

Content Preparation and Management for E-Commerce Web Sites

**ROBERT W. PROCTOR, KIM-PHUONG L. VU,
LAWRENCE J. NAJJAR, MISHA W. VAUGHAN, AND
GAVRIEL SALVENDY**

One of many opportunities afforded by the Internet is that of e-commerce, which includes transactions between businesses and customers, as well as within and between businesses. For companies to take full advantage of the potential offered by the Web, it is essential that their sites be prepared and organized in a highly usable manner. This often is not the case, as shown, for example, by the finding that users looking for an item on e-commerce Web sites cannot find it 36% of the time [6]. Companies that are able to organize and structure information in a way that promotes efficient and effective retrieval will save time and money, promote customer satisfaction and continued business, and gain an advantage over their competitors. Consequently, effective content design is crucial to the success of e-commerce sites. To help ensure that users' interactions with these sites are successful, the primary goals of content preparation should be to determine:

- What information is needed and how to elicit it
- How to structure and organize this information
- Which methods are best for retrieving the information, and
- How the information should be presented to the user.

ROBERT W. PROCTOR (proctor@psych.purdue.edu) is a professor of Psychological Sciences at Purdue University, West Lafayette, IN.

KIM-PHUONG L. VU (kvu@psych.purdue.edu) is a postdoctoral researcher in the Department of Psychological Sciences at Purdue University, West Lafayette, IN.

LAWRENCE J. NAJJAR (lawrence_najjar@yahoo.com) is a user interface designer at BMC Corporation, Austin, TX.

MISHA W. VAUGHAN (mvaughan@acm.org) is a principal usability engineer in the Usability and Interface Design Department at Oracle Corporation, Rosewood Shores, CA.

GAVRIEL SALVENDY (salvendy@ecn.purdue.edu) is a professor in the School of Industrial Engineering at Purdue University, West Lafayette, IN, and chair professor and head of the Department of Industrial Engineering at Tsinghua University, Beijing, China.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Knowledge pertinent to achieving these goals can be found in a variety of domains, including basic research on human information processing [7], applied research in computer science, management, and industrial engineering [9], and design applications of usability engineers and information architects [5]. Because this knowledge is dispersed and presented in different terminologies and contexts, cross-domain integration of the knowledge is necessary for its full potential to be realized. As a step toward such integration, an interdisciplinary white paper session on the topic of content preparation and management was organized for HCI International 2001. The participants included researchers and practitioners in many of the areas relating to the topic (see Figure 1). The benefit of the session is that the viewpoints of individuals with different perspectives and experiences could be brought to bear on the problems relating to content preparation. The resulting white paper supplied a detailed assessment of the generic problems facing content preparation, the methods available for minimizing these problems, and directions for future research [8]. This article provides an overview of the methods and recommendations from the white paper, focusing on four areas relating to content preparation and management: knowledge elicitation, organization and structure of information, retrieval of information, and presentation of information. Although our emphasis is on e-commerce, many of the methods are applicable to Web site design in general.

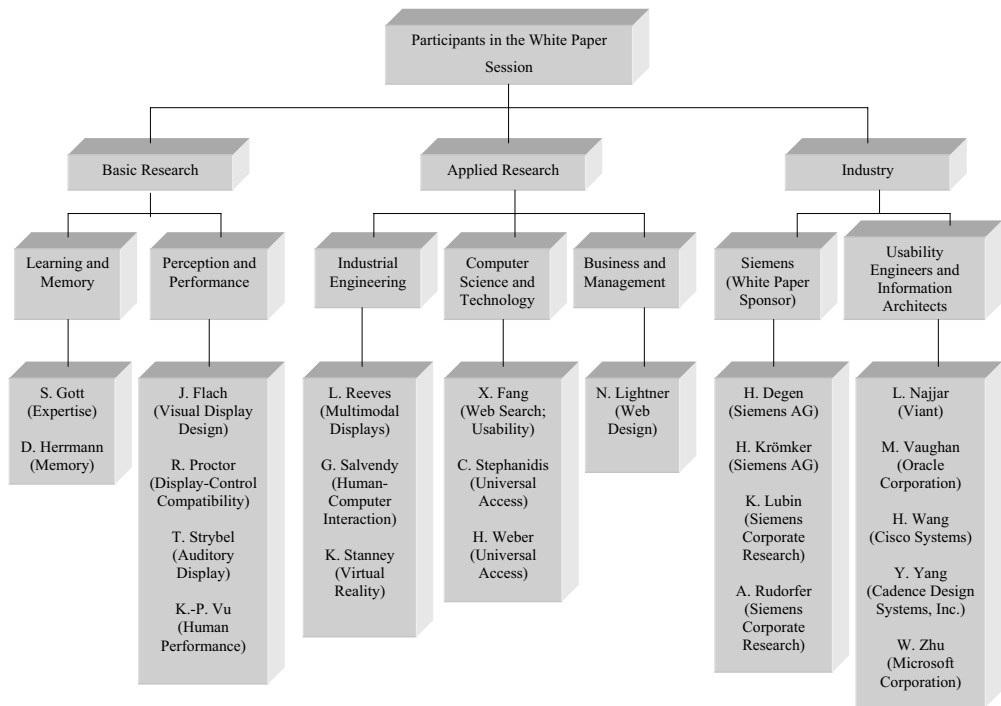


Figure 1. White paper participants.

Knowledge Elicitation

To structure the information in a Web site appropriately, one must first know what information from the domain of interest is needed for various purposes. One way to accomplish this goal is to elicit knowledge from domain experts [2]. Experts can indicate what information needs to be conveyed, how it should be structured for easy access, and what strategies are used to search and retrieve information.

Knowledge can also be elicited from a broader group of users to determine what information should be conveyed to them and how it should be conveyed. The knowledge elicitation task in this case is one of trying to understand the knowledge level of users so that information can be supplied to them effectively. Developers need to consider the knowledge that users bring for interacting with the system because this knowledge will likely predict their actions. When designing sites based on the user experience, the process commences with exploratory activities. Stanney, Maxey, and Salvendy identified key factors that should be elicited during such exploratory activities [10]:

- User capabilities and limitations;
- User requirements, roles, and responsibilities;
- Organizational factors;
- Task requirements;
- Equipment/system specifications; and
- Environmental factors.

Knowledge elicitation is not a simple task. It is often difficult to identify which individuals will provide the knowledge appropriate for a specific purpose. Because experts in different domains or areas within a domain may have dissimilar knowledge representations, care must be taken to ensure that the knowledge elicited is appropriate to its end-use. Consequently, knowledge elicitation should usually involve several experts and a range of users.

Knowledge elicitation methods are summarized in Table 1. Some methods are aimed at eliciting knowledge from experts (for example, verbal protocol analysis and critical incident reports), whereas others are more suited to novices and end-users (such as questionnaires and focus groups). Moreover, some methods are based on observations of behavior, whereas others are based on self-reports. Although there are many methods available for knowledge elicitation, a single best method cannot be identified because of their different strengths and weaknesses. Thus, it is recommended that several methods be used to increase the quantity and quality of pertinent information that is elicited.

Vaughan, Candland, and Wichansky provide an example of how different knowledge elicitation techniques can be used for specific purposes [12]. They collected data from users to assist in the design of an order-management customer application that provides a Web-accessible view of the customer's order and account status. Vaughan et al. conducted a Wants-and-Needs analysis to determine what types of information different users wanted in the Web application, a card-sorting task to determine information organization, and a questionnaire to capture the frequency with which each group performed certain tasks to determine which features should be presented at which levels of the user interface. The three types of information gathered using these techniques were successfully incorporated into the interface design.

Method	Brief Description
Interviews	Interviewer asks the expert or end-user questions relating to a specific topic
Verbal Protocol Analysis	Experts report thought-processes involved in performing a task or solving a problem
Group Task Analysis	A group of experts describes and discusses processes pertaining to a specific topic
Narratives, Scenarios, and Critical Incident Reports	The expert or end-user constructs stories to account for a set of observations
Questionnaires	User groups report information or preferences relating to a topic
Focus Groups	A group of users discusses requirements for a system
Wants and Needs Analysis	User groups/experts brainstorm about content they want and need in a system
Observation and Contextual Inquiries	Observe as users interact with a product in a natural environment
Ethnographic Studies	Observe user culture and work environment
User Diary	Users record and evaluate actions over a period of time
Concept Sorting	Users/experts establish relations among a fixed set of concepts
Log Files	Users' behaviors are logged to understand the users' interactions with the system

Table 1. Brief descriptions of knowledge elicitation methods. A more detailed version of this table that includes major strengths and weaknesses of each method can be found at www.psych.purdue.edu/~proctor/ACMtables.

There are no current standard recommendations for using specific methods of knowledge elicitation to attain different goals, although there are studies that have evaluated different knowledge techniques for specific purposes (see, for example, [1]). Moreover, the same methods may not be best for eliciting knowledge from different user profiles, such as from experts and novices. More specific evidence is needed about which combinations of knowledge elicitation methods yield the most complete sets of information for specific applications.

Organizing and Structuring Information

Organizing and structuring information is central to the success of content preparation. Creation of a highly usable navigation structure requires knowledge of the business context; content that is to be provided; and tasks and information needs of the end-users.

Each business has different goals, activities, and organizational characteristics that need to be reflected in the structure and organization of its information base. The content may differ in terms of the types of documents, the degree and nature of the structure of the information, and the format in which that information is stored. The primary users can vary in terms of computer expertise, domain expertise, information

needs, and so forth. The information structure must be designed to meet the needs of the business and the intended users. General principles include:

- *Prevent users from getting lost.* The information needs to be structured in a manner that is simple and intuitive. Users must be able to comfortably and confidently navigate through the site.
- *Communicate the structure of the site to the users.* Techniques exist to help communicate a site's structure (for example, obvious major section navigation, obvious sub-section navigation, navigational breadcrumbs, site map help).
- *Satisfy customers.* The organization and structure of the information must satisfy the needs, objectives, and preferences of users. For example, links to related Web sites or to a table comparing prices or descriptions of similar items can be added to promote users' confidence regarding the accuracy of the information.

Once information is posted to the Web, it can be modified, moved to a different location, or even deleted. This ability to constantly change and add to the Web sites results in the need to design an interface that structures the information in a manageable way, so that it can be reorganized and updated easily.

Table 2 summarizes methods used for structuring and organizing information. Some of these methods are designed specifically for improving Web sites, whereas others involve characterizing the information and facilitating end-users' interactions navigating through the site. It is important to know the mode of behavior in which the user will be engaged (such as execution of highly over-learned skills, retrieval of previously learned rules and knowledge, solving novel problems) because the nature and type of information needed varies as a function of the behavior mode.

Progress has been made in developing methods for organizing and structuring content. A major problem that remains is how to determine the fundamental elements of information for specific tasks. Perhaps the most promising approach is the statistical approach exemplified by latent semantic analysis, which represents words and larger text units as vectors in a high-dimensional semantic space constructed from the frequencies with which words occur in various contexts [4]. The value of this approach for human-computer interaction is that analysis can be performed automatically on demand. Initial tests of its usefulness for applications have been positive, but much remains to be examined.

End-users should always be involved in the design of the information architecture of a Web site because the success of the site is often measured by the ability of the end-users to interact with it to achieve goals or tasks. Usability testing should occur throughout the design process, and input from users should be incorporated. For example, card-sorting techniques, mentioned earlier, can be used initially to determine items that should be grouped and categorized together. Once these categories are incorporated into the Web site, usability tests can be performed to determine whether some items were overlooked or should be moved to different locations. As the design process continues, iterative testing cycles on prototypes can be used to determine whether modifications to the structure and organization of the site affect performance.

Method	Brief Description
Hyperlinks	Take users to a different section of the Web site or to a different Web page
Extensible Markup Language (XML)	Provides a tool to specify the content and structure of information
Semantic Web	Allows information to be linked based on its properties
Interactive Navigation Display	Uses navigation to inform users where they are in the site, how they got there, and how to get back
User Interface Standards	Provide suggestions for how a user interface should be designed and behave
Objects/Actions Interface Model	Decomposes information into a hierarchy of manageable units
Ecological Interface Design	Represents constraints in the task environment relevant to the user
Information Theory	Quantifies information in bits
Discourse Processing and Propositional Representation	Provide a way to characterize the ideas represented in sentences and paragraphs
Latent Semantic Analysis	Analyzes the meanings of words and sets of words through statistical analyses performed on a large body of text
Multivariate Analysis	Provides a representation of the dimensions for concepts
Concept Sorting	Establishes relations among a fixed set of concepts

Table 2. Brief descriptions of methods for structuring and organizing information on the Web. A more detailed version of this table that includes major strengths and weaknesses of each method can be found at www.psych.purdue.edu/~proctor/ACMtables.

Improving Web Search

The growing number of e-commerce sites available on the Web and the large amount of information available within specific sites make a powerful search engine an essential component of a Web site. Users no longer need to browse a large number of sites to find specific items. Instead, they enter search terms into a general search engine (such as Google.com) or the site's search engine (for example, Verity), and a list of Web sites or pages that are likely to contain the desired item or a list of matched items is returned. If the search is effective, the user will find the desired item. However, if the search does not return the desired hit or returns too many irrelevant hits, the user must repeat the process until the item is found. The number of searches that users have to make to find what they are looking for will affect the success of the Web site. According to Nielsen et al., 20% of e-commerce users gave up after one search query and an additional 13% after three queries [6]. If too many searches are required, the user will abandon the site and use another site to find the item.

Although search engines are important components of the Web, current search engines are, for the most part, inadequate and inconsistent with each other. For example, simple search engines may return too many irrelevant hits, but advanced search options are often hard to use. Furthermore, different general search engines return different hits for the same keyword queries. One reason why search engines are not optimally designed is that there is a lack of knowledge regarding how users search for information.

The problems associated with search engines have resulted in a tremendous amount of research looking at specific methods to remedy the problems. These methods can be classified as statistical methods, agent-based technologies, indexing tools, integration techniques, and search tools. Some methods are directed primarily at improving the design of search engines (either general Web search engines like Yahoo! or a site's search feature), and others are directed at improving users' interactions with the search engine. The methods (briefly described in Table 3) that seem to have the most promise for returning relevant hits are latent semantic analysis; agent-based technologies; meta tags; development of powerful search engines; and thesauruses.

Although methods have been developed to help improve Web searching, there is simply not enough known about how users search for information and whether their

Method	Brief Description
Natural Language Processors	Analyze meaning based on semantic relations
Latent Semantic Analysis	Analyzes the meanings of words and sets of words through statistical analyses performed on a large body of text
Adaptive Search/Agent-Based Technologies	Profile a users' past behaviors and uses this information to predict their goals
Meta Tag Tools	Index the words in a Web page
Database Search Engines	Process an indexed database for a Web site
Meta-search Engines	Submit keywords to multiple search engines
OminiSearch	Extracts useful data objects from dynamic Web pages
Powerful Search Engines	Provide fast and accurate results
Thesauruses, Dictionaries, and Alternate Spellings	Assist users to find useful search terms
Search Categories	Provide users with predefined categories to narrow search

Table 3. Brief descriptions of methods for improving Web search. A more detailed version of this table that includes major strengths and weaknesses of each method can be found at: <http://www.psych.purdue.edu/~proctor/ACMtables>.

search techniques and strategies can be generalized across different types of sites. It would be helpful to determine the strategies users adopt to help them get a relevant hit within the first few pages, especially when they have a specific search goal. It would also be helpful to determine how the design features needed to support searching for information could be integrated with those needed for completing the business transactions.

Presentation of Information

Even if software engineers can elicit the appropriate information needed to present products to users, organize and structure this information in a logical and manageable way, and design search engines that are effective and efficient in retrieving this information, they will still need to present the information in a manner compatible with the users' goals. Because the way in which the information is presented can affect the usability and success of the site, the relevant question is, what is the best way to present information?

The way that information is presented can be intuitive for designers, but may not be intuitive for end-users, making it difficult for users to interact with the site. User feedback can improve the way information is presented on Web sites, and designers can use a series of iterative design-feedback-redesign steps to create pages that are easy to use and understand. According to Karat, every dollar spent to fix a problem during early design stages saves up to \$100 to fix the same problem in later stages [3]. Despite this statistic and the numerous demonstrations of the cost-effectiveness of incorporating user input throughout the design cycle, most user interfaces are designed without significant user input.

Figure 2 summarizes different methods for designing information presentation. These methods include usability testing of products, methods for ensuring that critical items are included (see [11]), self-adaptation techniques to personalize the presentation of information, and so on. Although usability testing is often conducted late in the design process, many of the methods available focus on gathering user input earlier in the design phase. Thus, detailed evaluations of the contributions of the alternative methods applicable at early stages are needed, and emphasis should be placed on developing and evaluating methods applicable to earlier stages, when it is easier, quicker, and cheaper to make design improvements.

Because most Web sites are designed for the global market, there is an increasing tendency to allow users to customize the way information is presented (such as language) or to program the system to adapt to users' preferences (for example, product promotions) to promote personalized interactions with the system. However, users are concerned about the privacy of the personal information used to tailor a user interface, and they prefer to control the interaction. The quality of the adaptations made by the system may vary. As a result, the degree of adaptability the system can perform should be evaluated to determine the cost and benefits of these options.

Conclusion

Content preparation and management have become central concerns affecting efficient use of the Web and organizational intranets. The challenge is finding ways to prepare and manage content so that the full range of users will be able to access the information that they need quickly and easily. Content preparation is critical to the

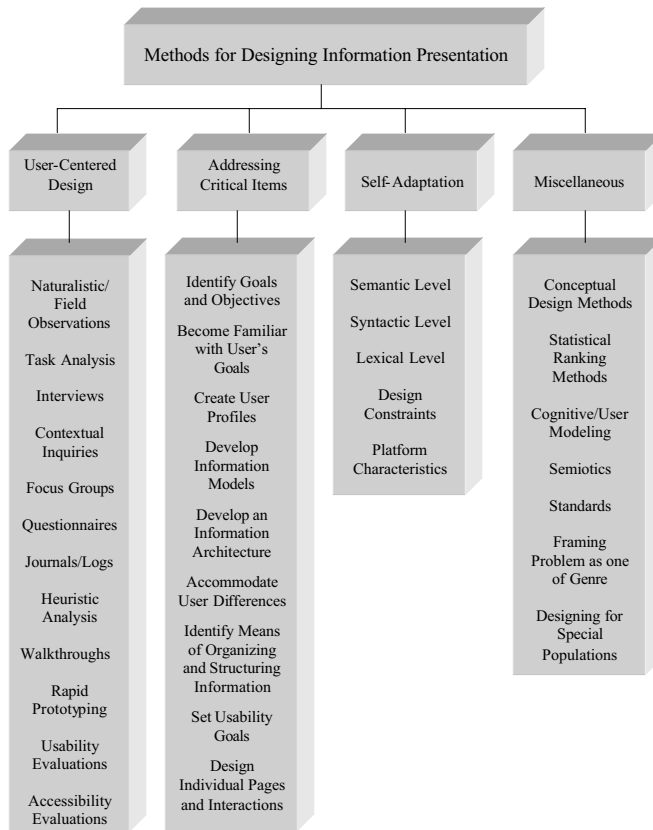


Figure 2. Summary of different methods for designing information presentation.

success of companies' electronic endeavors, resulting in the growth of numerous businesses devoted to providing this service.

In the present article, we reviewed four areas of concern with respect to content preparation and management. The first is how to elicit the knowledge and information that needs to be contained in a particular Web site or application. If the appropriate information is not identified, then the content-preparation process cannot possibly succeed. Fortunately, many knowledge elicitation techniques are available that allow identification of the information that needs to be included in the site. Assuming that this information is identified, then the second area involves organizing and structuring the content for the Web. The organization and structure of the information should reflect the context, content, and users. Numerous methods have been developed for assisting developers in determining the elements of information for specific tasks and how these elements should be combined.

Not only must an appropriate information architecture be developed, but it must be paired with an effective search engine that allows users to easily retrieve the information they desire. Much is known about ways to improve the effectiveness of search, with several specific methods showing considerable promise for further development. Finally, the user interface must be designed to provide accessibility to all possible users

and to maximize usability. The structure and organization of information has to be mapped to the interface display, and the information needs to be conveyed in a manner that promotes successful interactions with users.

At present, there is inadequate consideration of users throughout the content development and management process. There is a wealth of knowledge in many disciplines that bears on usability issues pertaining to content preparation and management. Tables 1-3 and Figure 2 summarize many of the methods and theories that can benefit content design. Figure 3 provides guidelines regarding each of the four areas covered in this article. Integration of the information from the respective disciplines in a way that will significantly advance the state-of-the-art in content preparation and management requires a concerted multi-disciplinary effort involving academia and industry, of which the work described in this article is a first step.

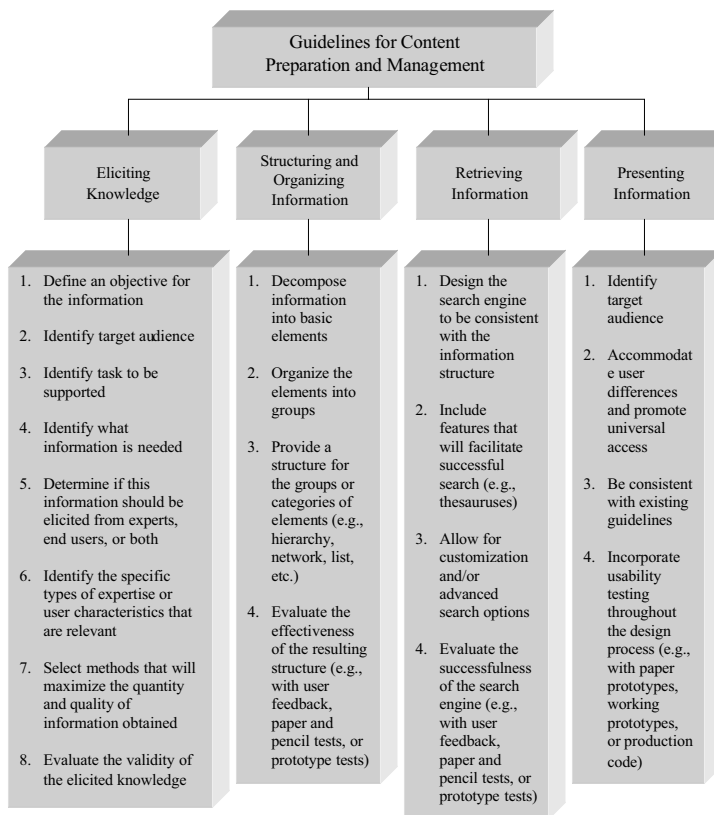


Figure 3. Content preparation and management guidelines.

References

1. Chao, C.-J., Salvendy, G., and Lightner, N. J. Development of a methodology for optimizing elicited knowledge. *Behaviour & Information Technology* 18 (1999), 413–430.
2. Cooke, N. J. Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies* 41 (1994), 801–849.

3. Karat, C.-M. Cost-benefit analysis of usability engineering techniques. *Proceedings of the Human Factors Society*, Santa Monica, CA (1990), 839–843.
4. Landauer, T. K. Learning and representing verbal meaning: The latent semantic analysis theory. *Current Directions in Psychological Sciences* 7 (1998), 161–164.
5. Najjar, L. E-commerce user interface design for the Web. In M. J. Smith, G. Salvendy, D. Harris, & R. J. Koubek (eds.), *Usability Evaluation and Interface Design: Cognitive Engineering, Intelligent Agents, and Virtual Reality 1*, 843–847. Erlbaum, Mahwah, NJ, 2001.
6. Nielsen, J., Molich, R., Snyder, C., and Farrell, S. *E-Commerce User Experience*. Nielsen Norman Group, Fremont, CA, 2000.
7. Proctor, R. W., and Vu, K.-P. L. Human information processing: An overview for human-computer interaction. In J. A. Jacko & A. Sears (eds.), *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*, 35–51. Erlbaum, Mahwah, NJ, 2003.
8. Proctor, R. W., Vu, K.-P. L., Salvendy, G., and 19 other authors. Content preparation and management for Web design: Eliciting, structuring, searching, and displaying information. *International Journal of Human-Computer Interaction* 14 (2002), 25–92.
9. Salvendy, G. (ed.). *Handbook of Industrial Engineering* (3rd ed.). Wiley, New York, NY, 2001.
10. Stanney, K. M., Maxey, J., and Salvendy, G. Socially-centered design. In G. Salvendy (ed.), *Handbook of Human Factors and Ergonomics* (2nd ed.), Wiley, New York, NY, 1997, 637–656.
11. Stanney, K. M., Smith, M. J., Carayon, P., and Salvendy, G. Human-computer interaction. In G. Salvendy (ed.), *Handbook of Industrial Engineering* (3rd ed.). Wiley, New York, NY, 2001, 1192–1236.
12. Vaughan, M. W., Candland, K. M., and Wichansky, A. M. Information architecture of a customer Web application: Blending content and transactions. In M. J. Smith, G. Salvendy, D. Harris, & R. J. Koubek (eds.), *Usability Evaluation and Interface Design: Cognitive Engineering, Intelligent Agents, and Virtual Reality 1*, Erlbaum, Mahwah, NJ, 2001, 833–837.