choice test and an essay task in order to measure learner’s knowledge and understanding. Essay tasks were analysed by a propositional analysis based on the suggestions of Kintsch and van Dijk (1978). We analysed participants’ essays for propositions also included in the learning material. Each correct essay proposition was counted by 1 point. In order to compare the four different hypertext versions a knowledge acquisition value was computed for each participant by dividing their amount of correct essay propositions by the overall score of propositions in the corresponding learning material.

Fig. 2. Hypothesized effects of text type and linearity on cognitive load and learning outcome.
5.2.2. Cognitive load

Two instruments assessing cognitive load have been used in this study: First, the mental effort rating scale (MERS; Paas, Van Merriënboer, & Adam, 1994). Second, a slightly adapted form of the NASA-TLX (task load index) developed by Hart and Staveland (1988). NASA-TLX consists of five subscales each represented by a one-item self-report question (i.e., Task requirements, Effort in understanding content, Expectation of success, Effort in navigation and Stress). Both measures were used in post-test (NASA-TLX immediately followed by the MERS).

An analysis of internal consistency provided a Cronbach’s Alphas of 0.49 for the NASA-TLX. In order to increase internal consistency we dropped the item “expectation of success” which lead to an acceptable value of Cronbach’s Alpha with 0.75. The four remaining items were summarized in a NASA-TLX overall score.

5.3. Material

For the narrative text type we chose a rather unknown fairy tale from Hans Christian Andersen “The Girl Who Trod on the Loaf”. The rationale for the choice of a fairy tale as a script-based text type was that this genre typically has a chronological plot with a clear beginning and end as well as several protagonists, their descriptions and interactions. Furthermore, it is a typical text genre for moral education and, thus, provides a learning scenario.

The original text has been divided into 35 locally coherent nodes with 3083 words. The encyclopaedia text type was build from several articles within the area of bacteria, viruses and genes. The chosen contents were articles from molecular biology (19 nodes with 2538 words).

The encyclopaedia articles had a higher complexity than the narrative text. In the narrative text there were more bridging sentences in order to guarantee local coherence. By reducing the amount of words in the encyclopaedia text, we balanced task difficulty. Each text was transformed into computer-based learning environments with either linear or a non-linear navigation access (see Fig. 3). In the linear condition, navigation was possible

Fig. 3. Sample screens from the narrative text condition. Linear program screen is presented left and hypertext program is presented right.
by means of a “next”, “back” and “starting page” button. Sequence in the narrative text was determined by its chronological narrative structure; in the linear encyclopaedia conditions, we arranged the nodes in alphabetical order depending on the topic of each node. For the non-linear hypertext conditions, we linked the nodes associatively. In addition, we added buttons leading to the starting page and back to the last visited page.

5.4. Procedure

The experiment started with oral and written instructions about the task and the handling of the learning programs followed by the pre-test. In the learning phase participants had 35 min time to read the texts provided at single desktop PCs. During this phase they were not allowed to take notes. Afterwards they had to complete the post-test with the same essay task and multiple-choice test as in pre-test as well as NASA-TLX and the MERS (pre- and post-test lasted up to 30 min each).

6. Results

A two-factorial analysis of variance revealed a significant simultaneous effect of text type on all dependant variables ($F(3;54) = 40.01, p < 0.001, \varepsilon^2 = 0.69$) and a significant interaction of both factors text type and presentation format (linear vs. non-linear) ($F(3;54) = 3.57, p < 0.05, \varepsilon^2 = 0.17$). There was no significant simultaneous effect of presentation format ($F(3;54) = 2.46, p = 0.07, \varepsilon^2 = 0.12$). Table 1 provides means and standard deviations of dependant variables of each experimental condition.

6.1. Learning outcome

In order to measure knowledge acquisition we computed an overall score for the multiple-choice test performance by dividing the individual scores by the maximum possible score ($\text{Multiple Choice \%}$; each for pre-test and post-test). In addition, we computed a relative score of propositions found in the pre-test and in the post-test essay task by dividing participants’ overall proposition score by the maximum possible number of propositions available in the learning material ($\text{Propositions \%}$; in the narrative text the maximum was

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Means and standard deviations of learning outcome and cognitive load</td>
</tr>
<tr>
<td>Dependent variables</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Multiple Choice %</td>
</tr>
<tr>
<td>Propositions %</td>
</tr>
<tr>
<td>NASA-TLX</td>
</tr>
<tr>
<td>MERS</td>
</tr>
</tbody>
</table>

Multiple Choice (recall) % and Propositions (essay) % = knowledge acquisition in percent. High mean values in NASA-TLX and MERS indicates a high cognitive load.
There were no significant pre-test differences. Univariate significance testing showed a significant main effect of text type for the dependant variables *Propositions %* \( (F(1;56) = 51.39, \ p < 0.001, \ \varepsilon^2 = 0.47) \) and *Multiple Choice %* \( (F(1;56) = 89.26, \ p < 0.001, \ \varepsilon^2 = 0.61) \). Gain in knowledge scores was significantly higher in both narrative text conditions than in conditions with the encyclopaedia texts. In post-test participants were able to reproduce about 17.67% of possible proposition on average. In the encyclopaedia text conditions the corresponding average value was 9.08%. In addition, the main effect of presentation format on *Propositions %* was significant \( (F(1;56) = 4.1, \ p = 0.048, \ \varepsilon^2 = 0.07) \). This effect was not replicated for *Multiple Choice %* \( (F(1;56) = 1.55, \ p = 0.22, \ \varepsilon^2 = 0.03) \).

An ANOVA with repeated measurement calculated on *Propositions %* showed a significant effect between pre- and post-test \( (F(1;56) = 500.83, \ p < 0.001, \ \varepsilon^2 = 0.9) \), a significant effect of text type and time of measurement \( (F(1;56) = 52, \ p < 0.001, \ \varepsilon^2 = 0.48) \), a marginally significant effect of presentation format and time of measurement \( (F(1;56) = 4.03, \ p = 0.05, \ \varepsilon^2 = 0.07) \) and a significant interaction of both independent variables and time of measurement \( (F(1;56) = 8.83, \ p < 0.01, \ \varepsilon^2 = 0.14) \); see Fig. 4 for increase of *Propositions %*.

### 6.2. Cognitive load

Univariate significance testing revealed a significant main effect of text type on experienced cognitive load measured by the MERS \( (F(1;56) = 17.64, \ p < 0.001, \ \varepsilon^2 = 0.24) \). Learning with the encyclopaedia articles required more mental effort than the narrative texts. Non-linear information presentation lead to slightly increased cognitive load compared to linear access, although not significantly \( (F(1;56) = 3.24, \ p = 0.08, \ \varepsilon^2 = 0.06) \). There was no significant interaction effect on the MERS-value \( (F(1;56) = 2.25, \ p = 0.14, \ \varepsilon^2 = 0.04) \). Nevertheless, the mean value of each condition showed that the linear presentation of the fairy tale required less mental effort than the non-linear presentation. This difference did not occur within the encyclopaedia text conditions (see Fig. 5).

![Fig. 4. Gain in knowledge acquisition (Propositions %).](image-url)
There was also a significant main effect of text type on the NASA-TLX scores. ($F(1;56) = 14.31$, $p < 0.001$, $\varepsilon^2 = 0.20$), but not of presentation format ($F(1;56) = 1.60$, $p = 0.21$, $\varepsilon^2 = 0.03$). The interaction of both independent variables exceeded the 0.05 level marginally ($F(1;56) = 3.79$, $p = 0.056$, $\varepsilon^2 = 0.06$).

The mean values of NASA-TLX items (see Table 2) showed a difference between items “task requirements” and “navigation”. While task requirements (TR) and navigation (N) led to a higher experienced CL in the non-linear narrative text condition (TR = 7.60; N = 5.93) than in the linear presentation of the same text material (TR = 5.7; N = 3.50), the values in both encyclopaedia conditions did not differ that far (non-linear: TR = 8.33; N = 5.83 vs. linear: TR = 8.50; N = 5.17).

6.3. Cross-validation of dependant variables

In order to cross-validate results the Pearson correlation between all dependent variables were computed (see Table 3).

![Fig. 5. Mean values of NASA-TLX.](image)

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Means and standard deviations of NASA TLX items</th>
</tr>
</thead>
<tbody>
<tr>
<td>NASA-TLX item</td>
<td>Experimental treatment</td>
</tr>
<tr>
<td></td>
<td>M</td>
</tr>
<tr>
<td>Task requirements</td>
<td>5.87</td>
</tr>
<tr>
<td>Mental effort</td>
<td>4.60</td>
</tr>
<tr>
<td>Navigation</td>
<td>3.50</td>
</tr>
<tr>
<td>Stress</td>
<td>4.07</td>
</tr>
</tbody>
</table>
Results reveal a high correlation of MERS and NASA-TLX with $r = 0.75^{**}, p < 0.01$. Correlations between measures of cognitive load and knowledge acquisition are also significant. High subjective ratings of cognitive load during hypermedia learning are associated with reduced learning outcome. A linear regression analysis using NASA-TLX score as independent and relative amount of increase in mentioned propositions ($\text{Propositions \%}$) resulted in a significant model ($F(1;58) = 10.58; p = 0.002$) $\text{Propositions \%} = 25.6 - \text{Cognitive Load} \times 2.6$ explaining 14% of variance.

7. Discussion

Research on learning with text and hypertext and its comparisons has focused differences in knowledge acquisition mostly without referring to basic mechanisms of reading comprehension. In this study we specified the global assumption of cognitive overhead in hypertext learning by means of cognitive load theory and models of text comprehension. This study focused on the influence of text type and its interaction with linear and non-linear navigation access on cognitive load and learning outcomes. Results confirm the hypothesized ordinal interaction stating that non-linear presentation of narrative text leads to increased cognitive load and decreased knowledge acquisition compared to linear presentation of the same text. Furthermore, this effect was only restricted to narrative text, whereas the same experimental manipulation with a less-structured encyclopaedia text did not show differences as was expected.

Text type significantly influenced outcomes of dependant variables. Overall, participants estimated learning with encyclopaedia texts as more difficult. Learning outcomes were also not as high as in narrative text conditions. Interestingly, there is still no satisfying explanation in text comprehension research why narrative texts are more easily to understand and to remember (cf. Graesser et al., 2002). An explanation here could be here that the content of the encyclopaedia text were presumably more complex and difficult as expected and, thus, led to difficulties in understanding. This might have contributed to intrinsic cognitive load which influenced learning negatively.

Furthermore, results confirm the hypothesis that learning from narrative text with a non-linear navigation access is unbenefficial. Participants’ essays reproduced less than a fourth of proposition than their counterparts with linear navigation access to the narrative text. Lesgold, Roth, and Curtis (1979) state that thematic change (as given in hypertext navigation) leads to fewer propositions available in working memory and, thus, to reduced building of inferences. The findings from our propositional essay analysis were only par-

<table>
<thead>
<tr>
<th>Table 3: Correlation of dependant measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>NASA-TLX overall score</td>
</tr>
<tr>
<td>-------------------------</td>
</tr>
<tr>
<td>NASA-TLX overall score</td>
</tr>
<tr>
<td>Mental effort rating</td>
</tr>
<tr>
<td>scale</td>
</tr>
<tr>
<td>Multiple Choice %</td>
</tr>
<tr>
<td>Propositions %</td>
</tr>
</tbody>
</table>

^{**} p < 0.01.
tially supported by the additional applied multiple-choice test. This is in line with other research showing that deeper information processing during learning, which corresponds to high GCL, does not affect recall performance (e.g. Bannert, 2002).

7.1. Effects on cognitive load

Results from the NASA-TLX also confirm the interaction hypothesis that a missing correspondence of presentation format and text type leads to an increased cognitive load. The non-linear access to the narrative text was considered as much more difficult than linear access to this text type. In both encyclopaedia text conditions the subjectively experienced cognitive load was comparable. Interestingly, NASA-TLX subscale “navigation” reached the highest values in condition with non-linear access to the narrative text type. This is an outstanding result because in all other CL-scales the encyclopaedia text conditions reached higher values. We attribute the high values in CL in the encyclopaedia text conditions to an increased Intrinsic Load caused by the content complexity of the learning material. Obviously, the effort in navigation did influence cognitive load negatively in a less complex domain and decreased learning performance. This finding limits Sweller’s (1994) postulation that with a low complexity of content, i.e., ICL, the influence of instructional format should not have a negative impact on CL. Considering the learning environments of our study, an instructional control of ECL seems appropriate although intrinsic cognitive load seems to be low or moderate.

7.2. General discussion

The main findings of this study can be explained by basic cognitive mechanisms, especially by schema theoretical approaches. Providing a certain text type should lead to activation of prior knowledge and, thus, activate schemata for (partially) automated information processing purposes. Information presentation incongruent with conventional presentation format (as here the narrative text presented as hypertext) reduces availability of automated text processing. Thus, additional mental effort is necessary in order to select and acquire information. We assume that working with the less linearly structured encyclopaedia texts did not automatically activate a corresponding reading schema that could have interfered with the presentation format. Consequently, we were not able to find an effect of this text type and its different presentation format on learning outcomes and cognitive load.

Learning outcomes as measured here with essays and analysis of propositions and measures of cognitive load are highly correlated. We were able to show by means of correlations and regression analysis that cognitive load measures seem to be an appropriate and valid estimator for predicting learning outcomes.

Limitations of this study are mainly associated with the learning material. First, the text material used in this study did not directly meet expected comparability. Despite our approach to adjust difficulty of the texts by means of adapting the length of each text type, outcomes reveal a slightly severed difficulty of the encyclopaedia texts. Second, the choice of a fairy tale as narrative text leaves the area of non-fiction instructional text. A follow-up experiment will address both issues by replicating these findings with a non-fiction script-based text that is comparable in length and difficulty with the encyclopaedia texts used here.
In general, the study contributes to a contemporary model of text comprehension by integrating dimensions of text type, linearity and non-linearity as well as its implications on cognitive load. Other studies (e.g., Zumbach & Pixner, 2006) will provide further evidence of the proposed model of reading text comprehension by integrating aspects of task and text difficulty and its implications on knowledge acquisition processes and outcomes.

Acknowledgement

We thank Maria Bannert for helpful comments on an earlier version of this paper.

References


Graesser, A. C., Olde, B., & Klettke, B. (2002). How does the mind construct and represent stories. In M. C. Green, J. J. Strange, & T. C. Brock (Eds.), *Narrative impact: Social and cognitive foundations* (pp. 231–263). Mahwah NJ: Lawrence Erlbaum Ass.


