Running head: EFFECTS OF THE METACOGNITIVE COMPUTER-TOOL MET.A.WARE

Effects of the metacognitive computer-tool *met.a.ware* on the web search of laypersons

Marc Stadtler & Rainer Bromme

University Münster, Germany

Correspondence concerning this article should be addressed to Marc Stadtler. University Münster, Psychologisches Institut III, Fliednerstr. 21, D-48149 Münster, Germany. Email: stadtlm@uni-muenster.de

Please cite this article in press as:

Stadtler, M. & Bromme, R. (in press). Effects of the metacognitive computer-tool on the web search of laypersons, *Computers in Human Behavior*.

Abstract

Today, laypersons often consult the Internet to inform themselves about health-related issues. However, the competent use of these often complex and heterogeneous information provisions cannot be taken for granted, because many Internet users are lacking the necessary metacognitive prerequisites. Therefore, we developed the metacognitive computer-tool *met.a.ware*, which supports laypersons' Internet research for medical information by the means of metacognitive prompting and ontological classification. In an experimental investigation of *met.a.ware* a total of 118 participants with little medical knowledge were asked to conduct an Internet research on a medical topic. Participants were randomly assigned to four experimental groups that worked with *met.a.ware* and either received evaluation prompts, monitoring prompts, both types of prompts, or no prompts. All experimental conditions were additionally provided with ontological classification. One control group took paper and pencil notes. A further control group took notes using a blank text window. Results showed that laypersons receiving evaluation prompts outperformed controls in terms of knowledge about sources and produced more arguments commenting on the source of information in an essay task. In addition, laypersons receiving monitoring prompts acquired significantly more knowledge about facts, but did not perform better on a comprehension test than the controls. The availability of ontological categories helped to structure the notes laypersons in the conditions working with ontological classification took during Internet research. Analyses of the notes further demonstrated that the availability of ontological categories guided information search in direction of the selected categories. It is concluded, that *met.a.ware* is an effective tool that supports laypersons' Internet research.

Effects of the metacognitive computer-tool *met.a.ware* on the web search of laypersons

Our modern industrial societies make it increasingly necessary for laypersons to deal with expert information in many areas of everyday life, as a patient, client or customer. Laypersons often retrieve information from the Internet, which used to be available to experts exclusively (Bromme, Jucks & Runde 2005). From a psychological perspective, this trend provides an intriguing scenario to study and support processes of informal learning in complex information environments, i.e. learning that takes place outside of academic or school contexts and mainly fulfils the purpose of making an informed, i.e. knowledge-based decision.

One example of this trend is the widespread use of medical information on the WWW (Fox & Fallows, 2003). Here, laypersons have immediate access to a vast amount of health-related web sites, which are only a mouse-click away. The information they retrieve may help laypersons to understand the assets and drawbacks of different therapeutical alternatives and eventually make a knowledge-based decision about their compliance with a suggested therapy (Morahan-Martin, 2004).

However, research has shown that Internet users often experience considerable difficulties when dealing with complex search tasks on the WWW (e.g., Eveland & Dunwoody, 2000; Hofer, 2004; Hölscher, 2002; Klatt et al., 2001; Lazonder, 2000). The immediate availability of vast amounts of information entails that laypersons dispose of numerous options for action at a time. This 'metacognitively complex' situation (Veenman, Wilhelm & Beishuizen, 2004) requires a well-directed use of metacognitive strategies. Laypersons have to plan their search process, decide which information is relevant, and thoroughly monitor the progress of their own comprehension to be efficient (Bannert, 2003). An accurate comprehension monitoring may be of particular importance when learning about complex scientific information because it enables laypersons to regulate their study process, for instance by re-reading certain parts of the text or slowing down their speed of reading (Baker, 1984; Thiede, Anderson & Therriault, 2003). In the context of Internet research, where nearly unlimited masses of information are available, an accurate comprehension monitoring may further aid laypersons to decide when to stop searching for further information.

Apart from this, – especially in case of conflicting information – laypersons should evaluate information in terms of reliability, because severe quality deficits in medical information provisions have been documented repeatedly (for a review, see Eysenbach, Powell, Kuss, & Eun-Ryoung, 2002). Effectively dealing with such information would imply that laypersons critically evaluate information and apply strategies of epistemic metacognition (Hofer, 2004), such as corroborating information from different web sites, actively searching for the author of a website and finding out about his or her credentials as well as possible affiliations with third-party-sponsors. Thus, in an optimal case, laypersons would not only form a representation of contents during Internet research, but also add an additional level of text representation, which contains information about sources (Perfetti, Rouet & Britt, 1999). The resulting integrated representation of content information and source information would enable laypersons to take aspects such as the motive or the expertise of an author into account, when weighing up opposing standpoints.

It should be noted that the research situation described so far differs considerably from learning about scientific matters in school, where relevant information is typically presented in a condensed format, using a single textbook-like document, which is preselected by an authority (i.e. the teacher) and presents a reliable account of a given topic. Taking this discrepancy into consideration, it should not astonish that the competent use of scientific information on the WWW cannot be taken for granted, especially because the training of metacognitive skills is still rather a desideratum than a regular and explicit part of most school curricula (Wilson, 1999). Indeed, many studies revealed that the proficient use of metacognitive strategies seems to be rare, even among university students (Bannert, 2004; Gerjets & Scheiter, 2003; Rouet & Eme, 2002; Stadtler & Bromme, 2004). For instance, Stadtler and Bromme

(2004) found that laypersons searching the WWW for information on the topic of cholesterol showed only limited use of epistemic metacognition. Analysis of verbal protocols revealed that participants who were university students with little medical knowledge used few and often inadequate criteria to assess the credibility of a given website. They mainly relied on 'predictive judgments' (cfr. Hogarth, 1987) uttered before opening a website as well as superficial 'ad-hoc' impressions about the layout of a website uttered immediately after accessing a website. Furthermore, laypersons seldom searched for author information or explored possible affiliations with commercial sponsors. In a similar vein, in a study using focus groups Eysenbach and Köhler (2002) found that when explicitly asked, adult laypersons could indicate adequate criteria for assessing a website's reliability. However, they did not actually use them when conducting an Internet research on a medical topic.

So, it can be concluded that external support is needed to let laypersons fully benefit from the WWW as a resource of scientific health-related information. This led us to develop the computer tool *met.a.ware*, which is meant to support laypersons' research for health-related information on the Internet through fostering the use of metacognitive strategies. In *met.a.ware*, we focus on fostering two types of metacognitive processes, evaluating information and monitoring one's own comprehension. Thereby, we seek to support laypersons in gaining knowledge about contents as well as sources when dealing with complex and often controversial information of varying quality on the WWW. This decision is also in line with our previous research findings, which revealed substantial correlations between these two metacognitive processes and knowledge acquisition as well as the soundness of laypersons' credibility judgments in a research scenario similar to the one used in the current study (Stadtler & Bromme, 2004). In what follows, we shall describe the rationale underlying the design of *met.a.ware*, i.e. the prompting of metacognition and the ontological classification of information.

Prompting metacognition to foster strategy use and knowledge acquisition

Met.a.ware supports laypersons' Internet research using the method of metacognitive prompting. This term denotes support measures, which are embedded in the learning context and prompt the learner to execute specific metacognitive processes on a regular basis (Bannert, 2004). The method of prompting has been classified as belonging to the so-called indirect support measures as opposed to direct interventions, such as explicit metacognitive strategy trainings (Friedrich & Mandl, 1992), which are mainly used to teach basic metacognitive strategies to learners who essentially lack these skills.

The goal of metacognitive prompting is to focus the learners' attention towards their own cognition during the learning process (Brown, 1997). The repeated prompting is meant to elicit metacognitive processes, which learners wouldn't show spontaneously. Thus, we assume that metacognitive prompting is especially suitable in cases where learners are generally capable of executing metacognitive processes, but do not or only seldom apply these strategies spontaneously.

Empirical evidence for the effect of different types of prompts on metacognitive strategy use as well as knowledge acquisition has been found in a number of experimental studies. These studies have been carried out in different content domains such as mathematics (Gerjets, Scheiter & Schuh, 2005; Kramarski & Gutman, 2006), psychology (Bannert, 2003), software programming (Schmidt & Ford, 2003), physics (Veenman, Elshout & Busato, 1994), educational measurement (Kauffman, 2004) and biology (Lin & Lehman, 1999). For example, the learners in the study of Gerjets, Scheiter and Schuh (2005) were prompted to reflect on the structural similarities between tasks and worked-out examples in a learning environment about combinatorics. Furthermore, they were prompted to monitor their comprehension of the worked-out examples after each example had been studied. The results revealed that prompts encouraged learners to deal more intensively with the examples as documented in logfile-analyses. This led to an improved problem solving performance in learners with low prior knowledge, whereas learners with high prior knowledge did neither benefit from the prompting procedure, nor were they hampered. In a similar vein, Stark, Tyroller and Krause (2004) who had learners to give reasons for their selections in the learning process by typing them into 'prompting windows' found that particularly learners with low metacognitive prerequisites benefited from metacognitive prompts. There was no effect on different forms of knowledge acquisition for learners who already disposed of these strategies. Prompts thus compensated for the lack of metacognitive skillfulness.

Further evidence for the assumption that prompts indeed impact on the metacognitive processes of learners has been found in studies using think-aloud methodology (Bannert, 2004; Veenman, 1993; Veenman, Elshout & Busato, 1994). For instance, Veenman, Elshout and Busato (1994) reported that metacognitive prompting improved the working method of novices in a physics simulation environment. More precisely, metacognitive prompts improved the occurrence and quality of orientation activities in the planning phase of an experiment, systematical orderliness while executing a plan and monitoring one's comprehension.

To sum up, there is evidence that metacognitive prompting is an effective and economical means to foster the use of metacognitive strategies and knowledge acquisition in web-based learning environments. Thus far, metacognitive prompts have been employed in sheltered learning environments with limited and reliable information. We assume that metacognitive regulation becomes even more important when laypersons deal with scientific health-information on the WWW, because the structure of online health-information outlined above increases the situation's metacognitive complexity. However, to our knowledge there are no studies that have used metacognitive prompts to support processes of dealing with scientific (health-) information on the WWW so far.

Providing representational guidance through ontological classification.

The concept of representational guidance has been introduced into instructional psychology by Daniel Suthers (e.g., Suthers, 2001, Suthers & Hundhausen, 2003). Representational guidance refers to the way in which specific design features of external representations, which allow for the externalization of information, constrain and structure both individual learning and the discourse between learners in collaborative learning settings (Suthers, 2001).[1] For instance, the availability of specific ontological categories such as 'causes of a disease', 'consequences of a medical condition' or 'therapy' in a computer software to which learners can assign information about a medical topic enhances the salience of the selected categories, whereas categories which are not included become less salient. Suthers and Hundhausen (2003) argue that providing learners with categories, which are elements of an underlying ontology, also help to structure the task domain. Furthermore, learners will seek to find new information as instances of the salient categories. Thereby, the external representation may guide information search and help learners to internalize and eventually think in terms of the underlying ontology.

In *met.a.ware* ontological classification serves two main goals. Firstly, it is meant to enhance the structuredness of verbatim notes laypersons take during their Internet research and secondly, it serves to focus laypersons on relevant categories of the search topic.

Enhancing the structuredness of notes: Taking notes is a frequent and important strategy of students who are engaged in text comprehension. By taking notes readers select information from texts and thereby make it available for later use (Vidal-Abarca et al., 2004). However, in previous research we have found that laypersons made use of the possibility to take free-form notes while searching the WWW for medical information in a rather unsystematic way (Stadtler, 2006). In many cases laypersons simply listed facts in the order of their access during

Internet research. Notes did seldom contain references to the author of a document or inserted headings indicating a specific subtopic. This was true both for handwritten notes and for electronical verbatim notes, which were taken using the copy & paste-function. However, laypersons taking notes electronically, recorded much more information than laypersons taking handwritten notes. This even stresses the need to structure information externalized during an Internet research. Assigning copied information to different ontological categories may support laypersons in making sense out of one's notes after Internet research.

Focusing laypersons on relevant categories of the search topic: As outlined before, providing laypersons with a set of ontological categories may guide laypersons' search for information. Having a set of ontological categories, laypersons can be focused on aspects of the search topic, which are relevant to the search task. For instance, in previous research Stadtler (2006) found that laypersons searching the WWW for a medical topic without external support focused on practical information about therapeutical alternatives and consequences of a medical condition and paid only little attention to more fundamental information, e.g. about the causes of a medical condition. This may be an understandable approach for laypersons who do not pursue an explicit learning goal during Internet research. However, to be able to make an informed decision about a health-related question it is necessary to widen one's focus and also take more fundamental information about a medical condition into account (O'Connor, 1995). Therefore, it is an interesting question, whether it is possible to guide laypersons' Internet research in the desired direction by providing them with ontological categories.

Although the primary goal of including ontological categories in *met.a.ware* was to enhance the structuredness of notes and to provide guidance for laypersons' Internet research it should be noted that Suthers and Hundhausen (2003) also expect the availability of ontological categories to impact on knowledge acquisition in collaborative settings. Suthers and Hundhausen argue that externalized information becomes highly salient and initiates negotiations of meaning between discourse partners. The intensive elaboration of information should result in improved knowledge about the subject matter. In individual web search, however, it is questionable whether laypersons actually elaborate on externalized information. We rather assume, that laypersons will turn their attention back to the materials on the WWW, once they have externalized information. However, as there are arguments for both assumptions, it is currently an open empirical question, whether ontological classification also fosters learning in an individual setting.

Research questions and predictions

The main goal of this study was to examine the effects of the metacognitive tool *met.a.ware* on laypersons' Internet research for medical information. With *met.a.ware* we sought to foster the acquisition of knowledge about contents as well as sources. Following Perfetti, Rouet and Britt (1999), this integrated representation enables laypersons to take information such as the motive or the expertise of an author into account when elaborating on controversial information.

More precisely, we expected that prompting laypersons for comprehension monitoring would foster the acquisition of content knowledge (content knowledge hypothesis). We further predicted that prompting for evaluation of information would foster the acquisition of knowledge about sources (source knowledge hypothesis). Finally, we predicted that laypersons receiving evaluation prompts would produce more arguments to justify their credibility judgments (justification of credibility rating hypothesis).

A second aim of this study was to explore the effects of providing laypersons with ontological classification during Internet research. Our analyses focus on qualitative investigations of the structuredness of notes and the question of how far the ontological classification has guided laypersons' information search. We expected that laypersons working with ontological classification would use this means effectively to structure their notes. In contrast to this, we expected that laypersons working without ontological classification would structure their notes insufficiently. Given the exploratory nature of ontological classification in supporting Internet research, we did not specify any hypotheses with respect to the effect of ontological classification on knowledge acquisition.

Method

Participants

Participants were 120 undergraduate students at the University of Muenster, Germany (86 female). Participants' age ranged from 19 to 38 with an average of 23.81 (*SD* = 3.74). Prior knowledge about the subject matter of the Internet search (cholesterol) was tested before the Internet search to ensure their layperson status in this domain. Two students scored more than 12 out of 24 test points (50%), and were dropped from all further analyses. The remaining 118 participants scored an average of 4.76 (*SD* = 2.28) out of 24 possible test points.

Materials

Preselected websites. For their Internet search, laypersons were given a set of 15 preselected websites on the topic cholesterol. These web sites were accessible via a list of links presenting the URLs in alphabetical order. When selecting web sites, we took care to ensure that the resulting pool of information reflected the given heterogeneity of information available online in terms of information providers and their perspectives on this controversially discussed topic. Thus, we included web sites hosted by universities, companies from the food and pharmaceutical industries, and nutritionists or journals in the field of medicine. Web pages were displayed on a standard 17-in. computer screen and could be browsed using Microsoft Internet Explorer 6.

Task. Laypersons were instructed to conduct an Internet search to gather information for a fictitious friend. This friend had been diagnosed with a high level of cholesterol and needed to decide whether to consent to medical treatment. Laypersons were further informed that the friend was particularly interested in information about the possible consequences of a high level of cholesterol and about the pros and cons of medical treatment. The primary reason for choosing the topic cholesterol was the fact that it is a controversially discussed medical topic with a high dissemination on the Internet. Furthermore, it was expected that participants' prior knowledge on the topic was low.

The metacognitive tool met.a.ware. To support laypersons' Internet search we developed the metacognitive tool *met.a.ware*. *Met.a.ware* enables laypersons to store the information they have found on the WWW systematically (ontological classification). They do this by assigning information to six different tabs labeled with ontological categories underlying the topic cholesterol (see upper part of Figure 1). The categories are: (1) *function of cholesterol*, (2) *causes of a high level of cholesterol*, (3) *therapy of a high level of cholesterol*, (4) *consequences of a high level of cholesterol*, (5) *risk factors for developing a coronary heart disease* and (6) *information about threshold values*. Ontological categories were chosen based on a review of specialist literature and in cooperation with medical experts. The ontological categories form a comprehensive set of categories with respect to the topic cholesterol, however they are not exhaustive. This decision is based on the finding that providing learners with too many categories proved to confuse learners rather than to support them (Suthers & Hundhausen, 2003). As depicted in figure 1, laypersons paste their notes into blank text slots. The number of text slots available is not fixed, i.e. each time a user pastes information into a text slot, the system automatically creates a new blank text slot.

--- Insert Figure 1 about here ---

The metacognitive prompting in *met.a.ware* focuses on the processes of monitoring comprehension and evaluating information. As an evaluation prompt, laypersons are required to indicate the source of information each time they paste it into the system. They also have to rate the author's credentials, the author's interest to present biased information, as well as their confidence in the information on 5-point scales (see lower part of Figure 1). Thus, evaluation prompts mainly focus laypersons on the author of a document. As a monitoring prompt, laypersons are requested to assess how well they have comprehended information, how much they currently know about the specific aspect of cholesterol, and how much information they still need to search for regarding this aspect of cholesterol. Once again, they answer these items by rating them on 5-point scales (see right part of Figure 1).

All ratings are attached permanently to the specific contents and can be retrieved at all times by the user of *met.a.ware* during future Internet research. This means that the laypersons add an additional layer of meta-information to the contents stored in *met.a.ware*. Both ontological categories and metacognitive prompts were tailored to support laypersons in dealing with online information on the topic cholesterol for this study. The technical architecture of *met.a.ware*, however, allows for a flexible adaptation of the tool towards other domains of inquiry where a different set of ontological categories may be needed and where researchers might want to prompt other metacognitive processes.

Demographic variables (four items), *computer- and Internet experience* (four items), and *interest in the topic cholesterol* (four items) were measured with a questionnaire comprising a total of 12 self-developed items. *Need for cognition*, which is defined as the tendency to engage in and enjoy effortful cognitive endeavors (Cacioppo, Petty & Chuan, 1984), was measured with a German version of the original questionnaire devised by Bless et al. (1994). The measure, which comprises 16 Likert-type items, demonstrated good internal consistency (Cronbach's α = .79). Additionally, laypersons were required to indicate their subjectively experienced *time pressure* during Internet search using one Likert-type item.

Instruments measuring content knowledge. Laypersons were asked to complete a self-developed 24-item multiple-choice test to measure their *factual knowledge* about the topic cholesterol. The measure's internal consistency was good as indicated by Cronbach's α = .78. *Comprehension* of the subject matter was ascertained with four open questions each requiring laypersons to compose a written statement about one of three central concepts of the subject matter, namely the risk-factor concept, the development of threshold values and the concept of relative and absolute risk reduction.

Instrument measuring source knowledge. Source knowledge was assessed with a set of four multiple-choice-items requiring participants to recall facts about the source of a web site. These included information such as the author's position, his or her affiliations, or the presence of commercial sponsors. The questions had to be answered for each web site visited during Internet search.

Justification of credibility judgments. To measure laypersons' ability to justify their credibility judgments after Internet research laypersons were requested to rate three selected web sites in terms of credibility and subsequently give reasons for their judgments.

All measures were presented in a computerized form. Sample items for the measures used are given in the appendix. For the sake of completeness, it should be noted that we collected data on further variables (such as the epistemological beliefs of participants and their ability to relate arguments to sources), which we do not report in this paper, because they fall outside the scope of this article.

Design

Participants were randomly assigned to one of four experimental groups who worked with different versions of *met.a.ware* and two control groups. The availability of metacognitive prompts was varied systematically across the four experimental groups. Participants received either evaluation prompts (*evaluation* group, n = 20), monitoring prompts (*monitoring* group, n = 19), both types of prompts (*evaluation + monitoring* group, n = 20) or no prompts (*no prompts* group, n = 20). All *met.a.ware* conditions were provided with tabs for ontological classification.

Participants in the first control group took notes using paper and pencil (*paper and pencil* – control group, n = 19). Neither prompts nor ontological classification were available in the *paper and pencil* – control group. To examine the effects of ontological classification, a second control group was introduced. This group worked with a plain text window that allowed them to take notes electronically by copying and pasting information from the WWW into text slots, but provided neither ontological classification nor metacognitive prompts (*text window* – control group, *n* = 20).

Procedure

Data were collected in group sessions, with a maximum of seven participants per session. Prior to the Internet research, participants completed the questionnaires measuring their computer and Internet experience, as well as their interest in the topic cholesterol. In addition, participants' factual knowledge on the topic cholesterol was measured before Internet research. Participants in the *met.a.ware* conditions were then instructed how to work with *met.a.ware*. The standardized video-instruction included information about the meaning of the ontological categories as well as information about the function and value of the metacognitive strategies participants were intended to execute. Research in training metacognition has shown that this is an important element to ensure that participants act in line with the metacognitive support provided (Bannert, 2003).

After 40 min had elapsed, the experimenter requested participants to finish their Internet research. Searching time was limited in order to avoid time on task effects. Additionally, participants were asked to rate the perceived time pressure after they had finished Internet research. After their Internet research, laypersons repeated the same multiple-choice test on factual knowledge and answered the four open questions measuring comprehension of the subject matter. Additionally, knowledge about sources was assessed and participants were asked to rate the credibility of the three most appreciated websites and to produce arguments to justify their ratings. Notes taken during the Internet research were not available in the post-tests. The whole session lasted about 100 min on average.

Data analyses

Factual knowledge. Gain scores, i.e. the difference between factual knowledge posttest and pretest scores, were chosen as the unit of analysis, because they offer a better interpretation of change between pretest and posttest than an analysis of covariance (ANCOVA) with prior knowledge as a covariable does (Rogosa, 1988).

Comprehension scores. The written answers to the four open comprehension questions were scored qualitatively in terms of soundness and detailedness. Participants could reach a maximum of 12 points on the four comprehension questions. To determine the reliability of the scoring procedure, two judges rated 10% of the answers independently from each other. Interrater-reliability as determined according to the formula of Holsti (1969) proved to be high, *CR* = 94%.

Credibility judgments. To analyze the number and type of arguments laypersons produced to justify their credibility judgments, we used a self-developed categorization scheme. Laypersons' arguments were classified using a set of mutually exclusive categories, which were called *Layout* (e.g. the professionalism, availability of pop-up ads), *Content* (e.g. internal consistency, agreement with information from other websites), and *Source* (e.g. the author's expertise, her perceived motives) (cfr. Wittwer, Jucks & Bromme, 2004). Interrater-agreement for the coding process was high, *CR* = 95%.

Structuredness of notes. To examine the structuredness of notes in the conditions with ontological classification, we analyzed in how far laypersons classified their notes consistently with the predefined ontological categories. To do so, we rated the degree of consistency for each unit of pasted information, i.e. for each text slot, as follows.

1. The information pasted relates exclusively to the category chosen (level 1).
2. The text slot contains both information consistent with the given category and information, which is not consistent with the category chosen (level 2).
3. The text slot contains only information inconsistent with the category chosen (level 3).

To determine the reliability of the coding procedure, two judges rated the notes of 10% of the participants working with ontological classification independently from each other. Inter-rater-reliability according to Holsti (1969) was high, *CR* = 92%. Disagreements between raters were resolved through discussion.

In contrast, laypersons in the conditions that worked without ontological classification had to structure their notes themselves. One way to accomplish this was to insert headings, which relate the pasted information to a particular topic. This is why we analyzed the notes taken in the *text window* – control group and the *paper and pencil* – control group with respect to the number of headings inserted as well as the proportion of pasted information that was preceded by a heading. For each text slot in the *text window* – control group we decided whether (a) it contained only information that was preceded by a heading, (b) contained both information that was preceded by a heading and information that was not preceded by a heading or (c) contained only information that was not preceded by a heading. In the *paper and pencil* – control group we analyzed the handwritten notes with respect to the amount of words that were preceded by a heading.

Effects of ontological classification on information search. To examine in how far the availability of ontological categories guided laypersons' search for information, we analyzed the notes laypersons took during Internet research. Notes taken can be considered as a good indicator of the contents laypersons are dealing with, particularly because participants had been encouraged to store information they consider important in *met.a.ware*. Thus, we compared the *no prompts* group and the *text window* – control group with respect to the amount of notes they took relating to the different ontological categories. For the *no prompts* group the number of words in each category was adjusted taking into account the erroneous classifications made by laypersons. Thus, information was recategorized, so that it was consistent with the predefined meaning of the categories.

Because the *text window* – control group worked without ontological classification, we had to categorize the notes taken in this condition with respect to their content focus, first. In a rating procedure we decided for each text slot to which ontological category the inserted information belongs. To determine the procedure's reliability, two independent raters judged 15% of the text slots of the *text window* – control group. Interrater-reliability according to Holsti (1969) was high, *CR* = 93%. Disagreements between raters were resolved through discussion.

Statistical analyses. We conducted planned contrasts between each of the experimental groups working with *met.a.ware* and the *paper and pencil* – control group to test all a priori specified hypotheses in this paper. Thereby we wanted to take a theory-driven approach, which bears the advantage of having a greater statistical power than post-hoc comparisons conducted in reaction to a significant omnibus F-test in an analysis of variance (Hays, 1988; Rosenthal & Rosnow, 2000). This is accomplished by reducing the probability that an existing effect is obscured by variation that is not of theoretical interest (Weinfield, Sroufe & Egeland, 2000). Because planned contrasts do not require a significant omnibus F-Test as a precondition, no omnibus F-tests are reported when planned contrasts were conducted (Czienskowski, 1996).

T-tests and a MANOVA were calculated to analyze the effects of ontological classification on knowledge acquisition and information search. An alpha-level of .05 was chosen for all statistical tests unless otherwise indicated.

Results

Covariables. Separate ANOVAs were conducted for each of the four covariables Internet-/computer experience, interest in the topic, need for cognition and time pressure to find out whether there were any differences between groups on these variables. Because we did not expect to find any differences, an alpha-level of .20 was considered as statistically significant. However, none of the ANOVAs yielded a significant result (all Fs < 1.36, *ns*) showing that groups did not differ on any of the covariables. As a consequence, the covariables were dropped from all further analyses.

Content Knowledge. Mean gain scores as well as standard deviations for the six groups are presented in Table 1. Planned contrasts between each experimental group and the *paper and pencil* – control group showed a significant difference between the *monitoring* group and the control group, $F(1, 93) = 10.35$, $p < .01$, $\eta^2_{part} = .10$, and a trend between the *evaluation + monitoring* group and the control group, $F(1, 93) = 3.75$, $p = .056$, $\eta^2_{part} = .04$.

--- Insert Table 1 about here ---

As expected, no significant difference could be found between controls and the *no prompts* group, $F(1,93) = .70$, *ns*. However, we also obtained a significant difference between controls and the *evaluation* group, which was not predicted by our hypothesis, $F(1, 93) = 4.69$, $p < .05$, $\eta^2_{part} = .05$. Finally, we did not find an effect of ontological classification on the acquisition of factual knowledge. A *t*-test between the *no prompts* group and the *text window* – control group did not yield a significant difference, $t(38) = -.62$, *ns*.

Means and standard deviations with respect to the comprehension of the subject matter are shown in Table 2.

--- Insert Table 2 about here ---

Planned contrasts revealed a marginally significant trend between the *monitoring* group and the paper and pencil – control group, $F(1, 93) = 3.36$, $p < .10$, $\eta^2_{part} = .04$. As predicted, no significant differences were found between the *no prompts* group and the paper and pencil –control group, $F(1, 93) = .25$, *ns*, or between the *evaluation* group and the paper and pencil – control group, $F(1, 93) = 1.09$, *ns*. However, we did not obtain a significant difference between the *evaluation + monitoring* group and the paper and pencil – control group either, which was not predicted by our hypothesis, $F(1, 93) = 1.13$, *ns*. Finally, we did not find an effect of ontological classification on the comprehension of the subject matter. A *t*-test between the *no prompts* group and the *text window* – control group did not yield a significant difference, $t(38) = -.63$, *ns*.

Source Knowledge. Table 3 depicts the mean percentage of correct answers on the source test. Percentages of correct answers were used instead of the total number of correct items, because participants were free to choose which web sites they visited. As a result, not all participants accessed all 15 web sites. To test our hypothesis that evaluation prompts would promote the acquisition of source knowledge, we performed planned contrasts between each of the four experimental groups working with *met.a.ware* and the *paper and pencil* – control group. In line with our hypothesis, the *evaluation* group, $F(1, 93) = 9.48$, $p < .01$, $\eta^2_{part} = .09$, and the *evaluation + monitoring* group, $F(1, 93) = 8.87$, $p < .01$, $\eta^2_{part} = .09$, significantly outperformed the *paper and pencil* – control group. Furthermore, no significant differences were found between the *paper and pencil* – control group and either the *no prompts* group, $F(1, 93) = 1.6$, *ns*, or the *monitoring* group, $F(1, 93) = .06$, *ns*.

---Insert Table 3 about here ---

Justification of credibility judgments. Using multivariate planned contrasts, each of the four experimental conditions working with *met.a.ware* was compared with the *paper and pencil* – control group with respect to the number of arguments in each of the three categories *Content*, *Layout* and *Source*. As expected, neither the *no prompts* group nor the *monitoring* group differed significantly from the *paper and pencil* – control group, all *F*s < .58, *ns*. However, contrary to our hypothesis, planned contrasts did not reveal a significant difference between the *evaluation* group and controls and the *evaluation + monitoring* group and controls, either, all *F*s < 2.06, *ns*. Thus, laypersons in the experimental conditions did not differ from the *paper and pencil* – control group with respect to the overall number of arguments produced.

However, univariate contrasts revealed that members of the *evaluation* group produced significantly more arguments classified as belonging to the category *Source* than the *paper and pencil* – control group, $F(1, 93) = 4.33$, $p < .05$, $\eta^2_{part} = .04$. The *evaluation + monitoring* group showed a tendency towards more arguments in the category *Source* than the *paper and pencil* – control group, $F(1, 93) = 2.96$, $p = .089$, $\eta^2_{part} = .03$. As expected, no significant differences could be found between the monitoring group and the *paper and pencil* – control group and the *no prompts* group and the *paper and pencil* – control group with respect to the number of arguments in the category *Source*, all *F*s < .39, *ns*. Also, none of the planned comparisons between the experimental groups and the controls with respect to the categories *Content* and *Layout* yielded any significant differences, all *F*s < 1.38, *ns*.

Thus, the results confirm our hypothesis only partially. Laypersons receiving evaluation prompts produced more arguments focusing on the author of a website than controls. However, they did not produce more arguments with regard to content and the website's layout.

Effects of ontological classification

Structure of laypersons' notes. The results with respect to the structuredness of notes laypersons took in the four experimental conditions working with ontological classification in *met.a.ware* are depicted in Table 4. Laypersons filled a total of 1459 text slots with information from the WWW. On average 73% (*SD* = 12.80) of text slots contained exclusively information consistent with the respective category (level 1). Furthermore, an average of 19% (*SD* = 11.17) of the text slots were classified as containing both information, which is consistent, and information, which is not consistent with the selected category (level 2). Finally, on average 9% (*SD* = 9.33) of text slots contained only information that was not consistent with the category chosen (level 3).[2]

---Insert Table 4 about here ---

Notes were further analyzed to examine the reasons why laypersons in some cases classified information not in line with the respective category. These analyses suggest three different types of deviation from the intended classification.

1. Laypersons insert whole paragraphs that relate to more than one ontological category.

Using the copy & paste procedure it is easy to store large parts of text in computer-tools such as *met.a.ware*. Thus, in some cases laypersons pasted whole paragraphs into *met.a.ware* without breaking them down into smaller units of meaning, which relate to different aspects of the topic. This can be a reasonable procedure especially when laypersons want to preserve the meaning of a whole paragraph, for instance when the paragraph relates information about the causes of a high blood level of cholesterol to therapeutical alternatives. Apart from this, the procedure may be appealing because it is less time-consuming to copy and paste whole paragraphs of information.

2. Laypersons insert information which is not covered by the ontological categories available

As outlined earlier, the set of ontological categories provided in *met.a.ware* was designed to be comprehensive but not exhaustive. Therefore, laypersons could paste information into the system, which could not be properly assigned to any of the given categories. This accounted for 26% of the cases in which text slots contained only information, which did not relate to the respective category (level 3). Related to the total amount of text slots used by laypersons, this type of classification is rather rare (2%).

3. Laypersons interpret a category differently from its intended meaning

Although laypersons were familiarized with the meaning of the ontological categories in the instruction, they sometimes interpreted a category differently from its intended meaning. This accounted for 74% of the cases in which text slots contained only information, which did not relate to the respective category (level 3) and for 6% of all classifications. Erroneous interpretation of a category may be due to the fact that only short versions of the categories' labels were displayed on the tabs, which may have been misleading in some cases (The full category name was displayed when the mouse lingered on the tab). For instance, the most frequent classification error of this type was that laypersons assigned information about undesirable medical side effects to the category labeled 'risk factors', which by definition meant 'risk factors for developing a coronary heart disease'. To summarize, the large majority of classifications is consistent with the intended meaning of the predefined categories. Thus, laypersons successfully used the availability of ontological categories to structure their verbatim notes taken from the Internet.

In contrast, participants in the *text window* – control group who worked without ontological classification mainly relied on the inherent structure of the information copied from the Internet, as they did not actively structure their notes by inserting self-written headings (only one participant inserted a self-written heading). The mean number of headings per 100 words was rather low (M = .49; SD = .37). This led to the fact that on average 65% (SD = 25.31) of text slots in the *text window* – control group contained only information that was not preceded by a heading and thus could not be easily recognized as belonging to a certain topic. Furthermore, on average 13% (SD = 25.01) of the text slots contained both information that was preceded by a heading and information that was not preceded by a heading. Finally, on average 22% (SD = 19.98) of the text slots contained only information that was preceded by a heading and thus could be easily recognized as belonging to a certain topic.

In the *paper and pencil* – control group, in which laypersons had to write down their own headings, the mean number of headings was 4.71 (SD = 3.86) per 100 words. On average 54% (SD = 36.02) of the words contained in the note sheets were not preceded by a heading and thus could not be easily recognized as belonging to a certain aspect of the topic cholesterol. Thus, neither laypersons in the *text window* – control group nor in the *paper and pencil* – control group structured their notes consequently by means of inserting self-written or electronically copied headings.

Effects of ontological classification on information search. The average number of words pasted into *met.a.ware* across all ontological categories was higher in the *no prompts* group working with ontological classification (M = 367.48; SD = 611.59) than in the *text window* – control group working without ontological classification (M = 203.49; SD = 357.02), as revealed by a MANOVA, $F(6, 33) = 6.96$, $p < .001$, $\eta^2_{part} = .56$.

--- Insert Table 5 about here ---

The univariate tests depicted in Table 5 show that the multivariate effect can be attributed to the fact that laypersons in the *no prompts* – group pasted significantly more information into the categories *'function of cholesterol'*, *'causes of a high level of cholesterol'*, *'threshold values'* and “*risk factors for a coronary heart disease*”. It can be concluded that the salience of ontological categories led laypersons to seek new information about these aspects of the topic cholesterol and pasted them into *met.a.ware*. No significant differences could be found however with respect to the categories *'therapy'* and *'consequences of a high level of cholesterol'*. It should be noted that these categories were introduced in the instruction as search goals. Therefore, we have focused laypersons on these categories, which as a result may have evened out effects of making categories salient through the provision of ontological categories.

To sum up, the results reveal that ontological categories had an impact on laypersons' information search as indicated by the notes stored in *met.a.ware*. This was true for ontological categories not introduced as search goals in the instruction.

Discussion

The goal of the present study was to test the possibility of supporting laypersons' Internet research for medical information with the metacognitive tool *met.a.ware*. Drawing on findings from the literature on metacognition we hypothesized that applying the metacognitive strategies of monitoring and evaluation would help laypersons to deal with medical information on the WWW successfully. Furthermore, we wanted to examine the effects of providing laypersons with representational guidance through ontological classification on notes taken during Internet research.

Our hypotheses concerning the acquisition of knowledge about contents were supported partially. Compared with controls, participants receiving monitoring prompts acquired significantly more factual knowledge on the topic cholesterol. This result is in line with findings from the literature on text comprehension research, which has shown that an accurate comprehension monitoring contributes substantially to learning from text (e.g., Thiede, Anderson & Therriault, 2003). The repeated prompting assisted laypersons in detecting comprehension failures and inconsistencies in their text representation and thereby enabled them to regulate their information processing accordingly, for instance by re-reading difficult parts of the text or slowing down their reading speed. Furthermore, laypersons were repeatedly prompted to determine how much information they still needed with respect to the ontological categories specified in *met.a.ware.* Thereby laypersons were enabled to regulate their subsequent search process according to their information needs, which may be of particular importance in search scenarios with limited time resources and unmanageable masses of information being available.

However, results with respect to the comprehension of the subject matter were less conclusive. Laypersons receiving only monitoring prompts performed slightly better on a comprehension test. Nonetheless, the *evaluation + monitoring* group did not differ significantly from controls in their comprehension of the subject matter. This may be because developing a high-level understanding within the given time limit of 40 min had been a rather demanding task for laypersons who initially had very little knowledge about the subject matter. Results of previous research (Stadtler & Bromme, 2004) has shown that when laypersons were confronted with a comparable scenario, their first goal was to gather rather basic information such as what is an acceptable level of cholesterol in the blood or which diseases may result from high cholesterol levels. After they had gathered some basic knowledge on the topic, they then moved on to explore more complex issues such as the relationship between cholesterol and other factors causing a coronary heart disease. These practical constraints may well explain why the mean scores on the comprehension task in the present study did not differ between groups.

Alternatively, the results may not be due to a simple lack of search time, but rather reflect the specific 'learning goal' laypersons pursue during Internet research. Because laypersons do not want to become experts in the area of their inquiry, they may be satisfied with a metonymic, i.e. partial understanding of concepts and relations (Bromme, Jucks & Runde, 2005). This may include basic knowledge about facts as measured in the factual knowledge test, but not a deeper understanding of more complex issues such as the relationship between different risk factors for developing a coronary heart disease. In this case, enhancing the search time would not result in deeper understanding of the subject matter. One would rather expect that laypersons terminate their web search process after a subjectively sufficient level of understanding had been achieved. Further research is needed to address this issue and clarify which explanation accounts for the current findings.

As Hofer (2004) has suggested, evaluating the credibility of a document is particularly crucial when dealing with information on the WWW, because single documents may contain faulty or biased information and not always provide a reliable account. This is why we gave laypersons evaluation prompts requiring them to rate information in terms of its credibility. Results on testing source knowledge revealed that evaluation prompts improved performance: members of both the *evaluation* group and the *evaluation + monitoring* group recalled more information about sources than the *paper and pencil* – control group. With regard to the explicitly given arguments for credibility judgments the *evaluation* group and the *evaluation + monitoring* group outperformed the *paper and pencil* – control group. However, this effect was only observed for arguments relating to the author of a document. No differences between conditions were found with respect to the number of arguments relating to the quality of information or the layout of the website. Taking into account that the evaluation prompts mainly focused laypersons' attention to the source of a document, this is a plausible result. The prompting procedure obviously did not trigger comprehensive processes of information evaluation but impacted in a more specific way on the evaluation behaviour. This result implies that the concrete selection of prompts is of particular importance in the design of metacognitive support measures. At the same time, this raises the question whether it would be equally possible to design metacognitive prompts, which focus laypersons on other indicators of a websites' reliability, such as the up-to-dateness or the consistency of information. Further research is needed to address this issue.

Finally, ontological classification was incorporated in *met.a.ware* to help laypersons structure their notes during Internet research. Analyses revealed that laypersons used this means efficiently. The large majority of notes were classified in accordance with the predefined categories. In contrast, laypersons working without ontological classification structured their notes insufficiently. This was particularly true for the *text window* – control group. The relative ease of copying and pasting information electronically led laypersons to store a great deal of information in different text slots. However, laypersons relied on the inherent structure of pasted information. Thus, a considerable proportion of information was not preceded by a heading indicating to which specific aspect of the topic the pasted information related. This result is in line with previous research (Stadtler, 2006) in which laypersons stored information in blank text windows that did not provide any text slots. Therefore, it is unlikely that the low degree of structuring electronical notes with headings observed in the current study is due to the sheer availability of text slots, which themselves could be used to structure information. Taken together, the results once again underline the need to assist laypersons in structuring their notes taken during Internet research. This can be accomplished by providing laypersons with ontological classification. Future research could focus on the question of how far the structuredness of notes supports laypersons in extracting information from their notes after their Internet research. Because we decided to make notes not available on the post-tests, we were not able to answer that question in our study.

Furthermore, we found that laypersons stored more information in *met.a.ware* when provided with ontological classification. This difference could be attributed to those categories, which were not introduced as search goals in the instruction. This finding indicates that the salience of ontological categories stimulated laypersons to seek new elements of the categories at hand, which would be predicted by representational guidance theory (Suthers & Hundhausen, 2003). However, it is still unclear in which way the availability of ontological categories actually affects the cognitive processing of information. Does the provision of ontological categories really lead laypersons to elaborate more intensively on information from salient categories? Or did the representational guidance lead laypersons to fill salient text-slots with information without really scrutinizing them. The present study was not designed to answer this question. However, the fact that we did not find an effect of ontological classification on knowledge acquisition does not bolster the idea that ontological classification leads to deep information processing. Further studies using think-aloud methodology are needed to shed some light on the effect of representational guidance on the cognitive processing of information during Internet research.

With respect to the generalization of our findings it should be noted, that for the purpose of the present study, *met.a.ware* has been tailored towards the needs of laypersons that were university students, familiar with the Internet and had developed at least some basic Internet search skills. Apart from this, it is likely that laypersons at this educational level had a general command of metacognitive strategies that could be triggered successfully by the prompting procedure in *met.a.ware*. Taking this into consideration, the results from the present study should not be generalized without care to the Internet research of laypersons that differ substantially with respect to their Internet experience and (meta-)cognitive prerequisites. In this case, it would be advisable to combine *met.a.ware* with a strategy training, in which basic search skills as well as metacognitive strategies are taught explicitly before being practised with *met.a.ware* in the context of Internet searching.

In sum, the present study provides evidence that *met.a.ware* is an economic means to support laypersons' Internet research for medical information. Using metacognitive prompts for monitoring and evaluation increased knowledge on both contents and sources. Ontological classification helped to structure notes taken during Internet research and focused laypersons on salient ontological categories. Apart from their practical value the current findings also have theoretical implications as they underline the role of metacognition in dealing with complex scientific information on the Internet. Whereas the importance of metacognition has already been emphasized with respect to comprehending traditional print texts and hypertexts, the role of metacognition in dealing with scientific online-information has not been addressed so far. Our study fills this void and provides an optimistic picture as to the support of laypersons through fostering metacognitive strategy use. The results open up the possibility of designing larger intervention programs, which combine the use of *met.a.ware* with a more direct training of metacognitive skills to support laypersons in dealing with complex scientific information on the Internet.

References

Baker, L., & Brown, A. L. (1984). Metacognitive skills and reading. In P. D. Pearson (Ed.), *Handbook of reading research* (pp. 353-394). New York: Longman.

Bannert, M. (2003). Effekte metakognitiver Lernhilfen auf den Wissenserwerb in vernetzten Lernumgebungen [Effects of metacognitive learning aids in networked learning environments]. *Zeitschrift für Pädagogische Psychologie, 17*, 13-25.

Bannert, M. (2004). Designing metacognitive support for hypermedia learning. In H. Niegemann, D. Leutner & R. Brünken (Eds.), *Instructional Design for Multimedia-Learning.* (pp. 19-30). Münster: Waxmann.

Bless, H., Wänke, M., Bohner, G., Fellhauer, R. F., & Schwarz, N. (1994). Need for Cognition: Eine Skala zur Erfassung von Engagement und Freude bei Denkaufgaben. [Need for cognition: A scale measuring engagement and happiness in cognitive tasks]. *Zeitschrift für Sozialpsychologie, 25*, 147-154.

Britt, M. A., & Aglinskas, C. (2002). Improving student's ability to identify and use source information. *Cognition and instruction, 20*, 485-522.

Bromme, R., Jucks, R., & Runde, A. (2005). Barriers and biases in computer-mediated expert-layperson-communication. In R. Bromme, F. W. Hesse, & H. Spada (Eds.), *Barriers, biases and opportunities of communication and cooperation with computers- and how they may be overcome* (pp. 89-118). New York: Springer.

Brown, A. L. (1997). Transforming schools into communities of thinking and learning about serious matters. *American Psychologist, 52*, 399-413.

Cacioppo, J. T., Petty, R. E., & Chuan, F. K. (1984). The Efficient Assessment of Need for Cognition. *Journal of Personality Assessment*, 48, 306-307.

Czienskowski, U. (1996). *Wissenschaftliche Experimente: Planung, Auswertung, Interpretation*. [Scientific experiments: planning, analysis and interpretation] Weinheim: Psychologie Verlags Union.

Eveland, W. P., & Dunwoody, S. (2000). Examining information processing on the World Wide Web using think aloud protocols. *Media Psychology, 2*, 219-244.

Eysenbach, G., & Köhler, C. (2002). How do consumers search for and appraise health information on the World Wide Web? Qualitative studies using focus groups, usability tests, and in-depth interviews. *British Medical Journal*, 324, 573-577.

Eysenbach, G., Powell, J., Kuss, O., & Eun-Ryoung, S. (2002). Empirical studies assessing the quality of health information for consumers on the World Wide Web. A systematic review. *JAMA*, *287*, 2691-2700.

Fox, S., & Fallows, D. (2003). *Internet health resources*. Retrieved February 7th, 2006, from http://www. pewinternet.org.

Friedrich, H. F., & Mandl, H. (1992). Lern- und Denkstrategien - ein Problemaufriss. In H. Mandl & H. F. Friedrich (Eds.), *Lern- und Denkstrategien. Analyse und Intervention* [Learning and thinking strategies. Analysis and intervention] (pp. 3-54). Göttingen: Hogrefe.

Gerjets, P., & Scheiter, K. (2003). Goal configurations and processing strategies as moderators between instructional design and cognitive load: Evidence from hypertext-based instruction. *Educational Psychologist*, 38, 33-41.

Gerjets, P., Scheiter, S., & Schuh, J. (2005). Instruktionale Unterstützung beim Fertigkeitserwerb aus Beispielen in hypertextbasierten Lernumgebungen [Instructional Support for Skill Acquisition from Examples in Hypermedia-Based Learning Environments]. *Zeitschrift für Pädagogische Psychologie, 19*, 23-38.

Hays, W. L. (1988). *Statistics* (4th ed.). Fort Worth, TX: Holt, Rinehart & Winston Inc.

Hogarth, R. M. (1987). *Judgment and choice. The psychology of decision.* New York: John Wiley & Sons.

Hölscher, C. (2002). *Die Rolle des Wissens im Internet. Gezielt suchen und auswählen* [The role of knowledge on the Internet]. Stuttgart: Klett-Cotta.

Holsti, O. R. (1969). *Content Analysis for the Social Sciences and Humanities*. USA: Addison Wesley Publishing Company.

Hofer, B. K. (2004). Epistemological understanding as a metacognitive process: Thinking aloud during online searching. *Educational Psychologist, 39*, 43-55.

Kauffman, D. F. (2004). Self-regulated learning in Web-based environments: Instructional tools designed to facilitate cognitive strategy use, metacognitive processing, and motivational beliefs. *Journal of educational computing research*, *30*, 139 - 161.

Klatt, R., Gavriilidis, K., Kleinsimlinghaus, K., & Feldman, M. (2001). *Nutzung elektronischer wissenschaftlicher Information in der Hochschulsausbildung* [Utilization of electronical scientific information in university teaching]. Dortmund: Sozialforschungsstelle Dortmund Landesinstitut.

Kramarski, B., & Gutman, M. (2006). How can self-regulated learning be supported in mathematical E-learning environments? *Journal of Computer Assisted Learning, 22,* 24-33.

Lazonder, A. W. (2000). Exploring novice users' training needs in searching information on the World Wide Web. *Journal of Computer Assisted Learning, 16*, 326-335.

Lin, X., & Lehman, J. D. (1999). Supporting learning of variable control in a computer-based biology environment: Effects of prompting college students to reflect on their own thinking. *Journal of Research in Science Teaching, 36*, 837-858.

Morahan-Martin, J. M. (2004). How Internet Users Find, Evaluate, and Use Online Health Information: A Cross-Cultural Review. *CyberPsychology & Behavior, 7*(5), 497-510.

O'Connor, A. M. (1995). Validation of a decisional conflict scale. *Medical decision making, 15,* 15-30.

Perfetti, C. A., Rouet, J.-F., & Britt, M. A. (1999). Toward a theory of documents representation. In H. v. Oostendorp & S. R. Goldman (Eds.), *The Construction of Mental Representations During Reading* (pp. 99-122). Mahwah, NJ: Lawrence Erlbaum Associates.

Rogosa, D. (1988). Myths about longitudinal research. In K. W. Schaie, R. T. Campbell, W. M. Meredith & S. C. Rawlings (Eds.), *Methodological issues in aging research* (pp. 171-209). New York, NY: Springer Publishing Company.

Rosenthal, R., & Rosnow, R. L. (1985). *Contrast analysis: Focused comparisons in the analysis of variance*. Cambridge, New York: Cambridge University Press.

Rouet, J.-F., & Eme, E. (2002). The role of metatextual knowledge in text comprehension. In P. Chambres & M. Izaute (Eds.), *Metacognition: Process, function and use* (pp. 121-133). Kluwer Academic Publishers.

Schmidt, A. M., & Ford, J. K. (2003). Learning within a learner control training environment: The interactive effects of goal orientation and metacognitive instruction on learning outcomes. *Personnel Psychology, 56*, 405-429.

Stadtler, M. (2006). *Auf der Suche nach medizinischen Fachinformationen. Metakognitionen bei der Internetrecherche von Laien* [Searching for medical information. Metacognition in laypersons' web search]. Münster: Waxmann.

Stadtler, M., & Bromme, R. (2004). Laypersons searching for medical information on the web: The role of metacognition. In K. Forbus, D. Gentner, & T. Regier (Eds.). *Proceedings of the 26th Annual Conference of the Cognitive Science Society*, (p. 1638), Mahwah, NJ: Erlbaum.

Stark, R., Tyroller, M., & Krause, U.-M. (2004). Effects of a metacognitive prompting procedure in the context of a computer-based learning environment: Practical relevance and explanation by metacognitive and motivational processes. In P. Gerjets, P. A. Kirschner, J. Elen & R. Joiner (Eds.), *Instructional design for effective and enjoyable computer-supported learning* (CD-ROM). Tübingen: Knowledge Media Research Center.

Suthers, D. (2001). Towards a Systematic Study of Representational Guidance for Collaborative Learning Discourse. Journal of Universal Computer Science, *7*. (Electronic publication: http://www.jucs.org/jucs_7_3/towards_a_systematic_study).

Suthers, D., & Hundhausen, C. D. (2003). An experimental study of the effects of representational guidance on collaborative learning processes. *The Journal of the Learning Sciences, 12*, 183-218.

Thiede, K. W., Anderson, M. C. M., & Therriault, D. (2003). Accuracy of Metacognitive Monitoring Affects Learning of Texts. *Journal of Educational Psychology, 95,* 66-73.

Veenman, M. V. J. (1993). *Intellectual ability and metacognitive skill: determinants of discovery learning in computerized learning environments*. Amsterdam: University of Amsterdam.

Veenman, M. V., Elshout, J. J., & Busato, V. V. (1994). Metacognitive mediation in learning with computer-based simulations. *Computers in Human Behavior, 10,* 93-106.

Veenman, M. V. J., Wilhelm, P., & Beishuizen, J. J. (2004). The relation between intellectual and metacognitive skills from a developmental perspective. *Learning & Instruction, 14,* 89-109.

Vidal-Abarca, E., Gil, L., Cerdán, R., Gilabert, R., Martínez, T., & Romero, A. (2004, September). Integrating scientific information from multiple documents: The role of note taking and strategies. *Paper presented at the EARLI SIG 2 Meeting, Comprehension of text and graphics: basic and applied issues*, Valencia, Spain.

Weinfield, N. S., Sroufe, L. A., & Egeland, B. (2000). Attachment from Infancy to Early Adulthood in a High-Risk Sample: Continuity, Discontinuity, and Their Correlates. *Child Development, 71,* 695-702.

Wilson, J. (1999). *Defining Metacognition: A step towards recognising metacognition as a worthwhile part of the curriculum.* Paper presented at the AARE Conference, Melbourne.

Wittwer, J., Bromme, R., & Jucks, R. (2004). Kann man dem Internet trauen, wenn es um die Gesundheit geht? Die Glaubwürdigkeitsbeurteilung medizinischer Fachinformationen im Internet durch Laien. [Is the Internet a reliable source of health-related information? Laypersons' credibility judgements about medical information on the Internet]. *Zeitschrift für Medienpsychologie, 2*, 48-56.

Appendix:

Sample question for the multiple choice test on factual knowledge:

For what purpose does our body need cholesterol?

- To transport oxygen in the blood
- To build cell membranes
- To break down carbohydrates
- To synthesize vitamin C
- Our body doesn't need cholesterol
- I don't know

Note: Each of the 24 items consisted of four distractors, one attractor and one “I don't know”-option, which was included to reduce the effect of guessing.

Sample questions for the open comprehension questions:

- Why do some researchers criticize the reduction of threshold values for the blood cholesterol level?
- Is it reasonable to assess the individual risk for coronary heart disease solely on the basis of the blood cholesterol level?

Sample questions for the need for cognition questionnaire:

- “I really enjoy a task that involves coming up with new solutions to problems.”

- "I would prefer complex to simple problems"

Note: Agreement was rated on a 7-point-Likert scale, in which 1 was labeled "totally agree" and 7 was labeled "totally disagree".

Sample questions for the multiple-choice test on knowledge about source information

What is the profession of the author of the information on this web site?

- Physician
- Scientist
- Nutrionist
- Journalist
- Layperson
- There are no information about the author available on the web site
- I don't know

Is there any advertisement for a cholesterol-related product on the web site?

- Yes
- No
- I don't know

Acknowledgements

The research reported in this paper has been supported by a scholarship from the *Deutsche Forschungsgemeinschaft* (DFG). The scholarship has been granted to the first author in the context of the Virtual Ph.D. Program "Knowledge Acquisition and Knowledge Exchange with New Media".

Footnotes

[1] In his work, Suthers focuses on the role of representational guidance in *collaborative* learning settings while acknowledging that representational tools can also guide *individual* learning interactions (Suthers, 2001). The following brief description of the theoretical rationale underlying the representational guidance effect focuses on the mechanisms working on an individual level. For a comprehensive description of the framework, see Suthers and Hundhausen (2003).

[2] Owing to rounding, percentages do not sum up to 100%.

Table 1

Mean Gain Scores and Standard Deviations for Factual Knowledge

	M	*SD*
Monitoring	11.00	3.04
Evaluation + Monitoring	9.95	2.72
Evaluation	10.20	3.71
No prompts	8.75	4.36
Paper and pencil – control group	7.84	2.83
Text window – control group	8.05	2.63

Table 2

Mean Scores and Standard Deviations for Comprehension

	M	*SD*
Monitoring	6.33	1.99
Evaluation + Monitoring	5.80	2.15
Evaluation	5.79	2.30
No prompts	5.43	2.19
Paper and pencil – control group	5.09	1.69
Text window – control group	5.00	2.10

Note: A maximum score of 12 points could be reached on the comprehension test

Table 3

Mean Percentage Correct on Source Test

	M	*SD*
Evaluation	45.33	10.80
Evaluation + Monitoring	44.92	13.82
Monitoring	33.83	12.60
No prompts	37.97	13.43
Paper and pencil – control group	32.79	12.72

Table 4

Degree of Structuredness of Notes in Met.a.Ware

Degree of structuredness (*dos*)	Mean percent-age of text slots	*SD*
level 1	73%	12.8
level 2	19%	11.17
level 3	9%	9.32

Note: The number of text slots has been collapsed across the four groups working with onto-logical classification.

Table 5

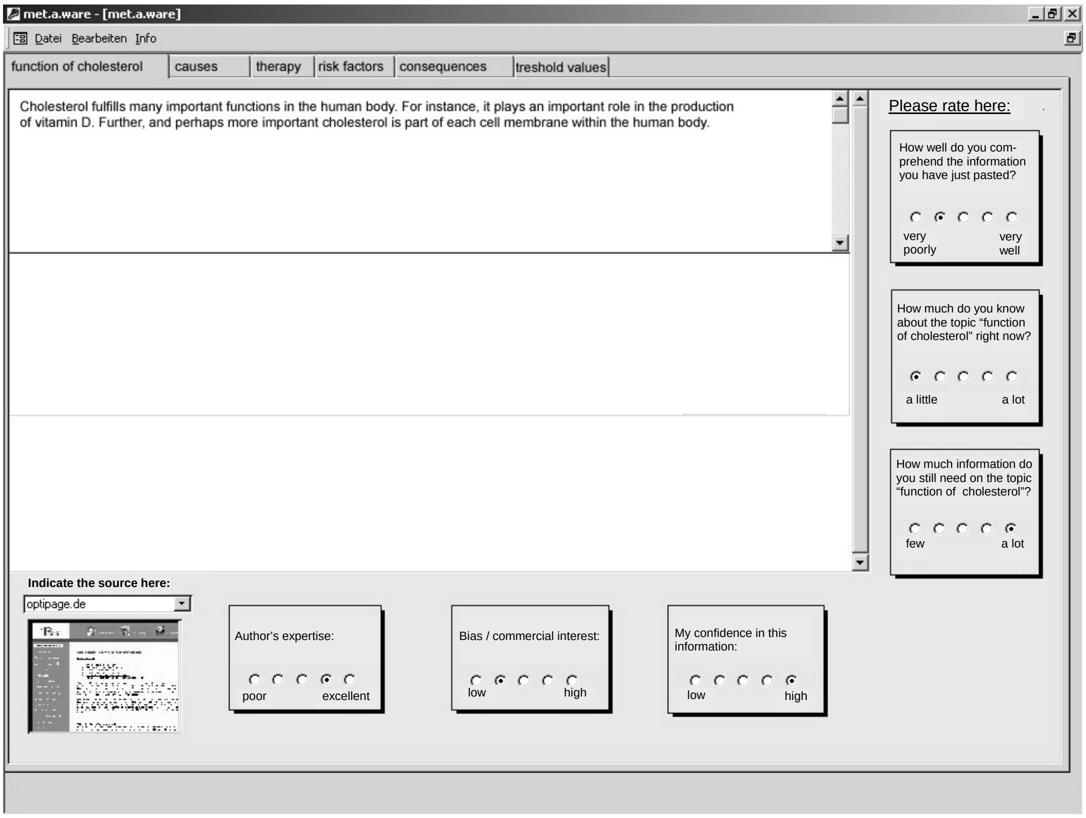
Univariate Tests of Number of Words Relating to the Ontological Categories

Ontological category	**F(1, 38)**	***p***	η^2_{part}
Therapy	1.43	*ns*	.04
Consequences	1.37	*ns*	.04
Function	13.71	< .001	.27
Causes	4.10	< .05	.10
Threshold values	8.96	< .01	.19
Risk factors	5.85	< .05	.13

Figure captions

Figure1: Screenshot of the metacognitive tool *met.a.ware*

Figure 1.


met.a.ware - [met.a.ware]
Datei Bearbeiten Info
function of cholesterol
causes
therapy
risk factors
consequences
treshold values
Cholesterol fulfills many important functions in the human body. For instance, it plays an important role in the production of vitamin D. Further, and perhaps more important cholesterol is part of each cell membrane within the human body.
Please rate here:
How well do you comprehend the information you have just pasted?
very poorly
very well
How much do you know about the topic "function of cholesterol" right now?
a little
a lot
How much information do you still need on the topic "function of cholesterol"?
few
a lot
Indicate the source here:
optipage.de
Author's expertise:
poor
excellent
Bias / commercial interest:
low
high
My confidence in this information:
low
high