Automated Computational Cognitive-Modeling: Goal-Specific Analysis for Large Websites

PAUL VAN SCHAIK, Teesside University
RAZA HABIB MUZAHIR, Teesside University
MIKE LOCKYER, Teesside University

The information architecture of websites is the most important remaining source of usability problems. Therefore, this research explores automated cognitive computational analysis of information architecture of large websites, as a basis for improvement. To support goal-specific analysis, an enhanced model of web navigation was implemented with a novel database-oriented approach. Web navigation was simulated on the information architecture of two large sites. With improved the labeling system of the information architecture, simulation results showed a significant reduction in navigation problems. The results of two experiments demonstrate that sites with improved information architecture results in better outcomes of users' information retrieval. Our database-oriented approach is extensible, also allowing non-goal-specific analysis, modeling of non-text media content, and analysis of the organization and navigation systems of information architectures.

Categories and Subject Descriptors: **H.5 [INFORMATION INTERFACES AND PRESENTATION (e.g., HCI)]**: Model Development; H.1.2 [Models and Principles]: User/Machine Systems—Human information processing; H.5.2 User Interfaces; H.5.4 [Information Interfaces and Presentation]: Hypertext/Hypermedia—Theory; **I.6.5 [Simulation and Modeling]**

General Terms: web navigation, simulation, usability

Additional Key Words and Phrases: computational cognitive modeling, website, information architecture, goal-specific analysis, CoLiDeS

ACM Reference Format:

Paul van Schaik, Raza Habib Muzahir, and Mike Lockyer, 2015. Automated computational cognitive-modeling: goal-specific analysis for large websites. *ACM Trans. Comput.- Hum. Interact.* xx, x, Article xx (<month> 2015), xx pages.
DOI:http://dx.doi.org/xx.xxxx/xxxxxxx.xxxxxxx

1. INTRODUCTION

When using Internet sites or intranet sites, most of us have experienced 'information pain' [Morville & Rosenfeld, 2006] in situations where the required information was hard or impossible to find, which sometimes can be just annoying and at other times can be the fundamental obstacle to complete a task. A central factor contributing to this problem is poor design of information (in particular the organization and labeling of information elements, and navigation through the information structure). Indeed, Nielsen [2009] reports that the information architecture of websites is the most important remaining factor causing usability problems for site users.

1.1 Information architecture

Information architecture is an important area of contemporary research [Jacob & Loehrlein, 2009] in psychology and computer science. Although no single accepted definition exists, the main aim of information architecture is to provide effective access to relevant online information resources (usually delivered through a website). Information architecture focuses on systems for the organization, labeling and navigation of information [Morville & Rosenfeld, 2006] for the benefit of end-users who need to find and use information. The role of an information architecture is to provide the design of information for building a website, but does not address information presentation in a site in terms of, for example, layout and the use of color and graphics.

The success of online information systems will therefore to a large extent be positively influenced by the extent to which their information architectures support end-users in finding and using information. This requirement demands a new research effort into information architecture, as online information systems are increasingly being used as a mechanism to enhance human task performance and are accessed by a variety of end-users; moreover, information architecture may influence users' web navigation in different ways.

Organization. Exact *organization schemes* (e.g., alphabetical, chronological and geographical) are useful when users know exactly what they are looking for [e.g., name, date or geographical location; Rosenfeld & Morville, 2006] and can make an exact match between their goal (e.g., information need) and major headings or links. Otherwise and in most cases, ambiguous schemes [e.g., topic, task, audience and metaphor; Rosenfeld & Morville, 2006] can be useful. In these schemes items are grouped in meaningful ways. Organization by topic requires that items are arranged in a conceptual structure matching that of target users, so users can determine whether or to what extent particular items match their information need, based on the users' conceptual structures. Organization by task has a similar requirement, but now the procedural structure needs to match that of target users. McDonald et al. [1990] give examples of creating a topic-based organization scheme from users' conceptual structures and task-based organization scheme from users' procedural structure in the same domain. *Organization structures* include hierarchy (top-down) and database model (bottom-up) [Rosenfeld & Morville, 2006]. Because a top-down structure is deliberately designed, users' success in finding information will depend on the extent to which the structure matches users' conceptual structures. A bottom-up structure is not always visible in a website with this underlying structure. However, information items are tagged with controlled vocabulary metadata to support searching and browsing. As a result, this approach is most useful when content is relatively homogeneous. In this case, the challenge is to make sure that the metadata match the users' semantic knowledge of the domain. Note that organization concerns the grouping of content, irrespective of the exact language (labeling) that is used.

Labeling. All (linguistic) elements (e.g., headings, links) of an information architecture need to be labeled. The requirement here is to make sure that labels match content that they represent are differentiated according to their content and are meaningful, all according to the users' semantic knowledge. Existing research [e.g., Blackmon et al., 2002] has examined how labeling of headings and links affects web navigation.

Navigation. *Embedded navigation* systems are integrated with website content. In a typical website, the global navigation system may be a set of links in the top menu bar of every web page in a site, giving direct access to major areas or functions in a conceptual hierarchy. The local navigation system may be a set of links in the left menu bar, with specific links depending on the choice made in the global navigation system, giving access to lower-order areas. Contextual navigation system may consist of various links in the content area (below the top menu bar and to the right of the left menu bar), giving access to specific related content at the lowest level of the conceptual hierarchy. These three navigation systems may reinforce the organization structure; therefore, as with organization structure, it is important that the conceptual structure represented by these systems matches the conceptual structure of target users. *Supplemental navigation* systems reside outside content pages. (Large) websites that lend themselves to hierarchically organization can benefit from a sitemap, showing this organization (typically at two or more levels) as a conceptual structure, which should match that of target users. For other websites an index can be appropriate, consisting of alphabetically presented keywords or phrases organized on only one or two levels. An index can be effective for users who can exactly articulate their information need(s) in terms of the item(s) to find on a particular website. Therefore, the design of an effective index requires an accurate description of information needs in terms that target users would use.

Existing research using cognitive computational modeling in relation to information architecture has mostly focused on relatively small websites that may appear to lack realism in representing real-world sites that serve large populations of end-users. In terms of the organization of information, research has focused on, for instance, the trade-off between depth (number of hierarchical levels of information) and breadth (number of information items per level) of information hierarchies without considering the most appropriate scheme for organizing the information (e.g., organization by time, place or a particular theme) to start with [e.g., Parush & Yuviler-Gavish, 2004]. Furthermore, the creation of comparable alternative organization schemes for the same information may have arguably led to artificial information hierarchies that may not generalize to information architectures in the real world. Given the disparate and patchy nature of existing information-architecture research and shortcomings reviewed above, our program of research set out to develop a body of knowledge about effective information architecture for large websites, with at least one hundred web pages, for end-users.

1.2 Cognitive computational modeling for information architecture

The use of established empirical psychological techniques for eliciting information structures, labeling schemes and navigation structures (e.g., card-sorting and rating) has been proposed to inform the information-architecture design of sites [e.g., Katsanos et al., 2008], but their use becomes unwieldy and impractical, and therefore does not scale up, for large information architectures. Reasons for this are the immense cognitive complexity of the elicitation tasks (in the case of card-sorting) not matched by humans' limited information-processing capacity and the massive amount of time humans would require to complete these tasks (both card-sorting and rating). Therefore, cognitive computational techniques with a cognitive-psychological basis to mimic human capabilities to process the information contained in large websites should be explored to analyze and assist in the analysis and design of the information architecture of large websites. For example, Latent Semantic Analysis (LSA) is an established machine-learning technique with a strong psychological basis, corroborated by the results of experiments in various areas of cognitive psychology [e.g., Landauer, 2007]. LSA analyzes large volumes of text. Extracted text is represented in matrix format. Singular value decomposition is first performed, resulting in a more compact representation with hundreds rather than thousands of dimensions. The condensed representation is then used to calculate the similarity between represented words, short texts (e.g., sentences and phrases) and whole texts, expressed as a cosine value. For example, when words are very similar, their cosine value is close to 1, but when words are very dissimilar the value close to 0. Existing research has addressed how redesigned information architecture for mainly relatively small sites based on psychological models, using LSA, can improve users' performance on these sites [e.g., Blackmon et al., 2005].

Our research therefore uses cognitive computational techniques to aid in the analysis and design of information architectures for large websites. The intended beneficial effect of the use of these techniques on users' task performance is then tested. The knowledge and guidelines produced by this research can help software developers to analyze and design websites that support end-users better in finding information.

1.3 Current study

In relation to cognitive computational modeling, Blackmon et al. [2002] distinguish two types of analysis: goal-specific and non-goal-specific. The first involves the simulation of web navigation by users in pursuit of particular goals ("information about a users' understanding of their tasks and underlying motivation", p. 463). The second involves the analysis of the familiarity of web page elements (e.g., headings and links), the distinctiveness of elements that need to be distinct in meaning (e.g., links under a particular heading need to be distinct) and of the similarity of elements that need to be similar in meaning (e.g., a link under a particular heading needs to be similar to the heading). This paper focuses on goal-specific analysis, while other work addresses non-goal-specific analysis [e.g., Blackmon et al., 2002; Muzahir, 2013]. This research builds on existing work on computational cognitive modeling of web navigation that focuses on labeling as a crucial element of information architecture [Blackmon et al., 2002, 2003, 2005; Van Oostendorp & Juvina, 2007]. Other work may address the organization system and navigation system as other elements of information architecture. For example, research may study how human perception of semantic similarity is related to human comprehension of the structure/organization of the website [e.g. Resnick & Sanchez, 2004].

We favor a usability engineering approach to the application of models of web navigation, based on two related ideas: (1) perfect knowledge is neither possible nor necessary (any model is an imperfect representation) to create useful results [Cudeck & Henley, 2003; Nielsen, 1993] and (2) imperfect knowledge that can be used to make predictions for improving the outcomes of web navigation is sufficient. The first aim of this paper is to provide automated cognitive computational goal-specific analysis of the information architecture of large websites, as a basis for improvement. The second aim is to validate this analysis through experiments with end-users. Novel aspects of this work include the automated analysis of large real-world websites with an enhanced web navigation model and, crucially, the separation of a website from its persistent representation through a database-oriented approach. The paper is organized as follows. Section 2 presents an enhanced model for goal-specific analysis and a novel implementation of the model. In Sections 3 and 4 this implementation is applied to analyze two large websites and experiments are reported to validate the model results. Section 5 provides a general discussion of this work and directions for future work.

2. ENHANCED GOAL-SPECIFIC ANALYSIS AND NOVEL IMPLEMENTATION

2.1 Models of web navigation and the role of spatial ability

According to Pirolli and Card [1999], users take into consideration the cost and value of choosing a particular action on a web page such as clicking on a link. In doing so, they select an element having the highest value of information scent [Chi et al., 2000, p. 162]: "Foragers use these proximal cues (snippets; graphics) to assess the distal content (page at the other end of the link). Information scent is the imperfect, subjective, perception of the value, cost, or access path of information sources obtained from proximal cues, such as Web links, or icons representing the content sources." More narrowly, information scent has been described as "the degree of semantic similarity between representative user goal statements [100-200 words] and heading/link texts on each web page" [Blackmon et al., 2002, p. 463] Several cognitive models of web navigation have been developed [e.g., CoLiDeS, Kitajima et al., 2000, 2005; CoLiDeS+, Van Oostendorp & Juvina, 2007; SNIF-ACT, Fu & Pirolli, 2007; and MESA, Miller & Remington, 2004; for a review see Katsanos et al., 2010], motivated by the concept of information scent.

The two most important models of web navigation are SNIF-ACT and CoLiDeS. A strength of SNIF-ACT is that its origins lie in information foraging theory and ACT-R. CoLiDeS and its successors has the advantage of the more faithful and detailed representation of human comprehension, including comprehension of the high-level organizational structure of a text according to Kintsch's [1998] construction-integration theory. For example, the attention phase and the action selection phase in CoLiDeS are inspired by Kintsch's construction phase and integration phase, respectively.

Existing research on web navigation has established that spatial ability is an influential factor in web navigation. For example, Juvina and Van Oostendorp [2006] theoretically argue and report empirical evidence for the idea that spatial ability is a strong positive predictor of task performance in web navigation. Therefore, although the focus of the current research is on information architecture, spatial ability was studied as an additional factor in web navigation to allow for the main effects of both factors and their potential interaction effect to be analyzed.

2.2 Limitations of existing models

A limitation of models such as MESA and SNIF-ACT and, consequently, tools that are based on these models [e.g. CogTool; Teo & John, 2008] is the way they address the concept of information scent; when a semantic or literal match of the goal statement against individual elements is attempted, context (e.g., headings, in addition to links) and the history of links chosen is ignored. However, research demonstrates that users' navigation decisions take into consideration the relevance of links encountered in previous steps (navigation history), which help reaching the current page [Howes et al., 2002]. This limitation is addressed by the Comprehension based Linked model of Deliberate Search (CoLiDeS), taking into account both headings and links in web navigation [e.g., Kitajima et al., 2000], and CoLiDeS+, taking into account navigation history [e.g., Van Oostendorp & Juvina, 2007].

According to CoLiDes, web navigation takes place in three steps. In Step 1 (goal formation), a goal to meet a particular information need is formed. In Step 2 (attention phase), the current web page is parsed into subregions. The action of attending to a subregion having the highest information scent (semantic similarity of heading with goal statement) is selected. In Step 3 (action-selection phase), elements in the selected subregion are elaborated. A link is selected, having the highest information scent, and traversal to the linked page occurs. Steps 2 and 3 are repeated on subsequent web pages until the desired information matching the goal statement is found on a page or the navigation process stops without success. In CoLiDeS and CoLiDeS+, information scent is calculated as the LSA cosine, a measure of similarity between two texts (e.g., the goal statement and a link label).

CoLiDeS+ addresses CoLiDeS' limitation of not taking into account navigation history; CoLiDeS+ augments CoLiDeS with the concept of path adequacy, the semantic similarity between goal statement and link labels on a navigation path. Therefore, path adequacy represents the goal relevance of selected links on previous pages leading to the current page, whereas information scent captures goal relevance of elements on the current page. According to CoLiDeS+, users take into account both information scent and path adequacy in selecting a link on the current page. CoLiDeS+ accounts for backtracking (going back to the last visited page), which occurs if neither information scent nor path adequacy increase. However, CoLiDeS+ ignores the intermediate scent emitted from heading label and links without heading (e.g., links in top menu bar, side bar and bottom bar) and only takes into account links with heading.

2.3 Overcoming limitations

The research reported here extends CoLiDeS [Kitajima et al., 2000] and CoLiDeS+ [Juvina & Van Oostendorp, 2006, 2008] by proposing an enhanced combined CoLiDeS/CoLiDeS+ (abbreviated as CoLiDeS/+). The novel implementation of the enhanced model automates the manual simulation process and therefore scales up the analysis process to large websites. A database-oriented approach for implementing the model is used, separating websites from their persistent database representation, thereby facilitating goal-specific (and non-goal-specific) analysis on large sites.

Building on CoLiDeS and CoLiDeS+, this work can be used to improve a website's information architecture, based on the simulation of a human user's cognitive processes and behavior as a software program. If, according to enhanced CoLiDeS/+ simulation, web navigation for a particular goal fails, the headings and/or links on the path leading to the goal in the website can be improved (reworded) and the simulation program run again to establish if a user with this particular goal would be able to successfully navigate the site after improvement.

Until recently, few published automated simulations of CoLiDeS or CoLiDeS+ existed. However, an attempt has been made to develop a software prototype of CoLiDeS [Kitajima et al., 2005; Karanam et al., 2011], but this work has several limitations (Online Appendix A, Section A1). Thus, although Kitajima et al.'s work is significant in its attempt to automatically analyze a large website (online encyclopedia), there is a need for a better implementation of CoLiDeS and CoLiDeS+. This research proposes an enhanced goal-specific analysis with a novel database-oriented approach to simulate web navigation.

2.4 Enhanced goal-specific analysis

A recursive algorithm was developed comprising of the following six steps; the algorithm selects web elements on the basis of the highest cosine similarity with the goal statement in forward search and allowing backtracking when forward search fails (for details, see Online Appendix B), building on CoLiDeS and CoLiDeS+.

Step 1: segment page.
Step 2: select region.
Step 3: segment region.
Step 4: check if region contains the target of goal statement.
Step 5: select action (link).
Step 6: process next page.

This algorithm was implemented as software to support automated goal-specific analysis of websites. The following are enhanced features of this work.

2.4.1. Paragraphs as basis of goal-directed search. According to Brown [2005], "… a user goal statement … should be 100-200 words in length, and it should represent the user's main goal and sub goal" (p. 27). Instead of manually compiling a set of realistic user-goals by the analyst as proposed by Blackmon et al. [2002], in our work paragraphs identified by <p>-tag in HTML source code on web pages are selected as goals. Paragraphs with the length of 1400 characters (200 words) or more are considered as potential user-goals in simulations. Pages containing the text of these goals become destination pages. This approach adds an element of practicality in selecting representative goals, constrained by the available content in a particular website.

2.4.2. Accounting for links without heading. In order to overcome limitations of CoLiDeS/CoLiDeS+ (Online Appendix A, Section A2), in enhanced CoLiDeS/+, parsing is achieved by dividing the page into sections with headings and a set of headingless sections of links (all links from the top bar, the side bar and the bottom bar). Identified elements are links with heading, paragraphs with heading, and images with heading. Other elements are links without heading and images without heading.

Enhanced CoLiDeS/+ selects the link having the highest cosine with the goal statement from the section having the highest heading cosine with the goal statement on a particular web page. Enhanced CoLiDeS/+ identifies this link and also selects the link having the highest cosine with the goal statement from all links without heading on the same page. The algorithm then selects from these two the link having the highest cosine value. This link is then traversed because it has the highest information scent on a particular web page and thus is more likely to be selected by a user. If neither of these two identified links has an increasing cosine value then the algorithm returns to the previous page and checks the next link on that page in descending cosine order.

Modeling headingless links and including these in simulations is an advance over previous research using CoLiDeS and CoLiDeS+ and necessary for many present-day information architectures in which such links are present. This advance should be seen as a first step towards more sophisticated work that may differentiate frequently used links in the top bar from those in the left sidebar and less frequently used links in the bottom bar.

2.4.3. Including heading label in the evaluation of link label. Theories of human text comprehension [Kintsch, 1998; Rapp & Van den Broek, 2005] provide a motivation for adding heading text to link text when computing the scent of links nested within a region with a heading. For example, according to the landscape model [Rapp & Van den Broek, 2005], through subsequent cycles in the reading process residual information from the preceding cycle (among other factors) influences the activation of concepts. Therefore, in the case of web navigation the heading labels that were processed in previous cycles would influence the activation of concepts in the current cycles. Consequently, enhanced CoLiDeS/+ includes heading label along with link label in calculating a cosine between the link under a heading and the goal statement. Thus, the text of a particular link under heading now becomes heading label + the link label. As links under a heading should extend the meaning of heading label, this is a way to enhance the scent of links under a heading. As a result, when the scent of link labels is evaluated in a simulation, the cosine will be at least as high as it would be without the heading label included. Therefore, all else being equal, the simulation would be more likely to succeed on a particular goal.

2.4.4. Selecting link with cosine higher than heading cosine. In contrast to CoLiDeS+ (Online Appendix A, Section A3), in enhanced CoLiDeS/+, at every stage when the focused-on area is a link under a heading, the heading cosine value is set as a cut-off value for all the links under that heading. Therefore, the cosine of link label (heading label + link label; see previous section) should be higher than, or at least as high as, the cosine of its corresponding heading.

2.4.5. Automated reporting of entire traversed path. Enhanced CoLiDeS/+ reports the entire path followed to reach the target page after any backtracking, when the simulation terminates matching a particular goal statement with the web page content. Moreover, this traversed navigated path is presented for comparison with the shortest path to the target page.

2.5 Novel implementation

For flexibility, control, efficiency of analysis and extensibility, it was important to separate a to-be-analyzed website from the representation of its information architecture. This was achieved by taking a database-oriented approach to create a persistent representation. In terms of flexibility, this allows the creation and comparative goal-specific analysis of numerous information architecture versions of the same website without actually building different website versions. By contrast, the information architecture of a live website, rather than its representation as a persistent database, that is analyzed will normally change over time and is not subject to experimental control by researchers; therefore, different analyses will, in fact, analyze different versions of the same information architecture without experimental control. In terms of efficiency, the approach avoids overhead of parsing web pages1 during simulation in goal-specific analysis; moreover, various other types of analysis have been developed [e.g., non-goal-specific analyses; Muzahir, 2013] or can be developed (see Section 5.2), all using the same database. In terms of extensibility, the approach allows for other aspects of web page design to be represented and analyzed (e.g., images).

A database schema was designed consisting of tables for (a) a website's information architecture and (b) goal-specific analysis.2 For (a) there were tables for the following entities: web page, heading, text paragraph, link with heading, link without heading, headingless section of links, image with heading, image without heading, modified heading, modified link with heading and modified link without heading. For (b) there were additional tables for the following entities: heading cosine, path-adequacy cosine for heading, cosine for link under heading, path-

---

[1] Parsing (of a particular website) is done once and the resulting representation is stored in a database for further use in any analysis of the represented website. Any improvements made to the website's labeling system can be stored in the database without the need for further parsing.

[2] Additional tables were also created for non-goal-specific analysis (Muzahir, 2013).

adequacy cosine for link under heading, cosine for link without heading, path-adequacy cosine for link without heading, text paragraph cosine, solution path and simulation success.

The algorithm (Online Appendix B) was implemented in Visual C#.NET using Visual Studio 2008. Data were stored in Microsoft SQL Server and accessed using SQL Server Management Studio for manual access and using Subsonic 3.0 (as Object Relational Mapper) for access through the C# code. The implementation was coded using a Windows form application. Enhanced CoLiDeS/+ simulation used the AutoCWW2 LSA tools (*http://autocww2.colorado.edu/HomePage.html*, in particular *http://autocww2.colorado.edu/OneToMany.html*) to calculate cosines. The semantic space for first-year college (*http://lsa.colorado.edu/spaces.html*) was used, in agreement with participants' education level in the two studies that are reported here (see also Section 3.1.1). The simulation was implemented on Windows 7 Enterprise, using 2 dual-core processors of 2.53 GHz each with 12 GB of RAM on a 64-bit operating system.

3. STUDY 1: LARGE INTERNET SITE

3.1 Goal-specific analysis and improvement

3.1.1. Simulation (original information architecture). A university Internet site was captured, the 11000 web pages were parsed and the site was recreated and stored as a database model [Muzahir, 2013]. Because of extremely large requirements of memory (estimated 40-45 GB RAM) and processing power, for the purpose of goal-specific analysis, a subsite (300 pages) for one of the academic schools was randomly selected. In the simulation, LSA cosines were computed using the first-year-college semantic space (TasaALL) in both studies. In terms of reading level, this was deemed to be the most appropriate available semantic space for users of the websites. However, a limitation is that the space is based on American English, although the research used British English websites and users. Moreover, in common with other semantic spaces, the space does not represent proper names (see also Section 3.1.2). Another shortcoming of TasaALL is that it is likely neologisms (such as 'Twitter' and 'Facebook') are not included. Enhanced CoLiDeS/+ simulation was run on the subsite with the original information architecture for 41 goals that had the required length (100-200 words; see Section 2.4.1), with two possible outcomes: success (the simulation finds the correct page matching the goal) or failure (the simulation does not find the correct page). Twenty-two goals (54%) succeeded and 19 (46%) goals failed.

3.1.2. Improvement of original information architecture. Various factors that influence users' success in finding information have been identified in the labeling of web page elements [e.g., Swierenga et al., 2011]. According to Nielsen [2000], clear and elaborate labels help users to precisely predict the information on following pages. Spool et al. [1999] demonstrated that a positive correlation exists between the length of link labels (7-12 words) and success rates in finding information. A possible reason to this is that lengthy link labels generally carry more information and are less confusable than short labels.

According to Olston and Chi [2003], browsing cues (link labels) have a limited scent and should be improved. Users can experience difficulty while browsing due to low information scent of web elements. This limitation can be attributed to at least three causes. "First, poor link labeling can lead to inappropriate cues. Second, since each web page tends to contain a large number of potential destinations, the cues are typically short and thus cannot convey a large amount of information. Third and most importantly, browsing cues are usually not customized based on each user's information goal" (p. 182).

The approach used in this research was to improve the labeling system of an information architecture in such a way that the cosine of correct labels of headings and links would increase in relation to the goal. As a result, when the scent of heading labels and link labels is evaluated in a simulation both will be increased; the scent of link labels benefits even more because their scent is calculated as the scent of the combined heading and link labels. Therefore, all else being equal, the simulation would be more likely to succeed on a particular goal. The labels of competing goal-specific links were reworded to degrade their cosine values. Degrading the goal-link cosine value of incorrect links may improve success rate on one task but could also reduce success rate on other tasks. However, our results of Study 1 show no evidence of such a reduction and in Study 2 only for 1 task was there an adverse effect.

We applied Blackmon et al.'s [2003] approach to identifying goal-specific competing links, based on the following compound criterion: the competing link label (1) must be under the same heading as the correct link, (2) must have a goal-link cosine value that equals at least 80% of the goal-link cosine for the correct link label and (3) should not be judged by the analyst as a false alarm, in other words, a link that real users would probably not select.

Furthermore, as a relaxation of the principle of increasing information scent along the navigation path to the destination page [Juvina & Van Oostendorp, 2007], a strictly non-decreasing-cosine strategy for rewording was accomplished for web page elements on the navigation path to the target page. This was important, as it was observed some of the analyzed web pages on the path to the destination page did not have an increasing cosine, but the cosine remained unchanged. Accordingly, enhanced CoLiDeS/+ was adapted by allowing page traversal with non-decreasing, rather than strictly increasing, cosine value.

For increasing similarity (cosine value) of goal statement with heading- and link labels, the most common type of improvement made for goals that failed in the simulation was the elaboration (expansion) of heading and link labels. Specifically, in many cases link labels consisted of proper names that had to be elaborated to make them meaningful; in a similarity evaluation, proper names will have cosines of 0 (complete lack of similarity), as they have no intrinsic meaning and are therefore not part of the LSA corpus. Note that elaboration is not the only or always the best strategy for improvement. Some individual words carry clear meaning for a heading compared to long strings of words that may be more diffuse in meaning, and users are more likely to read heading labels containing 1-3 words than heading labels containing 5 or more words. Therefore, 'less can be more'.

Changes were made manually by editing the modified tables in the database for the simulation to access the data representing the improved information architecture of the subsite from the database automatically. These changes were then reflected in an improved version of the subsite by editing the HTML code manually.

3.1.3. Simulation (improved information architecture. After improvements had been made to the information architecture, enhanced CoLiDeS/+ simulation was run again. Of the 41 goals 36 (88%) succeeded and 5 (12%) failed; all goals that succeeded with the original information architecture also succeeded with the improved information architecture (see Table I). Analysis of the results from the two simulations (original versus improved information architecture) shows a positive effect of information architecture on successful navigation, $phi = 0.38$, $p < 0.001$, with improved architecture resulting in more goals that succeeded. In a subsequent experiment, simulation results were validated.

3.2 Validation

3.2.1. Method. Design. Given that the simulation had demonstrated theoretically that navigation would be more successful with improved information architecture, the aim of the experiment was to empirically demonstrate that actual navigation by users was also more successful employing the site with improved information architecture. Therefore, (3 practice and 31 main) tasks (see Table I) were selected for which, according to the simulation results, navigation failed with the original information architecture, but succeeded with the improved information architecture. In the experiment, the independent variable was information architecture, with two levels: original and improved. The dependent variables were task completion, correctness of answers, length of navigation (number of page loads), time-on-task, perceived disorientation [an important indicator of problematic web navigation; Van Oostendorp et al., 2009], and task performance [logarithmically transformed correctness/time; Van Oostendorp & Juvina, 2007]. Spatial ability was a potential covariate, as this is a predictor of task performance in web navigation [Ahmed & Blustein, 2006; Van Oostendorp & Juvina, 2007].

Participants. There were 94 participants (22 female; 92 university students), with mean age of 27.86 years ($SD = 8.36$). They received £10 for taking part. Forty-six used the subsite with original information architecture (control group) and 48 with modified information architecture (experimental group). All participants had experience using the Web and the vast majority had been using websites for more than two years. None had experience with using the site that was tested.

Materials and equipment. A bespoke software program to run the experiment was developed and coded as a Windows Form Application using Visual Basic. The experiment used two locally saved site versions (original and improved; see Section 3.1 and Figure 1).

The experiment ran on personal computers (Intel(R) Core(TM)2 Duo CPU, 3GHz processing power, 2 GB RAM, Microsoft Windows 7 Enterprise, 32-bit operating system). The screen dimensions were 1280×1024 with a refresh rate of 75Hz. Each monitor had an active-matrix TFT LCD Screen with a 19-inch viewable image. Contrast and brightness were set to optimal levels.

Participants completed Ahuja and Webster's [2001] disorientation scale on screen as a measure of disorientation; internal-consistency reliability was good, Cronbach's alpha = 0.81; therefore, an average disorientation score was calculated per participant. A mental-rotation test was used to measure spatial ability; this consisted of 3 practice- and the first 10 main problem items from Vandenberg and Kuse's [1978] version of the Shepard and Metzler [1971] three-dimensional rotation test. Each item consisted of a target figure, two correct alternatives (rotations of the criterion), and two incorrect figures or 'distracters'. Participants had to identify the two correct figures matching the target. Each response to the mental-rotation test was scored as correct (score = 1) if two correct figures were chosen and as an incorrect (score = 0) otherwise. The scores were summed to create an overall spatial-ability score.

Table I. Simulation results

a. Study 1

Set of 41 randomly selected and analyzed goals

	Simulation result after rewording		Total
Simulation result before rewording	Failure	Success	
Failure	5	17	22
Success	0	19	19
Total	5	36	41

Goals used in validation (experiment), all successfully simulated after rewording

Simulation result before rewording	Practice	Main	Total
Failure	2	15	17
Success	1	16	17
Total	3	31	34

b. Study 2

All goals not targeted by rewording

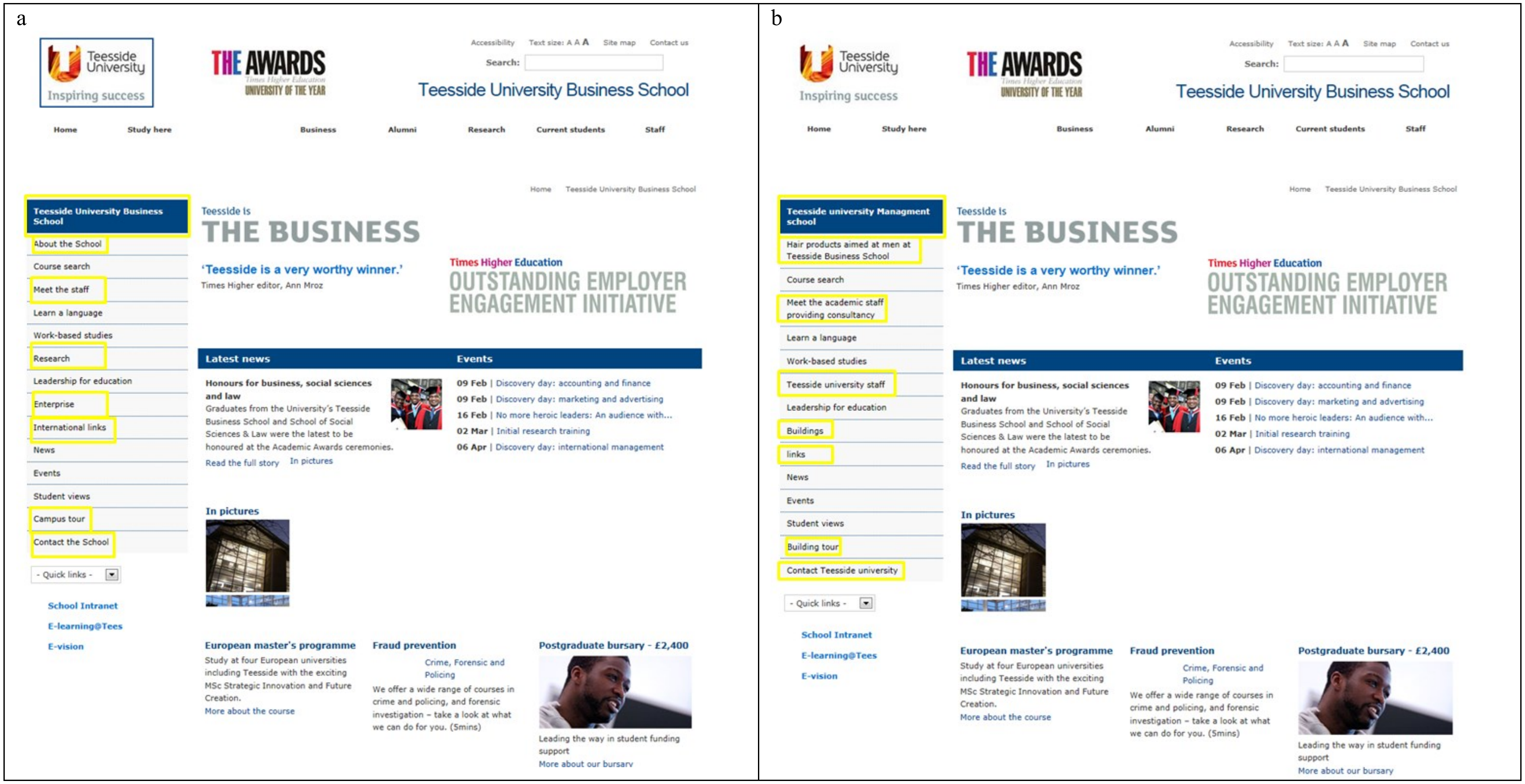
	Simulation result after rewording		Total
Simulation result before rewording	Failure	Success	
Failure	39	16	55
Success	1	18	19
Total	40	34	74

All goals targeted by rewording

	Simulation result after rewording		Total
Simulation result before rewording	Failure	Success	
Failure	7	26	33
Success	0	0	0
Total	7	26	33

Goals used in validation (experiment)

Simulation result after rewording	Practice	Main	Total
Failure	0	3	3
Success	3	10	13
Total	3	13	16



'Teesside is a very worthy winner.'

Times Higher editor, Ann Mroz

Honours for business, social sciences and law

Graduates from the University's Teesside Business School and School of Social Sciences & Law were the latest to be honoured at the Academic Awards ceremonies.

Read the full story In pictures

09 Feb | Discovery day: accounting and finance

09 Feb | Discovery day: marketing and advertising

16 Feb | No more heroic leaders: An audience with...

02 Mar | Initial research training

06 Apr | Discovery day: international management

European master's programme

Study at four European universities including Teesside with the exciting MSc Strategic Innovation and Future Creation.

More about the course

Fraud prevention

Crime, Forensic and Policing

We offer a wide range of courses in crime and policing, and forensic investigation – take a look at what we can do for you. (5mins)

Postgraduate bursary - £2,400

Leading the way in student funding support

More about our bursary

'Teesside is a very worthy winner.'

Times Higher editor, Ann Mroz

Honours for business, social sciences and law

Graduates from the University's Teesside Business School and School of Social Sciences & Law were the latest to be honoured at the Academic Awards ceremonies.

Read the full story In pictures

09 Feb | Discovery day: accounting and finance

09 Feb | Discovery day: marketing and advertising

16 Feb | No more heroic leaders: An audience with...

02 Mar | Initial research training

06 Apr | Discovery day: international management

European master's programme

Study at four European universities including Teesside with the exciting MSc Strategic Innovation and Future Creation.

More about the course

Fraud prevention

Crime, Forensic and Policing

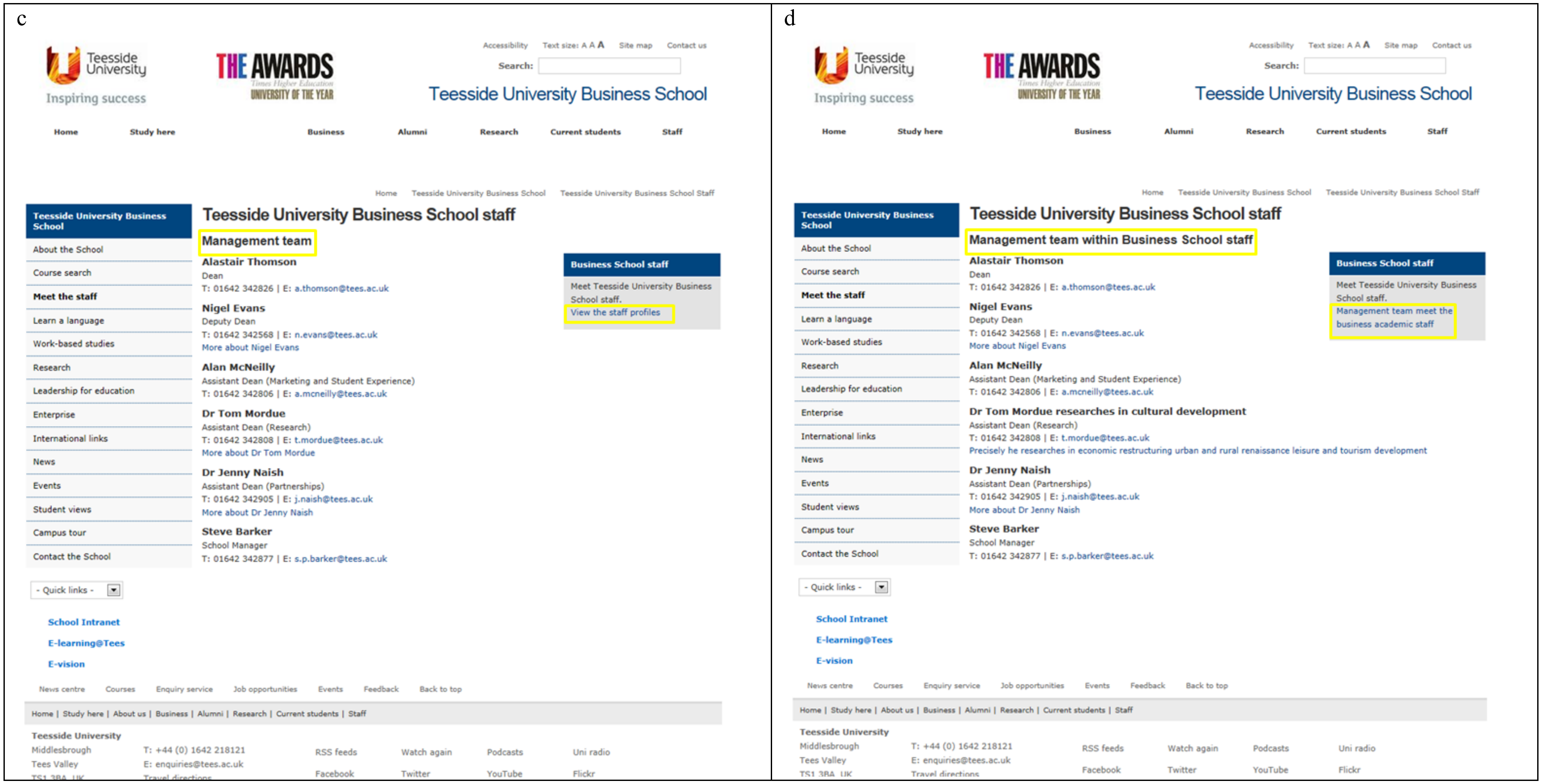
We offer a wide range of courses in crime and policing, and forensic investigation – take a look at what we can do for you. (5mins)

Postgraduate bursary - £2,400

Leading the way in student funding support

More about our bursary

Fig. 1. Web pages, Study 1 (a, c original; b, d improved).



Teesside University Business School

Teesside University Business School staff

Management team

Alastair Thomson
Dean
T: 01642 342826 | E: a.thomson@tees.ac.uk

Nigel Evans
Deputy Dean
T: 01642 342568 | E: n.evans@tees.ac.uk
More about Nigel Evans

Alan McNeilly
Assistant Dean (Marketing and Student Experience)
T: 01642 342806 | E: a.mcneilly@tees.ac.uk

Dr Tom Mordue
Assistant Dean (Research)
T: 01642 342808 | E: t.mordue@tees.ac.uk
More about Dr Tom Mordue

Dr Jenny Naish
Assistant Dean (Partnerships)
T: 01642 342905 | E: j.naish@tees.ac.uk
More about Dr Jenny Naish

Steve Barker
School Manager
T: 01642 342877 | E: s.p.barker@tees.ac.uk

Business School staff
Meet Teesside University Business School staff.
View the staff profiles

Teesside University
Middlesbrough
Tees Valley
TS1 3BA UK
T: +44 (0) 1642 218121
E: enquiries@tees.ac.uk
Travel directions

Teesside University Business School

Teesside University Business School staff

Management team within Business School staff

Alastair Thomson
Dean
T: 01642 342826 | E: a.thomson@tees.ac.uk

Nigel Evans
Deputy Dean
T: 01642 342568 | E: n.evans@tees.ac.uk
More about Nigel Evans

Alan McNeilly
Assistant Dean (Marketing and Student Experience)
T: 01642 342806 | E: a.mcneilly@tees.ac.uk

Dr Tom Mordue researches in cultural development
Assistant Dean (Research)
T: 01642 342808 | E: t.mordue@tees.ac.uk
Precisely he researches in economic restructuring urban and rural renaissance leisure and tourism development

Dr Jenny Naish
Assistant Dean (Partnerships)
T: 01642 342905 | E: j.naish@tees.ac.uk
More about Dr Jenny Naish

Steve Barker
School Manager
T: 01642 342877 | E: s.p.barker@tees.ac.uk

Business School staff
Meet Teesside University Business School staff.
Management team meet the business academic staff

Teesside University
Middlesbrough
Tees Valley
TS1 3BA UK
T: +44 (0) 1642 218121
E: enquiries@tees.ac.uk
Travel directions

Fig. 1 (continued).

Procedure. The experiment ran in a computer lab with 10 to 12 participants working independently and randomly allocated to one of the two site versions. Participants first undertook the spatial-ability test, with a time limit of three minutes for the 10 main problems. They then completed a series of information retrieval tasks. Finally, participants completed the disorientation scale and answered demographic questions.

In each information retrieval task, participants first completed a practice run of 3 trials. The main trials followed with a maximum of a 31 further trials. In every trial, a paragraph containing 100-200 words of text (representing a goal; see Section 2) was presented at the top of the screen.[3] Once participants had read the paragraph they had to click a button labeled 'Show Website'. The homepage appeared on the screen and they had to look for the paragraph in the site. Participants were told to take the most direct route possible to locate the paragraph in the site. Once they found the paragraph, they clicked on a button labeled 'Your answer', which opened a dialog box at the bottom of the screen. In this answer box, participants had to type the title of the web page containing the paragraph. They were instructed to type "Not found" in the answer dialog box if they had not found the paragraph after searching for five minutes. After clicking on 'OK' they moved to next trial. The experiment took approximately 55 minutes to complete.

3.2.2. Results. Initial analysis showed that, although the groups had been randomly allocated to experimental conditions, the groups differed statistically significantly on spatial ability, t (92) = -3.08, $p < 0.01$, $r = 0.31$ (mean [*SD*] = 2.37 [2.30] for the control group and 4.10 [3.09] for the experimental group). Consequently, the assumption of independence of the independent variable (information architecture) from the potential covariate (spatial ability) was violated; thus, data analysis through analysis of covariance was precluded. Therefore, the data were analyzed by blocking [recommended by Tabachnick & Fidell, 2001] on spatial ability (with low and high spatial ability, respectively), using a median split.

Descriptives (Table II; Figure 2) indicated that the experimental group was superior on outcome measures. The results of 2-by-2 analysis[4] of variance (ANOVA) (Table III) demonstrate that the experimental group outperformed the control group statistically significantly on task completion, correctness per completed task, time-on-task, task performance (logarithmically transformed correctness/time) and number of page loads, with effect sizes ranging from moderate to very large (according to conventions for effect size for estimates of explained variance for each tested effect[5]); the effect on perceived disorientation was approaching significance.

There was a significant effect of spatial ability on time-on-task; the effect was approaching significance on task completion, time-on-task (for correctly completed tasks) and task performance; the effect sizes were small to moderate. The interaction effect was not significant.

---

[3] In the subsite, destination pages with text paragraphs matching the goal statement were located 1, 2 or 3 levels deep from the homepage.

[4] Sample sizes were 31 for original/low spatial ability, 19 improved/low spatial ability, 15 original/high spatial ability and 29 improved/high spatial ability.

[5] 0.01 for small, 0.059 for moderate and 0.138 for large (Clark-Carter, 2009), which Clark-Carter converted from Cohen's (1988) recommendations for effect-size measure *f*

Table II. Descriptives as a function of information architecture and spatial ability (Study 1)

		Information architecture			
		Original		Modified	
Outcome measure	Spatial ability	Mean	*SD*	Mean	*SD*
Task completion (pct)	Low	12.70	6.97	24.62	13.61
	High	18.71	8.11	26.92	12.59
Correct/complete (pct)	Low	45.41	29.78	76.78	20.76
	High	42.29	22.92	83.09	12.05
Time (average) (s)	Low	241	97	180	102
	High	193	99	151	62
Task performance	Low	0.0022	0.0018	0.0051	0.0022
	High	0.0027	0.0020	0.0063	0.0024
Page loads (average)	Low	24.12	16.09	14.11	9.61
	High	19.11	7.35	16.63	10.19
Disorientation	Low	3.66	0.91	3.04	1.28
	High	3.59	1.40	3.33	1.33

Note. Low: spatial-ability score ≤ 2. High: spatial-ability score ≥ 3. Task performance: ln(correct/time). Correct/complete: percentage of tasks with correct answers out of completed number of tasks.

A detailed analysis was conducted for tasks by simulation success (for full details see Online Appendix C). For each task, success was higher with the improved information architecture. Furthermore, the success rate by participants on the site with the original information architecture was 24% of that on the site with the improved information architecture over the tasks for which the simulation on the site with the original information architecture succeeded and 32% over the tasks for which the simulation failed. These results show that task success was relatively low for participants with the original information architecture compared to the improved information architecture. The results also demonstrate the advantage of the improved information architecture (shown in the previous results over all tasks) in the results per task and even more so when the simulation failed on the original information architecture.

Participants' written comments about positive and negative aspects of the site were categorized in relation to information scent or layout and other aspects. The site with the original information architecture received 10 positive and 22 negative comments on scent, but the site with the improved information architecture received 25 positive and 5 negative comments, a statistically significant pattern of findings, $phi = 0.52$, $p < 0.001$. The original information architecture received 23 positive and 18 negative comments on layout and other aspects, but the improved architecture received 25 positive and 35 negative comments, a statistically non-significant pattern of findings, $phi = -0.15$, $p > 0.15$. These results indicate that the improved information architecture was experienced as having significantly better information scent than the original architecture, but was not experienced as being significantly worse on layout and other aspects.

Table III. Testing the effect of information architecture on outcome measures (Study 1)

Outcome measure	Source	*df*	*SS*	*MS*	*F*	*p*	ε^2
Task	IA	1	2179	2179	19.32	< 0.001	0.16
completion	SA	1	372	372	3.30	.073	0.02
	IA by SA	1	74	74	0.66	.420	0.00
	Error	90	10148	113			
Correct/	IA	1	28001	28001	55.04	< 0.001	0.37
complete	SA	1	54	54	0.11	.745	0.00
task	IA by SA	1	479	479	0.94	.335	0.00
	Error	90	45786	509			
Time	IA	1	55944	55944	7.04	.009	0.06
(average)	SA	1	32264	32264	4.06	.047	0.03
	IA by SA	1	1981	1981	0.25	.619	0.00
	Error	90	714986	7944			
Task	IA	1	0.000228	0.000228	51.16	< 0.001	0.35
Performance	SA	1	0.000015	0.000015	3.41	0.068	0.02
	IA by SA	1	0.000002	0.000002	0.46	0.500	0.00
	Error	90	0.000401	0.000004			
Page loads	IA	1	840	840	5.77	0.018	0.05
(average)	SA	1	34	34	0.23	0.633	0.00
	IA by SA	1	305	305	2.10	0.151	0.01
	Error	90	13096	146			
Disorientation	IA	1	4.18	4.18	2.87	0.094	0.02
	SA	1	0.27	0.27	0.19	0.668	0.00
	IA by SA	1	0.74	0.74	0.51	0.479	0.00
	Error	90	131.16	1.46			

Note. Correct/complete: percentage of tasks with correct answers out of completed number of tasks. IA: information architecture. SA: spatial ability.

3.2.3. Discussion. The results of the experiment with the large Internet subsite demonstrate that improved information architecture, based on the findings of enhanced CoLiDeS/+ simulation, resulted in better outcomes of information retrieval in terms of task completion, correctness, speed, efficiency, task performance and experience of information scent. Both those with low and high spatial ability benefitted from improved information architecture.

4. STUDY 2: LARGE INTRANET SITE

The rationale for designing Study 2 is one of replication [Hornbæk et al., 2014], within our paper, in a different domain (restricted information-oriented intranet website for staff in Study 2 versus public higher-education Internet site in Study 1). Note that, because the original websites in the two studies were real-world live website and not created as part of an experiment, the websites will have differed in various ways, not only in terms of the domain.

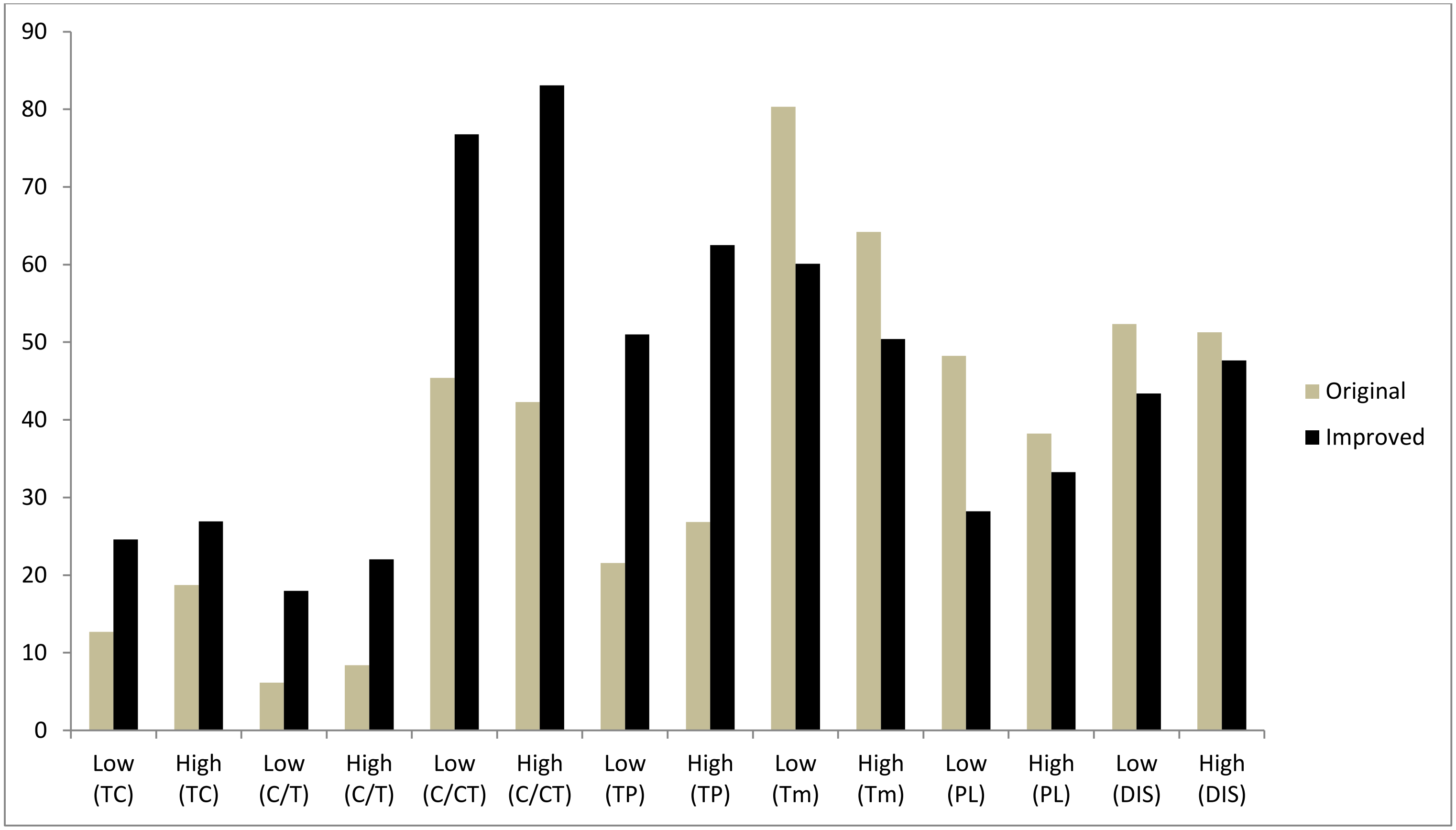


Fig. 2. Mean values (rescaled as a percentage) as a function of information architecture and spatial ability (Study 1)

Note. Low: low spatial ability. High: high spatial ability. TC: task completion. C/CT: correct/completed task. TP: task performance. Tm: time (average). PL: page loads. DIS: disorientation.

4.1 Goal-specific analysis and improvement

4.1.1. Simulation (original information architecture). A university intranet site was captured, the 33000 web pages were parsed and the site was recreated and stored as a database model [Muzahir, 2013].[6,7] Because of extremely large requirements of memory (estimated 40-45 GB RAM) and processing power, for the purpose of goal-specific analysis, a subsite (500 pages) of the university intranet site for one of the departments was randomly selected. CoLiDeS/+ simulation was run on the site with the original information architecture for the 107 goals that had the required length (100-200 words; see Section 2). The results showed that 19 goals (18%) succeeded and 88 (82%) goals failed.

4.1.2. Improvement of original information architecture. The same approach to making improvements was followed as in Study 1. In order to increase similarity (cosine value) of goal statement with heading- and link labels, the following common types of improvement (or combinations of these) were made for goals that failed in the simulation. Domain-specific abbreviations were written in full. Heading and link labels were elaborated (expanded). Non-specific link labels (e.g., 'Click here') were made specific.

4.1.3. Simulation (improved information architecture). After improvements had been made to the information architecture, enhanced CoLiDeS/+ simulation was run again. Of the 107 goals 60 (56%) succeeded and 47 (44%) failed and only 1 goal that succeeded with the original information architecture did not with the improved information architecture (see Table I). Analysis of the results from the two simulations (original versus improved information architecture) shows a positive effect of information architecture on successful navigation, $phi = 0.40$, $p < 0.001$, with improved architecture resulting in more goals that succeeded. In a subsequent experiment, simulation results were validated.

4.2 Validation

4.2.1 Method. Design. There were 64 participants (university students; 18 female), with mean age of 26.95 years ($SD = 9.61$). They received £10 for their participation. Thirty-two used the site with original information architecture (control group) and another 32 with modified information architecture (experimental group). All participants had experience using the Web and the vast majority had been using websites for more than two years. None had experience with using the site.

Design, materials and equipment, and procedure. The same design, materials and equipment, and procedure were used as in Study 1, with the following exceptions. The disorientation scale was reliable, Cronbach's alpha = 0.91. The experiment used two locally saved site versions (original and improved; see Section 4.1 and Figure 3). (Three practice and 13 main) tasks (see Table I) were selected for which, according to the simulation results, navigation failed with the original information architecture, but succeeded with the improved information architecture.[8] Participants had to type the page number as their answer in each information retrieval task.

---

[6] As much of the content of this website did not reside in web pages, but linked content files (e.g., in .html, .cfm, .doc, .docx or .pdf formats), these files were captured and represented as text paragraphs.

[7] Non-goal-specific was conducted on the entire website and is reported elsewhere (Muzahir, 2013).

[8] In the subsite, destination pages with text paragraphs matching the goal statement were located 2, 3, 4 or 6 levels deep from the homepage.

4.2.2. Results. Initial analysis showed that, although the groups had been randomly allocated to experimental conditions, the groups differed statistically significantly on spatial ability, t (62) = 3.64, $p < 0.001$, $r = 0.42$ (mean [*SD*] = 2.53 [1.68] for the control group and 4.56 [2.68] for the experimental group). Therefore, as in Study 1, the data were analyzed by blocking on spatial ability (with low and high spatial ability, respectively), using a median split.

Descriptives (Table IV; Figure 4) indicate that the experimental group was superior on outcome measures. The results of 2-by-2 ANOVA[9] (Table V) demonstrate that the experimental group outperformed the control group statistically significantly on task completion, correctness per completed task, time-on-task and task performance (logarithmically transformed correctness/time), with effect sizes ranging from moderate to very large; the effect on perceived disorientation was approaching significance.

There was a significant interaction effect on correctness per task (completed or not), correctness per completed task and task performance. Because of the three significant interaction effects, simple-effect tests were conducted as follow-up. The effect of information architecture was significant for each of these outcome measures for both low- and high-spatial-ability participants, with $p < 0.001$; the effect sizes were somewhat higher for high- (r = 0.81, 0.85 and 0.71, respectively) than for low-spatial-ability (r = 0.68, 0.65 and 0.68, respectively) participants.

A detailed analysis was conducted for tasks by simulation success for full details see Online Appendix D). For each the task success was higher with the improved information architecture. Furthermore, the success rate by participants on the site with the original information architecture was 6% of that on the site with the improved information architecture over the tasks for which the simulation on the site with improved information architecture succeeded and 5% over the tasks for which the simulation failed. These results show that task success was relatively low for participants using the original site compared to the improved site. The results also show that the advantage of the improved site shown in the previous results over all tasks was reflected in the results per task.

The site with the original information architecture received 6 positive and 24 negative comments on scent, but the site with the improved information architecture received 25 positive and 12 negative comments, a statistically significant pattern of results, $phi = 0.47$, $p < 0.001$. The original information architecture received 19 positive and 13 negative comments on layout and other aspects, but the improved architecture received 7 positive and 21 negative comments, a statistically significant pattern of results, $phi = -0.35$, $p < 0.01$. These results indicate that the improved information architecture was experienced as having significantly better information scent than the original architecture, but was also experienced as being significantly worse on layout and other aspects.

4.2.3. Discussion. The results of the experiment with the large intranet subsite demonstrate that improved information architecture, based on the findings of enhanced CoLiDeS/+ simulation, resulted in better outcomes of information retrieval in terms of task completion, correctness, speed, task performance and experience of information scent. Both those with low and high spatial ability benefitted from improved information architecture, but the effect of information architecture on

---

[9] Sample sizes were 23 for original/low spatial ability, 12 improved/low spatial ability, 9 original/high spatial ability and 20 improved/high spatial ability.

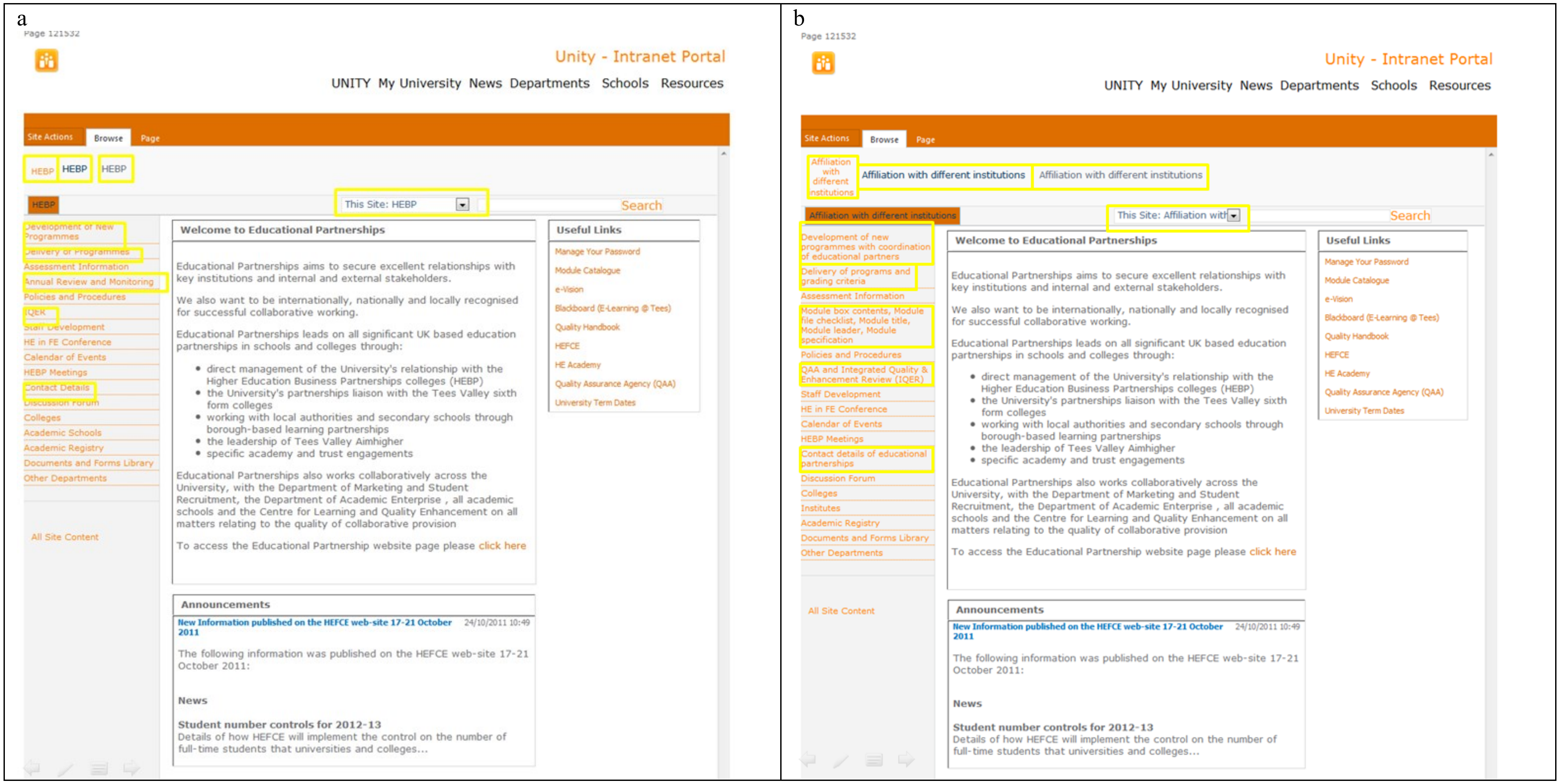
correctness and task performance was stronger for high-spatial ability users and there was a trend approaching significance for high-spatial-ability users to benefit more from improvement in terms of overall page loads.

5. DISCUSSION

5.1 Main results and limitations

In relation to Aim 1 (automated cognitive computational goal-specific analysis of the information architecture of large websites), the current research successfully automated the simulation of web navigation by implementing an enhanced model combining CoLiDeS and CoLiDeS+, with several advances on these models (Online Appendix E). Goal-specific analysis was conducted by automated simulation on the information architecture of two large real-world Internet- and intranet subsites (consisting of hundreds of web pages). Improvements made to the information architecture were demonstrated by higher success rate in simulation results for both subsites. In general, there is no guarantee that the 'improvement strategy' is guaranteed to always monotonically improve all paths to all targets (and not just some at the expense of others). However, our results for Study 1 show that in our goal-specific analysis no improvements for particular targets were at the expense of others. Furthermore, in Study 2, only 1 of the intended improvements was at the expense of another target.

A crucial factor in realizing the automated analysis of large real-world sites with improved web navigation was the separation of website from its persistent representation by way of a database-oriented approach, conferring benefits of flexibility, control, efficiency and extensibility. In contrast, previous work using automated tools analyzed live websites without a database-oriented approach [e.g., Chi et al., 2003] and therefore could not achieve these benefits. For example, in order to evaluate the positive effect of improvements made to the information architecture of a website based on simulation results, the actual live website would have to be altered, which would be more labor-intensive and might practically be impossible if the analyst had no control over the website. Our approach allows for many types of further automated analysis, by 'interrogating' the database, that are beyond the scope of the work reported here. The enhanced CoLiDeS/+ implementation considers links without heading, but could also easily accommodate images without or with heading. Similarly, the analysis could be extended with other types of non-text medium.



Welcome to Educational Partnerships

Educational Partnerships aims to secure excellent relationships with key institutions and internal and external stakeholders.

We also want to be internationally, nationally and locally recognised for successful collaborative working.

Educational Partnerships leads on all significant UK based education partnerships in schools and colleges through:

- direct management of the University's relationship with the Higher Education Business Partnerships colleges (HEBP)
- the University's partnerships liaison with the Tees Valley sixth form colleges
- working with local authorities and secondary schools through borough-based learning partnerships
- the leadership of Tees Valley Aimhigher
- specific academy and trust engagements

Educational Partnerships also works collaboratively across the University, with the Department of Marketing and Student Recruitment, the Department of Academic Enterprise , all academic schools and the Centre for Learning and Quality Enhancement on all matters relating to the quality of collaborative provision

To access the Educational Partnership website page please click here

Announcements

New Information published on the HEFCE web-site 17-21 October 2011 24/10/2011 10:49

The following information was published on the HEFCE web-site 17-21 October 2011:

News

Student number controls for 2012-13
Details of how HEFCE will implement the control on the number of full-time students that universities and colleges...

Welcome to Educational Partnerships

Educational Partnerships aims to secure excellent relationships with key institutions and internal and external stakeholders.

We also want to be internationally, nationally and locally recognised for successful collaborative working.

Educational Partnerships leads on all significant UK based education partnerships in schools and colleges through:

- direct management of the University's relationship with the Higher Education Business Partnerships colleges (HEBP)
- the University's partnerships liaison with the Tees Valley sixth form colleges
- working with local authorities and secondary schools through borough-based learning partnerships
- the leadership of Tees Valley Aimhigher
- specific academy and trust engagements

Educational Partnerships also works collaboratively across the University, with the Department of Marketing and Student Recruitment, the Department of Academic Enterprise , all academic schools and the Centre for Learning and Quality Enhancement on all matters relating to the quality of collaborative provision

To access the Educational Partnership website page please click here

Announcements

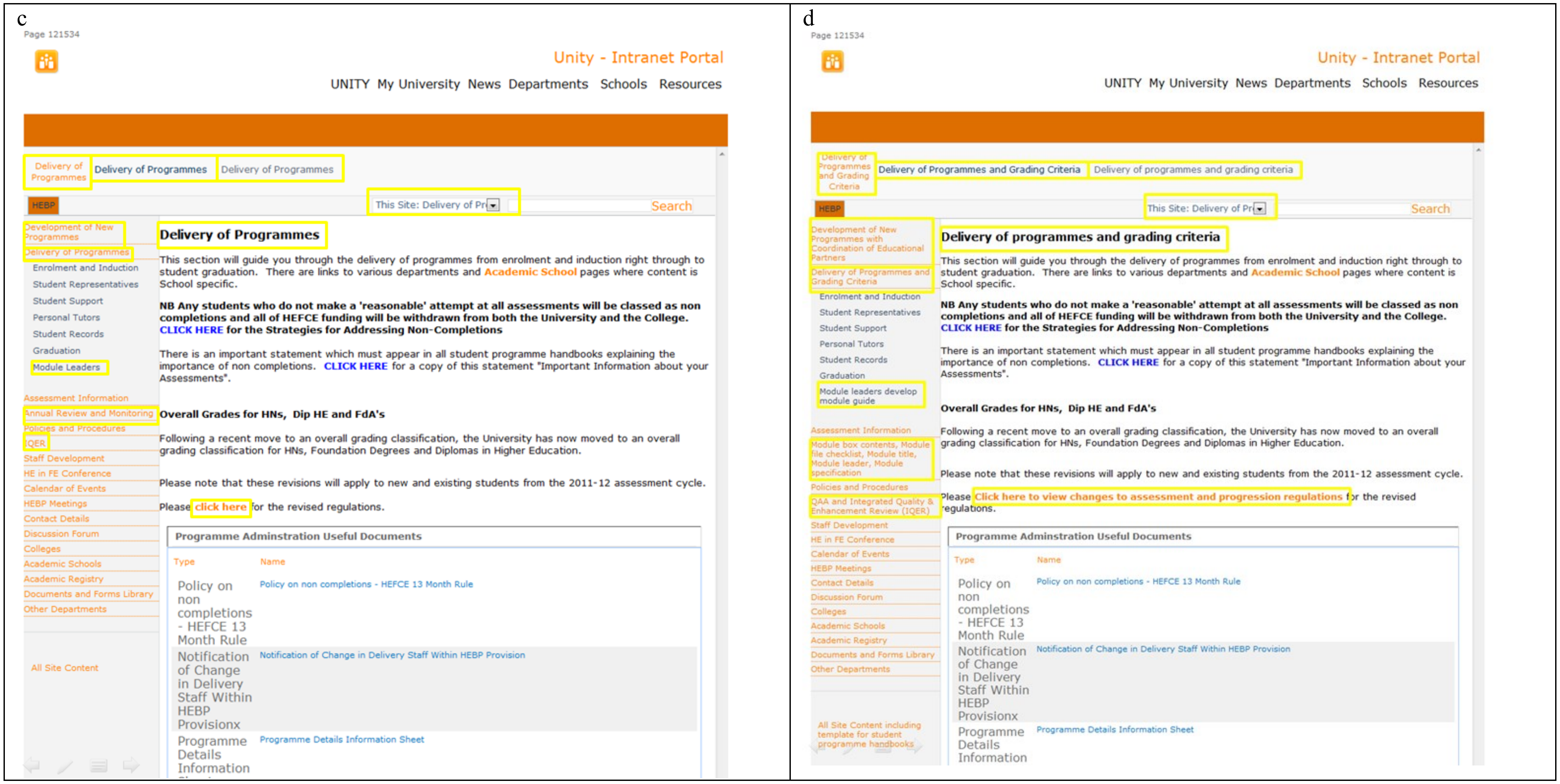
New Information published on the HEFCE web-site 17-21 October 2011 24/10/2011 10:49

The following information was published on the HEFCE web-site 17-21 October 2011:

News

Student number controls for 2012-13
Details of how HEFCE will implement the control on the number of full-time students that universities and colleges...

Fig. 3. Web pages, Study 2 (a, c original; b, d improved)



Delivery of Programmes

This section will guide you through the delivery of programmes from enrolment and induction right through to student graduation. There are links to various departments and Academic School pages where content is School specific.

NB Any students who do not make a 'reasonable' attempt at all assessments will be classed as non completions and all of HEFCE funding will be withdrawn from both the University and the College. CLICK HERE for the Strategies for Addressing Non-Completions

There is an important statement which must appear in all student programme handbooks explaining the importance of non completions. CLICK HERE for a copy of this statement "Important Information about your Assessments".

Overall Grades for HNs, Dip HE and FdA's

Following a recent move to an overall grading classification, the University has now moved to an overall grading classification for HNs, Foundation Degrees and Diplomas in Higher Education.

Please note that these revisions will apply to new and existing students from the 2011-12 assessment cycle.

Please click here for the revised regulations.

Programme Adminstration Useful Documents

Type	Name
Policy on non completions - HEFCE 13 Month Rule	Policy on non completions - HEFCE 13 Month Rule
Notification of Change in Delivery Staff Within HEBP Provisionx	Notification of Change in Delivery Staff Within HEBP Provision
Programme Details Information	Programme Details Information Sheet

Delivery of programmes and grading criteria

This section will guide you through the delivery of programmes from enrolment and induction right through to student graduation. There are links to various departments and Academic School pages where content is School specific.

NB Any students who do not make a 'reasonable' attempt at all assessments will be classed as non completions and all of HEFCE funding will be withdrawn from both the University and the College. CLICK HERE for the Strategies for Addressing Non-Completions

There is an important statement which must appear in all student programme handbooks explaining the importance of non completions. CLICK HERE for a copy of this statement "Important Information about your Assessments".

Overall Grades for HNs, Dip HE and FdA's

Following a recent move to an overall grading classification, the University has now moved to an overall grading classification for HNs, Foundation Degrees and Diplomas in Higher Education.

Please note that these revisions will apply to new and existing students from the 2011-12 assessment cycle.

Please Click here to view changes to assessment and progression regulations for the revised regulations.

Programme Adminstration Useful Documents

Type	Name
Policy on non completions - HEFCE 13 Month Rule	Policy on non completions - HEFCE 13 Month Rule
Notification of Change in Delivery Staff Within HEBP Provisionx	Notification of Change in Delivery Staff Within HEBP Provision
Programme Details Information	Programme Details Information Sheet

Fig. 3. (continued)

Table IV. Descriptives as a function of information architecture and spatial ability (Study 2)

		Information architecture			
		Original		Modified	
Outcome measure	Spatial ability	Mean	*SD*	Mean	*SD*
Task completion (pct)	Low	35.33	24.03	51.04	22.90
	High	40.28	31.11	63.75	13.39
Correct/completed (pct)	Low	11.01	18.60	50.26	29.03
	High	2.78	8.33	68.44	22.22
Time (average) (s)	Low	264	131	226	104
	High	273	100	180	46
Task performance	Low	0.0005	0.0008	0.0026	0.0015
	High	0.0001	0.0004	0.0042	0.0023
Page loads (average)	Low	29.69	22.38	32.62	23.68
	High	41.25	37.69	23.48	10.90
Disorientation	Low	4.32	1.26	3.54	1.57
	High	4.35	1.68	3.77	1.64

Note. Low: spatial-ability score ≤ 3. High: spatial-ability score ≥ 4. Task performance: ln(correct/time). Correct/complete: percentage of tasks with correct answers out of completed number of tasks.

In relation to Aim 2 (validation of automated analysis through experiments with end-users), improvements of the information architecture, according to simulation results, were incorporated in locally saved sites, and tested; we found empirical evidence for the improvements in experiments with end-users against original sites without improvements. Our results are consistent with those of Blackmon et al. [2002, 2003, 2005], showing that a model-based approach using CoLiDeS can be employed to improve websites with measurably better navigation outcomes for users. However, Blackmon and colleagues analyzed experimental websites that worked like the full version but with fewer terminal-node web pages, whereas we analyzed large real-world websites and subsites of similar size. Both users with low spatial ability and users with high spatial benefitted from improved information architecture. The effect information architecture was equally strong on most measures in both studies and stronger for high-spatial-ability users on correctness and task performance in Study 2. Therefore, in contrast to Van Oostendorp and Juvina's [2007] findings of their Study 1, our low-spatial ability users did not benefit more on the task performance measure. The difference in findings may be explained by the following difference between our experiments and their work: they provided users with navigation support (suggesting which link to select), whereas we provided improved information scent without this type of support.

Table V. Testing the effect of information architecture on outcome measures (Study 2)

Outcome measure	Source	*df*	*SS*	*MS*	*F*	*p*	ε^2
Task completion	IA	1	5334	5334	10.80	.002	0.13
	SA	1	1083	1083	2.19	.144	0.02
	IA by SA	1	209	209	0.42	.518	0.00
	Error	60	29621	494			
Correct/ complete	IA	1	38223	38223	85.51	< 0.001	0.56
	SA	1	344	344	0.77	.384	0.00
	IA by SA	1	2424	2424	5.42	.023	0.03
	Error	60	26819	447			
Time (average)	IA	1	59573	59573	5.80	.019	0.07
	SA	1	4634	4634	0.45	.504	0.00
	IA by SA	1	10904	10904	1.06	.307	0.00
	Error	60	616435	10274			
Task performance	IA	1	0.00013	0.00013	54.83	< 0.001	0.44
	SA	1	0.00001	0.00001	2.21	.142	0.01
	IA by SA	1	0.00001	0.00001	5.78	.019	0.04
	Error	60	0.00015	0.00000			
Page loads (average)	IA	1	764	764	1.49	.227	0.01
	SA	1	20	20	0.04	.843	0.00
	IA by SA	1	1487	1487	2.89	.094	0.03
	Error	60	30816	514			
Disorientation	IA	1	6.47	6.47	2.86	.096	0.03
	SA	1	0.24	0.24	0.11	.747	0.00
	IA by SA	1	0.15	0.15	0.07	.796	0.00
	Error	60	135.87	2.26			

Note. Correct/complete: percentage of tasks with correct answers out of completed number of tasks.

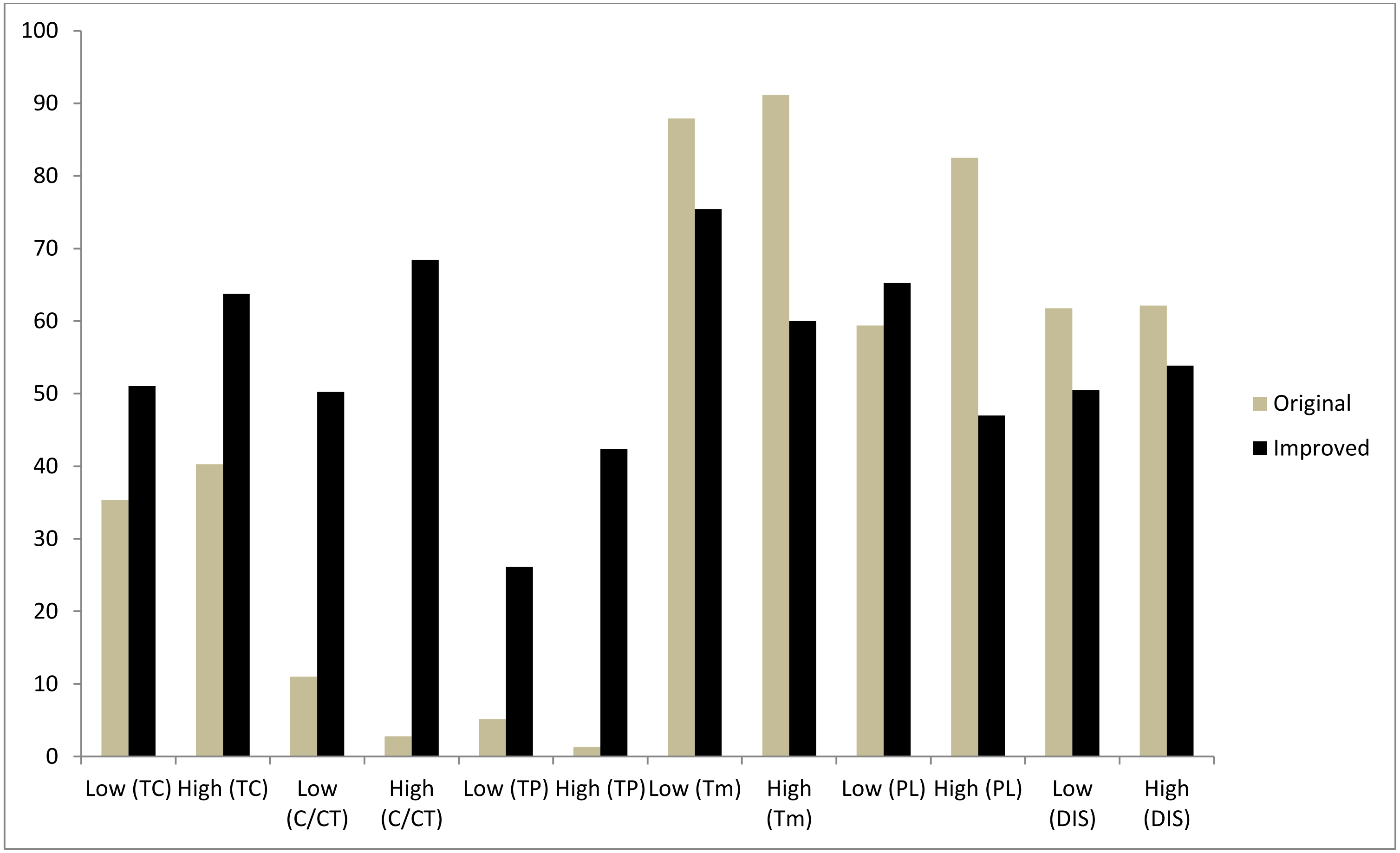


Fig. 4. Mean values (rescaled as a percentage) as a function of information architecture and spatial ability (Study 2).

Note. Low: low spatial ability. High: high spatial ability. TC: task completion. C/CT: correct/completed task. TP: task performance. Tm: time (average). PL: page loads. DIS: disorientation.

More detailed analysis per task showed, as expected, that in Study 1 participants' odds of task success were higher when the simulation succeeded on the improved information architecture and specific improvements in information scent for the particular task had been made. However, perhaps surprisingly, the odds were also higher in the site with improved information architecture for which no improvements had been made (because of simulation success with the original information architecture). Nevertheless, in Study 2 detailed analysis per task showed, as expected, that when with improved information architecture the simulation passed, the odds of task success were higher than when improvements had been made but the simulation failed. However, perhaps surprisingly, the latter odds were still relatively high compared to those on the original site where the simulation failed. Collectively, these results lead to the conjecture, to be explored in the future, that the context provided by the results of task performance on previous tasks may affect performance on the current task. In particular, in the situation were an information architecture has poor information scent for some tasks and better for others (as opposed to an information architecture with good information scent for almost all tasks), poor scent apparently also (negatively) affects performance on tasks for which scent is better (Study 1). Moreover, if the information architecture has poor information scent for some tasks and better for others, performance on tasks for which scent is worse may still be better than when information scent is poor for all tasks (Study 2). Similar types of sequential context effect, where experience or judgment on previous trials influences what happens on the next trial, have been observed in web navigation [David et al., 2007], psychophysics [Lockhead, 2004] and frequency estimation [Kusev et al., 2011]. For example, David et al. [2007] found evidence for a 'virtuous cycle' in web navigation. The successful execution of information-seeking goals in one cycle enhanced self-efficacy. As a result, perceived difficulty of information goals in the following cycle is reduced. Moreover, as a result of self-efficacy from previous cycles, more challenging goals are formulated in subsequent cycles. Similarly, in our work, better information architecture for some tasks, and consequently more successful web navigation, apparently 'carries over' to more successful navigation on tasks on for which information architecture is poorer.

Although goal-specific analysis was successfully only conducted on subsites (consisting of hundreds of pages), because of extremely large requirements of memory and processing power, goal-specific analysis of entire large websites (consisting of thousands of pages) was precluded. Nevertheless, non-goal-specific analysis was successfully conducted on entire large websites [consisting of thousands of pages; Muzahir, 2013]. However, with increasing memory capacity and processing power, automated goal-specific analysis of sites is expected to become increasingly feasible for larger websites.

Our approach of using content (here, text paragraphs) as potential user-goals in simulations is important because it is necessary that all goals for which information is available in a website can be found by a user. These goals represent all the information needs that the site can meet, whether these goals are representative of 'users' goals in the wild' or not. Although we do not assume that these are the only goals that users may have, it is essential that the available information can be found by users who may have needs that this information meets. Of course, other research may complement this work by focusing on users' goals in the wild, which may uncover new information needs that a particular website currently does not meet. However, this work will then be addressing two issues at the same time. The first is whether users can find available information in a website which meets their information needs. The second is whether the site does not meet users' particular information needs; by definition then, users will not be able to find that information. Our work focuses on the first of these issues.

5.2 Future work

5.2.1. Support for global design decisions. As shown in this paper, simulation results can be used as a basis for detailed design work improving the information architecture of a site. However, the results may also be used to decide whether or not the labeling system of a website's information architecture should be redesigned from scratch. A cut-off point (e.g., 90% successful simulation results) may be set to decide whether it is worth 'repairing' the site. If the cut-off is achieved or exceeded then making improvements to the existing information architecture and otherwise a complete redesign may be considered. Based on our work and related research [e.g., Blackmon et al., 2002, 2003, 2005] guidelines for designing a new information architecture are presented in Online Appendix F.

As all of a website's information-architecture-related information is stored in a database, our database-oriented approach can easily facilitate further analyses, in addition to goal-specific analysis (this paper) and non-goal-specific analysis [Muzahir, 2013]. For example, this allows evaluation of the extent to which a real-world information architecture is balanced, that is having little variation in the level of depth of pages with content (text paragraphs or other). This is important, as content on pages at extreme depth from the homepage will be exceedingly prone to navigation errors in reaching these pages [Van Schaik & Ling, 2012]. A balanced information architecture would spread the risk of navigation errors more evenly. The results of a balance analysis may be used to decide whether or not the organization system of a website's information architecture should be redesigned from scratch. Again a cut-off point may be set to decide whether it is worth 'repairing' the site. Similarly, another analysis could evaluate the navigation system of an information architecture in terms of the degree of connectedness of individual pages, as a basis for decisions on improving parts of this system or redesign from scratch.

5.2.2. Cognitive computational modeling beyond information architecture. Given that the information architecture of websites is the most important remaining source of usability problems [Nielsen, 2009], the current research builds on existing work by analyzing the information architecture of websites. Therefore, other aspects of web page design, such as web page layout are not addressed. Questions then arise regarding (1) the extent to which the automated analysis of large websites would benefit from modeling additional aspects of web design, in particular page layout, and (2) the feasibility of this additional work.

First, Teo and John [2008] demonstrate that, in the case of a two-column layout, including layout in the modeling of web navigation improves the prediction of page loads and correct first page loads. However, this work did not compare identical layouts with different labeling schemes (problematic and improved) and only studied two-column layouts. Moreover, in other research, Blackmon [2012] found that "the distribution of attention among available information patches was strongly determined by the rank ordering of semantic similarity between user goal but was *not* [emphasis added] influenced by website designs with very different visual layouts" (p. 3). In particular, most of the variance in human performance on websites is caused by the pattern of semantic similarity of a user's goal with headings (information patches). Specifically, when the user is pulled by high semantic similarity to focus on links nested under the correct heading, the task is easy, but users flounder and often encounter task failure when high goal-heading semantic similarity pulls the user towards an incorrect heading(s) rather than the correct heading. In the current research, both heading labels and link labels were improved. The purpose was not to establish the independent contributions of heading labels and link labels to the enhancement of web navigation [but see Resnick & Sanchez, 2004]. Future work may examine these separate contributions.

Furthermore, Resnick and Sanchez [2004] found no advantage of either of two organization schemes as long as link labels were not poor. In addition, attempts have been made to include images in modeling of web navigation [CoLiDeS+ Pic; Van Oostendorp et al., 2012], but this work has not studied large websites and has not developed automated support.

Second, as part of the modeling of web navigation, Teo and John [2008] automatically created a device model representing two-column web page layouts from imported web pages [see also Teo et al., 2012]. However, the automatic capture of more complex layouts and identical page layouts created with different coding techniques or (combinations of) coding languages (HTML, JavaScript, style sheets) may not be trivial. In addition, visual aspects such as the use of color and animation on web pages may also affect web navigation. In sum, in order to establish the feasibility and benefits of including other aspects of web page design than information architecture in computational cognitive modeling of web navigation for large websites further work seems to be needed.

Our results show that some improvements in information architecture are not without downsides, as in both studies with improvements the number of negative comments was higher for layout than without improvements. This suggests that changes to information architecture to improve information scent can be at odds with subjective evaluations of layout, despite improvements in task performance. This is an interesting and unique finding worthy of further exploration.

5.2.2. Cognitive computational modeling for personal information architecture. Personal information management is an important area of contemporary research in psychology and computer science: “an activity in which an individual stores his/her personal information items in order to retrieve them later on” [Bergman, 2012, p. 55]. An important aspect of this management is *personal information architecture*, information architecture created, maintained and used by the same individual. As presumably all people (whether they are computer users or not) create and use their own personal information architecture and may suffer information pain in our personal information management, personal information management affects us all.

Given the nature of personal information management, the success of us using our own (personal) information architectures will to a large extent be positively influenced by the extent to which they support this activity. This demand is not new, but research into personal information management is relatively sparse and dispersed. Moreover, there appears to be a lack of work that approaches personal information management from the perspective of information architecture, where both human cognition and computational modeling can be crucial.

On the one hand, personal-information management research has studied empirical information-finding [e.g., Bergman, 2012] and computational tools to support personal information management [e.g., Jones & Anderson, 2011] of personal information collections. Increasingly, these collections can reside in various types of system, including standalone and intranet file systems, e-mail, Web-based systems and mobile-based systems [Jones & Anderson, 2011]. On the other hand, information architecture research has focused on navigation in websites through computational cognitive modeling and usability engineering, including non-goal-specific and goal-specific analysis [e.g., Blackmon et al., 2002, 2003, 2005]. In the case of large websites, this work becomes feasible by automating the capture, modeling and analysis, where a database-oriented approach is advantageous (this paper). Although the application of cognitive computational modeling to personal information management and personal information architecture appears to be lacking, this application would seem to hold great promise. However, the problem in website analysis that the labels of headings and links could be proper names rather than words (this paper) would be compounded by naming conventions using non-word labels (e.g., abbreviations, initials and dates) in the analysis of personal information architectures. Words, but not non-word labels, are included in the semantic space that represents a particular language; therefore, only words can be used in non-goal-specific and goal-specific analysis. Thus, analysis of personal information architectures that relies on non-word labels would be incomplete if conducted the same way as the analysis of a website's information architecture. However, a database-oriented approach still seems indispensable, as this allows analysts to 'interrogate' the information architecture efficiently through database queries, relate characteristics of computer users' information architectures to their strategy for creating, maintaining and retrieving information from it, and define information retrieval tasks to be used in experiments.

6. CONCLUSION

The work reported here demonstrates the feasibility of automating cognitive computational analysis of the information architecture of large websites, as a basis for improvement, and the validity of analysis results through experiments with end-users. With further advances in computing technology in terms of internal and external memory, and processing power this work is likely to scale up allowing the analysis of increasingly larger sites. Our flexible database-oriented approach allows, for example, goal-specific analysis, non-goal-specific analysis, modeling of non-text media content, and analysis of the organization and navigation systems of information architectures. We look forward to future work exploiting this approach in usability engineering, as a basis for improving web navigation.

ELECTRONIC APPENDIX

The electronic appendix for this article can be accessed in the ACM Digital Library.

ACKNOWLEDGMENTS

The authors are grateful for support from the Teesside University's Research Doctoral Scholarship scheme. The authors also thank Dr. Marilyn Blackmon (Institute of Cognitive Science [ICS], University of Colorado at Boulder) and Richard Brown for access to and advice regarding the use of the AutoCWW2 LSA tools.

REFERENCES

AHMED, I. & BLUSTEIN, J. 2006. Influence of spatial ability in navigation: using look-ahead Breadcrumbs on the web. *International Journal of Web Based Communities* 2, 183-196.

AHUJA, J.S. AND WEBSTER, J. 2001. Perceived disorientation: An examination of a new measure to assess web design effectiveness. *Interacting with Computers* 14, 15-29.

BERGMAN, O. 2012. The user-subjective approach to personal information management: From theory to practice. *Studies in Computational Intelligence* 396, 55-81.

BLACKMON, M.H. 2012. Information scent determines attention allocation and link selection among multiple information patches on a webpage. *Behaviour and Information Technology* 31, 3-15.

BLACKMON, M.H., KITAJIMA, M. AND POLSON, P.G. 2003. Repairing usability problems identified by the cognitive walkthrough for the web. In *The CHI 2003 New Horizons Conference Proceedings: Conference on Human Factors in Computing Systems,* Ft. Lauderdale, FL, 5 April 2003 through 10 April 2003, BELLOTTI V, ERICKSON T, COCKTON G AND KORHONEN P, Eds. , 497-504.

BLACKMON, M.H., KITAJIMA, M. AND POLSON, P.G. 2005. Tool for accurately predicting website navigation problems, non-problems, problem severity, and effectiveness of repairs. In *CHI 2005: Technology, Safety, Community - Conference on Human Factors in Computing Systems,* Portland, OR, 2 April 2005 through 7 April 2005, KELLOGG W., ZHAI S., GALE C. AND VEER G., Eds. , 31-40.

BLACKMON, M.H., POLSON, P.G., KITAJIMA, M. AND LEWIS, C. 2002. Cognitive walkthrough for the Web. In *Conference on Human Factors in Computing Systems,* Minneapolis, MN, 20 April 2002 through 25 April 2002, TERVEEN L., WIXON D., COMSTOCK E. AND SASSE A., Eds., 463-470.

BROWN, R. 2005. ACWW: Adding Automation to the Cognitive Walkthrough for the Web (CWW). UNF Theses and Dissertations. Paper 195. University of North Florida, Jacksonville.

CHI, E.H., PIROLLI, P. AND PITKOW, J. 2000. Scent of a site: a system for analyzing and predicting information scent, usage, and usability of a Web site. In *CHI 2000 - Conference on Human Factors in Computing Systems 'The Future is Here',* The Hague, Netherlands, 1 April 2000 through 5 April 2000, ACM, New York, NY, United States, 161-168.

CHI, E.H., ROSIEN, A., SUPATTANASIRI, G., WILLIAMS, A., ROYER, C., CHOW, C., ROBLES, E., DALAL, B., CHEN, J. AND COUSINS, S. 2003. The bloodhound project: Automating discovery of web usability issues using the InfoScent™ simulator. In *The CHI 2003 New Horizons Conference Proceedings: Conference on Human Factors in Computing Systems,* Ft. Lauderdale, FL, 5 April 2003 through 10 April 2003, BELLOTTI V, ERICKSON T, COCKTON G AND KORHONEN P, Eds. , 505-512.

CLARK-CARTER, D. 2009. Quantitative psychological research. Psychology Press, Hove.

COHEN, J. 1988. Statistical power analysis for the behavioral sciences. Erlbaum, Hillsdale, NJ.

CUDECK, R. AND HENLY, S.J. 2003. A realistic perspective on pattern representation in growth data: Comment on Bauer and Curran (2003). *Psychological methods* 8, 378-383.

DAVID, P., SONG, M., HAYES, A. AND FREDIN, E.S. 2007. A cyclic model of information seeking in hyperlinked environments: The role of goals, self-efficacy, and intrinsic motivation. *International Journal of Human-Computer Studies* 65, 170-182. .

FU, W.-T. AND PIROLLI, P. 2007. SNIF-ACT: A cognitive model of user navigation on the World Wide Web. *Human-Computer Interaction* 22, 355-412.

HOWES, A., PAYNE, S. AND RICHARDSON, J. 2002. An instance-based model of the effect of previous choices on the control of interactive search. In *Proceedings of the 24th Annual Meeting of the Cognitive Science Society,* pp. 476-481. http://csjarchive.cogsci.rpi.edu/Proceedings/2002/CogSci02.pdf

JACOB, E.K. AND LOEHRLEIN, A. 2009. Annual Review of Information Science and Technology. 43, 147-186. DOI: 10.1002/aris.2009.1440430110.

HORNBÆK, K., SANDER, S.S., BARGAS-AVILA, J. AND SIMONSEN, J.G. 2014. Is once enough? On the extent and content of replications in human-computer interaction. In *32nd Annual ACM Conference on Human Factors in Computing Systems, CHI 2014,* Toronto, ON, 26 April 2014 through 1 May 2014, Association for Computing Machinery, 3523-3532.

HOWES, A., PAYNE, S. AND RICHARDSON, J. 2002. An instance-based model of the effect of previous choices on the control of interactive search. In *Proceedings of the 24th Annual Meeting of the Cognitive Science Society,* 476-481.

JONES, W. AND ANDERSON, K.M. 2011. Many views, many modes, many tools... one structure: Towards a non-disruptive integration of personal information. *HT 2011 - Proceedings of the 22nd ACM Conference on Hypertext and Hypermedia* 113-122.

JUVINA, I. AND VAN OOSTENDORP, H. 2006. Individual differences and behavioral metrics involved in modeling web navigation. *Universal Access in the Information Society* 4, 258-269.

JUVINA, I. AND VAN OOSTENDORP, H.. 2008. Modeling semantic and structural knowledge in Web navigation. *Discourse Processes* 45, 346-364.

KARANAM, S., VAN OOSTENDORP, H. AND INDURKHYA, B. 2011. Towards a fully computational model of web-navigation. In *24th International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems – Part 1, IEA/AIE 2011*, Syracuse, NY, 28 June 2011 through 1 July 2011, Springer, Berlin, 327-337.

KATSANOS, C., TSELIOS, N. AND AVOURIS, N. 2008. Automated semantic elaboration of web site information architecture. *Interacting with Computers* 20, 535-544.

KATSANOS, C., TSELIOS, N. AND AVOURIS, N. 2010. A survey of tools supporting design and evaluation of websites based on models of human information interaction. *International Journal on Artificial Intelligence Tools* 19, 755-781.

KINTSCH, W. 1998. *Comprehension: a paradigm for cognition*. Cambridge University Press, Cambridge, MA.

KITAJIMA, M., BLACKMON, M.H. AND POLSON, P.G. 2000. A comprehension-based model of web navigation and its application to web usability analysis. In *People and Computers XIV - Usability or*

Else! (Proceedings of HCI 2000), S. MCDONALD, Y. WAERN AND G. COCKTON, Eds. Springer, Berlin, 357-373.

KITAJIMA, M., BLACKMON, M.H. AND POLSON, P.G. 2005. Cognitive architecture for website design and usability evaluation: comprehension and information scent in performing by exploration. In *Proceedings of 11th International Conference on Human-Computer Interaction, HCI-International*.

KITAJIMA, M., KARIYA, N., TAKAGI, H. AND ZHANG, Y. 2005. Evaluation of website usability using Markov chains and latent semantic analysis. *IEICE Transactions on Communications* E88-B, 1467-1475.

KUSEV, P., AYTON, P., VAN SCHAIK, P., TSANEVA-ATANASOVA, K., STEWART, N. AND CHATER, N. 2011. Judgments Relative to Patterns: How Temporal Sequence Patterns Affect Judgments and Memory. *Journal of Experimental Psychology: Human Perception and Performance* 37, 1874-1886.

LANDAUER, T. 2007. LSA as a theory of meaning. In *Handbook of latent semantic analysis,* T. LANDAUER, D. MCNAMARA, S. DENNIS AND W. KINTSCH, Eds. Erlbaum, Mahwah, NJ, 3-34.

LOCKHEAD, G.R. 2004. Absolute judgments are relative: A reinterpretation of some psychophysical ideas. *Review of General Psychology* 8, 265-272.

MCDONALD, J.E., PAAP, K. AND MCDONALD, D.R. 1990. Hypertext perspectives: Using Pathfinder to build hypertext systems. In *Pathfinder associative networks*, R. SCHVANEVELDT, Ed. Ablex, Norwood, NJ, 197-211.

MILLER, C.S. AND REMINGTON, R.W. 2004. Modeling information navigation: Implications for information architecture. *Human-Computer Interaction* 19, 225-271.

MORVILLE, P. AND ROSENFELD, L. 2006. IA for the World Wide Web (3rd ed.). O'Reilly, Sebastopol, Calif.

MUZAHIR, R.H. 2013. Computational cognitive modeling of information architecture and web navigation. Unpublished PhD thesis. Teesside University, Middlesbrough. Available at *http://sss-studnet.tees.ac.uk/psychology/staff/Paul_vs/PhDprojects/Muzahir2013.pdf*

NIELSEN, J. 1993. *Usability engineering*. Academic Press, USA.

NIELSEN, J. 2000. *Designing web usability*. New Riders, Indianapolis, IN.

NIELSEN, J. 2009. IA task failures remain costly. *http://www.nngroup.com/articles/ia-task-failures-remain-costly/*

OLSTON, C. AND CHI, E.H. 2003. ScentTrails: Integrating browsing and searching on the Web. *ACM Transactions on Computer-Human Interaction* 10, 177-197.

PARUSH, A. AND YUVILER-GAVISH, N. 2004. Web navigation structures in cellular phones: The depth/breadth trade-off issue. *International Journal of Human Computer Studies* 60, 753-770.

PIROLLI, P. AND CARD, S. 1999. Information foraging. *Psychological review* 106, 643-675.

RAPP, D.N. AND VAN DEN BROEK, P. 2005. Dynamic text comprehension: An integrative view of reading. *Current Directions in Psychological Science* 14, 276-279. .

RESNICK, M.L. AND SANCHEZ, J. 2004. Effects of organizational scheme and labeling on task performance in product-centered and user-centered retail web sites. *Human factors* 46, 104-117.

SCHAIK, P. VAN AND LING, J. 2012. An experimental analysis of experiential and cognitive variables in web navigation. *Human-Computer Interaction* 27, 199-234.

SHEPARD, R.N. AND METZLER, J. 1971. Mental rotation of three-dimensional objects. *Science* 171, 701-703.

SPOOL, J.M., SCANLON, T., SCHROEDER, W., SNYDER, C. AND DEANGELO, T. 1999. Web Site Usability: a designer's guide. Morgan-Kaufmann, Los Altos, CA.

SWIERENGA, S.J., SUNG, J., PIERCE, G.L. AND PROPST, D.B. 2011. Website design and usability assessment implications from a usability study with visually impaired users. In *6th International Conference on Universal Access in Human-Computer Interaction, UAHCI 2011, Held as Part of HCI International 2011,* Orlando, FL, 9 July 2011 through 14 July 2011, 382-389.

TABACHNICK, B. AND FIDELL, L. 2001. *Computer-assisted research design and analysis* Allyn and Bacon, Boston.

TEO, L. AND JOHN, B.E. 2008. Towards a tool for predicting goal-directed exploratory behavior. In *52nd Human Factors and Ergonomics Society Annual Meeting, HFES 2008,* New York, NY, 22 September 2008 through 26 September 2008, 950-954.

TEO, L., JOHN, B.E. AND BLACKMON, M.H. 2012. CogTool-Explorer: A model of goal-directed user exploration that considers information layout. In *30th ACM Conference on Human Factors in Computing Systems, CHI 2012,* Austin, TX, 5 May 2012 through 10 May 2012, 2479-2488.

VAN OOSTENDORP, H. AND JUVINA, I. 2007. Using a cognitive model to generate web navigation support. *International Journal of Human Computer Studies* 65, 887-897.

VAN OOSTENDORP, H., KARANAM, S. AND INDURKHYA, B. 2012. CoLiDeS+ Pic: A cognitive model of web-navigation based on semantic information from pictures. *Behaviour and Information Technology* 31, 17-30.

VAN OOSTENDORP, H., MADRID, I. AND PUERTA MELGUIZO, M.C. 2009. The effect of menu type and task complexity on information retrieval performance. *The Open Ergonomics Journal* 2, 64-71.

VANDENBERG, S. AND KUSE, A. 1978. Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills* 47, 599-604.Rafal Ablamowicz and Bertfried Fauser. 2007. CLIFFORD: a Maple 11 Package for Clifford Algebra Computations, version 11. (2007). Retrieved February 28, 2008 from http://math.tntech.edu/rafal/cliff11/index.html

Received December 2013; revised June 2014 and November 2014; accepted March 2015

Online Appendix to:
Automated Computational Cognitive-Modeling: Goal-Specific Analysis for Large Websites

PAUL VAN SCHAIK, Teesside University
RAZA HABIB MUZAHIR, Teesside University
MIKE LOCKYER, Teesside University

A. SOME LIMITATIONS OF COLIDES AND COLIDES+

A1 Implementation of CoLiDeS

Kitajima et al.'s [2005] implementation of CoLiDes has several limitations. It is not clear whether or how pages are parsed and modeled. Modeling does not use a separate attention phase and separate action phase, in other words does not distinguish between headings and links, and therefore not between links without headings and links with headings. Although backtracking is included in the modeling of task sequence length, but without the use of path adequacy, backtracking is not included in the modeling of usability problems. The implementation does not seem to model information architecture or goal-specific analysis in a persistent database, so every analysis would have to start from scratch. Moreover, the reported research does not provide a comparison of simulation results with the empirical results of web navigation. Furthermore, this work models a system with restricted information architecture, having three fixed levels. In addition, it is not explained how exactly link labels are repaired (e.g., by elaboration or another technique).

Karanam et al.'s [2011] software prototype of CoLiDeS has various shortcomings. This starts with the homepage and a goal statement, and predicts the link on each page that a user would select. This approach to implementing CoLiDeS is incorrect, or at least incomplete, as it does not consider the headings on a page before choosing a particular link. Besides, it elaborates the hyperlinks using near-neighbor analysis before computing the semantic similarity between a user-goal and a hyperlink text. However, not every user can necessarily elaborate all links correctly. Elaborating a text using a near-neighbor strategy before LSA analysis ignores the possibility that users will ignore difficult-to-understand text [Blackmon et al., 2002, 2003, 2005]. Furthermore, no details are provided regarding the stop condition and navigation history is not taken into account.

A2 Only accounting for links with heading

In CoLiDeS/CoLiDeS+, functioning on a single web page screen object, navigation is regarded as the outcome of a multi-step process algorithm that divides any given page in four processes [Kitajima et al., 2000]: "parsing, focusing on, comprehension, and selection" (p. 3). A shortcoming of this model is that it takes into consideration only the headings and links with heading.

A3 Not taking into account headings

CoLiDeS+ does not take into consideration the headings under which links are organized. This implies that users do not take into account the scent of the corresponding heading when evaluating the scent of a particular link. To overcome

DOI:http://dx.doi.org/10.1145/0000000.0000000

this limitation in a simulation, on each selection, the cosine value of the to-be-selected element against the goal statement should increase.

B. ALGORITHM OF ENHANCED COLIDES/+

```
Input
Description of goal and start web page
Algorithm
Step 1: segment page
Determine the regions on the page
Each region is defined by its heading, except the set of headingless sections of links, which is treated as a separate ‘headingless’ region.
Step 2: select region
Calculate cosine of the heading of each region with goal statement
IF highest cosine ≥ previously highest cosine
THEN
    Select region with heading with the highest cosine
    Add heading of selected region to path
ELSE
    Calculate path-goal relevancy for each region/heading
    IF highest path-goal relevancy ≥ previously highest path-goal relevancy
    THEN
        Select the region with the highest path-goal relevancy
        Add heading of selected region to path
    END IF
END IF
Step 3: segment region
IF region selected
THEN
    Identify elements of selected region – page elements under the heading defining the region
END IF
Step 4: check if region contains the target of goal statement
IF region selected
THEN
    Calculate cosine of the each element of the selected region with goal statement.
END IF
IF highest cosine ≥ 0.85 is found
THEN
    Stop
    Report path to solution and matching target
END IF
Step 5: select action (link)
Calculate cosine of each link without heading
Identify maximum of highest cosines from Step 4 (links with heading) and Step 5 (links without heading)
IF maximum ≥ previously highest cosine
THEN
    Select the link with the highest cosine
    Add selected link to path
ELSE
    Calculate path adequacy for each link
    IF highest path adequacy ≥ previously highest path adequacy
        Select the link with the highest path adequacy
        Add selected link to path
    ELSEIF there are still remaining regions
        Remove last heading from path
```

```
        Select region with second-highest goal relevancy
        Add heading of selected region to path
        Go to Step 3
    ELSE (impasse – decreasing cosine)
        IF previous page exists
        THEN
            Go back to previous page
            Go to Step 2
        ELSE
            Stop
            Report: “no matching goal statement with a target”
        END IF
    END IF
END IF
Step 6: process next page
Identify next page by following the selected link
Go to Step 1
```

The implementation of the algorithm of enhanced CoLiDeS/+ is available at the following location:
https://github.com/pvschaik/CoLiDeS-slash-plus

C. Success per task in Validation Experiment (Study 1)

Original information architecture

Practice Task	Simulation success, website with original information architecture	Pct correct [C]	Pct incorrect [IC]	Pct not completed	Pct completed
1	fail	39	61	0	100
2	pass	30	70	0	100
3	fail	70	30	0	100
Main Task					
1	pass	61	39	0	100
2	fail	59	35	7	93
3	fail	4	72	24	76
4	fail	46	13	41	59
5	pass	17	33	50	50
6	pass	2	35	63	37
7	pass	7	11	83	17
8	pass	7	2	91	9
9	fail	4	0	96	4
10	pass	2	2	96	4
11	fail	2	0	98	2
12	pass	2	0	98	2
13	fail				
14	pass				
15	pass				
16	pass				
17	pass				
18	fail				
19	fail				
Main	mean, pass, Tasks 1-12	13.98			
	mean, fail, Tasks 1-12	23.04			
Practice	mean, pass	30.43			
	mean, fail	54.35			

Improved information architecture

Practice Task	Simulation success, website with original information architecture	Pct correct [C]	Pct incorrect [IC]	Pct not completed	Pct completed
1	fail	75	25	0	100
2	pass	90	10	0	100
3	fail	90	10	0	100
Main Task					
1	pass	96	4	0	100
2	fail	83	15	2	98
3	fail	88	8	4	96
4	fail	81	6	13	88
5	pass	77	4	19	81
6	pass	25	50	25	75
7	pass	21	40	40	60
8	pass	48	0	52	48
9	fail	35	4	60	40
10	pass	13	23	65	35
11	fail	13	10	77	23
12	pass	17	0	83	17
13	fail	10	4	85	15
14	pass	6	2	92	8
15	pass	8	0	92	8
16	pass	6	0	94	6
17	pass	2	2	96	4
18	fail	2	0	98	2
19	fail	2	0	98	2
Main	mean, pass, Tasks 1-12	42.26			
	mean, fail, Tasks 1-12	60.00			
Practice	mean, pass	89.58			
	mean, fail	82.29			

Original (O) versus improved (I) information architecture

Practice Task	Simulation success, website with original information architecture	Odds (pct correct, O/I)
1	fail	0.52
2	pass	0.34
3	fail	0.78
Main Task		
1	pass	0.64
2	fail	0.70
3	fail	0.05
4	fail	0.56
5	pass	0.23
6	pass	0.09
7	pass	0.31
8	pass	0.14
9	fail	0.12
10	pass	0.17
11	fail	0.17
12	pass	0.13
13	fail	
14	pass	
15	pass	
16	pass	
17	pass	
18	fail	
19	fail	
Main	mean, pass, Tasks 1-12	0.24
	mean, fail, Tasks 1-12	0.32
Practice	mean, pass	0.34
	mean, fail	0.65

D. Success per task in Validation Experiment (Study 1)

Original information architecture

Practice Task	Simulation success, website with improved information architecture	Pct correct [C]	Pct incorrect [IC]	Pct not completed	Pct completed
1	pass	3	97	0	100
2	pass	25	75	0	100
3	pass	16	84	0	100
Main Task					
1	fail	3	97	0	100
2	pass	22	41	38	63
3	pass	0	47	53	47
4	pass	3	34	63	38
5	pass	0	25	75	25
6	pass	0	13	88	13
7	fail	0	6	94	6
8	pass	0	3	97	3
Main	mean, pass, Tasks 1-7	5			
	mean, fail, Tasks 1-7	10			
Practice	mean, pass	15			

Improved information architecture

Practice Task	Simulation success, website with improved information architecture	Pct correct [C]	Pct incorrect [IC]	Pct not completed	Pct completed
1	pass	59	41	0	100
2	pass	88	13	0	100
3	pass	84	16	0	100
Main Task					
1	fail	63	38	0	100
2	pass	94	3	3	97
3	pass	53	44	3	97
4	pass	59	16	25	75
5	pass	13	50	38	63
6	pass	16	13	72	28
7	fail	0	13	88	13
8	pass				
Main	mean, pass, Tasks 1-7	47			
	mean, fail, Tasks 1-7	31			
Practice	mean, pass	77			

Original (O) versus improved (I) information architecture

Practice Task	Simulation success, website with improved information architecture	Odds (pct correct, O/I)
1	pass	0.05
2	pass	0.29
3	pass	0.19
Main Task		
1	fail	0.05
2	pass	0.23
3	pass	0.00
4	pass	0.05
5	pass	0.00
6	pass	0.00
7	fail	
8	pass	
Main	mean, pass, Tasks 1-7	0.06
	mean, fail, Tasks 1-7	0.05
Practice	mean, pass	0.17

E. Comparison of Enhanced CoLiDeS/+ Model with CoLiDeS and CoLiDeS+

E1 Similarities

CoLiDeS, CoLiDeS+ and CoLiDeS/+ share the following characteristics.

1. The models are based on Kintsch's theory of text comprehension.
2. In the models, processing takes place in two phases: parsing the page into regions followed by focusing attention on the region selected because it is semantically most similar.
3. While focusing on the selected region, the models evaluate the links within that region and select an action (link) in that region.

E2 Differences

Criteria	CoLiDeS/+	CoLiDeS	CoLiDeS+
Considers heading, link with heading and link without heading, along with link relevance (cosine)	Yes	No	No
Uses heading label to elaborate link label is forward search	Yes	No	No
Checks that link cosine higher than heading cosine	Yes	No	No
Increasing (or at least non-decreasing) similarity in path to solution	Yes	No	Yes
Backtracking (in case of impasse)	Yes	No	Yes
Automated and simulated implementation	Yes	No	No
Predicts correct path	Yes	No	No
Automated goal statement identification	Yes – uses paragraphs (100-200 words long)	No	No
Simulated path and shortest analysis for comparison	Yes	No	No
Outputs all paths to reach target web page	Yes	No	No

F. Guidelines for Designing Information Architecture

The information architecture of a website can be defined as the structure of its organization and navigation, and its labeling scheme [Morville & Rosenfeld, 2006] for the benefit of end-users who need to find and use information. Here, we present a small but powerful set of guidelines for improving the information architecture of websites, irrespective of size. Although it is possible to apply these guidelines without specific tool support, their application can be more effective when appropriate tools are used (for example tools that compute cosine[10] and term-vector length[11]). The guidelines do not replace other credible guidelines for web design. Moreover, guidelines in general (including these) are not a substitute for, but rather complement, usability-testing of websites with target users.

Goal-specific analysis and design per goal. For each piece of content (for example text[12], image, audio clip, video clip) as a goal, define its goal statement [Blackmon et al., 2002]. Make sure that information scent (the similarity of the goal statement with each web page element leading to the destination page) is sufficient for target users [Blackmon et al., 2003].

Non-goal-specific analysis and design per web page.

1 Make sure that each heading label has a distinct meaning for target users [Blackmon et al. 2002, 2003].
2 Under each heading on each page, make sure that each link label has a distinct meaning for target users [Blackmon et al. 2002, 2003].
3 Make sure that the label of each link that in a headingless region has a distinct meaning for target users.
4 Make sure that the meaning of each link label under a particular heading has sufficient similarity with that of the heading label, and enhances the sense and meaning of its heading for target users[13].
5 Make sure that the meaning of each heading label is familiar to target users [Blackmon et al., 2002].
6 Make sure that the meaning of each link label is familiar to target users [Blackmon et al., 2002].
7 Make sure that the meaning of each image label or its 'alternate' text is familiar to target users.

---

[10] https://autocww2.colorado.edu/OneToMany.html
[11] https://autocww2.colorado.edu/unfamiliar.html
[12] paragraphs of text on a web page
[13] and the meaning of the heading label enhances the sense and meaning of its links for target users