COMPARING THE USE OF DIRECTED WORLD WIDE WEB NAVIGATION GUIDANCE WITH CONVENTIONAL TOOLS

DAVID RIBEIRO LAMAS

O projecto **Computer Aided Information Navigation** utiliza um novo método de navegação adaptativa que tem como objectivo aumentar o valor da World Wide Web como ferramenta pedagógica. Para testar se a abordagem utilizada consegue ou não influenciar a compreensão de determinado assunto, foi realizada uma experiência que comparou o **CAIN** com outras ferramentas convencionais em três medidas de desempenho: compreensão, tempo gasto e apreciação de utilização. Os resultados mostram que o **CAIN** proporciona um aumento da compreensão na ordem dos 30%, sugerem que o tempo gasto também pode ser reduzido e que, em geral, os utilizadores gostam de usar a Web com não importa qual ferramenta.

The **Computer Aided Information Navigation** project provides a new method of adaptive navigation support that aims to increases the World Wide Web's value as a pedagogical tool. To test if the system's approach improves comprehension under specific conditions, an experiment was devised to compare the use of **CAIN** with conventional tools on three measures of performance: Comprehension, Time taken and System usage appreciation. Results show that **CAIN** improves user's comprehension by 30%, suggest that the time taken can also be reduced and that, as far as enjoying the system, users enjoy using the Web with whatever tools they use.

KEYWORDS

Empirical study, adaptive navigation, world wide web, pedagogical tool

INTRODUCTION

A large volume of information, such as the one available over the World Wide Web (Web) has an inherent application potential in an educational setting. Unfortunately, besides inheriting and enhancing most of hypertext's qualities, the Web also incorporates and stresses common hypertext problems (Brown 1990; Shum 1996; Shum & McKnight 1997).

It is the Web's disorganized nature that mainly prevents its most effective use as a pedagogical tool and, as such, research is needed to develop the means that will help to transform the available data into meaningful information able to support learning or initial research activities (Eklund & Ziegler 1996; Marshall 1995).

The **Computer Aided Information Navigation** project (**CAIN**) attempts to provide adaptive navigation support as a way of increasing the Web's value as a pedagogical tool. **CAIN**'s approach is a reactive one as the effort goes into providing a way to deal with the vast amount of useful information available on the Web as is, rather than to try to improve or alter the Web's infrastructure in any way.

In order to help users reach their goals, **CAIN** provides direct guidance navigation support as a form of non-obtrusive *weak hypertext linearization* enabling the user to follow a context specific ranked sequence of selected Web pages without ever needing to perform any search or follow any link if they do not wish to. This approach does not intend to prevent goal-oriented

exploration but to provide a sound thread or guideline to help users to retain their focus in the way a **Rough Guide** does.

The system's representation of the Web, crucial to the success of its adaptive features, builds on the Dublin Core (Weibel & Lagoze 1997) and on the Resource Description Framework (Lassila & Swick 1998).

The users' representation combines three user modeling techniques resulting in a hybrid solution that uses stereotypes, overlays and attribute-value pairs (Benyon & Murray 1993; Brusilovsky 1996).

Navigation support is accomplished by a basic route-finding heuristic which selects context specific Web model items, sorts them using associated qualitative ratings and presents them to the user, one at the time, based on the attributes of the user's model (Lamas 1998).

The next pages present and discuss the results of an experiment undertaken to test if the directed World Wide Web navigation guidance provided by **CAIN** improves the performance of Web users under appropriate conditions. Section 2 describes the methodology whereas section 3 presents and discusses the experiment's results. Finally, section 4 highlights some related empirical work and section 5 discusses the validity and completeness of the experiment and provides future work guidelines.

METHOD

It is hypothesised that contextualized direct guidance over the information available on the World Wide Web improves comprehension under the specific condition that:

- Users are either inexperienced in Web use or have little expertise in the target subject topic or both.

- Users use the system pursuing learning or initial research goals.

For this purpose, a standard task was devised in which volunteers were required to use the WWW to learn as much as they could about two given subject topics — cryptography and poetry — using either **CAIN** or alternative tools the users chose to use.

The experiment compared the use of **CAIN** with conventional tools on three measures of performance:

Comprehension as measured by a multiple-choice questionnaire;

Time taken to complete the tasks; and

System usage appreciation assessed by a like/dislike rating at the end of each task.

Performance was examined as function of:

Tool used, either *CAIN* or *Free*;

Web expertise, the Web experience of the users, either *Low* or *High*; and

Subject expertise, the users' expertise in the target subjects topic, either *Low* or *High*.

In order to implement the experiment, volunteers were recruited from the university's population[1], specialist help was required to build the domain model and to provide the subject abstracts and questionnaires and a specially designed environment was prepared to implement the experiment.

DESIGN AND PROCEDURE

The devised task consisted of carrying out two learning assignments performed in sequence and assessed at each one's completion.

The task could be accomplished using either **CAIN** or any other Web navigation tool.

As shown in the diagram presented in figure 1, 40 volunteers were used as test subjects. Out this 40, 20 volunteers performed the test using **CAIN** whereas the other 20 accomplished it using whatever tools they wanted to.

Further, each of the using **CAIN** and not using **CAIN** groups were divided into two 10 volunteers groups each one doing the learning assignments in an opposite order.

---

[1] Universidade Fernando Pessoa, Porto, Portugal

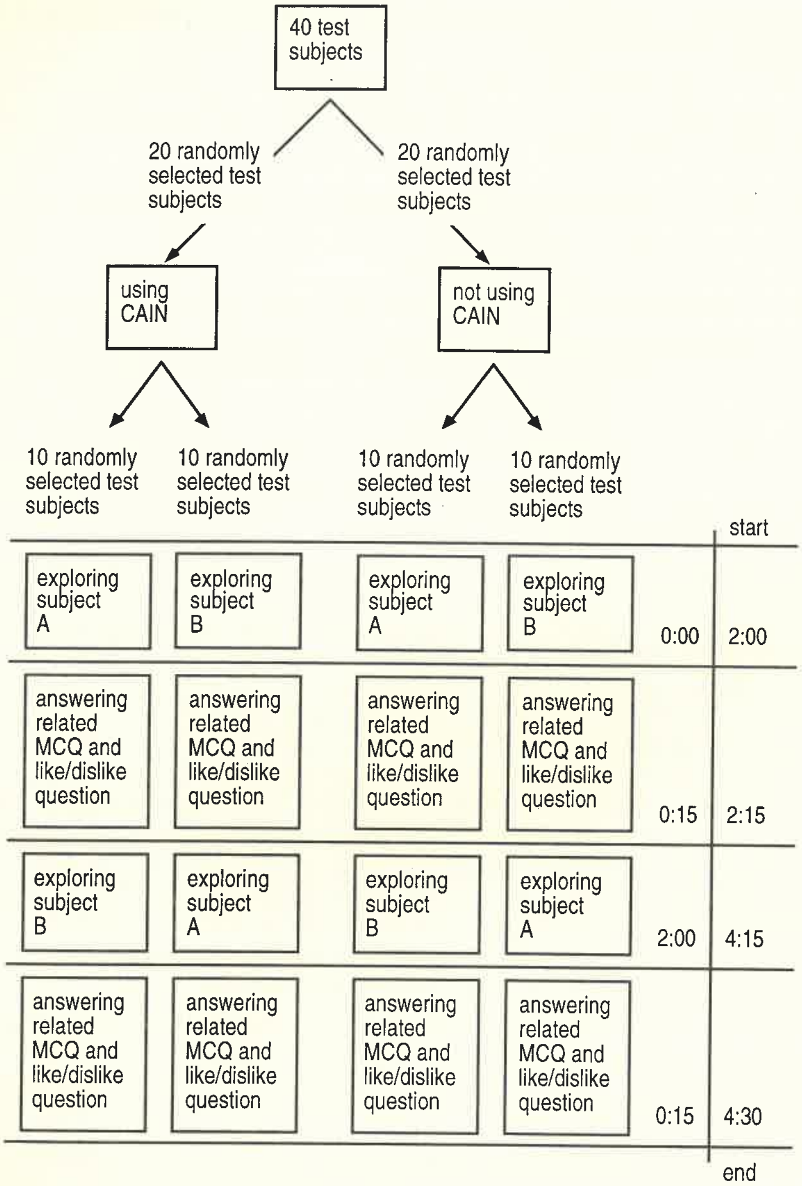

Figure 1: Simple design and procedure diagram

The result was fully counterbalanced with the used tool as a between subjects independent variable.

Each 10 volunteers group was then assigned a task variation either *CAIN AB*, *CAIN BA*, *Free AB* and *Free BA* depending on the used tool and on the learning assignment order (see figure 1).

To collect data, a set of multiple choice questions were set up to be answered[2]...

Before the test started — A self-assessment on Web expertise ranging from 1 to 7 was conducted.

Before each of the learning tasks started — A self-assessment on the learning task's subject expertise ranging from 1 to 7.

After completion of each of the learning tasks — A ten question multiple choice questionnaire, with four possible answers, to assess comprehension over the given subject plus a system usage appreciation question ranging from 1 to 7.

Further, the entire procedure time boundaries was previously set, as shown in figure's 1 framework, and each volunteer's learning tasks performance time was measured with a 60 seconds resolution[3].

Volunteers were tested individually in a specially prepared and isolated environment and were only interacted with before the test started, before each of the learning tasks started and right after each of the learning tasks completion for data collection purposes.

MATERIALS AND APPARATUS

As the learning assignments and respective multiple choice questionnaires play such an important role in this experience, two specialists[4] were invited to define them and to provide all relevant data from the **CAIN**'s domain model to the assignment descriptions and the comprehension assessment multiple choice questionnaires.

Setting up the prototype

Building the domain model was done by hand, working with each assignment's specialist at a time. While the specialist searched and browsed the Web for relevant resources, every selected Web resource was added, with the relevant metadata and ratings, to the system's domain model.

The user model was set up as if the system believed that the volunteers did not know anything about the learning assignments subjects.

---

[2] these are, in the same order, the *Web expertise*, *Expertise A*, *Expertise B*, *Score A*, *Score B*, *System A* and *System B* variables
[3] these are the *Time A* and *Time B* variables
[4] professors Estela Pinto Ribeiro Lamas and Feliz Ribeiro Gouveia, both from Universidade Fernando Pessoa

In fact, two separate but identical users and respective user models were prepared so that two tests could be run in parallel.

All these system initialisation procedures were tied together enabling system resetting between distinct test runs thus ensuring the same conditions to each and every volunteer.

The questionnaires

Task scripts learning assignment descriptions and the various questions were specially prepared to cater for each of the four experience task variations.

Everything was written in Portuguese as the experiment was conducted in Portugal.

Apart from ensuring the specific multiple choice questionnaires, the subject specialists also provided marking grids so that data could be gathered as soon as the tests were accomplished.

The test site

The test site was designed so that volunteers could enjoy suitable silence and isolation.

The test site included:

- A 17 inch colour monitor;
- A standard keyboard;
- A standard two button mouse;
- A personal computer running Windows 95 and Netscape Communicator; and
- A comfortably sized table and chair.

In fact, two of such test sites were made available and used in order to accommodate the different user's schedule preferences.

Finally, a pilot experiment was run before the main one with four extra volunteers to tune up a number of the experiences parameters such as the time limits, the learning assignments descriptions, the questionnaires and the overall set-up but no considerable changes were needed.

Volunteers

An opportunity sample of 40 volunteers was used in the experiment. All volunteers belong to the university's population and were the first 40 ones responding to the call for volunteers broadcasted throughout the university.

The sample consisted of 24 undergraduate students, 12 postgraduate students and 4 staff members of which 28 males and 12 females.

All had English reading skills and none suffered from any form of perceivable impairment.

No financial inducements were offered and written consent was obtained from each of the volunteers.

Results and discussion

The results of the experiment were further processed as follows and the resulting three new independent variables are used instead of the original ones:

- As this experiment has a counterbalanced design, the order in which the volunteers underwent the conditions does not have to be separately analysed. Thus, the initial *Task* variable with four possible values *CAIN AB*, *CAIN CA*, *Free AB* and *Free BA* was converted in the *Tool* grouping variable with only two values *CAIN* and *Free*.

- In order to class volunteers as either *High* or *Low* in Web expertise, a mean was taken of the overall assessed Web expertise and volunteers who scored above the mean were assigned to the high Web expertise group and volunteers who scored below were assigned to the low Web expertise group resulting in the new *Web expertise* grouping variable.

- The same was done with subject expertise resulting in the new *Subject expertise-grouping* variable.

Finally, comprehension assessment scores were collapsed across subject A and subject B multiple choice questionnaires' scores resulting in the new Average score variable.

The same technique was applied to the time taken to accomplish the learning assignments and to the system usage appreciation resulting in the two new variables *Average time* and *Average system*.

The new variables meanings and usage conventions are explained in table 1.

Variable	Meaning
Tool	The tool used by the volunteer to perform the task classified as either CAIN or Free
Web expertise	The volunteer's self assessed Web expertise classified as either Low or High
Subject expertise	The volunteer's self assessed learning task subject's expertise classified as either Low or High
Average score	Average comprehension score achieved by a volunteer after completing a learning task ranging from 0 — minimum — to 10 — maximum
Average time	Average time taken by a volunteer to complete a learning task measured in minutes
Average system	Average used system appreciation by volunteer rated from 1 — minimum — to 7 — maximum

Table 1: Variables' meanings

The following sections assess the three dependent measures of performance:

Comprehension — As measured by the *Average score* variable.

Time taken — Measured by the *Average time* variable.

System usage appreciation — Measured by the *Average system* variable.

COMPREHENSION

As the results of the self assessment conducted before each learning task started showed that subject expertise was equivalent for the *CAIN* and *Free* groups ($t(38)=0.276$, $p=0.784$), a difference of only 2.5% (approximately), a meaningful comparison between the comprehension performance of the *CAIN* and *Free* groups was possible.

A *t*-test comparing the comprehension score means shows that volunteers comprehension scores are significantly higher in the *CAIN* group than in the *Free* group ($t(38)=2.9$, $p=0.006$).

In fact, taking the comprehension score group statistics, presented in table 2, and the mean subject expertise and comprehension scores contrast, presented in table 3, it is possible to see that **CAIN** produces an improvement in comprehension scores of nearly 30% compared with the other approaches used by the volunteers.

Tool	*Volunteers*	*Mean*	*Standard deviation*	*Standard error mean*
Free	20	3.8750	1.2126	0.2711
CAIN	20	5.0500	1.3367	0.2989

Table 2: Comprehension score group statistics

	Tool	
Mean	Free	CAIN
Subject expertise	3.03	3.10
Average score	3.88	5.05

Table 3: Mean subject expertise and comprehension score contrast

Figure 2 illustrates comprehension score results depicted in table 2.

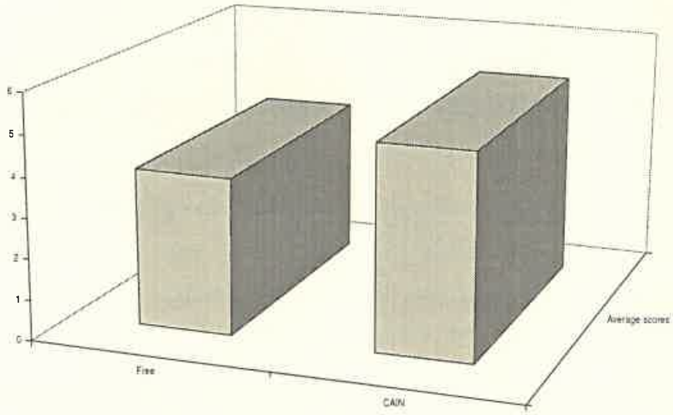

Figure 2: Comprehension scores

In order to understand the circumstances behind these results an analysis of variance of comprehension score means, with *Subject expertise*, *Tool* and *Web expertise* as the three independent variables was conducted.

This analysis shows a significant main effect of *Tool* such that comprehension scores were higher after completing a learning assignment with *CAIN* than after completing a learning assignment with *Free* ($F(1,32)=4.85$, $p=0.035$). This repeats the result from the *t*-test and in this case, no other variable or interaction between variables has enough statistical significance ($p>0.05$) to be regarded as influential.

Separate analyses of variance were further conducted on comprehension scores for subjects *A* and *B* which revealed that the main effect of *Tool* was highly significant on comprehension scores for subject *B* ($F(1,32)=8.554$, $p=0.006$) but not for subject *A* ($F(1,32)=0.709$, $p>0.05$).

Subject *A* was *cryptography* and subject *B* topic was *poetry*. These results may reflect the fact that cryptography Web sites have better structure or are easier to find and use than poetry Web sites and so benefit little from the use of **CAIN**.

The separate analyses of variance also revealed a significant main effect of Web expertise for subject *A* ($F(1,32)=5.979$, $p=0.020$) and subject *B*

($F(1,32)=5.353$, $p=0.027$) which did not appear in analysis of variance of comprehension score means.

This situation results from a crossover interaction between Web expertise and comprehension scores for subjects *A* and *B* (see table 4 and figure 3).

Web expertise	*Overall mean*	*Subject A mean*	*Subject B mean*
Low	4.38	4.29	4.47
High	4.52	5.83	3.22

Table 4: Comprehension scores means by *Web expertise* level

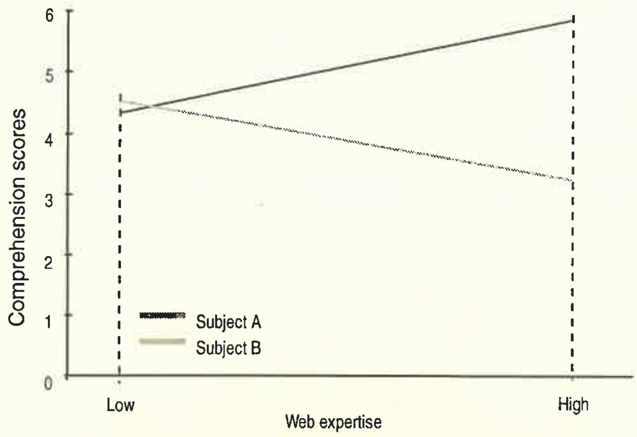

Figure 3: Crossover interaction between Web expertise and comprehension scores

Although unexpected, these results suggest that the subject topics used — cryptography and poetry — are at opposite ends of a qualitative range for Web sites and that by using them both, this experiment neutralised their differences and provided a fair testing environment.

Finally, table 5 and figure 4 provides a straightforward comparison between the mean comprehension scores under every condition combination.

		Subject expertise		
Web expertise	*Tool*	*Low*	*High*	*Mean*
Low	Free	3.80	3.50	3.65
	CAIN	4.25	5.33	4.79
High	Free	2.66	4.25	3.46
	CAIN	5.87	4.75	5.31
Mean		4.15	4.48	

Table 5: *Web expertise* X *Tool* X *Subject expertise* comprehension scores contrast

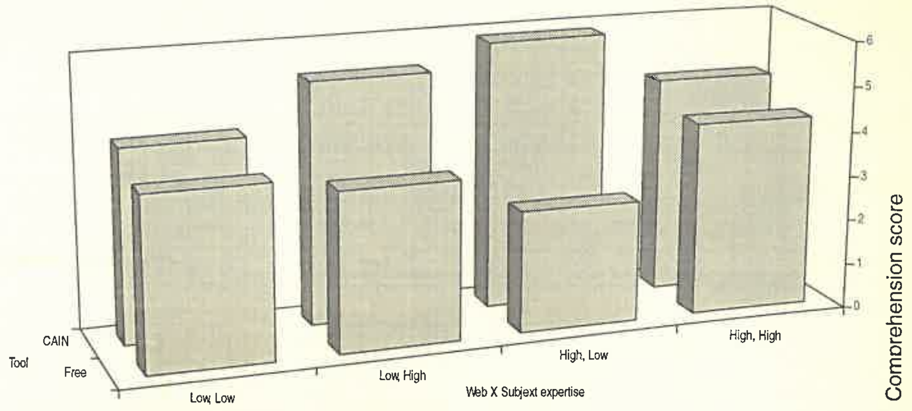

Figure 4: Comprehension scores contrast

TIME TAKEN

Next, an analysis of variance of time performances was conducted with the same *Subject expertise*, *Tool* and *Web expertise* independent variables.

Although overall learning assignment completion times are quicker with *CAIN* than with *Free*, this difference is *not* statistically significant ($F(1,32)=2.05$, $p>0.05$).

Nevertheless, several interesting remarks can be made:

- There is a significant main effect of Web expertise such that highly experienced Web users performed faster than the low experienced Web users ($F(1,32)=9.02$, $p=0.05$). As all the experiment was conducted over the Web, it is natural that highly experienced Web users performed faster because they were already used to the environment.

- There is an especially interesting 3-way interaction between *Subject expertise*, *Tool* and *Web expertise* ($F(1,32)=4.61$, p=0.039).

Table 6 and figure 5 illustrate this last result highlighting the fact that *CAIN* produces faster performance than *Free* under all combinations of levels of the independent variables except when the volunteer has *High* Web expertise and *High* subject expertise, *i.e.* it suggests that:

- Time taken may be slower with **CAIN** when the user is a highly experienced Web user and an expert in the learning assignment field, but;

- **CAIN** is faster when users are inexperienced Web users and inexperienced in the study subject or inexperienced in the Web and experienced in the subject or experienced in the Web and inexperienced in the subject.

Web expertise	Tool	Subject expertise Low	High	Mean
Low	Free	91.70	86.50	89.10
	CAIN	86.17	68.50	77.34
High	Free	74.67	61.50	68.09
	CAIN	50.50	75.44	62.97
Mean		75.76	72.99	

Table 6: *Web expertise* X *Tool* X *Subject expertise* performance times contrast

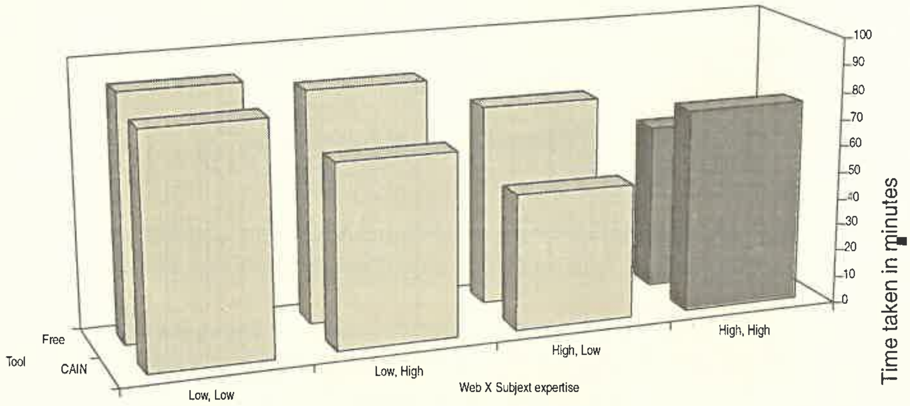

Figure 5: Performance times contrast

Please note that, as shown in figure 6, Web expertise alone does not slow performance when using **CAIN**. In fact, highly experienced Web users performances are slightly faster under *CAIN* than *Free*.

It is the particular combination of *High* Web expertise and *High* subject expertise that slows performance-using **CAIN** compared to using other tools on the World Wide Web.

Further analyses of variance, conducted separately on time performances for subjects *A* and *B*, revealed that the main effect of *Tool* was specially significant for subject *B* ($F(1,32)=5.783$, $p=0.022$) but not significant for subject *A* ($F(1,32)=0.017$, $p>0.05$).

These results provide further evidence that cryptography Web sites may differ in quality or quantity from poetry Web sites and that **CAIN** may be more effective when addressing topics whose Web sites have characteristics similar to those of poetry Web sites.

The same analysis of variance confirmed the significance of the main effect of Web expertise on both subject *A* ($F(1,32)=4.819$, $p=0.036$) and subject *B* ($F(1,32)=5.362$, $p=0.027$) performance times.

This only indicates that although **CAIN** improves time performance in general, Web expertise still plays an important role as novice Web users spent more time than experienced Web users. This does not come as a surprise since **CAIN** is itself Web based and it does not replace existing navigation tools, it just provides non-obtrusive navigation support.

SYSTEM USAGE APPRECIATION

As for system appreciation, the *t*-test for the equality of system usage appreciation mean results does not show any significant difference between the *CAIN* and the *Free* groups ($t(38)=0.698$, $p=0.489$).

In fact, table 7 shows that *CAIN* rated 5.7 out of 7 while *Free* only rated 5.475 out of 7. This is only a 4% increase.

Average system	*Volunteers*	*Mean*	*Standard deviation*	*Standard error mean*
Free	20	5.475	1.0696	0.2392
CAIN	20	5.700	0.9652	0.2158

Table 7: System usage appreciation group statistics

Table 8 and figure 6 provides a straightforward comparison between the mean like/dislike scores under every condition combination.

		Subject expertise		
Web expertise	*Tool*	*Low*	*High*	*Mean*
Low	Free	6.00	5.50	5.75
	CAIN	5.75	6.00	5.88
High	Free	4.75	5.75	5.25
	CAIN	4.86	5.92	5.39
Mean		5.34	5.79	

Table 8: *Web expertise* X *Tool* X *Subject expertise* system usage appreciation contrast

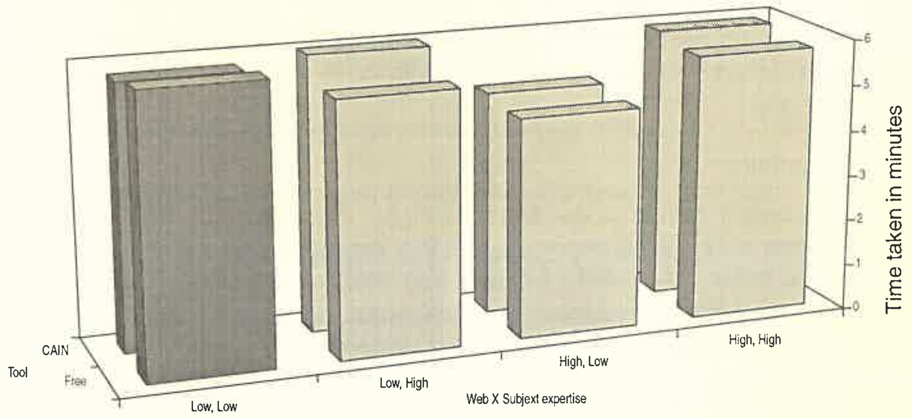

Figure 6: System usage appreciation

As shown, satisfaction was higher for **CAIN** under almost every condition combination except for users ranking low in both Web and subject expertise. This was probably due to the extra effort required to cope with new systems, the Web and **CAIN**, and a new subject at the same time.

Finally, in this case there was no significant difference between the independent analysis of system usage appreciation for subjects *A* and *B*.

RELATED EMPIRICAL EVIDENCE

 Although to date empirical studies on the value of adaptivity in hypertext based learning environments are limited, they are critically important to validate the different approaches in this research area.

Although most of the existing empirical evidence, reviewed by Eklund and Brusilovsky (1998), fails to show if adaptive hypertext navigation contributes at all to improve the student's or researcher's performance, it provides an intuition that adaptive navigation could be valuable.

Nevertheless, in experiments measuring the number of hypertext nodes used to accomplish some predefined goal, existing empirical evidence favours *adaptive curriculum sequencing*[5] versus adaptive link annotation over a closed information corpus. Although not directly related with **CAIN**, these results do not contradict this project's approach.

Further, on the InterBook[6] experiment reported in the same review, apart from other results, it is also stressed that the direct guidance mechanism was chosen by test subjects in over 90 % of the transactions. Again, this system relies on a closed information corpus but nonetheless, the experiment's results provide added support to **CAIN**'s approach.

This is however an under-investigated area and there is clearly a need for continued studies on the value of adaptivity in hypertext based learning environments.

FINAL REMARKS

The experiment herein presented examined the nature of the interaction between the users' Web expertise, Subject expertise and the used tool and its results show that **CAIN**, as specified and implemented, does indeed improve user's comprehension by 30 % when supporting learning or initial research tasks performed over the World Wide Web.

---

[5] a direct guidance approach

[6] an adaptive hypertext system supporting adaptive link annotation and direct guidance in the form of a *continue* link

This is probably due to the way **CAIN** guides the user through a sequence of previously selected and relevant Web resources.

Though not in a very significant way, the experiment's results also suggests that the time taken to accomplish such tasks can also be reduced when using **CAIN** except when there is no point in using it *i.e.* when the user is already an expert using the World Wide Web and an expert on the task's subject.

As far as enjoying the system, the results do not show any significant difference between using **CAIN** or any other approach to navigate the World Wide Web. In fact, the results show that users in general enjoy using the Web with whatever tools they use.

Although encouraging, this experiment's results are not sufficient to understand all the implications of a system like **CAIN** and further studies must be undertaken. In fact, the current investigation only measures the adaptive guidance support leaving untouched issues related to all the other relevant areas brought together in this work such as:

- The adequacy of the metadata records and associated rating strategy;
- The soundness of the implicit co-operation model;
- The reliability of the relevance feedback mechanism; and
- The validity of adaptivity itself.

Specifically, it fails to understand other success factors such as:

- How relevant is the order in which Web resources are presented to the user?
- How do users cope with the domain building process?
- Are the rating criteria adequate and easy to apply?
- Is such Web representation successful for all its users?
- How does the system cope with long term users?
- How does physical distance affect the system's collaborative nature?

This is an under-investigated area and there is clearly a need for continued studies on the value of adaptivity in hypertext based learning environments.

In conclusion, the **Computer Aided Information Navigation** seams to be in the right direction and research will continue in order to achieve this collaborative environment's long term goals: To *guide* the user while *learning*, *memorising* and *forgetting* things in order to increase *motivation*, *attention* and *empathy*, reducing the environment's noise and improving the knowledge acquisition process.

ACKNOWLEDGEMENTS

The author whishes to acknowledge Universidade Fernando Pessoa for funding and resources, its Multimedia Resource Center for work environment and scientific support and all who read and help mature this study and report.

BIBLIOGRAPHY

BENYON, D. & MURRAY, D. "Adaptive systems: from intelligent tutoring to autonomous agents", *Knowledge-Based Systems*, 6.4, (1993): pp. 197-219.

BROWN, P. "Assessing the quality of hypertext documents", *Proceedings of the ECHT'90*, UK, 1990.

BRUSILOVSKY, P. "Methods and techniques of adaptive hypermedia", *User Modeling and User-Adapted Interaction*, 6, (1996): pp. 87-129.

EKLUND, J. & BRUSILOVSKY, P. "The value of adaptivity in hypermedia learning environments: a short review of empirical evidence", *On-line proceedings of the Second Workshop on Adaptive Hypertext and Hypermedia*, http://wwwis.win.tue.nl/ah98/Eklund.html, Agosto 29, 1998.

EKLUND, J. & ZIEGLER, R. "Navigating the Web: possibilities and practicalities for adaptive navigational support", *AusWeb 96: Second Australian World Wide Web Conference Proceedings*, Southern Cross University, Queensland, Australia, 1996.

LAMAS, D. "Adaptivity on the Internet", *University of Portsmouth research seminars*, University of Portsmouth, UK, 1998.

LASSILA, O. & SWICK, R. "Resource description framework RDF model and syntax", http://www.w3.org/Tr/WD-rdf-syntax/, Janeiro 9, 1998.

MARSHALL, A. "Developing hypertext courseware on the World Wide Web", *Educational multimedia and hypermedia: proceedings of Ed-Media 95*, Association for the Advancement of Computing in Education, 1995.

SHUM, S. "The missing link: hypermedia usability research & the Web", *ACM SIGCHI Bulletin*, 28.4, (1996): pp. 221-248.

SHUM, S. & MCKNIGHT, C. "World Wide Web usability: introduction to this special issue", *International Journal of Human-Computer Studies*, 47.2, (1997).

WEIBEL, S. & LAGOZE, C. "An element set to support resource discovery: the state of the Dublin Core: January 1997", *International Journal on Digital Libraries*, 1.2, (1997): pp 176-186.