The Blank-Page Technique: Reinvigorating Paper Prototyping in Usability Testing

—BRIAN STILL, SENIOR MEMBER, IEEE, AND JOHN MORRIS

Abstract*—*Arguably, usability testing is most effective when integrated into the user-centered design process. One way to encourage this integration is to reemphasize the value of paper prototyping. In a recent test of a university library website, we married low-fidelity paper prototyping with medium-fidelity wireframe prototyping. When user navigation led to nonexisting pages or dead ends, users were encouraged to create what they thought should be where there was nothing. This blank-page technique gave us insights into users' mental models regarding site content and design, providing developers with useful data concerning how users conceptualized information they encountered.

Index Terms*—*Medium-level prototyping, paper prototyping, usability testing, user-centered design, wireframes.

Lane Becker has suggested that usability evaluation is not especially effective—or consistent—when employed in a single instance, as a validation tool, late in the design process. Testing like this, Becker writes, "results in a thick document that outlines everything that's wrong with a web application . . . that can't be fixed in the few weeks left before launch" [1].

Becker asserts usability testing is most effective when integrated into the user-centered design process. Rather than occurring once, used as a quality-assurance tool to evaluate a "snapshot" of the product, it offers more positive feedback when employed as early and often as possible. Genov echoes this, writing that if we think of a product's design as a controlled system, such as a thermostat, requiring constant feedback throughout its process to ensure effective production, then usability testing is more effective when broken into a series of smaller, iterative tests occurring throughout system development and even afterward when it is live or available for its intended audience(s) [2, p. 23].

One way to encourage this kind of integration is to reemphasize the value (and flexibility) of paper prototyping. As Snyder states, "One benefit of paper prototyping is that it imposes relatively few constraints on a design" [3, p. 50]. In fact, it enables users to be proactive producers and, therefore, key contributors to the design process. Paper prototyping allows users to fully express their conceptualizations and, thus, reveal their mental models of a proposed design. It also helps mitigate risk because big changes can be made without a significant cost investment [4]. Team building is encouraged, promoting communication among various stakeholders, including users [5], [6]. Further, because nothing is hard coded yet, designers are willing to make changes, and users are more comfortable suggesting them [7], [8]. The result is that real users have a substantial say, beginning early on, about the products designed for them, and this almost always translates to better usability, fostering an environment where users are a key part of the design process.

Of course, the speed of the development process and limited funding available for testing have relegated paper prototyping to the category of underused usability methods. Also, some regard paper prototyping as less useful than more developed prototypes because a paper-based design does not accurately reflect the final product [9]. The actual process of playing the role of computer, of shuffling paper around based upon different user requests while testing with paper, has also been described as too time-consuming and awkward to be effective. Leone, Gillihan, and Rauch note,

> Presenting and updating the mock-ups after the user makes a navigational decision needs to be done in a timely fashion or the users lose their flow of concentration. [8, p. 233]

Paper prototyping, they argue, results too often in "activity overload" [8, p. 233]. Consequently, a useful tool for conceptualizing is increasingly being left out of testing; this is reflected in the limited recent literature discussing paper prototyping.

Medium-level prototypes, sometimes referred to as wireframes, are used frequently as a compromise between low and high fidelity [10]. They are more tangible than paper prototypes because they are

Manuscript received April 08, 2008; revised October 31, 2008. Current version published May 21, 2010.
The authors are with Texas Tech University, Lubbock, TX 79409 USA (email: Brian.still@ttu.edu; John.morris@ttu.edu).
Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

IEEE 10.1109/TPC.2010.2046100

often presented on the screen and are clickable. But at the same time, they are not yet substantially coded, so they can be generated quickly, and changes can still be made to them that would be too expensive or time intensive for higher-fidelity prototypes. Still, for all of the positive benefits of medium-fidelity prototyping, users are primarily relegated to a reactive role when interacting with them. Any natural, user-driven navigational paths are often met with dead ends where no content has been built or links do not exist. In addition, because wireframes appear to mirror the final product (albeit without everything wired together), the ability of the design to grow based upon user input is lessened. Developers have made many of the decisions already and, thus, have created what Arnowitz, Arent, and Berger call an "overachieving prototype" really meant to be evaluated, not designed, by the user [11, p. xxix].

But if the real strength of usability evaluation lies in the multiple opportunities—at all stages of the process—users have to contribute to a product's development, how can we enable users to contribute to a product's design beyond the earliest stages of that process? How can we extend the effective depth of medium-fidelity prototyping so that dead ends can become productive means for users to aid in the design of content? How can we effectively tap into users' mental models within the context of a wireframe mockup? In a case study, we explored these issues by marrying low-fidelity paper prototyping with medium-fidelity wireframe prototyping. Since the terms "paper prototype" and "wireframe" have become somewhat ambiguous over the years, it is useful to clarify that our definitions of these concepts are consistent with those expressed by Snyder, with one reservation [3]. The wireframes also contained visual elements similar to those of composite designs (comps), such as color, artwork, and some use of nonsense words and pictures to illustrate potential layout. However, while comps are generally nonworking depictions focusing on building the graphical design elements of an interface, the prototype in the case study was primarily a working wireframe, so we will adhere to this terminology throughout.

The focus of our research was a university library website. Fifteen users representing three client-defined groups (faculty, graduate students, and undergraduate students) performed a series of tasks using a wireframe prototype of the new library website. In instances where user navigation led to nonexistent pages or dead ends, users were encouraged to create, either by writing or drawing, what they thought should be included in these blank areas. Importantly, this creation process was carried out within the context of the wireframe, which increased the relevance of user suggestions to the current application (for a discussion on the importance of context in general, see Maguire [12]). In addition, we are not proposing a strictly participatory design process in this reporting, although it can be quite useful (for a seminal paper, see Muller [13]), and it typically overlaps conceptually with paper prototyping [3]. The potential benefits of the proposed blank-page technique can be evident in even one phase of testing, as in the following case study, or throughout development, as would be the case in a participatory design approach. How this blank-page technique extended medium-fidelity and repurposed low-fidelity testing will now be addressed, followed by some of the results illustrating the technique's effectiveness.

Background

The university library formed a web redesign team in 2006 to begin overhauling its website. The redesign team was primarily interested in obtaining a better understanding of the site's users, investigating existing content and design, and gathering ideas for content not yet developed in the prototype. At that time, the team concluded usability testing would be necessary but did not know how to implement it. The team also anticipated the need for only one round of testing, planned toward the end of the process.

A meeting with library representatives at an open house held every semester in our usability research lab led to subsequent discussions and an eventual contract for services in 2007. The agreement was that we would conduct three rounds of testing, beginning first with the medium-fidelity wireframe prototype of the newly constructed library website. We would have preferred to begin testing earlier in the process, starting with a paper-prototype test. However, as previously noted, paper prototyping is not widely relied upon. Also, as providers of testing services, we had to work with the client's schedule and meet the client's goals. Still, we felt that it was valuable to solicit user feedback on the proposed design even when working with a wireframe prototype. We discussed our plans for using paper prototyping within the context of the existing wireframe prototype to gather data on areas where navigation or content had not yet been

developed. The clients showed immediate interest and agreed to the proposal.

Previous surveys and other pertinent data such as web-use logs, focus-group results, and a heuristic analysis of comparable university library sites were all available to us in the discovery stage of the testing process to determine representative users and tasks. As already noted, the client insisted we test three distinct groups: faculty, graduate students, and undergraduate students. To gain a better understanding of these groups beyond the data already provided to us, we also conducted site visits. With user profiles and task analyses completed, we began user recruitment and formulated a test plan.

There were several goals of the case study. First, the client wanted to investigate whether the labeling, search feature, and navigation elements on the prototype were usable. In addition, the client was interested in testing the usefulness of two secondary homepages, which were created to represent the user groups: one for faculty and the other for students. Another goal was to help the client gather information for portions of the prototype that had not yet been developed (i.e., the blank pages, new search feature, and so on). The blank-page technique was employed for this goal. A final goal was to identify benefits and possible drawbacks of using the blank-page technique.

METHODS

Participants We began our study, approved by our institution's Institutional Review Board, with a prescreening questionnaire used to collect demographic data (see Table I). There were 15 participants (10 female) from a variety of disciplines, including English, Information Systems, Psychology, Agriculture, Natural Resources, and Technical Communication. Participants, all of whom gave written informed consent, were volunteers solicited by university-wide, departmental, and classroom announcements. Refreshments and small gifts (i.e., pens and T-shirts) were offered after testing. Faculty members (three female and two male) ranged in age from 31 to 40 (one participant), 41 to 50 (three participants), and 51 and over (one participant). Graduate students (three female and two male) ranged in age from 21 to 30 (one participant), 31 to 40 (three participants), and 41 to 50 (one participant). Undergraduate students (four female and one male) ranged in age from 21 to 30 years.

TABLE I
PRESCREENING RESPONSES

		Faculty	Graduates	Undergrads
Age	21–30	0	1	5
	31–40	1	3	0
	41–50	3	1	0
	≥ 51	1	0	0
Gender	Female	3	3	4
	Male	2	2	1
Frequency of use of library website	Almost never	0	0	2
	< Once/ month	0	1	0
	Monthly	2	1	1
	Weekly	3	3	2
Ability to find information[a]	1	0	0	0
	2	0	1	0
	3	1	0	2
	4	3	1	3
	5	0	3	0
	6	1	0	0

[a]1 = no experience; 6 = very experienced

Regarding frequency of library website use, prescreening responses from all user groups ranged from "daily use" to "almost never" (Table I). Faculty used the site slightly more often than graduate students, and both groups used it more than undergraduates. Ratings for the ability to find information on the current site ranged from 2 to 6 on a scale of 1 to 6; faculty and graduate students averaged 4.2, and undergraduates averaged 3.6. While ratings of the ability to find information appear to correlate somewhat with frequency of site use, responses overall tended to hover in the middle of the scale, from 3 to 5.

Wireframe Prototype Wireframes provided by the client are shown in Figs. 1 and 2. The proposed design included one primary homepage (Fig. 1) and three secondary homepages. The secondary homepages were tailored for three prospective user groups: faculty, students, and distance learners. The faculty homepage displayed in Fig. 2 is an example. Distance learners were not tested during this phase, and the homepage for this user group was excluded.

To remain consistent with the main university website, which does not include a distinction for graduate students, a graduate student homepage was not created. Each of the secondary homepages prominently displayed information thought to be the most relevant for each user group.

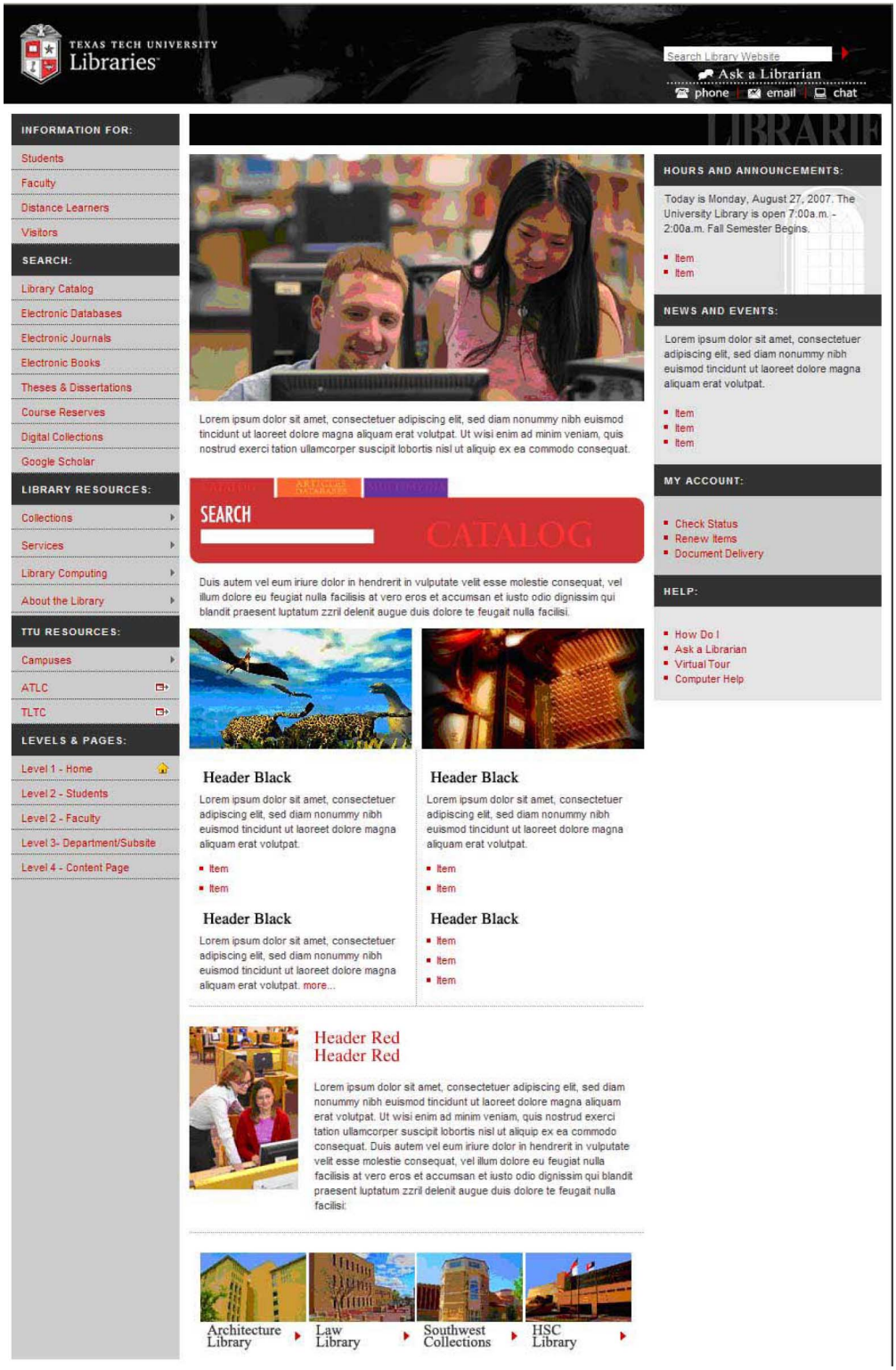

Fig. 1. Library homepage prototype.

From the client's perspective, the goal of the proposed design was to facilitate the acquisition of user-specific information through the construction of subsites.

The wireframes possessed a limited degree of functionality. Users could click on all main navigation links in menus on the left and right sides of the pages. However, links off the main page, often

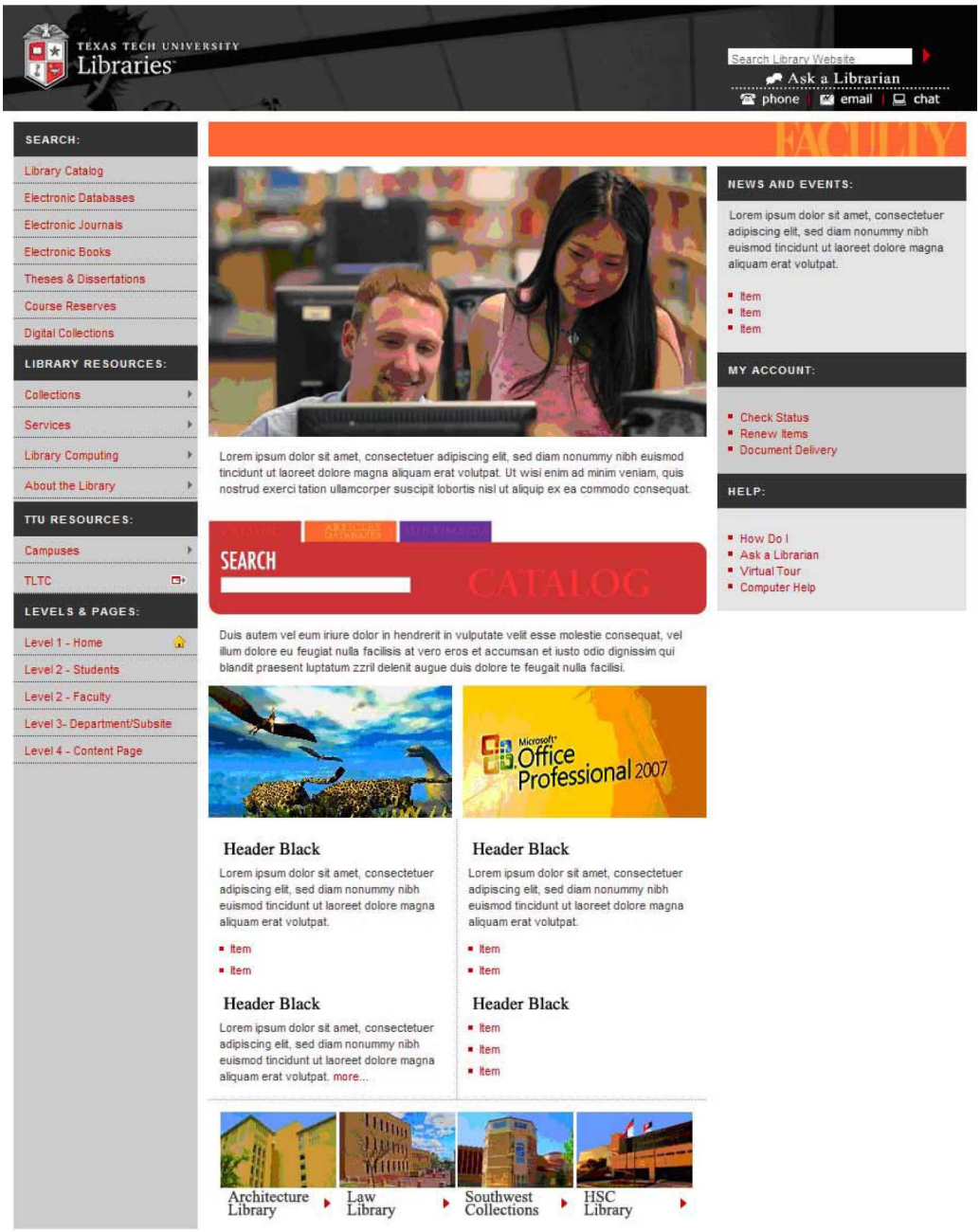

Fig. 2. Faculty homepage prototype.

appearing in dynamic drop-down menus, were in many cases not active. The search box, featured prominently in the center of the screen, was also nonfunctional. For all instances when users clicked on inactive links or attempted to perform searches, the typical dead-link page or "404 error" screen was replaced. Rather than redirecting the user, which is typical in wireframe testing, the users saw the following message on the screen:

> This section of the site you have now selected hasn't been developed.
>
> We would like your help. Please tell or show us, writing or drawing, what you would expect to find on this page.

The facilitator then provided the test participant with an 8.5×11-inch blank sheet of paper and a writing instrument. The participant was encouraged to create the page as they expected it to look; in most cases, this included what he or she thought the page would look like as well as recommendations for what the page should look like. For example, participants might explain, "I

would expect the page to look like this. But I would rather the page look like this."

Think-aloud protocol, adhering to Boren and Ramey's approach, was encouraged through all phases of testing [14]. When participants encountered the message screen, they provided commentary of their thinking while they created the page they expected to see. Then, at the conclusion of their page creation, they were navigated back to the most recent active page with content. There was no limit on the number of blank pages participants could encounter. Users were allowed to navigate the wireframe prototype according to the paths they determined.

Tasks Five scenarios and tasks were created for each user group (Table II). Some tasks overlapped between user groups to allow for group comparison. There was no time limit, since there was no optimal task-completion time at this beginning stage of the testing process. We also did not wish to interfere with the thought processes of participants as they created pages.

Testing The testing personnel included faculty, graduate students, and two staff members from the library. We elected to administer only one test per user. A facilitator worked with participants inside the testing area while other team members observed discreetly from the control room behind a one-way mirror. As participants completed tasks, video and audio data were gathered.

Once testing was completed, users were given a brief posttest interview (video and audio were also recorded during the interview). The interview consisted of 10 questions (four ratings scales and six open-ended) focusing on issues such as satisfaction, ease of use, and navigation. Users were also asked if they felt the testing process was too time-consuming or "activity intensive," since previous research suggested that one drawback to paper prototyping was that it was often too overwhelming for participants as well as administrators in terms of the complexity of operations. Interestingly, no test participants considered the process "activity overloaded," which counters conclusions drawn by Leone, Gillihan, and Rauch [8].

Pilot Study In a pilot study conducted before testing, data from six participants provided two important pieces of information. First, participants in the pilot study did not express any difficulty in understanding the nature of the blank-page creation process or the tasks to be completed. This helped verify that the methodology of the proposed technique was not too difficult to implement with real users. Second, each of the six participants completed the full round of testing in less than one hour. This confirmed that it was not necessary to impose a time limit on tasks and also facilitated the scheduling of participants for the actual study.

RESULTS

The results from testing are broken into two major sections: wireframe testing and blank-page data. After a brief presentation of the overall task times, we will relay some of the principal findings from the wireframe testing, followed by the data generated from the blank-page technique.

As displayed in Table III, task completion times ranged from about 3.7 minutes to 6 minutes across the five tasks and three user groups. Since the tasks are not necessarily the same across user groups, it is not meaningful to compare the task times directly. However, it is useful to recognize that users in the faculty, graduate, and undergraduate groups took, on average, very close to the same amount of time to complete all five tasks, at 21.6, 23.5, and 22.1 minutes, respectively. The total testing time, including introductions and posttest interviews, fell between 45 minutes and 1 hour for all participants.

Wireframe Testing Task analysis, observations, and think-aloud protocol data, as well as posttest interview results, allowed us to identify several key usability issues for the client. For example, we learned that participants experienced difficulty with some headings and labels. Several headings, such as "Digital Collections," "Document Delivery/ILL," and "Library Computing," were not intuitive. Users did not know how to use these links to complete the navigation required for successful task completion.

Hours of operation were also not sufficiently detailed, and users were further confused by different online help prompts, namely "Ask a Librarian," "How do I?," and "Computer Help." They had problems differentiating among the terms. Photos were not considered helpful,especially at levels beneath the homepage. Users wanted more details that provided answers to task prompts on deeper-level pages, although they believed the homepage—despite its successful branding of the library's mission—should also have fewer photos. As one user commented, "I know what a university library website is for and who uses it. Why do I

TABLE II
SCENARIOS AND TASKS

	Scenarios	Tasks
	Faculty	
1	You would like to include all of the library hours on your syllabi for your students, but are unsure what they are.	Locate this information on the library website.
2	You gave books checked out to you to your graduate student to return to the library. You want to verify that the books were returned.	Use the library website to find out if they have cleared your record
3	You are interested in knowing the entire list of journals the library holds in your discipline, in both print and electronic formats.	Locate that information on the library website.
4	You have submitted an article to a journal in your discipline and are told that you need to make significant revisions. The revisions require you to find additional information to back up your findings.	Use the library website to assist you in making these necessary revisions.
5	While reading an article in your discipline you find a citation for a resource that would back up your research. You find that the Texas Tech Libraries do not have this specific resource, but you have heard that the library can get this from another library.	Using the library website, locate information on how you would obtain this item.
	Graduate Students	
1	You have submitted an article to a journal in your discipline and are told that you need to make significant revisions. The revisions require you to find additional information to back up your findings.	Use the library website to assist you in making these necessary revisions.
2	While reading an article in your discipline you find a citation for a resource that would back up your research. You find that the Texas Tech Libraries do not have this specific resource, but you have heard that the library can get this from another library.	Using the library website, locate information on how you would obtain this item.
3	The professor that you TA for has heard of a database in the library that would be useful to your students. He would like you to schedule a demonstration of this database for the students.	Using the library website, schedule a demonstration for your class.
4	You have heard from other grad students that you can reserve a study carrel in the library for the entire semester, and you're interested in doing that.	Locate information on the library website about reserving a carrel.
5	You have written a report for a class, which must be posted on WebCT. You want to change the format from Word to PDF.	Use the website to see if this can be done at the library.
	Undergraduate Students	
1	You must write a research paper for a Business Administration class. The assignment requires 8–10 references from the scholarly journal literature on an administration-related topic of your choice.	Use the library website to begin your research for this assignment.
2	For your Introduction to Social Psychology class, you are responsible for reading three articles corresponding to different units within the class. The instructor states in the syllabus that the articles are available electronically through the University library website.	Locate the information on the website.
3	You have heard from your roommate that the library has copies of textbooks for large freshman- and sophomore-level classes since the bookstore is out of the text for your Linear Algebra class.	Use the library website to see if the library has a copy of it.
4	A group of students from one of your classes is getting together to study at the library.	Use the website to find out what hours the library is open on Sunday.
5	You have written a report for a class, which must be posted on WebCT. You want to change the format from Word to PDF.	Use the website to see if this can be done at the library.

need so many photos of people showing me what I already know about?"

Not all users believed the library should divide its site into role-driven sections, namely, the faculty and student areas, and three of the graduate students were unsure which section they should be using. However, if secondary homepages were necessary, some strongly suggested that content in these areas be aimed directly at fulfilling the interests of those represented by the area. In particular, two of the faculty participants expressed that they wanted more functionality to be exposed, and they did not like that so much of

TABLE III
TASK AND TEST COMPLETION TIMES, ROUNDED TO THE TENTH OF AN HOUR

Task	Faculty	Graduates	Undergrads	Avg. Task Time
1	4.0	5.1	5.4	4.8
2	3.8	3.7	4.1	3.9
3	4.8	6.0	4.0	5.0
4	4.0	5.0	3.7	4.2
5	5.0	3.7	4.9	4.5
Total	21.6	23.5	22.1	22.4

the content space was filled with pictures—what one faculty participant referred to as "nonpertinent information." (See Fig. 2.)

Interestingly, the client design team was surprised to learn that users, in their search for information, rarely began navigation on secondary homepages, which were intended to cater to specific roles. (See Fig. 2.) Further, when users did take these navigational paths, they quickly became confused. They then returned to the main homepage to reorient themselves and continue the task. Occasionally, users avoided traditional navigation with links and instead used the site's search box, even though they were informed that it was not functional. The majority of the participants indicated they would have done this right away, but they expressed frustration that the main search feature was not easy to find, hidden by too many photos or other content clutter. They also regarded the tabs above the search feature, which included catalog, articles/database, and multimedia, as confusing. (See Figs. 1 and 2.) In all, twelve participants (five faculty, four undergraduate, and three graduate students) initiated searches using one of the search engines (the main search feature in the center or the "Search Library Website" feature in the upper right-hand corner) during testing. At the client's request, we rerouted participants who used the "Search Library Website" feature to the homepage because this element was controlled by a third party.

TABLE IV
BLANK PAGES ENCOUNTERED BY USER GROUP

No.	Blank pages	Faculty	Graduate	Undergrads	Totals
1.	Electronic Databases	1	2	1	4
2.	Electronic Journals	2	0	2	4
3.	Hours	5	0	3	8
4.	How Do I?	1	1	3	5
5.	Computing Help	0	1	0	1
6.	Workstations	0	1	1	2
7.	Articles/Database Tab	2	3	3	8
8.	Check Status in My Account	5	0	1	6
9.	Collections	0	0	1	1
10.	Printing in Library Computing	0	0	2	2
11.	Document Delivery	3	1	0	4
12.	Instruction Tours	0	4	0	4
13.	Interlibrary Loan	2	2	0	4
14.	Software Page of Library	0	1	0	1
15.	Ask a Librarian	1	0	0	1
16.	Course Reserves	0	4	5	9
17.	Library Catalog	5	2	3	10
18.	Technology	0	1	0	1
19.	Theses and Dissertations	1	0	0	1
20.	Tour and Announcements	0	0	1	1
	Total	**28**	**23**	**26**	**77**

Blank-Page Technique As users clicked on the search box or navigated to a page that did not exist, they were directed to a blank page with the message indicating that the page was not complete but that they could help us design it through their comments or drawings. Given how participants navigated, often at odds with expected navigational routes that the wireframe allowed and relying heavily on search features, they encountered 77 blank pages during testing. Normally, of course, participants would be redirected back to a navigation path that would allow them to complete the task. But our method of asking users to provide their input on the blank pages yielded substantial insights.

The frequency of blank-page selection was comparable across the user groups (see Table IV), with faculty, graduate, and undergraduate students selecting blank pages on 28, 23, and 26 occasions, respectively. Not surprisingly, the blank pages selected most often ("Library Hours,"

"Articles/Databases" search option, "Check Status in My Account," and "Course Reserves") were directly related to tasks being conducted. Therefore, blank pages were selected most often only when users sought to perform tasks representative of those typically performed by all faculty and student users of the library website.

Much of the user-suggested content for the blank pages, especially when the pages were drawn by hand, cannot be effectively represented quantitatively. However, in some cases, we were able to tally items suggested as content. For example, the items listed in Table V represent user-suggested content for the "Articles/Databases" search option.

In many cases, more than one user had the same suggestion. For example, two users suggested a search box for the "Article/Databases" page, and three additional items were suggested regarding the search box. Some comments focused on how the search should be structured, suggesting it possess "Boolean capabilities." Others suggested how items should be displayed after searching: "First I would expect to see a list of databases resulting from the search." Others focused on the design of the page, writing that the search-returns page should "use color for links to signify format (i.e., PDF, HTML, etc.)."

In total, 189 user-suggested content items were generated by using the blank-page technique. These items focused on a variety of subjects, ranging from electronic journals ("If I login, it should recall my previous searches") and work stations ("What kinds of programs are installed?") to document delivery ("There needs to be some sort of tutorial to walk us through how to use it") and instructional tours ("These should be divided up and categorized according to discipline and/or role").

Yet another example of data generated from blank-page commenting can be found when focusing on the multimedia tab above the search box in the middle of the prototype screen. Developers had not yet determined what content should be placed there, so all of the insights provided through the blank-page technique helped the client conceptualize just what was meant by a multimedia search. Several items were repeatedly suggested as content for the multimedia search, including video (ten participants), equipment available (eight participants), availability of multimedia resources (four participants), video courses (four participants), lectures (four participants), and podcasts (three participants). We believe these suggestions were facilitated by the blank-page technique. That is, by having to draw their suggestions, participants were able to better express what information they felt would be provided on this page.

TABLE V
WRITTEN/VERBAL COMMENTS PROVIDED FOR ARTICLES/DATABASE TAB ON SEARCH

Information on how to find an article
Just like current webpage design
Participant assumes it has articles (x 2), journals (x 2), and databases
A full list of everything subscribed to
Organize journals according to search
Wants to select databases to search
Search box (x 2) • To search for databases based on a particular subject (i.e., keyword, search term) • More specific search box to narrow search terms • Boolean capabilities
Search Results (x 4) • "Your search for 'search term' returned…" • First would expect to see a list of databases resulting from the search • List of citations (x 4) with first few lines of text (x 2) • Linkable title (x 2) • Sort by author, date, title (x 2) • Bolded or highlighted search terms (x 2) • Ranked by relevance (in percentage) to search terms (x 2) • Use color for links to signify format (i.e., PDF, HTML, etc.) • Icons for full text results (x 2), PDF, HTML, etc. • Want to know readily if results are full text • Want to know if items are available in holdings • Separate results by full text or nonavailability • "Request" button for items not in holdings • Date

Clearly, the number and nature of comments reveal that as useful as medium-fidelity prototyping is, a great deal of user-driven information goes undiscovered in later stages of testing. Paper prototyping as an additional method, augmenting

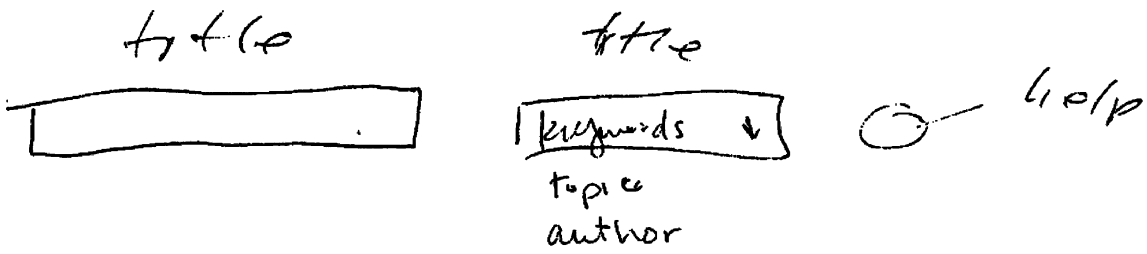

Fig. 3. Participant 9's (graduate student's) Library Catalog drawing.

the existing medium-fidelity wireframe testing, garners user insights. This not only offers developers additional useful data regarding how users conceptualize information, but it also encourages the positive, proactive user involvement that Becker, Genov, and others suggest is crucial for effective usability testing and, as a result, user-centered design.

User Drawings In addition to writing their suggestions, users sketched what they expected to see. These drawings were enlightening and allowed users to act as designers, since they not only offered verbal or written feedback but attempted to make their visions operational by showing what they felt the page should look like [15]. However, and consistent with what others have mentioned, we did not attempt to solicit these drawings as a design proposal [16]–[18]. Rather, our goal was to push users to express their thoughts and expectations, mental models, of what they expected to encounter. Granted, user expectations are likely affected by past experiences with sites of varying usability, but simply soliciting verbal responses, which are also likely affected by the same previous experience, would not necessarily compel users to realize their design suggestions fully. That is, responses could be made without having to realize their implications.

Figs. 3–5 show examples of participant drawings. These drawings were chosen to illustrate the varying levels of detail exhibited by participants. A graduate student's drawing of the Library Catalog blank page (Fig. 3) was much less detailed than an undergraduate participant's drawing for the same page (Fig. 4). Another drawing, created by a graduate student for the Instruction Tours blank page (Fig. 5), illustrates a high level of complexity. Of course, these distinctions are only subjective at this point since we have not yet instantiated a scheme to adequately and consistently categorize the drawings according to complexity or detail. This categorization process may be quite beneficial for future reporting, but it is beyond the present scope of introducing the blank-page technique.

Delivery to Client Key findings from the blank-page drawings, written and verbal comments, and the principal findings from the wireframe testing were analyzed using affinity diagramming. Recommendations on key issues were then given to the client. In addition, all written comments and drawings were included in the appendices of the final report.

DISCUSSION

The blank-page technique helps generate good, user-centered design, and providing data not normally available means that this technique also makes wireframe prototyping, as well as usability testing in general, more effective. The wireframes present a context within which the blank-page paper prototyping takes place. This method not only allows usability testing of elements of the proposed page design, but it also allows users to contribute to the design of pages that have not yet been created. Some benefits, as well as potential drawbacks, are discussed below.

Benefits The data-collection time was surprisingly fast. In fact, the greatest average amount of time expended on any one task in the case study did not exceed 6 minutes, and the average time spent on all five tasks across the three user groups was only 22.4 minutes (Table III). This left plenty of time for pre- and posttesting, introductions, and think-aloud training. The total time that participants spent at our facility was between 45 minutes and 1 hour, which is reasonable and suggests that using the blank-page technique will not dramatically inflate testing times.

In this short amount of time, users revealed aspects of their mental models through the blank-page drawings that likely exceeded what would have been captured from verbal responses. A major benefit of users creating (drawing) the pages is that it

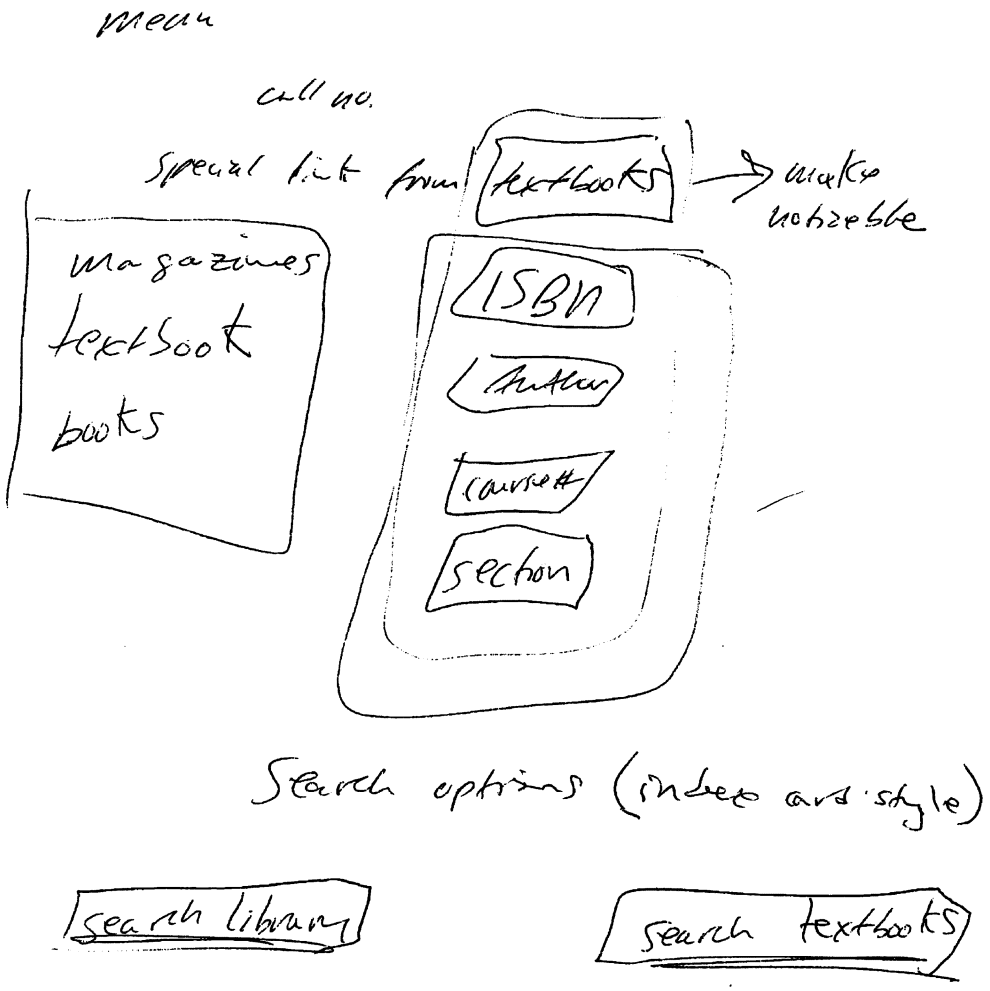

Fig. 4. Participant 3's (undergraduate's) Library Catalog drawing.

forces them to think more about their expectations. Simply asking participants what they would expect to find, on the other hand, does not necessarily put them in a position of having to realize their design suggestions. However, this is not to suggest that we should design pages exactly as the participants have drawn them. Rather, the information is useful for illustrating how a user perceives a particular concept, or linked page in this case. It might be that one or more users perceive a certain page to contain several layers, a hierarchical design, buttons or links with specific labels and locations, or certain items of content. The blank-page technique produced a large amount of data that informs this process. In all, 77 blank pages were encountered/created during testing, resulting in 189 items of user-suggested content. While other factors such as users' conceptualizations of page structures, navigational elements, and layout were not analyzed directly in this case study, the drawings were presented to the developers. These developers are not trained in usability evaluation. Consequently, even though they reported to us that the raw blank-page data were useful to understanding users and reconciling design issues, we suggested to them that any changes made based on these data should be further tested. In fact, we believe it is necessary to conduct additional usability testing on any newly designed pages if designers with limited or no usability training develop them based on blank-page data.

As a proof of concept, it is hoped that our study also suggests that paper prototyping, as realized through the blank-page technique, is not only a low-fidelity, early-stage method of usability evaluation. It can be used at times other than the earliest stage of the development process, enabling developers who are, for example, dependent on wireframe prototyping to obtain additional, user-centered data. This allows them to anticipate

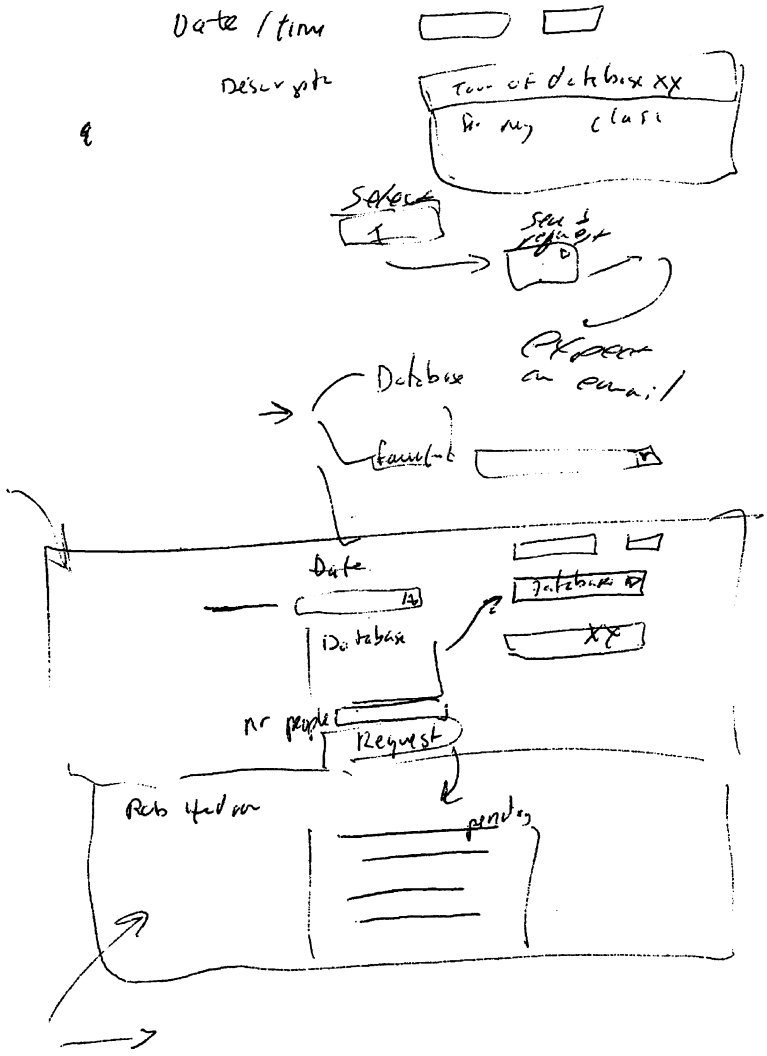

Fig. 5. Participant 12's (graduate student's) drawing of Instruction Tours.

and then design more usable products, such as websites.

Consequently, the blank-page technique is useful for more than just encouraging user feedback. Since this feedback is not ignored, it also prevents any type of cognitive tunneling among developers. When users have a strong voice in the process, it is difficult for developers to assume what users want and design accordingly, which is something they have a tendency to do. As Beaudouin-Lafon and Mackay write,

> In our experience, programmers often argue in favor of software prototypes even at the earliest stages of design. Because they already are familiar with a programming language, these programmers believe it will be faster and more useful to write code than to "waste time" creating prototypes. [19, p. 1008]

Paper prototyping, however, mitigates this. Developers cannot focus only on what they feel is important when user input is facilitated at every stage.

But what about late in the process? Can the blank-page technique be used effectively as an augmentation to higher-fidelity prototypes, ones that are far more developed and closer to live release? Certainly, any significant changes that users might suggest would be difficult to implement. But if we believe that websites, along with all sorts of other products, are complex systems, as Redish notes, requiring more nuanced approaches for evaluation, then allowing opportunities for users to continue contributing to the design of a product at later stages of prototype development should not be discouraged [20]. Where's the harm? Besides, as Morgan and Borns's work with eBay's website indicates, usability evaluation should

not be a one-time snapshot [21]. It starts early, continues through all phases of prototyping, and then continues past that as user data are gathered after a site is operational.

Drawbacks Even though no users indicated they were overwhelmed by playing the roles of designer and evaluator, it is likely that the time taken to comment on all blank pages, either writing or drawing, was longer than typical wireframe testing. Use of the blank-page technique provided a tremendous amount of data. Although this is generally a good thing, more data often means more analysis time. This could be problematic if there is little turnaround time in the project environment. We propose two ways to address this. First, if time is of the essence, the focus should be on the most critical areas of the site to be developed: those areas in which there are the most unknowns or questions. Second, this issue should be addressed when suggesting the method to the client. Conceptual research and development always takes time, and the blank-page technique is part of this conceptual process. Understanding this from the onset may alleviate some time pressures.

Besides, what can be gained from qualitative feedback (via think-aloud protocol)? There is also no way to measure how efficient a website is when using the blank-page technique. In that sense, this method functions much like any paper prototype. However, it does yield supplementary, proactive data that go beyond efficiency by showing how users conceptualize what they encounter.

Future Directions

At this point, using the blank-page technique for website design is promising. However, it is not known how or even if this technique would be useful to other realms, such as software design. Also, for the less technologically inclined, it is not clear exactly how the proposed technique would affect the end design. It might be expected that the technique would be useful for older users, for instance, because it would help them portray their mental models of the design, which could be quite different than the conceptualizations of other user groups.

Therefore, future research is needed to test the applicability of the technique in different realms and for different user groups. In addition, because only one case study is presented, it would be useful to compare these results to those from a follow-up study. This could reveal if there is any consistency in research findings. Finally, due to time constraints, we were not able to address the complexities in participants' drawings beyond the illustrations provided earlier (Figs. 3–5). It would be useful to establish a methodology for quantifying the varying levels of complexity by previously discussed factors such as hierarchical design elements, buttons or links with specific labels and locations, and the number of content items. It might then be possible to compare varying levels of complexity with user characteristics, such as age and experience, as well as with task characteristics.

The use of paper prototyping within the context of wireframes poses some specific, interesting questions. For instance, how does the wireframe context actually affect the paper-prototyping process? To what extent does the given context shape user responses? Should blank pages include navigational elements common throughout the wireframe, such as the banner and navigation column? Or should the pages be truly blank?

Conclusion

Perhaps paper prototyping cannot help eliminate the inconsistency in results that the Molich et al. article on the comparative evaluation of Microsoft's Hotmail so eloquently demonstrates [22]. But its frequent use, beginning early and employed throughout as a means of proactive user involvement, does encourage a user-focused design and evaluation process—a process not meant to cure all of a product's ills, but rather one that places users front and center, making them active participants. This technique makes testing as much about design as evaluation. It also promotes multiple, early tests as well as continuous system monitoring to keep it working effectively, not only with the guidance of developers, but with sustained, in-depth feedback from those who matter most—the users.

Acknowledgments

The authors would like to thank the anonymous reviewers for their feedback.

References

[1] L. Becker. (2004, Jun.). 90% of all usability testing is useless. *Adaptive Path.* [Online]. Available: http://www.adaptivepath.com/ideas/essays/archives/000328.php

[2] A. Genov, "Iterative usability testing as continuous feedback: A control systems perspective," *J. Usability Studies*, vol. 1, no. 1, pp. 18–27, 2005.
[3] C. Snyder, *Paper Prototyping: The Fast and Easy Way to Design and Refine User Interfaces*. London, UK: Morgan Kaufmann, 2003.
[4] C. Snyder. (1996). Using paper prototypes to manage risk. *User Interface Engineering*. [Online]. Available: http://www.uie.com/articles/prototyping_risk/
[5] M. Rettig, "Prototyping for tiny fingers," *Commun. ACM*, vol. 37, no. 4, pp. 21–27, 1994.
[6] T. Scanlon. (1998, May). Paper prototypes: Still our favorite. *User Interface Engineering*. [Online]. Available: http://www.uie.com/articles/paper_prototyping/
[7] H. M. Grady, "Web site design: A case study in usability testing using paper prototypes," in *Proc. IEEE Int. Professional Communication Conf.*, 2000, pp. 39–45.
[8] P. Leone, D. L. Gillihan, and T. L. Rauch, "Web-based prototyping for user sessions: Medium-fidelity prototyping," in *Proc. 44th Int. Technical Communications Conf.*, Toronto, ON, Canada, 1997, pp. 231–234.
[9] J. Rudd, K. Stern, and S. Isensee, "Low vs. high-fidelity prototyping debate," *Interactions*, pp. 76–85, 1996.
[10] D. Engelberg and A. Seffah, "A framework for rapid mid-fidelity prototyping of web sites," in *Proc. IFIP 17th World Computer Congr.*, Deventer, The Netherlands, 2002, pp. 203–215.
[11] J. Arnowitz, M. Arent, and N. Berger, *Effective Prototyping for Software Makers*. San Francisco, CA: Morgan-Kaufman, 2006.
[12] M. Maguire, "Context of use within usability activities," *Int. J. Human-Comput. Studies*, vol. 55, no. 4, pp. 453–483, 2001.
[13] M. Muller, "Participatory design: The third space in human-computer interaction," in *The Human Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*, J. A. Jacko and A. Sears, Eds. Mahwah, NJ: Lawrence Erlbaum Associates, 2003, pp. 1051–1068.
[14] M. T. Boren and J. Ramey, "Thinking aloud: Reconciling theory and practice," *IEEE Trans. Prof. Commun.*, vol. 43, no. 3, pp. 261–278, Sep., 2000.
[15] J. Löwgren and E. Stolterman, *Thoughtful Interaction Design: A Design Perspective on Information Technology*. Cambridge, MA: MIT Press, 2004.
[16] J. Nielsen, *Usability Engineering*. San Mateo, CA: Morgan Kaufmann, 1993.
[17] A. Druin, J. Montemayor, J. Hendler, B. McAlister, A. Boltman, E. Fiterman, C. Plaisant, A. Kruskal, H. Olsen, I. Revett, T. Plaisant-Schwenn, L. Sumida, and R. Wagner, "Designing PETS: A personal electronic teller of stories," in *Proc. SIGCHI Conf. Human Factors in Computing Systems*, 1999, pp. 326–329.
[18] J. Blomberg and A. Henderson, "Reflections on participatory design: Lessons from the Trillium experience," in *Proc. SIGCHI Conf. Human Factors in Computing Systems*, 1990, pp. 353–359.
[19] M. Beaudouin-Lafon and W. Mackay, "Prototyping tools and techniques," in *The Human Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*, J. A. Jacko and A. Sears, Eds. Mahwah, NJ: Lawrence Erlbaum Associates, 2003, pp. 1006–1031.
[20] J. Redish, "Expanding usability testing to evaluate complex systems," *J. Usability Studies*, vol. 2, no. 3, pp. 102–111, 2006.
[21] M. Morgan and L. Borns, "360 degrees of usability," in *Proc. CHI Extended Abstracts on Human Factors in Computing Systems*, Vienna, Austria, 2004, pp. 795–809.
[22] R. Molich, M. Ede, K. Kaasgaard, and B. Karyukin, "Comparative usability evaluation," *Behav. Inf. Technol.*, vol. 23, no. 1, pp. 65–74, 2004.

Brian Still (SM'08) is an assistant professor at Texas Tech University, Lubbock, where he teaches technical communication and directs the school's Usability Research Laboratory.

John Morris is currently pursuing the Ph.D. degree in human factors psychology at Texas Tech University, Lubbock. He is President of the Human Factors and Ergonomics Student Chapter at Texas Tech. His research interests include usability testing methodologies, website navigation, aerospace, and the psychology of perception.