A Deliberation Theory-Based Approach to the Management of Usability Guidelines

Fahri Yetim
Information Systems, University of Siegen, Germany

Fahri.Yetim@uni-siegen.de

Abstract

Designing interaction entails addressing multiple issues and challenges, ranging from the technical and economic to the legal and ethical. Usability guidelines recommend or prescribe courses of action and thus play a significant role in designing usable systems. This paper argues that approaches to guidelines need to support processes of deliberation and tradeoff and suggests a deliberation theory-informed model for the organization of guidelines. The model integrates concepts from Habermas' discourse theory and Toulmin's model of argumentation to categorize and represent guidelines. In addition, the paper presents two explorative studies conducted to understand the representational fit of the suggested categories to the domain of guidelines. The studies specifically consider the characteristics of coverage and encodability and also explore difficult cases. Finally, a brief summary of the usability evaluation results of the prototype that instantiated the proposed model is provided. This paper contributes to research and praxis by providing a theory-based model and a prototype for the management of guidelines.

Keywords: Usability Categories; Usability Guidelines; Human-Computer Interaction; Human Factors in Information Systems; Discourse Theory; Deliberation; Reflective Design; Tool.

Introduction

Guidelines and the Need for a Deliberative Approach

Capturing knowledge about the successful design of Human-Computer Interaction (HCI) is important for novice and experienced designers as well as usability engineers. Traditionally, this knowledge has largely been described in guidelines (Mayhew, 1992; Nielsen, 2000; Shneiderman, 1998). Guidelines present design experiences and support design practice with useful sets of recommendations or prescriptions. Thus, one of the important steps in helping designers to design usable systems is to provide them with useful guidelines. Many guidelines can be used for both design and evaluation of interactive systems. The formats and contents of guidelines vary in quality and level of detail, ranging from ill-structured commonsense statements to formalized rules ready for automatic guideline check. However, like any statement or communicative act (Habermas, 1984), guidelines themselves may be subject to criticism and may provoke discussions among designers and researchers as well as between designers, managers, and users, due to

possible conflicting differences in expertise, interests, and preferences. To mention a few: a guideline may be challenged with respect to its comprehensibility because of the jargon or specific vocabularies used in the source as coming from various disciplines (e.g., psychology, human factors, ethnography), or with respect to its validity because of the sources (empirical evidence versus personal experience). A guideline may also be contestable in application situations, e.g., when it competes with other guidelines applicable to the same situation or allows multiple interpretations of the way it is to be applied (consult Mariage, Vanderdonckt, & Pribeanu, 2005, for additional shortcomings of guidelines).

Thus, the construction of a useful and valid set of guidelines as well as the application of the most appropriate ones needs to be justified, balanced, and traded off. In practice, many approaches to the development and organization of guidelines involve communicative review processes and encourage consensus (e.g. Koyani et al., 2003; Scapin et al., 2000; "The Web Accessibility Initiative", 2008). However, there is a lack of a theoretically well-founded approach to dealing with guidelines in a deliberative way. In fact, inspired by the theoretical works of Schön (1983), Habermas (1984) and Toulmin (1958), many IS researchers have emphasized the relevance of deliberation for the development of information systems in general (e.g., Goldkuhl & Lyytinen, 1982; Hirschheim, Klein, & Lyytinen, 1996; Klein & Hirschheim, 2001; Ulrich, 2001). However, in the context of Human-Computer-Interaction (HCI) such deliberation-theoretical insights have not been thoroughly integrated in the approaches dealing with guidelines.

Objectives and Contribution of this Paper

Situated in this deliberative paradigm, the overall goal of our research is to fill this gap by developing a method and a tool for organizing and deliberating on guidelines. Our approach uses concepts from a discourse-based model for deliberation (Yetim, 2006a), which is mainly based on Habermas' (1984, 1996) discourse theory and additionally considers Toulmin's (1958) model of argumentation. As deliberation plays a central role in Habermas' and Toulmin's works, both theories provide concepts and guidance for the organization and representation of guidelines, particularly in building structures to enable reflections and critiques on guidelines in a systematic way compatible with rational practice. They allow us to bring a broad concept of rationality that also includes ethical and moral rationality into the discursive processes of developing and using guidelines. Thus, a guidelines management tool that is informed by these deliberation theories is assumed to be useful not only for researches and practitioners, but also for students in an educational context as it allows them to manage and access design wisdom as well as to practice critical thinking.

The current paper leaves out issues of deliberation on guidelines and focuses on works so far towards this goal. The objectives and structures of this paper are as follows. Firstly, it presents the theoretical background of this work, including the introduction of a theory-based model for deliberation and the articulation of its origin and usage options. With this background established, the subsequent section suggests a model for organizing guidelines that integrates a set of concepts from the deliberation model with Toulmin's argument schema. This section justifies and illustrates with examples why some concepts of the deliberation model can be viewed as usability categories and, thus, as appropriate for categorizing guidelines, and why Toulmin's argument schema is chosen for representing guidelines. Then, two sections follow that report on the evaluation of some aspects of the proposed model and its implementation. The first section compares two explorative studies on the usage of categories to categorize a set of guidelines and discusses lessons learned from these studies, and the second briefly summarizes the experiences gained through a usability evaluation of the system prototype that instantiated the proposed model. Neither the explorative studies nor the usability evaluation of the prototype claim to fully validate the model. Nevertheless, they support our initial understanding of some important aspects of the pro-

posed approach and enable us to reflect on the implications of the research so far. Based on these experiences, the remaining sections discuss the merits, limitations, and implications of this work for the proposed model, research and practice and, finally, offer some conclusions and further research directions.

All in all, this paper claims to contribute to HCI research by suggesting a theory-based approach to the management of usability guidelines. In particular, it provides researchers and practitioners with novel HCI artifacts, i.e., a model for organizing guidelines and a system prototype that instantiates the model. In addition, the explorative studies reveal some gaps in the set of research-based guidelines and point to some interesting areas for future research.

Theoretical Background

A Model for Deliberative Practice

According to Webster and Porter (1913), to deliberate is to weigh in the mind; to reflect; to consider the reasons for and against; etc. Deliberation can take place in an individual's mind but can also be dialogical. Deliberation is viewed here as a "cooperative" dialogue in which communicating actors pursue similar goals and strive to find a common solution to the satisfaction of all concerned. Argumentation is embedded in such dialogs.

Yetim (2006a) proposed a meta-communication model for deliberation about communication (Figure 1). Its architecture and concepts aim to ensure that communication about communication takes place in a structured and discursive way. The model provides actors with a set of basic issues (or potential communication breakdowns) that need to be systematically reflected on when analyzing or designing communication. The meta-communication model consists of two levels: the *conversation for clarification level* and the *discourse level*. The conversation level organizes a set of basic issues in the form of a staircase and provides a structure for conversations about communication. The discourse level provides different types of discourses to be used for argumentative examination of controversial positions that may arise at the conversation level. Depending on the type of controversy, different discourses are to be entered. In other words, the conversation level represents what should be talked about when analyzing or designing communications, while the discourse level represents how the possible conflicts on each issue should be resolved, namely with arguments in related discourses.

The staircase used at the conversation for clarification level is an extended version of Ulrich's (2001) *philosophical staircase* for reflective practice. In Ulrich's work the steps of the staircase were called semiotic steps (syntactic, semantic, and pragmatic clarity), epistemological steps (expressive, empirical, and normative validity) and practical-philosophical steps (instrumental, strategic, and communicative rationality). Whereas semiotic steps are theoretically grounded in Peirce's (1931-35) work, the epistemological and practical-philosophical steps are grounded in Habermas' (1984) theory of communicative action. The staircase has been revised and extended (Yetim, 2006a) by two additional steps (physical clarity and aesthetic rationality). In addition, a set of discourses proposed by Habermas (1984, 1993, 1996) are used at the discourse level. Consult Yetim (2006a) for a fuller description and for the rationale of the model.

Figure 1: A model for deliberation (adopted from Yetim, 2006a)
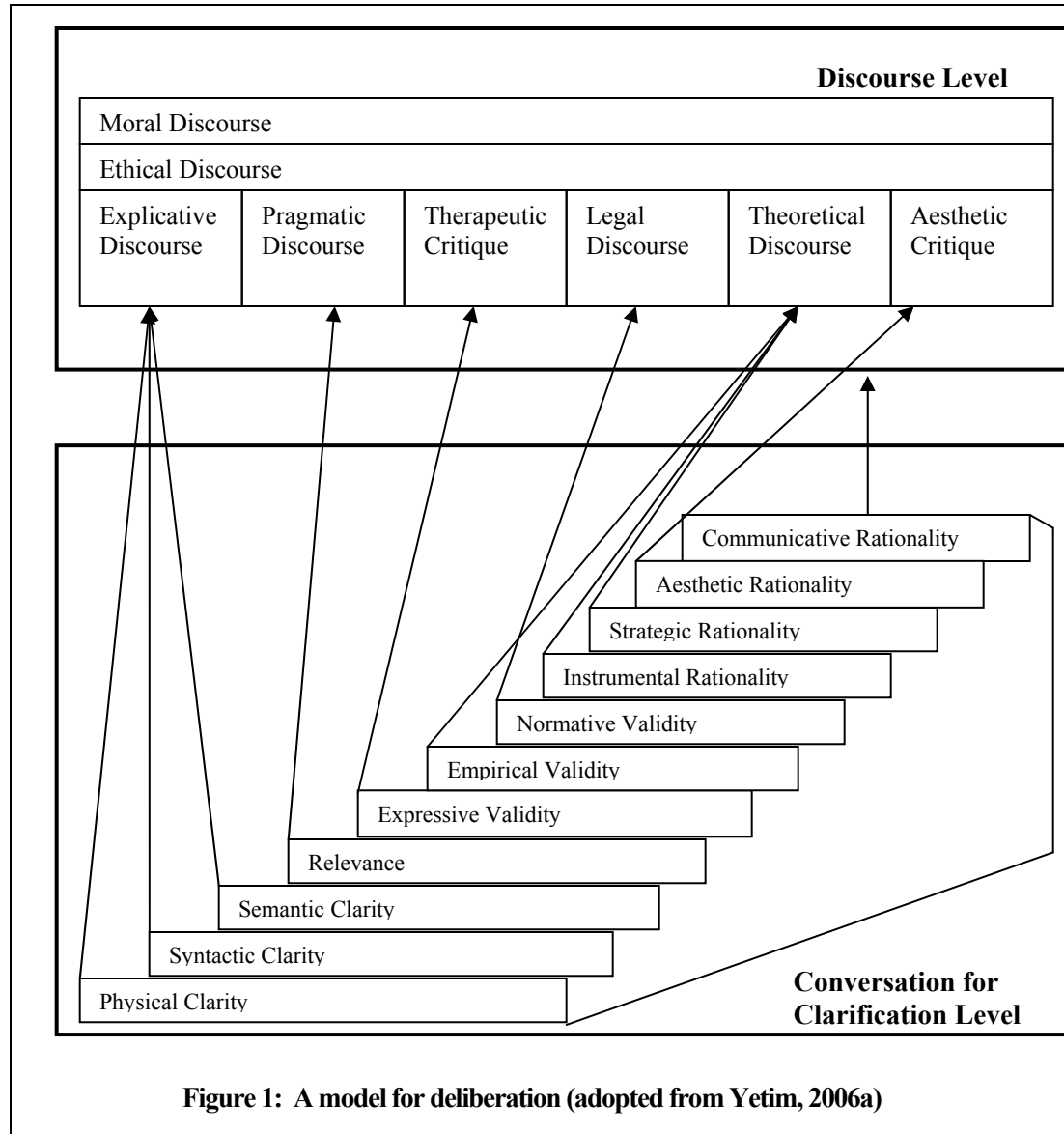
The Usage Options of this Model

For HCI researchers and practitioners, the model in Figure 1 offers at least three usage options. Firstly, it can be used to systematically deliberate about those issues that are organized by the staircase. They range from media-related technical issues through to syntactic, semantic, and relevance aspects of the communication content (or signs) as to validity, appropriateness, and effectiveness in an interaction situation. While clarifying these issues, possible disagreements can be resolved in related discourses, in which participants justify their positions with arguments. Communicative usage of the model has been theoretically investigated (Yetim, 2006a, 2008a). In addition, the model has been implemented within a prototype called DISCOURSIUM to promote critical examination of information in the web (Yetim, 2008b).

The second usage option – on which this paper focuses – deals with organizing guidelines. This option excludes the discourse level and uses only the categories of issues represented by the staircase for categorizing guidelines.

Finally, the third usage option – which our overall research strives for – is the combination of the two options mentioned, that is, the application of the staircase for organizing guidelines along the categories and the use of the whole model for systematically deliberating on guidelines organized. As guidelines recommend or prescribe courses of action, deliberating on guidelines can be done not only within the processes of guidelines development but also during their applications. The model offers structures to reflect, for example, on their comprehensibility, validity, and appropriateness or legitimacy in application contexts. In this usage option, the double role of the staircase is obvious: it provides categories for organizing guidelines as well as categories of issues that can be systematically reflected on while collaboratively assessing guidelines or discussing their applicability in a given context.

Leaving out the issue of deliberation on guidelines, this paper focuses on the management of guidelines and next describes a model for organizing guidelines that integrates the concepts of the staircase and Toulmin's (1958) argument schema.

A Model for Organizing Guidelines

When organizing guidelines, at least two issues need explicit attention: Firstly, how can or should they be categorized? And secondly, what information about them is relevant and should therefore be captured or represented? In the following subsections we first elaborate on why and how the steps of the staircase can serve as categories of guidelines and then discuss why and how Toulmin's argumentation schema can be used for representing relevant information on guidelines.

Regarding the Staircase as Usability Categories

There are different approaches to the organization of guidelines. Usually, guidelines are organized either around the media (e.g., text, graphics, video) or around the activities in the context of human-computer interaction or processes of information systems development, such as planning, design, production, and maintenance (e.g., Grammenos, Akoumianakis, & Stephanidis, 2000; "IBM web design guidelines", 2008; Scapin et al., 2000; Vanderdonckt,1995). In addition, some guidelines are organized in several dimensions along which the usability of interfaces can be evaluated. For example, Microsoft Usability Guidelines (Keeker, 1997) are organized in five major categories: content, ease-of-use, promotion, made-for-the-medium, and emotion.

In contrast, this paper claims that the staircase provides a conceptualization of usability with lower-level categories that are theoretically grounded in communication and action theories. Considering the communication and action perspective is relevant as most HCI guidelines – if not all – are about interface-mediated communication of information and/or concerned with the creation of action potentials on user interfaces. The categories represent communication and action specific criteria for grouping such guidelines. Table 1 lists the categories and briefly defines the characteristics of related guidelines.

Guidelines belonging to the first three steps are expected to make comprehensible signs or contents communicated. Habermas (1984) states that comprehensibility is a precondition of the acceptance of utterances. Accordingly, these steps deal with three different aspects of communicating comprehensibly, namely that the signs (contents) are perceivable/readable, syntactically well formed, and have clear references or word meaning. As global diversity in technological standards, syntactic conventions, and the meanings of signs may cause breakdowns or misinterpreta-

tions (Smith & Yetim, 2004; Yetim, 1998), these steps are viewed as basic usability categories, especially when designing multimedia information (Nielsen, 2000).

Table 1: Categories and characteristics of related guidelines	
Categories	**Characteristics of related guidelines**
Physical Clarity	Guidelines are related to communication media and make recommendations on actions for making signs (communication contents) perceivable/readable/hearable (e.g., "provide equivalent alternatives to auditory and visual content").
Syntactic Clarity	Guidelines deal with syntactic structures and rules, and aim at achieving syntactically clear and correct formulation of signs in order to avoid misinterpretations (e.g., "ensure that homepage panels are of a width that will make them recognizable as panels").
Semantic Clarity	Guidelines deal with the clarity of the meaning of signs and are recommendations on actions for making the meaning of signs (contents) comprehensible (e.g., "identify words which may have culture-specific meanings").
Relevance	Guidelines are concerned with how to communicate relevant signs (content) to the receivers (e.g., "explain the benefits users receive from sharing personal information").
Expressive Validity	Guidelines deal with the sincerity of expressions and recommend actions for achieving truthfulness and transparency when communicating contents (e.g., "show that there is a real organization behind your site").
Empirical Validity	Guidelines are related to the truth of the communication content and recommend actions to ensure that the content agrees with the fact (e.g., make it easy to verify the accuracy of the information on your site").
Normative Validity	Guidelines deal with norms and legal issues and recommend actions for communicating in accord with accepted social norms, values and laws (e.g., "provide access to a privacy policy from every page, and highlight it whenever users give personal information").
Instrumental Rationality	Guidelines deal with the efficiency of actions or means for achieving a given goal. They recommend what to do, or what not to do, to allow users efficient interaction with the communication content or the system (e.g., avoid requiring users to scroll to determine page contents").
Strategic Rationality	Guidelines are concerned with egocentric calculation of success for achieving a given goal and recommend effective strategic means to influence or persuade receivers to do what is desired by the provider/sender of the content (e.g., "provide different site paths to facilitate different shopping strategies").
Aesthetic Rationality	Guidelines deal with aesthetic aspects and recommend actions for communicating content in accord with culturally established standards of aesthetic values/preferences (e.g., "design in a style that will appeal to your audience's tastes").

The *relevance* step deals with pragmatic aspects of communication (Allwood, 2000; Schütz, 1970). Regarding relevance as a subcategory of content, Agarwal and Venkatesh (2002) demonstrate its importance for assessing the usability of websites (see also Venkatesh & Ramesh, 2006). Guidelines belonging to these categories are concerned with the communication of relevant signs (or contents) to the receivers.

The next three steps refer to the validity aspects of the signs presented and are based on Habermas' (1984) notion of validity claims. The argument is that the presentation of contents or action potentials on the user interface of an information system should not only be comprehensible and relevant but also represent sincere intentions, refer to the true (or inter-subjectively believed) state of the world, and be possible to perform in accordance with socially accepted norms. Accordingly, guidelines recommend what to do to satisfy the criteria of the three validity dimensions. These aspects are of importance as errors and violation of (privacy) norms can influence users' perception of the credibility of websites and the risk, which are viewed as important perceptual factors that impact online trust (Corritore, Kracher, & Wiedenbeck, 2003; Hoffman, Novak, & Peralta, 1999). The evaluation of the validity of signs at the user interface (including icons, labels, etc.) requires that they are visible and understandable.

The next higher steps of the staircase are concerned with the rationality aspects of communicated signs. These include Habermas' (1984) three types of rationality and, additionally, aesthetic rationality. Guidelines related to these categories recommend actions to meet the criteria of these different types of rationality. Instrumental rationality deals with whether the signs or their organization on the user interface allow users efficient actions or interactions with the system to achieve a given goal. In contrast, strategic rationality deals with whether the signs and their organization at the user interface are effective strategic means to influence or persuade users as social actors to do what is desired by the provider/sender of the content. It deals with egocentric calculation of success. Efficiency and effectiveness of interactions for achieving specific goals are fundamental usability criteria according to the ISO's definition of usability (Karat, 1997; see also Nielsen, 2000). In addition, considering effective strategic means, especially for e-commerce interfaces, is becoming an ingredient of design (Nielsen, Molich, Snyder, & Farrel, 2001).

Aesthetic rationality deals with whether signs presented are in accord with culturally accepted aesthetic values. Recent research has demonstrated the importance of creating aesthetically pleasing design for the usability of systems (Norman, 2004; Tractinsky, Katz, & Ikar, 2000). Visual aesthetics plays a role in users' satisfaction and pleasure and their attitudes toward interactive systems (Tractinsky, Cokhavi, Kirschenbaum, & Sharfi, 2006).

Finally, the concept of communicative rationality refers to whether a mutual understanding among actors is achieved on the basis of inter-subjectively recognized validity claims (Habermas, 1984). This step does not represent a category for organizing guidelines. Instead, its usage is limited to the deliberation on guidelines within the communicative usage option of the model (Yetim, 2006a).

In conclusion, the staircase presents the conceptualization of usability with 10 specific categories. They cover many usability aspects of user interfaces and relate to many usability principles ranging from the principle of appropriate presentation of content to its validity and adequacy in relation to users' actions. Moreover, they are more specific than some categories used in other studies. For example, "media use" is viewed as a subcategory of the content and defined as "the extent to which a website uses media appropriately and effectively to communicate the content" (Agarwal & Venkatesh, 2002, p.176). In contrast, the categories of this study provide further differentiation concerning the appropriateness of media use, such as its appropriateness for making something readable/perceivable, for allowing efficient communication and action, for making signs aesthetically appealing, etc. The categories represent purposes and specify what needs sys-

tem designers need to satisfy. They also make explicit what the guidelines are expected to communicate, namely, recommendations on what should be done to provide readable/perceivable, syntactically and semantically clear signs, to communicate relevant and valid (trustworthy, reliable, appropriate) information, and to act in an efficient and effective way.

Representing Guidelines by Argument Schema

In addition to determining which guidelines belong to which categories, representing information on guidelines is another significant aspect of the proposed approach. As mentioned earlier, guidelines are generally based on design experiences or empirical research, support a set of principles (i.e. fundamental ideals or beliefs), and are specific to a particular domain, such as the web. As they recommend and prescribe courses of action, they can be challenged and justified through argumentation, i.e. through a process of making claims and providing support and justification for these claims. This raises the issue of how information on guidelines should be represented in order to be consistent with the argumentative practice.

Argument schemas used in everyday communication are viewed as very useful tools for structuring and analyzing information (Newman & Marshall, 1992; Walton 1996). Schemas have already been successfully used within HCI for capturing design rationale or structuring the discourse in standardized, computationally manageable forms (e.g., Buckingham Shum & Hammond, 1994; Carroll & Moran, 1991). Toulmin's (1958) argument schema is an influential tool (Loui, 2005), which enriches the traditional premise-conclusion model of arguments by distinguishing additional elements. According to Toulmin, the logical structure of human argumentation and reasoning consists, implicitly or explicitly, of six categories: Claim, Ground, Warrant, Backing, Qualifiers, and Rebuttals. A *claim (*or conclusion) whose merits we are seeking to establish is based on some *ground* (or data). The relation between the two is characterized by a rule of inference, a *warrant*, i.e. a statement that justifies the link of the information set forth in the ground and claim. The warrant itself can be *backed* by some other facts or experiences. In addition, *qualifiers* are phrases expressing the degree of certainty placed on a claim, and *rebuttals* express extraordinary or exceptional circumstances that might defeat the warranted claim.

This paper uses Toulmin's schema as it provides an intuitively plausible set of categories and relations for analyzing and representing the logical structure of the information on guidelines: It particularly allows differentiating between elements in terms of functional or relationally defined roles, for example, between the inference rules (warrants) and knowledge sources (backing) underlying the recommendation/prescription of particular guidelines. This could help us to critically analyze and understand types of information involved in justifying guidelines (e.g. type of warrant or backing) and - when they are called into question - to collaboratively examine their plausibility, weaknesses, and strengths in related discourses. This is compatible with the philosophy of the deliberation model in Figure 1, which favors sorting issues/validity claims according to their logic and knowledge domains in order to examine them in corresponding discourses (see Habermas (1984) and also Yetim (2006a, 2008b) for additional discussion of this relationship).

Table 2 illustrates, with examples, how Toulmin's schema can be used to encapsulate knowledge on guidelines and represent them in relation to the categories from Table 1. In line with the argument schema, each category can be conceived of as a ground (or intended purpose) and each related guideline as a claim (or recommended action). It has the form: "IF you want to achieve X, then do Y". The knowledge of guidelines includes their justification or rationale (warrant) and supporting evidence (backing) such as empirical research or consensus among experts. In addition, optional information on the degree of strength/importance of the guidelines can indicate whether a content developer *must, should* or *can* satisfy the guideline. Finally, optional informa-

tion about contextual conditions or exceptions (if any) can be presented to inform the application of guidelines (e.g. specific tasks, systems, groups or cultures).

Table 2: Illustrative examples for organizing guidelines					
Categories (*Intended Purposes*)	**Guidelines** *(Recommended Actions)*	**Rationale** *(Warrant)*	**Supporting evidence** *(Backing)*	**Strength/ modality** *(Qualifier)*	**Context/ exceptions** *(Rebuttals)*
Physical Clarity	1. Provide equivalent alternatives to auditory and visual content.	Since many people cannot use video, images, or sound, but they can use equivalent information provided via other media.	Web content accessibility guidelines 1.0; W3C recommendation 5-May-1999	Must	Unless the intended user groups need a specific media.
	2. …	…	…	…	…
Syntactic Clarity	1. Ensure that homepage panels are of a width that will make them recognizable as panels.	Since the width of panels seems to be critical for helping users understand the overall layout of a website.	Much supporting research cited in (Koyani et al. 2003, p39)	Importance: 4 (out of 5) Strength of Evidence: 3 (out of 5)	
	2. …	…	…	…	…
Semantic Clarity	1. Identify words which may have culture-specific meanings.	Since some culture-specific meanings may cause misunderstandings (e.g., the word "faculty" could be interpreted to mean "subjects", "buildings" or "academic staff members").	Kukulska-Hulme (2000)		
	2. …	…	...	...	...
Relevance	1. Explain the benefits users receive from sharing personal information.	Since users will feel more inclined to provide information if the advantage of doing so is clear.	IBM web design guidelines (2008)		
	2. …	...	...	...	...
Expressive Validity	1. Show that there is a real organization behind your site.	Since this will boost the site's credibility.	Much supporting research cited in Fogg (2002)		
	2. …	…	…	…	…
Empirical Validity	1. Make it easy to verify the accuracy of the information on your site.	Since providing third-party support (citations, references, source material) for information can help to build website credibility.	Much supporting research cited in Fogg (2002)		
	2. …	…	…	…	…

Normative Validity	1. Provide access to a privacy policy from every page, and highlight it whenever users give personal information.	Since access to this policy helps engender trust.	IBM web design guidelines (2008)		
	2. …	…	…	…	…
Instrumental Rationality	1. Avoid requiring users to scroll to determine page contents.	Since users should be able to recognize immediately whether the subject of any given page interests them.	IBM web design guidelines (2008)		
	2. ...	...	...	...	...
Strategic Rationality	1. Provide different site paths to facilitate different shopping strategies.	Since sites that accommodate their users' strategies are more likely to succeed than those that force users to learn new strategies.	IBM web design guidelines (2008)		
	2. …	...	...	...	...
Aesthetic Rationality	1. Design in a style that will appeal to your audience's tastes.	Since people may prefer different styles (e.g., a reference site for a general corporate will need to convey a different image than a site which should appeal to restaurant managers and hobbyist connoisseurs interested in exotic fruit).	IBM web design guidelines (2008)		
	2. …	…	…	…	…

Finally, we should provide two notes on the usage of Toulmin's schema for representing guidelines: Firstly, the examples in Table 2 are chosen for illustrations as they best represent the purposes of the categories, and the references are used as backing. Citations can be seen to function as backing for an argument and should be understood as a placeholder for a line of argument establishing the warrant and/or its applicability to the current case (Newman & Marshal, 1992). Instead of references, one may provide the contents of the original texts. The credibility of the references or that of the original texts remains open to discursive evaluation and critical challenge when deliberation takes place on each guideline in the communicative usage mode of the model.

Secondly, there are also critiques on Toulmin's model that point to its deficiencies such as lack of a topic for organizing the detailed argumentation, inability to express argument summaries or to compare positions, etc. (Hitchcock & Verheij, 2005; Newman & Marshall, 1992). We contend that some of the deficiencies are related to the domain of argument mapping and, therefore, do not regard them as primarily relevant for the purpose of this paper. On the other hand, some missing elements or links leading to more detail or specific information, such as the owner of the information or the degrees of agreement, can be easily added to the current representation in online environments in order to compensate for the representational deficiencies of this schema (as done by many argumentation-related tools). That said, the next section focuses on explorative studies conducted to experiment with the proposed categories.

Evaluation of the Categories

In order to gain an initial understanding of the applicability of the categorization model and also to clarify potential problems, two explorative studies have been conducted by employing a set of research-based web design and usability guidelines. To present the experiences from the studies, this section first briefly describes the criteria used for evaluating the categorization model and also introduces the guidelines employed. It then presents each study and, finally, compares and discusses their results and the lessons learned from them.

Criteria for the Evaluation

Computer-mediated representational schemas participate in three distinct (though in practice interrelated) relations (Newman & Marshall, 1992): (1) a representational relation to the represented domain; (2) a relation to the cognitive, perceptual, and activity systems of the users; and (3) a relation to the system's computational capabilities. Each of these relations establishes a "design dimension" along which to evaluate representational schemas: (1) concerns the schema's *representational fit*, (2) concerns its *usability*, (3) concerns its *computational power*.

The explorative studies focus primarily on issues of representational fit, reserving a few words about usability for the discussion. Two general criteria related to the representational fit are coverage and encodability, which proved useful in understanding the limitations of categorization schemas (Newman & Marshall, 1992). In our context, these dimensions deal with examining how, on the one hand, the categories capture the set of guidelines studied (coverage) and, on the other hand, how the guidelines are mapped into the categories (encodability). In addition, we explore difficult cases. Specifically, the following issues are explored:

1. Do the categories cover the guidelines, i.e., can all guidelines be associated with them?

2. Do the guidelines studied cover the categories, i.e., do they support the purposes of categories?

3. How readily can guidelines, in practice, be mapped into the categories, i.e., are there any significant difficulties while re-categorizing guidelines?

As the relation between the representation (categories) and the represented domain (guidelines) is emphasized here, it is important to note that assessment of representational fit is fundamentally an interpretative enterprise (Newman & Marshal, 1992). That is, a human understander must decide questions with respect to both the categories and the guidelines. Concerning the categories, s/he must arrive at a stable and, hopefully, appropriate interpretation of the categories with respect to the guidelines. And concerning the guidelines, s/he must both arrive at some interpretation of the guidelines and decide questions such as what purposes they serve.

Guidelines Employed

As mentioned earlier, there are several guidelines and suggestions for effective interaction design. Among them, Koyani et al. (2003) present a set of 187 research-based guidelines that was constructed after extensive expert reviews of 500 guidelines. We use this set of guidelines for the following reasons: firstly, it represents the result of a rigorous selection; secondly, the description of each guideline has a clear structure and is to some extent compatible with the proposed representation schema. For example, it contains additional comments (including justifications) and information about their relevance as well as illustrative examples to make them comprehensible.

The set of guidelines was developed by the Communication Technology Branch of the National Cancer Institute in the USA to provide clear information in an efficient and effective manner to

cancer patients, health professionals, researchers, and the public. The guidelines aim to help those involved in the creation of information-oriented websites to base their decisions on the current and best available evidence. Primary audiences for the guidelines are website designers, usability specialists, managers, and others involved in the creation or maintenance of websites. A secondary audience is researchers investigating web design issues.

In Koyani et al. (2003), the guidelines are grouped according to web design issues, which are: (1) Design process and evaluation; (2) Optimizing the user experience; (3) Accessibility; (4) Hardware and software; (5) The Homepage; (6) Page layout; (7) Navigation; (8) Scrolling and paging; (9) Headings, titles, and labels; (10) Links; (11) Text appearance; (12) Lists; (13) Screen-based controls (Widgets); (14) Graphics, image, and multimedia; (15) Writing web content; (16) Content organization; (17) Search. In this set, guidelines are presented as statements along with additional comments; many guidelines are also illustrated with examples.

Study 1

The first study was conducted by the author in order to gain first-hand experience about the applicability of the categorization model and also to clarify potential problems with the categories in relation to guidelines. In this study, the whole set of guidelines was employed. Its results are documented (Yetim, 2006b) and are briefly summarized here to provide data for the discussion and comparison with the second study.

While assigning guidelines to the categories, the guiding questions have been, "What purpose(s) does this guideline serve?" or "What type of breakdown might occur if we did not follow the corresponding guideline?" With these questions in mind, the author interpreted the descriptions of guidelines in Koyani et al. (2003) and completed the categorization in a two-stage process. A worksheet was developed to record any problems in the first stage. In the second stage, they were clarified by looking again in the description of the guidelines.

Table 3 lists the distribution of guidelines across categories: it presents the number of guidelines assigned to each category, and the percentage of 187 guidelines considered. The results of this study indicated that firstly, concerning the coverage of categories, all guidelines could be associated with the categories. Secondly, concerning the guidelines support of the purposes of categories, the largest number of guidelines has been assigned to the category *instrumental rationality*, followed by *semantic clarity* and *strategic rationality* and no guideline could be assigned to *expressive validity*. Many guidelines could be assigned to multiple categories. Finally, the author experienced some difficulties – based on the descriptions of guidelines – in deciding whether a guideline served the purpose of *instrumental rationality* or of *strategic rationality*. These and other issues will be discussed after having introduced the results of the second study.

Table 3: Distribution of guidelines across categories in the first study

Categories	Number of guidelines assigned	Percentage (total 187)
Physical Clarity	18	9.62
Syntactic Clarity	16	8.55
Semantic Clarity	38	20.32
Relevance	38	20.32
Expressive Validity	0	0.0
Empirical Validity	7	3.74
Normative Validity	6	3.20
Instrumental Rationality	75	40.10
Strategic Rationality	43	22.99
Aesthetic Rationality	3	1.60

Study 2

Purpose, participants and guidelines

The second study aimed to overcome the perspective of the author of the first study and to consider other subjects to gain additional data for the evaluation issues. For that purpose, third year bachelor students were considered who had an understanding of web usability issues and a basic knowledge of web programming after taking classes on HCI. Students are viewed as potential users of the model for two reasons: Firstly, the general philosophy of the approach is to allow broad participation and not to restrict the usage of the model to experts, and, secondly, the guideline management tool that instantiates the model is also intended to be used in educational context, particularly in interaction design classes, to provide students with some design wisdom. The specific participants were identified after an introduction of the purpose of the study in a class session. As guidelines employed were in English, a sufficient knowledge of the English language was required. Four students (two male and two female) volunteered to participate.

For consistency with the first study the set of guidelines from Koyani et al. (2003) is used. However, in order to accommodate the background knowledge of the participants and also to reduce their workload, groups of guidelines that require specific technical knowledge and/or do not primarily deal with human communication have been withdrawn from the study. For efficiency reasons, instead of checking the comprehensibility of each guideline for each participant, the decision to exclude some groups of guidelines is made by the author based on his familiarity with both the technical descriptions of the guidelines and the technical background of the students. The groups of guidelines excluded are: Design Process and Evaluation, Hardware and Software, Screen-based Controls (Widgets), Graphics, Images, and Multimedia, Search. As a result, 69 guidelines were withdrawn from the study, and the remaining set of 118 guidelines was employed.

Procedure

To coordinate the group work for categorizing guidelines, a Delphi-like method is employed, an efficient and widely accepted method for structuring group communication to bring a wide variety of views to the surface (Linstone & Turoff, 1975). To correlate with the theoretical foundation of this approach, which values deliberativeness in the process (Habermas, 1996; Yetim & Turoff, 2004), the second stage of the two-stage Delphi processes was conducted with direct interactions in the form of face-to-face discussion without moderation. As argued in other Delphi studies, the methodological concern to make the interviews rational, well-argued discussions seems to be successful (Tapio, 2002). The role of the author was to provide the materials and to summarize the results of each stage in terms of the agreements and disagreements. In addition, the author interviewed the participants to clarify any difficulties that they had while classifying guidelines.

Steps taken were as follows: After an explanation of the purpose of the study, the following documents were provided to the participants in hard copy as well as online to support their work preferences:

- The description of categories and an example guideline for each category. As brief definitions and explanations are typically used for making system features comprehensible and usable, they are intended to simulate the use situations of categories in online environments.

- The reduced set of guidelines from Koyani et al. (2003)

- A template for assigning categories, which contains guidelines headings and a column to insert related categories

- Instructions as follows: first to read a guideline and related comments on it, and then to categorize it – but only if they believe they understand the guideline, otherwise to skip it. As the purpose of this study is not to test the comprehensibility of the guidelines, we preferred to withdraw such guidelines from the study rather than to provide inaccurate responses. Participants were also informed that for each guideline multiple categories could be considered if applicable. They were not told that every guideline can be categorized. Instead, they were instructed to make a question mark when a guideline was understood but not easy to assign to any category, in order to clarify such problems with categories in interviews.

The individual work of each participant was analyzed and the group result of the first stage computed by applying the majority rule, that is, a guideline was assigned to those categories that were recommended at least by three of four participants. The remaining guidelines were prepared for the second stage.

The second stage was conducted as face-to-face group discussion without a facilitator. The author first interviewed the participants to understand the difficulties that they had had in the first stage and informed them about the results of the first stage, showing where agreements and disagreements had been achieved. Then, they were instructed to discuss with the group each guideline remaining from the first study, and to re-categorize them individually. Within two hours, they finished this task in a discussion without facilitation. Finally, the results were analyzed and an interview conducted to clarify the difficult cases, as discussed later.

Results

As mentioned above, from the set of 187 guidelines only 118 were employed in the second study. The result of the first stage of this study showed agreement on the categorization of 66 guidelines out of 118 (55.94 %) whereas the second stage conducted with remaining guidelines led to agreement on 49 guidelines out of 52 (94.6 %). Through the two-stage process, participants achieved agreement on 115 guidelines out of 118 (98.80 %). Two guidelines were marked with questions and one guideline assigned to different categories. The details of this study's results are provided in the Appendix. Table 4 summarizes the distribution of guidelines across categories. Next, we discuss them in comparison with the first study.

Table 4: Distribution of guidelines across categories in the second study		
Categories	**Number of guidelines assigned**	**Percentage (total 115)**
Physical Clarity	16	13.91
Syntactic Clarity	30	26.08
Semantic Clarity	18	15.65
Relevance	14	12.17
Expressive Validity	0	0.0
Empirical Validity	0	0.0
Normative Validity	0	0.0
Instrumental Rationality	29	25.21
Strategic Rationality	15	13.04
Aesthetic Rationality	0	0.0

Comparison of Both Studies and Lessons Learned

To compare both studies, we consider only the results of 115 guidelines that are commonly applied and successfully categorized in both studies. The comparison in Table 5 indicates a significant number of similarities and differences. From 115 guidelines only 60 guidelines (52 %) have been assigned to shared categories in both studies. The disagreements between two studies can be

attributed to several factors. In what follows, we reflect on the results of the studies around the three specific evaluation issues that concern aspects of coverage, encodability, and difficult cases. We reserve some reflections on their implications for the model later for the discussion section, after having introduced the evaluation results of the prototype.

(1) With respect to the issue of coverage, the results show that categories are general enough to capture the guidelines. Coverage is regarded as a measure of the percentage of guidelines that are not labeled with question marks. The first study (Table 3) demonstrated that the categories can meaningfully be used to categorize the entire set of guidelines, including those guidelines that deal with the hardware and system design processes and thus have no direct link to human communication aspects. We conclude that this is because the purposes of technology- and process-related guidelines can be evaluated according to whether they contribute to aspects of comprehensibility, relevance, and validity as well as rationality of signs. The coverage was also high in the second study (Table 4). Excluding the two cases with question marks and considering the number of guidelines that fell into one of the categories (i.e., 116 out of 118 guidelines) resulted in a coverage of > 98%. Nevertheless, there was surprisingly low agreement in the categorization results of both studies (52 %), as shown in Table 5. Potential reasons will be discussed later in (3).

Table 5: Results of two studies based on the common set of 115 guidelines			
Categories	**Number of guidelines assigned in study 1**	**Number of guidelines assigned in study 2**	**Number of shared assignments**
Physical Clarity	11	16	10
Syntactic Clarity	12	30	7
Semantic Clarity	26	18	13
Relevance	19	14	8
Expressive Validity	0	0	0
Empirical Validity	1	0	0
Normative Validity	5	0	0
Instrumental Rationality	42	29	18
Strategic Rationality	29	15	4
Aesthetic Rationality	2	0	0

(2) Regarding the issue of whether the set of guidelines supports the purposes of categories, it was surprising that no guideline could be assigned to *expressive validity* in both studies even though this category relates strongly to the trustworthiness and credibility aspects of web contents. On the other hand, by using "Stanford credibility guidelines" (Fogg, 2002), many guidelines could be linked to this category (also shown in Table 2). Therefore, instead of questioning this category, we ascribe this result to the fact that guidelines on credibility and trustworthiness were not included in the set of guidelines studied. It was also surprising that no guidelines were assigned to the categories *empirical validity*, *normative validity*, and *aesthetic rationality* in the second study. One possible explanation is that some of the guidelines that are assigned to these categories in the first study were not considered in the second study. On the other hand, as shown in the Appendix, there were also guidelines within the common set of guidelines that are assigned to one of them in the first study. This re-directed our attention to the description of guidelines, in search of possible reasons.

Before discussing the reasons for difficulties, it should be noted that in both studies a significant number of guidelines has been assigned to the category *instrumental rationality*. This may be explained by the fact that the guidelines employed dealt with web design issues, including navigation and design processes. In addition, the results of both studies show that many guidelines could be assigned to multiple categories. Actually this is not surprising since – depending on the context – a guideline can serve several purposes. For example, guidelines for creating visual elements and layouts can serve both the purpose of *physical clarity* since they facilitate reading and perception

of the signs, and the purpose of *aesthetic rationality* since they also influence the aesthetic appearance of the signs. Even so, many approaches to classification of guidelines (including those mentioned earlier) restrict each guideline to only one category. This is in our view a limitation since it may inhibit the complex nature of the guidelines.

(3) Finally, concerning the difficulties with categorization, the first study (Yetim, 2006b) has already reported on difficulties in deciding – based on the descriptions of the guidelines – whether a guideline serves the purpose of *instrumental rationality* or that of *strategic rationality*. Both of them deal with the efficiency and effectiveness of the means for achieving a given purpose. As one distinguishing character is that *strategic rationality* is a social concept of rationality (Habermas, 1984), the decision was made by looking at where user cognitive aspects have been the focus in the description of guidelines. While explaining the reasons for question marks during the interview, two participants of the second study mentioned that they had difficulty finding a relation between the corresponding guidelines and the purposes expressed by the categories.

In the remainder of this section, we reflect on further difficulties observed and discuss potential reasons for some differences in both studies in more detail along with usability aspects of categories and guidelines, which we have neglected so far. To substantiate our arguments, it is worthwhile exploring the relevant aspects by means of an example guideline (Table 6).

Table 6: An example guideline (Koyani et al., 2003 p.23)

Heading: Comply with Section 508

Guideline: If a website is being designed for the United States government, ensure that it meets the requirements of Section 508 of the Rehabilitation Act. Ideally, all websites should strive to be accessible and compliant with Section 508.

Comments: Section 508 requires federal agencies to ensure that their procurement of information technology takes into account the needs of all users – including people with disabilities. About eight percent of the user population has a disability that may make the traditional use of a website very difficult or impossible. About four percent have vision-related disabilities, two percent have movement-related issues, one percent have hearing-related disabilities, and less than one percent have learning-related disabilities.

Compliance with Section 508 enables Federal employees with disabilities to have access to and use of information and data that is comparable to that provided to others. This also enhances the ability of members of the public with disabilities to access information or services from a Federal agency.

For additional information on Section 508 and accessibility:
- http://www.section508.gov
- http://www.w3.org/WAI/
- http://www.usability.gov/accessibility/index.html

The guideline in Table 6 was associated with *normative validity* in the first study and with *physical clarity* by each participant in the first stage of the second study. A closer look at the description of the guideline and the interviews revealed the following: The author of the first study regarded the reference to laws and regulations in the formulation of the guidelines as sufficient to assign it to *normative validity* without considering the details of the comments. In contrast, each participant of the second study paid more attention to the comments and regarded the guideline as serving the purpose of the physical clarity, i.e. the first category. The author of this paper discussed this case with the participants after the final interview. They agreed that multiple interpretations and thus multiple categories for this guideline would have been appropriate: *normative*

validity (due to conformance to a regulation*); physical clarity* (because of considering vision-/hearing-related disabilities) and *instrumental rationality* (because of dealing with movement-related issues).

With respect to the reasons for differences in the results of both studies, one might speculate that the sequence of the categories might have played a role for different choices. In this example case, after regarding the first category *physical clarity* as appropriate, participants of the second study might have paid less attention to the subsequent categories.

Further reasons for differences might be encoding problems such as implicitness and ambiguity. As the purposes of categories that guidelines serve do not show up in the description of guidelines, the participants must infer much implicit information in order to assign guidelines to the existing categories. Classificatory ambiguity might have also played a role. For example, the first study reported difficulties in distinguishing between "instrumental rationality" and "strategic rationality" as they deal with both the efficiency and effectiveness of the means for achieving a given purpose. Similarly, the second study pointed to difficulties with respect to the definition of these categories. However, we do not assume that other categorization models will not encountered similar encoding problems associated with implicitness and ambiguity as such problems will arise whenever people attempt to adapt their diversely and implicit structured understandings of their domain and task to standardized representational schema (Newman & Marshall, 1992).

Thus, we tend to interpret the results as a supporting argument for our deliberative approach as they illustrate the importance of the inclusion of many perspectives and interpretations when categorizing guidelines. In addition, the example guideline points to the importance of the method for representing information on guidelines in order to minimize the occurrence of classificatory ambiguity. The guideline specifies its context of use (e.g., "if a website is being designed for the United States government") and additionally recommends its application to all websites. Alternatively, one might choose a description that first specifies what should be done and then justifies why it should be done (e.g., "complying with section 508" in the US context). There are many guidelines that aim to highlight the same problem but are phrased differently or have different foci. There does not seem to be one single best way to describe guidelines and to avoid overlaps, although the separation of recommendation/prescription and its justification would enable one to deal with them in flexible ways, for example, providing purpose and/or context-specific justifications. Writing guidelines is important from the perspective of the proposed deliberative approach, and will be given explicit attention later when we discuss the implications of this approach.

Finally, differences might have also resulted from problems related to the usability of categories and guidelines. While focusing on issues of representational fit, this paper has so far ignored this dimension. However, the characteristics of the categories and the guidelines are not unrelated to their usability. Traditionally, Human-Computer Interaction researchers have addressed such questions from a cognitive perspective, using the concepts of "cognitive compatibility" or "cognitive fit" (e.g., Neuwirth & Kaufer, 1989; Te'eni, Carey, & Zhang, 2006). From this perspective, it seems reasonable to assume that much of the inconsistency in the results of both studies may have resulted from the cognitive demands placed on participants during categorization. HCI literature argues that c*ognitive complexity* – defined as a function of the incompatibility between representation and use of information – increases the demands on resources and the probability of error (Norman, 1990; Te'eni 2001). In our case, such complexity may result from the interpretation and the translation of the implicit information from the descriptions of the guidelines to the purposes of the categories. As the HCI-related background knowledge of the author differs from that of the participants of the second study, it is reasonable to assume differences between them with respect to the "semantic distance" of the categories and guidelines. Future research should explore in this direction.

On the other hand, relying purely on individual cognition and limiting the usability to the fit between the representations and the cognitive, perceptual, and activity structures of users may ignore the role of social negotiation and the relation between usability and learning, adaptation, and change. The importance of learning is inherent to our deliberative approach as emphasized at the beginning of this paper. Assuming that encoding difficulties are inherent in the enterprise of mapping, we conclude – in line with Newman and Marshall (1992) – that the resolution of these difficulties cannot be seen simply as an individual cognitive issue if representations (or categories) are to serve to structure cooperative work. Instead, a categorization practice needs to consider the social context characterized by differing initial understandings of the domain.

In sum, even though the explorative studies on mapping exercises do not intend or claim to validate the categorization framework empirically, they supplement the theoretical grounding of the categories by providing additional data and useful insights into the relationship between the proposed framework and the current practice for categorizing guidelines. The next section goes beyond the current practice perspective to summarize user experiences with a system prototype that instantiated the proposed model for organizing guidelines.

Evaluation of the System Prototype

We have developed a web-based prototype, a guidelines management tool, to instantiate the proposed model for organizing (i.e., categorizing and representing guidelines). The objective of the tool is to provide facilities for organizing a broad range of design and usability guidelines and to support access and retrieval of guidelines. The tool is intended to serve not only designers and usability engineers during design and evaluation of interfaces but also students interested in learning about and accessing the design experiences.

Figure 2 illustrates, by means of two screenshots, (a) how the proposed model for representation of guidelines based on Toulmin's schema is implemented as a template for entering guidelines into the system database (a similar template is used for presenting guidelines), and (b) how the categories are used to allow a category-based search of guidelines. The development and evaluation of the prototype have been published elsewhere (Bock & Yetim, 2008). Therefore, this section only briefly summarizes the evaluation results to supplement the explorative studies.

The purpose of the evaluation was to assess the usability and usefulness of the prototype as perceived by students as potential end-users and also to identify problems and elicit user comments for improving the current design. This means that the evaluation was not limited to the use of the model but also considered the whole system, including the functionalities such as entering, searching, deleting, and updating guidelines.

The participants of the evaluation were eight students with some experience in user interface design as they had already taken classes in the area of Human-Computer Interaction. Their ages ranged from twenty-three to thirty-one. The participants were first informed about the purpose of the system and of the categories. During the evaluation the subjects were provided with two tasks dealing with the usage of a system's functionality such as submitting, searching, updating, and deleting guidelines. In order to focus on the tool that implements the model, only a small sample of guidelines from Koyani et al. (2003) was used within these tasks. Moreover, as shown in Table 7, the guidelines were prepared for the users in the format that is advocated in this paper to relieve them from the burden of transforming guidelines between different formats and to eliminate potential problems (i.e. lessons learned from the explorative studies).

For the evaluation, IBM Computer Usability Satisfaction Questionnaires were used because they are available for public use and have also shown to be extremely reliable (0.94). In addition, they do not require any special software and require less time to analyze the results. Moreover, they
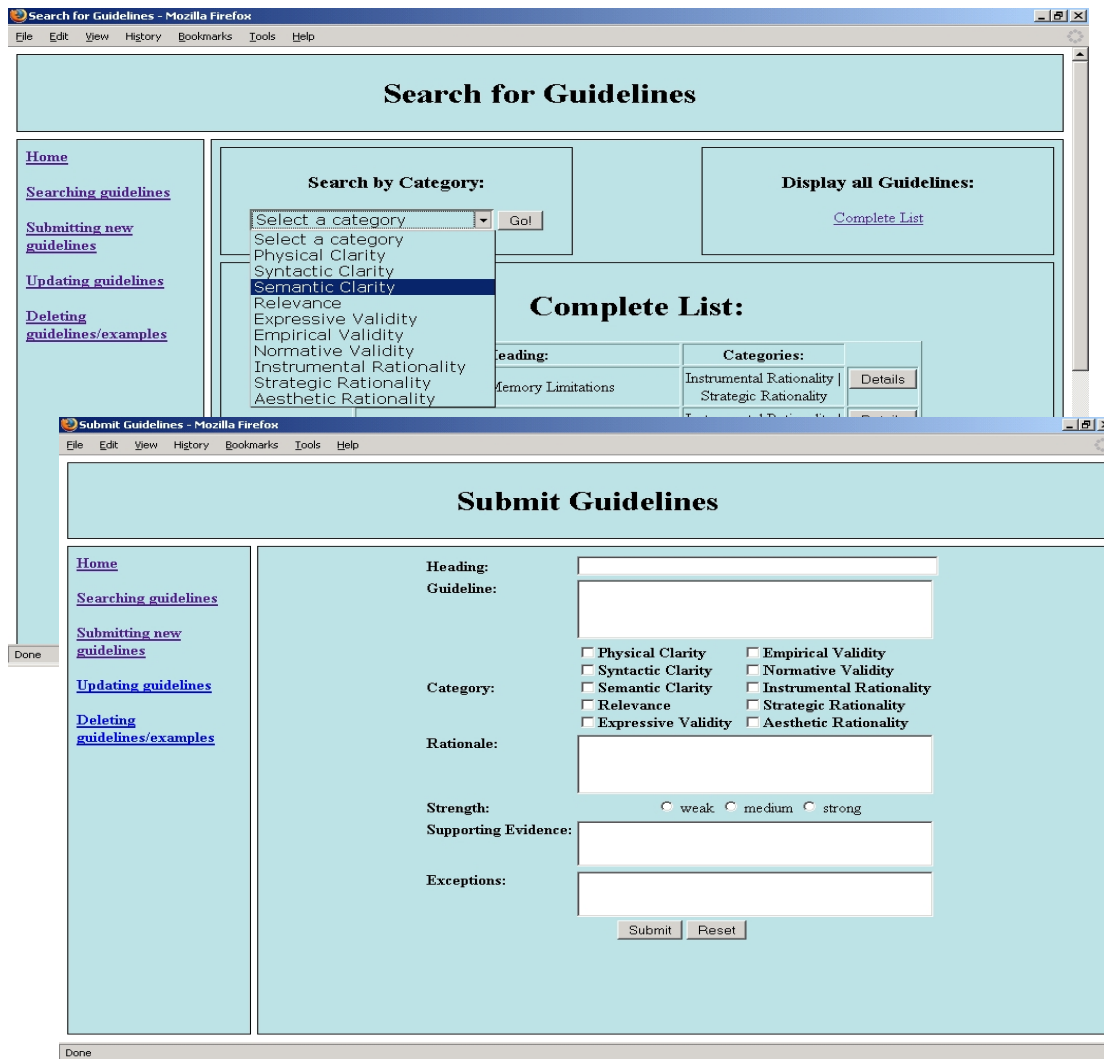
Figure 2: Screenshots for submitting and searching guidelines

have been in use for several years, also for the evaluation of a guidelines management system (Grammenos et al., 2000). For measuring the user's subjective opinion in a scenario-based situation, two types of questionnaires are typically used: an After-Scenario Questionnaire (ASQ), which is filled in by each participant after completing each task, and a Computer System Usability Questionnaire (CSUQ), which is filled in at the end of the evaluation.

Table 7: Example of a rewritten guidelines (adopted from Bock and Yetim, 2008)

Heading: Use a Single Data Entry Method

Guideline: Design data entry transactions so that users can stay with one entry method as long as possible.

Category: Instrumental Rationality

Rationale: Do not have users shift back and forth between data entry methods. Requiring users to make numerous shifts from keyboard to mouse to keyboard can substantially slow their entry speed.

Supporting Evidence: Czaja and Sharit, 1997; Engel and Granda, 1975; Foley and Wallace, 1974; Smith and Mosier, 1986.

Strength of Evidence: Strong

The details of the tasks, procedures, and results can be found in Bock and Yetim (2008). Table 8 provides a summary of the evaluation results. The result of the subjective evaluation with the questionnaires used is a set of psychometric metrics which can be summarized as follows:

- ASQ metric provides an indication of a participant's satisfaction with the system for a given scenario. The ASQ results of two tasks indicate the satisfaction of the users.

- SYSUSE metric provides an indication of the system's usefulness. The result conforms that this type of system is regarded as very useful. In addition, the users have commented that they would use it if it became available.

- INFOQUAL metric is the score for information quality. This was the best score among aspects measured in the evaluation. As the system does not provide online documentation and help-facilities, the quality of information in our prototype concerns the presentation of guidelines and system's warnings.

- INTERQUAL metric is the score for interface quality. The users found the system's interface to be simple and intuitive. Still, some users offered comments towards enhancing it.

- OVERALL metric provides an indication of the overall satisfaction score. The overall attitude of the users towards the system was positive.

Table 8: CSUQ results (range from 1, highest, to 7)	
	Means
ASQ	2,28 (first task)
	2,33 (second task)
SYSUSE	2,13
INFOQUAL	1,80
INTERQUAL	2,20
OVERALL	2,06

All in all, the evaluation of the system offered valuable insights into its functional and interactional characteristics. The results indicated ease of use and overall efficiency and effectiveness in completing the evaluation tasks. Users' comments also confirmed that there is a current need and demand for computerized tools for working with guidelines. The shortcomings identified concerned the limited documentation and online help facilities as well as aesthetic aspects of the user interface, which are to be considered in future versions of the system.

Discussion

The goal of this paper was to present a deliberation theory-based approach to the organization of guidelines, especially of those related to communication and action aspects of user interfaces or websites. For that purpose, a set of basic concepts from Habermas' (1984) discourse theory are regarded as usability categories along which guidelines can be purposefully categorized. Toulmin's argument schema is used as a model for organizing/representing guidelines, in which the categories and additional descriptions of guidelines have been integrated. To further our understanding of the model, two explorative studies have been conducted that focused on the representational fit of the categories to the domain of guidelines. Finally, the evaluation of a tool that instantiated the model for organizing guidelines provided additional insights about the users' views on the usability and usefulness of the tool; this implicitly informed us about the usability of the implemented model as well.

To discuss the research in detail, this section is organized as follows. We first discuss the merits and limitations of this research, together with a general reflection on the intricacies of theory-informed design, and then discuss its implications for the proposed model, research and practice.

Theory-Informed Design and Limitations of this Research

Theories play a significant role in design-oriented research, and theory-ingrained artifacts are valued by many researchers (e.g., Goldkuhl, 2004; Venable, 2006). It is often assumed that the application of theoretical knowledge results in an artifact with desired features. The application of one particular theory may yield a "better" artifact than the application of another theory. On the other hand, while a theory-based design is possible, a simple deductive bridge from theory into design cannot be hoped for. Moreover, the limited scope of theories may preclude adequate grounding for design decisions. Hence, Carroll and Kellogg (1989) argue for reconciling theory-based design and hermeneutics. Pointing out that HCI designs embody multiple psychological claims and that the usability aspects of an artifact can be overdetermined by independent psychological rationales, they recommend viewing artifacts not through the filter of an isolated theoretical abstraction nor without abstraction, but recognizing and analyzing the multiple, simultaneous psychological claims and theories embodied by the artifact.

Accordingly, the evaluation of an artifact may have different goals and focus on different aspects of the artifact. For example, one may test the theoretical claims involved or evaluate how well the artifact embodies the theoretical ideas or how faithfully the concepts are implemented in the models, methods, and instantiations (Venable, 2006). An artifact may also be tested to find out how well it meets the articulated requirements. Alternatively, it may be tested on how well it works in the whole context, including both articulated and agreed requirements and unarticulated and not agreed requirements in the real context (Baskerville, Pries-Heje, & Venable, 2007). A number of errors (including those related to evaluation process and social issues) may mar the evaluation of the artifact as well as the justification of the involved theories (Bakserville et al., 2007.).

Against this background, the purpose and scope of the evaluation studies conducted in this research may be viewed as narrowly defined. The artifacts presented are informed by Habermas' theory of communicative action and Toulmin's theory of argumentation. Language action theories generally provide theoretical constructs which can function as conceptual, explanatory, or value grounding for parts of design approaches (Goldkuhl, 2004). Each of the artifacts mentioned in this paper (i.e., the model for organizing guidelines and the system prototype) makes use of theoretical ideas and concepts and embodies multiple claims. Even so, the evaluations of the model artifact and the prototype were limited in that they were not comparative, that is, they did not measure whether the theories applied impart better quality to the artifact than alternative theories or whether they made a difference for the user experience of the design. Instead, the explorative studies assessed the representational fit of the categories and reported on difficulties with them when used for categorizing existing guidelines. In addition, the studies also left unanswered how well the representation schema as a whole can be used to translate the description of existing guidelines and whether its usage makes a difference. On the other hand, the evaluation of the prototype focusing on its usability and usefulness implicitly provided supporting evidence for the usability of the categories and the representation schema. Nevertheless, we encourage the choice of alternative evaluation approaches in future research in order to assess theoretical claims and/or compare the representation schema for guidelines with other schemas informed by other theories.

Beyond these deficiencies, there are also some limitations of the evaluation studies themselves. One of the limitations concerns the subjects involved in the explorative studies as well as in the evaluation of the prototype. The first explorative study considered only the author's own perspective in the classification of guidelines, whereas both the second explorative study and the evaluation of the prototype involved only the students. As a consequence, the set of guidelines has been cut down to accommodate the idiosyncrasies of the second study (i.e., students' familiarity and time).

Another limitation of the second explorative study was that participants were given less information and we did not train them. However, this limitation is somewhat alleviated in today's world of interacting with tools by relying on the self-explanatory features of the concepts.

Finally, it should be mentioned that the participants of the second study were students of the author and that this role of the author may have had an effect on the results of the interviews, conducted by the author. In addition, in the usability evaluation of the prototype, the attitudes of the participants towards the system may have been influenced by the fact that the evaluation was conducted by a student, a schoolmate of the participants, as part of his diploma thesis.

Implications for the Model

Even though the evaluations of the model for organizing guidelines and of the prototype focused on some aspects, the results allow us to formulate some implications. Firstly, the proposed categories embody many claims, such as that they cover the guidelines or are easy to use in categorizing existing guidelines. Whereas the explorative studies indicated no significant problems concerning the issue of coverage, they report on some difficulties in the use of the categories, such as distinguishing between strategic and instrumental rationality in some cases. Minimizing such difficulties may require improvements with respect to both the writing of guidelines and the categories. Concerning the categories, providing detailed descriptions or explanations of the categories, particularly in online environments, or refining the categories can make the assignment of those guidelines that are written in other formats easier. For example, the instrumental rationality would be described in more detailed form that may include different meanings of instrumentality. This would increase the chances for participants to relate the expressions in the description of guidelines to the descriptions of the categories. The same applies to other categories such as the relevance or strategic rationality.

Secondly, in our research we kept the original names of the categories in order not to lose the link to the theory. As explained to the participants, the studies did not indicate any comprehensibility problems with the wording of the categories. Nevertheless, for practitioners, the categories at the user interface level may be substituted by more user-friendly terminology (e.g., one may use truthfulness for expressive validity or efficiency for instrumental rationality).

Finally, even though the deliberation aspects of our approach are not addressed in this paper, the differences between the results of the explorative studies indicate the need for communication and the inclusion of different interpretations, particularly when categorizing guidelines that are not formatted in the way as proposed here. The second explorative study provided some support for the deliberative approach as its second stage showed that discussions among participants improved their understanding of the guidelines and led to higher agreement in the categorization results of the second stage (94.6 %).

Implications for Research

This approach has implications for research. Firstly, it provides a set of theory-based concepts that are likely to be significant factors for achieving usability, especially in the context of the Pragmatic Web (Schoop. de Moor, & Dietz, 2006). Using our online tool, researchers may conduct experiments with larger groups to further invest in the validation of the categorization framework and/or to investigate its usability.

Secondly, the most critical aspect of usability is contingent upon the actual system, and the importance of the categories is context-dependent. For example, *aesthetic rationality* might be a primary criterion for systems designed for entertainment, while efficiency and thus *instrumental rationality* is likely to be a major usability goal in the design of banking systems. Researchers

may integrate a mechanism for weighting the importance of categories such as proposed in (Venkatesh & Ramesh, 2006), add further categories, or refine the existing ones. For example, the category *relevance* can be divided into lower-level categories by distinguishing between thematic relevance, motivational relevance, etc. (Schütz, 1970).

Finally, a gap in the set of guidelines (Koyani et al., 2003) has been mentioned that relates to aspects of web credibility, namely, to *expressive validity* (in both explorative studies) and *empirical and normative validity* (in the second study). However, this does not imply an absence of such guidelines in HCI as many guidelines on the credibility of websites exist (see Table 2). On the other hand, one might speculate that such guidelines might be excluded from the set of guidelines studied, possibly due to lack of evidence, which were identified by extensive review processes during the development of this set. To fill this gap in the set of evidence-based guidelines, researchers may investigate guidelines that recommend how to design to achieve trustworthiness while communicating content.

Implications for Practice

This research has also implications for practice. Firstly, HCI is achieved with several guidelines and suggestions for effective interaction design. Although we employed a particular set of guidelines, practitioners may elaborate on the use of the model with alternative guidelines. While organizing guidelines, differentiating not only between domains (virtual communities, web pages, etc.) but also between the guidelines may be of value. For example, guidelines can be managed according to different stages of the system development, and this may ensure that explicit attention is paid to the usability concerns at each stage of the system development. As argued earlier, there are issues of comprehensibility, validity, and rationality at each stage of the system development. In this context, separation guidelines that require technical knowledge from those that are more user-oriented might help to make the participation of end users in the categorization process less difficult. In addition, guidelines can be grouped according to the characterizations, such as guidelines that relate to localization issues or are valid within specific cultural contexts can be separated from those guidelines that relate to internationalization of interfaces. This may make designers and managers aware of the diversity of cultural conventions/preferences, but also of the (in)validity or (un)acceptability of guidelines within and across contexts.

Secondly, guidelines can be expressed, communicated, and propagated effectively and efficiently. This paper has argued in favour of the use of an argument schema for representing information on guidelines. Instead of organizing guidelines as general statements around thematic blocks (e.g., navigation or homepage), practitioners may carefully distinguish between purposes (i.e., the categories), recommendations, and justifications as well as other contextual information to represent guidelines. As argued earlier, organizing information on a guideline in the form of an argument would be consistent with the discourse theory and allow rational critiques of its constituents. In addition, the general issue arises when developing guidelines, i.e., whether guidelines should be developed/written for general purposes or for more specific ones. As mentioned in the HCI literature (e.g., Scapin et al., 2000), general guidelines have the advantage that they are open to interpretation and contexts, but they are vulnerable to invalid interpretation and use in those contexts that may lack conditions under which the guidelines have been tested and validated. On the other hand, specific guidelines narrow the scope of interpretations but may prevent designers from applying them in other situations without any risk of invalidity. With regard to categorization, the general guidelines are more likely to be assigned to multiple categories as they are more open to interpretation.

Thirdly, even though the usage of the model for deliberation on guidelines has not been investigated in this paper, the approach presented so far should inspire managers to apply the model in

organizations. As guidelines continue to increase and diversify globally, organizations may profit from a collective intelligence and thus may value articulating and sharing usability wisdom that guidelines convey. Practitioners should implement a clear process of participation for the review of guidelines as well as for conversations on when and how they need to be applied. Using the deliberative model, practitioners may implement three situations for reflective conversations on guidelines (Yetim, 2006a):

- Conversations while developing and categorizing guidelines, which allow development teams to collaboratively specify or evaluate guidelines according to their comprehensibility, relevance, validity and rationality aspects.

- Conversations while applying guidelines, which allow design teams to re-assess the importance or appropriateness of guidelines for the use context.

- Conversations while using the system, which allow users to indirectly provide further critique and feedback on guidelines that the system built on.

While discussing guidelines, actors can enter discourses and provide arguments if they have at least two competing positions. The final decisions can, for example, be achieved through voting, which may provide the strength of evidence for the relevance of the guidelines in a given context. As the categories of the staircase and the discourses structure conversations on guidelines, they can serve as a memory and also be linked to further published literature on the related issue. This would allow identifying and reading the context-driven discussion and controversial positions on a particular suggestion. As Shneiderman (2003) remarks, to make a guidelines process effective, participants will have to be motivated to read it, think about it, discuss it – even complain about it. Organizations should produce an annual revision that improves the guidelines.

Finally, knowledge about the successful design of usable systems is not only described in guidelines but also in design patterns (e.g., Erickson, 2008; van Welie 2008), which have a lot in common with guidelines. Our approach can also be applied to the design patterns for interactive systems for supporting deliberation on and organizing patterns. A pattern captures design knowledge and guides the designer in using this knowledge. Although a number of different formats for writing design patterns have been suggested, a pattern typically describes a recurring problem, its context, the forces that are at play in the situation, and a solution to the problem. Reflection is considered as the most important activity in pattern writing. Practitioners may use the model for deliberation to clarify and reflect on a pattern's structure and content systematically. As each pattern represents the current understanding of what the best arrangement is for solving a recurring problem (Arvola, 2006), the empirical validity of the proposed solution (e.g., whether it is a proven solution to the stated problem) as well as its normative validity or acceptability can be challenged and negotiated. In addition, practitioners may explore the applicability of the proposed representation schema based on Toulmin's work to structure a pattern's content. The schema in Table 2 provides elements that are compatible with the basic elements of a pattern: the purposes or problems that the patterns aim to solve can be described along the categories; the solutions can be articulated as recommended actions; the justifications of the solutions as warrants, which can be backed by evidence; the descriptions of the contexts need to be added as a new category to the schema as rebuttals in Toulmin's schema express exceptions, or alternatively, one may use this field for describing the contexts (including the exceptions); and finally, one may provide optional information on the degree of strength/importance of the patterns. The tool developed for guidelines may also be of value to the pattern community as its functionalities can be easily modified and adapted in order to be employed for managing patterns.

Conclusions and Future Research

Capturing knowledge about the successful design of human-computer interaction is important for novice and experienced designers as well as usability engineers. Traditionally, this knowledge has largely been described in guidelines. Providing interface designers with guidelines is an important step in helping them to achieve the usability and accessibility of user interfaces or websites. This paper has argued that organizing and applying guidelines needs to support processes of deliberation and tradeoff for shared interaction contexts and that a deliberation theory based approach is needed to deal with guidelines in a rational way. The suggested approach uses concepts from Habermas' discourse theory and Toulmin's model of argumentation. The current paper presented the work done so far towards this overall goal.

Viewing our research from the perspective of design science research (Hevner, March, Park, & Ram, 2004; Peffers, Tuunanen, Rothenberger, & Chatterjee, 2008), we can conclude that current research has produced two designed artifacts, a model and a tool that instantiates the model. In addition to the evaluation of the model with respect to its representational fit, the initial experiences with the tool with respect to its usability indicate that the tool was valued by its potential users. The model provides research contributions to HCI literature by justifying and demonstrating the applicability of concepts from accepted deliberation theories for organizing guidelines, while the tool provides practical contributions as an artifact in helping to manage design experiences.

Concerning future research, we have already alluded to some research areas while discussing the implications of the study. The key research focus that we would like to pursue in our future work is to extend the current functionality of the prototype by adding online deliberation facilities. We will particularly explore the application of the discourses from the model in Figure 1 to allow deliberation on guidelines organized in the system. Another direction for future research is to further our understanding of the use of the framework in different contexts and among various user groups. Along the way, it is quite likely that modifications to the model will be necessary to be sensitive to application contexts (including languages). Additional definitions or refinement of categories may also need to be considered to fully understand what usability and guidelines mean to people in different countries as well as across user groups (consumer, investor, etc.).

Acknowledgements

References

Agarwal, R., & Venkatesh, V. (2002). Assessing a firm's web presence: A heuristic evaluation procedure for the measurement of usability. *Information Systems Research, 13*(2), 168-186.

Allwood, J. (2000). An activity-based approach to pragmatism. In H. Bunt & W. Black (Eds.), *Abduction, belief and context in dialog* (pp. 47-79). Amsterdam: John Benjamin Publ.

Arvola, M. (2006). Interaction design patterns for computers in sociable use. *International Journal of Computer Applications in Technology, 25* (2/3), 128-139.

Baskerville, R., Pries-Heje, J., & Venable, J. (2007). Soft design science research: Extending the boundaries of evaluation in design science research. *Proceedings of the 2nd International Conference on Design Science Research in Information Systems and Technology (DESRIST 2007)*, Pasadena, California, USA, 13-15 May 2007.

Bock, R., & Yetim, F. (2008). Entwicklung und evaluation eines web-basierten tools zum management von usability guidelines. [Development and evaluation of a web-based tool for the management of usability guidelines.] *Information Wissenschaft & Praxis,* 59(3), 175-180.

Buckingham Shum, S., & Hammond, N. (1994). Argumentation-based design rationale: What use at what cost? *International Journal of Human-Computer Studies*, *40*(4), 603-652.

Carroll, J. M., & Kellogg, W. A. (1989). Artifact as theory-nexus: Hermeneutics meets theory-based design. *ACM SIGCHI Bulletin--Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Wings for the Mind*, *20*(SI), 7-14.

Carroll, J. M., & Moran, T. P. (Eds.). (2001). Special issue on design rationale. *Human-Computer Interaction*, 6 (3 & 4), 197-442.

Corritore, C. L., Kracher, B., & Wiedenbeck, S. (2003). On-line trust: Concepts, evolving themes, a model. *International Journal of Human-Computer Studies*, *58*(6), 737-758.

Erickson, T. (2008). *The interaction design patterns page*. Retrieved May 30, 2008, from http://www.visi.com/~snowfall/InteractionPatterns.html

Fogg, B. J. (2002). *Stanford guidelines for web credibility. A research summary from the Stanford persuasive technology lab.* Stanford University. Retrieved May 22, 2009, from http://credibility.stanford.edu/guidelines/

Goldkuhl, G. (2004). Design theories in information systems - A need for multi grounding. *JITTA: Journal of Information Technology Theory and Application*, *6*(2), 59-72.

Goldkuhl, G., & Lyytinen, K. (1982). A language action view of information systems. *Proceedings of the Third International Conference on Information Systems (ICI*S), Ann Arbor, Michigan.

Grammenos, D, Akoumianakis, D., & Stephanidis, C. (2000). Integrated support for working with guidelines: The Sherlock guideline management system. *Interacting with Computers*, *12*(3), 281-311.

Habermas, J. (1984). *The theory of communicative action: Reason and the rationalization of society.* Boston, MA: Beacon Press.

Habermas, J. (1993). *Justification and application.* Cambridge, MA: MIT Press.

Habermas, J. (1996). *Between facts and norms*. Cambridge, MA: MIT Press.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly, 28*, 75-105.

Hirschheim, R., Klein, H. K., & Lyytinen, K. (1996). Exploring the intellectual foundations of information systems. *Accounting, Management and Information Technologies*, *6*, 1-64.

Hitchcock, D., & Verheij, B. (Eds.). (2005). Special issue: The Toulmin model today. *Argumentation*, *19*(3), 255-391.

Hoffman, D. L., Novak, T. P., & Peralta, M. (1999). Building consumer trust online. *Communications of the ACM*, *42*(4), 80-85.

IBM web design guidelines. (2008). Retrieved May 30, 2008, from http://www-3.ibm.com/ibm/easy/eou_ext.nsf/publish/572

Karat, J. (1997). Evolving the scope of user-centered design. *Communications of the ACM*, *40*, 33-38.

Keeker, K. (1997). Improving web-site usability and appeal: Guidelines compiled by MSN usability research. Retrieved May 30, 2008, from http://msdn.microsoft.com/en-us/library/cc889361.aspx

Klein, H., & Hirschheim, R. (2001). Choosing between competing design ideals in information systems development. *Information Systems Frontiers, 2*, 75-90.

Koyani, S. J., Balley, R. W., Nall, J. R., Allison, S., Mulligan, C., Bailey, K., & Tolson, M. (2003*). Research-based web design & usability guidelines.* NIH Publication NO. 03-5424.

Kukulska-Hulme, A. (2000). Communication with users: Insight from second language acquisition. *Interacting with Computers*, *12*, 587-599.

Linstone, H., & Turoff, M. (Eds.). (1975). *The Delphi method: Techniques and applications*. Addison Wesley Advanced Book Program.

Loui, R. P. (2005). A citation-based reflection on Toulmin and argument. *Argumentation*, *19*(3), 259-266.

Mariage, C., Vanderdonckt, J., & Pribeanu, C (2005). State of the art of web usability guidelines. In R. W. Proctor & K.-Ph.L.Vu (Eds.), *The handbook of human factors in web design*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Mayhew, D.J. (1992). *Principles and guidelines in software user interface design*. Englewood Cliffs, New Jersey: Prentice-Hall.

Neuwirth, C. M., & Kaufer, D. S. (1989). The role of external representation in the writing process: Implications for the design of hypertext-based writing tools. *Proceedings of Hypertext'89*. Baltimore, MD: Association for Computing Machinery, pp. 319-341.

Newman, S., & Marshall, C. (1992). *Pushing Toulmin too far: Learning from an argument representation scheme.* Xerox PARC technical report no, SSL-92-45. Retrieved May 30, 2008, from http://www.csdl.tamu.edu/marshall/toulmin.pdf

Nielsen, J. (2000). *Designing web usability. The art of the simplicity,* Indianapolis: New Riders Publishing.

Nielsen, J., Molich, R., Snyder, C., & Farrel, S. (2001). *E-commerce user experience.* Fremont, CA: Nielsen Norman Group.

Norman, D. A. (1990). *The design of everyday things,* New York; Doubleday.

Norman, D. A. (2004). *Emotional design. Why we love (or hate) everyday things*. Cambridge: Basic Books.

Peffers, K., Tuunanen, T, Rothenberger, M. A., & Chatterjee, A. (2008). A design science research methodology for information systems research. *Journal of Management Information Systems, 24*(3), 45–77.

Peirce, C. S. (1931-35). *Collected Papers*, Vol. I-V. C. Hartshorne & P. Weiss (Eds.). Cambridge: Harvard University Pres.

Scapin, D., Leulier, C., Vanderdonckt, J., Bastien, C., Farenc, C., Palanque, P., & Bastide, R. (2000). A framework for organizing web usability guidelines. *Proceedings of 6th Conference on Human Factors & the Web;* May 2000, Austin, Texas, USA.

Schoop, M., de Moor, A., & Dietz, J. (2006). The pragmatic web: A manifesto. *Communications of the ACM*, *49* (5), 75-76.

Schön, D. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.

Schütz, A., 1970. *Reflections on the problem of relevance*. New Haven: Yale University Press.

Shneiderman, B. (1998). *Designing the user interface*. Addison-Wesley.

Shneiderman, B. (2003). Forward to "research-based web design & usability guidelines". In Koyani et al., (2003). *Research-based web design & usability guidelines.* NIH Publication NO. 03-5424

Smith, A., & Yetim, F. (Eds.). (2004). Special issue: Global human computer systems: Cultural determinants of usability. *Interacting with Computers, 1*.

Tapio, P. (2002). Disaggregating policy Delphi using cluster analysis as a tool for systematic scenario formation. *Journal of Technological Forecasting and Social Change*, *70*, 83-101.

Te'eni, D. (2001). Review: A cognitive-affective model of organizational communication for designing IT. *MIS Quarterly*, *25*(2), 251–312.

Te'eni, D., Carey, J.M., & Zhang, P. (2006). *Human-computer interaction: Developing effective organizational information systems.* John Wiley and Sons.

Toulmin, S. (1958). *The uses of argument*. Cambridge: Cambridge University Press.

Tractinsky, N., Cokhavi, A., Kirschenbaum, M., & Sharfi, T. (2006). Evaluating the consistency of immediate aesthetic perceptions of web. *Int. Journal of Human-Computer Studies*, *64*(11), 1071-1083.

Tractinsky, N., Katz, A. S., & Ikar, D. (2000). What is beautiful is usable. *Interacting with Computers, 13*, 127-145.

Ulrich, W. (2001). A philosophical staircase for information systems definition, design and development. *Journal of Information Technology Theory and Application, 3*, 55-84.

Vanderdonckt, J. (1995). Accessing guidelines information with SIERRA. *Proceedings of Fifth International Conference on Human-Computer Interaction INTERACT'95.* London: Chapman & Hall, pp. 311-316.

van Welie, M. (2008). *Patterns in interaction design*. Retrieved May 30, 2008, from http://www.welie.com

Venable, J. (2006). The role of theory and theorising in design science research. *Proceedings of the 1st International Conference on Design Science Research in Information Systems and Technology*. Claremont, California, pp. 1-18.

Venkatesh, V., & Ramesh, V. (2006). Web and wireless site usability: Understanding differences and modeling use. *MISQ Quarterly*, *30*(1), 181-206.

Walton, D. (1996). *Argumentation schemes for presumptive reasoning*. Mahwah, N.J.: Erlbaum.

The Web Accessibility Initiative. (2008). Retrieved May 30, 2008, from http://www.w3.org/WAI/guid-tech.html

Webster, N., & Porter, N. (1913). *Webster's revised unabridged dictionary of the English language*. G.C. Merriam Company.

Yetim, F. (1998). Interkulturalität und informatische Gestaltung – Eine interdisziplinäre Annäherung. [Interculturality and design – An interdisciplinary approach]. *Informatik-Spektrum*, *21*, 203-212.

Yetim, F. (2006a). Acting with genres: Discursive-ethical concepts for reflecting on and legitimating genres. *European Journal of Information Systems*, *15*(1), 54-69.

Yetim, F. (2006b). Bridging diversity: A deliberative approach to the organization and application of usability guidelines. *Proceedings of European Conference on Information Systems (ECIS)*, June 12-14, 2006, Gothenburg, Sweden.

Yetim, F. (2008a). A framework for organizing justifications for strategic use in adaptive interaction contexts. *Proceedings of the European Conference on Information Systems (ECIS-2008).* June 9-11, 2006, Galway, Irland.

Yetim, F. (2008b). Critical examination of information: A discursive approach and its implementations. *Informing Science: The International Journal of an Emerging Transdiscipline*, *11*, 125-146. Retrieved November 11, 2008, from http://inform.nu/Articles/Vol11/ISJv11p125-146Yetim212.pdf

Yetim, F., & Turoff, M. (2004). Structuring communication processes and enhancing public discourse: The Delphi method revisited. *Proceedings of the Language-Action Perspective on Communication Modeling (LAP 2004)*, June 2-3, Rutgers University, NJ, USA.

Appendix: Categorization Results of Two Studies Concerning the Same Set of Guidelines Employed

Guideline Number & Heading	Related Categories *Note:* (1) The abbreviations are the first three letters of the first word and the first letter of the second word in the names of categories. (2) "XXX" points to guidelines excluded. (3) Shared categories of two studies are highlighted.	
	Study 1 (Yetim, 2006b)	**Study 2**
1. Design Process and Evaluation		
…		XXX
2. Optimizing the User Experience		
2.1 Display Information in a Directly Usable Format	InsR, SynC	SemC
2.2 Do not Display Unsolicited Windows or Graphics	**InsR**	**InsR**
2.3 Provide Assistance to Users	**StrR**	**StrR**
2.4 Provide Printing Options	**InsR**	**InsR**
2.5 Standardize Task Sequences	**InsR**	**InsR,** StrR
2.6 Minimize Page Download Time	InsR	*No agreement PhyC (#2), InsR (#1), StrR (#1)*
2.7 Warn of Times Outs	StrR, NorV	InsR
2.8 Reduce the Users Workload	StrR	InsR
2.9 Use Users Terminology in Help Documentation	StrR, **SemC**	**SemC**
2.10 Provide Feedback When Users Must Wait	StrR	InsR
2.11 Inform Users of Long Download Times	**StrR**	**StrR**
2.12 Do not Require Users to Multitask While Reading	StrR, NorV	InsR
2.13 Design for Working Memory Limitations	InsR	StrR
2.14 Develop Pages that Will Print Properly	InsR, **PhyC**	**PhyC**
3. Accessibility		
3.1 Comply with Section 508	NorV	PhyC
3.2 Design Forms for Users Using Assistive Technology	StrR	PhyC
3.3 Provide Text Equivalents for Non-Text Elements	**PhyC**, InsR	**PhyC**
3.4 Do Not Use Color Alone to Convey Information	**PhyC**	**PhyC**
3.5 Provide Equivalent Pages	**PhyC**	**PhyC**
3.6 Ensure that Scripts Allow Accessibility	**PhyC**	**PhyC**
3.7 Provide Client-Side Image Maps	**PhyC**	**PhyC**
3.8 Enable Users to Skip Repetitive Navigation Links	**InsR**	**InsR**
3.9 Provide Frame Titles	SemC	PhyC
3.10 Test Plug-ins and Applets for Accessibility	SemC, InsR	PhyC
3.11 Synchronize Multimedia Elements	SynC, InsR	PhyC
3.12 Do Not Require Style Sheets	**PhyC**, NorV	**PhyC**
3.13 Avoid Screen Flicker	**PhyC**	**PhyC**
4. Hardware and Software		
…		XXX

5. The Homepage		
5.1 Create a Positive First Impression of Your Site	StrR	Rel
5.2 Ensure the Homepage Looks like a Homepage	SemC	SynC
5.3 Show All Major Options on the Homepage	**Rel**	**Rel**
5.4 Enable Access to the Homepage	InsR	StrR
5.5 Attend to Homepage Panel Width	**SynC**	**SynC**
5.6 Announce Changes to a Website	Rel, StrR	?
5.7 Communicate the Website's Purpose	**Rel,** StrR	**Rel**
5.8 Limit Prose Text on the Homepage	Rel, StrR	SynC
5.9 Limit Homepage Length	Rel, StrR	InsR
6. Page Layout		
6.1 Set Appropriate Page Lengths	**InsR**	**InsR**
6.2 Use Frame When Functions Must Remain Accessible	**InsR**	**InsR**
6.3 Establish Level of Importance	**Rel**	**Rel**
6.4 Place Important Items at Top Center	StrR, Rel	SynC
6.5 Place Important Items Consistently	StrR	SynC
6.6 Structure for Easy Comparison	InsR	SynC
6.7 Use Moderate White Space	InsR	StrR
6.8 Align Items on a Page	**SynC**, AesR	**SynC**
6.9 Choose Appropriate Line Lengths	InsR	StrR
6.10 Avoid Scroll Stoppers	SemC	SynC
7. Navigation		
7.1 Provide Feedback on Users Location	InsR	SynC, StrR
7.2 Use a Clickable List of Contents on Long Pages	**InsR**, Rel	**InsR**, SynC
7.3 Do Not Create Pages with No Navigational Options	**InsR**	**InsR**
7.4 Differentiate and Group Navigation Elements	SemC	SynC
7.5 Use Descriptive Tab Labels	**SemC**	**SemC**
7.6 Present Tabs Effectively	**SemC**	**SemC**
7.7 Use Site Maps	**InsR**	SynC, **InsR**
7.8 Use Appropriate Menu Types	**InsR**, SynC	**InsR**, StrR
7.9 Keep Navigation – only Pages Short	InsR	SynC
7.10 Use Glosses to Assist Navigation	Rel, StrR	InsR
8. Scrolling and Paging		
8.1 Eliminate Horizontal Scrolling	**InsR**	**InsR**
8.2 Use Scrolling Pages for Reading Comprehension	StrR	InsR
8.3 Use Paging Rather Than Scrolling	**InsR**	**InsR**
8.4 Scroll Fewer Screenfuls	**InsR**	**InsR**
8.5 Facilitate Rapid Scrolling	StrR	InsR
9. Headings, Titles, and Labels		
9.1 Use Clear Category Label	**SemC**	**SemC**
9.2 Use Unique and Descriptive Headings	**SemC**	**SemC**
9.3 Use Descriptive Row and Column Headings	**SemC**	**SemC**
9.4 Use Descriptive Headings Liberally	StrR	Rel
9.5 Provide Descriptive Page Titles	SemC	Rel
9.6 Highlight Critical Data	**Rel**, StrR	**Rel**
9.7 Provide Users with Good Ways to Reduce Options	**InsR**	**InsR**
9.8 Use Headings in the Appropriate HTML Order	**SynC**, SemC	**SynC**
10. Links		
10.1 Provide Consistent Clickability Cues	SemC	SynC
10.2 Avoid Misleading Cues to Click	SemC	SynC
10.3 Use Text for Links	**SemC**	**SemC**
10.4 Use Meaningful Link Labels	**SemC**	**SemC**
10.5 Match Link Names with Their Destination Pages	SemC, SynC	?

10.6 Ensure that Embedded Links are Descriptive	**SemC**, SynC	**SemC**
10.7 Repeat Important Links	Rel, **StrR**	**StrR**
10.8 Designate Used Links	**InsR**	**InsR**
10.9 Link to Related Content	Rel, **InsR**	**InsR**
10.10 Link to Supportive Information	EmpV, Rel	SemC
10.11 Use Appropriate Text Link Lengths	NorV, SemC	InsR
10.12 Indicate Internal vs. External Links	InsR	SynC
10.13 Use Pointing – and – Clicking	**InsR**	**InsR**
10.14 Clarify Clickable Regions of Images	SemC, InsR	SynC
11. Text Appearance		
11.1 Use Black Text on Plain, High-contrast Backgrounds	**PhyC**	**PhyC**
11.2 Ensure Visual Consistency	AesR	SynC
11.3 Format Common Items Consistently	**SynC**	**SynC**
11.4 Use at Least 12-Point Font	**PhyC**	**PhyC**
11.5 Use Familiar Fonts	StrR	PhyC
11.6 Emphasize Importance	**Rel**	**Rel**
11.7 Use Attention-Attracting Features when Appropriate	**Rel**, StrR	**Rel**
12. Lists		
12.1 Order Elements to Maximize User Performance	InsR	StrR
12.2 Display Related Items in Lists	Rel	SynC
12.3 Introduce Each List	SemC	Rel, SynC
12.4 Format Lists to Ease Scanning	PhyC	InsR
12.5 Start Numbered Items at one	**SynC**	**SynC**
12.6 Place Important Items at Top of the List	StrR, **Rel**	**Rel**
12.7 Capitalize First Letter of First Word in Lists	**SynC**	**SynC**
12.8 Use Appropriate List Style	**SynC**	**SynC**
13. Screen-based Controls (Widgets)		
…		XXX
15. Writing Web Content		
15.1 Define Acronyms and Abbreviations	**SemC**	**SemC**
15.2 Use Abbreviations Sparingly	**SemC**	**SemC**
15.3 Use Familiar Words	Rel**, SemC**	**SemC**
15.4 Use Mixed Case with Prose	InsR	SynC
15.5 Avoid Jargon	**SemC**	**SemC**
15.6 Make First Sentences Descriptive	StrR	Rel
15.7 Use Active Voice	StrR	SemC
15.8 Write Instructions in the Affirmative	StrR	SemC
15.9 Limit the Number of Words and Sentences	InsR	SynC
15.10 Limit Prose Text on Navigation Pages	InsR	StrR
15.11 Make Action Sequences Clear	InsR, SynC	StrR
16. Content Organization		
16.1 Organize Information Clearly	InsR	SynC
16.2 Put Critical Information Near the Top of the Website	InsR	Rel
16.3 Facilitate Scanning	InsR	SynC
16.4 Group Related Elements	InsR	StrR
16.5 Display Only Necessary Information	**StrR**, Rel	**StrR**, InsR
16.6 Ensure that Necessary Information is Displayed	StrR, **Rel**	**Rel**
16.7 Format Information for Multiple Audiences	StrR	SemC
16.8 Design Quantitative Content for Quick Understanding	**InsR**	**InsR**
16.9 Use Color for Grouping	InsR, SemC	SynC
17. Search		
…		XXX

Biography

Fahri Yetim received his B.S. in Computer Science (1986), M.S. in Information Science (1989), and Ph.D. in Information Science (1994) in Konstanz, Germany. He held a DAAD Professorship (Langzeitdozentur) at the Marmara University Istanbul, German Department of Information Systems (1997-2000), was Visiting Professor at the New Jersey Institute of Technology, USA (2001-2004), and Deputy Professor (Vertretungsprofessor) at the Cologne University of Applied Sciences (2006 – 2007). Currently, he is a Senior Researcher at the University of Siegen.