Design implications from a usability study of GramStain-Tutor™

Sara Kim, Douglas M Brock, Adam Orkand, and Michael L Astion

Sara Kim is in the Department of Family Medicine, Douglas Brock is in the Department of Medical Education, and Adam Orkand and Michael Astion are in the Department of Laboratory Medicine, at the University of Washington. Address for correspondence: Dr Sara Kim, Department of Family Medicine, Box 356390, University of Washington, Seattle, WA 9819 USA. Tel: +1 206 543 9425; fax: +1 206 543 3821; email: sarakim@u.washington.edu

Abstract
To examine interface issues in an educational software program, we recruited twenty-two users in health sciences departments for a usability study. Using the GramStain-Tutor™ (GST) CD-ROM, these users performed tasks in navigating the program and using various interactive features. Their actions and comments were recorded with a digital video recorder and an audio recorder. The main results of our study are: (a) the majority of the users did not use the multiple navigational options available in the program; (b) navigational patterns differed based on the training background and content knowledge of the users; (c) several visual cues critical to program use were not discovered by the majority of the users; and (d) icons representing specific functions were not intuitive from users’ perspectives. We demonstrated how a usability study of a relatively small number of users can identify specific problems in interface design. We recommend these interface problems be addressed before conducting educational studies that examine how educational software programs affect student learning.

Usability testing
Usability testing is a collection of evaluation methods that is increasing in popularity. It is based on observing how users perform tasks, and obtaining feedback from users through formal tests or compilations of user comments (Nielsen, 1993). These results lead to a set of recommendations that are fed back into the design process with the end goal of improving the product. Thus, usability is an iterative process of collecting response data or user feedback on a product, re-designing the product in response to that feedback, and re-testing the product to examine whether changes have enhanced the usability.

© British Educational Communications and Technology Agency, 2001. Published by Blackwell Publishers, 108 Cowley Road, Oxford, OX4 1JF, UK and 350 Main Street, Malden, MA 02148, USA.
In a typical usability session, the following issues are explored (Nielsen, 1993; Salvemini, 1998):

- Can users easily learn a computer system supported by sufficient navigational control?
- Is the system efficient to use or is usability hindered by an excessive number of icons and menu options?
- Are system features easy to remember so that time to re-learn the system is minimized?
- Does the system have low errors and efficient error recovery mechanisms?
- Is the system pleasant to use?

These questions can be answered by having users complete tasks while using the computer system. Tasks refer to a set of representative actions routinely performed in a particular working context. Often, users of different levels of content expertise and computer skills are asked to carry out these tasks. Using as few as 4-5 subjects who participate in a task-based study, 80% of usability problems can be identified (Virzi, 1992).

Usability techniques include both qualitative and quantitative approaches, such as observations, think aloud protocols, questionnaires, interviews, logging actual use, and user feedback through debriefing (Nielsen, 1993). Typically, a checklist, audio recorders and video recorders are used in observation sessions. Users are encouraged to articulate their thought processes while performing tasks. This “think-aloud” process enables the observer to understand the nature of errors committed while performing tasks and the reasons for the errors (Lewis, 1982). User actions captured by a video recorder help examine the proportion of users who complete the tasks correctly, the time it takes them to complete the tasks, and the types and number of errors users commit. Questionnaires are popular usability instruments for collecting feedback from a large pool of users. These users rate the quality of the design layout and embedded features, including the navigational structure, the aesthetic appeal of the product, and the readability of text and images. These data can reveal specific interface elements that need to be re-designed.

Navigation in multimedia programs

Designing an effective and efficient navigational system is recognized as the most difficult challenge in developing multimedia products (Brickell, 1993; Jonassen, 1992; Koneman and Jonassen, 1994). The following list outlines key questions related to navigation in multimedia programs:

- How should the information space be organized so that a learner can navigate the content by following meaningful paths? (Lyardet, Rossi and Schwabe, 1998)
- To what extent should learners have flexibility in exploring a program’s contents (learner control) or should programs predetermine a learner’s navigational pathway? (Jonassen, 1990; Lawless and Brown, 1997)
- How should navigation be designed so that learners are aware of where they are within the system? (Brickell, 1993)
• What patterns of navigation emerge as learners explore a program? And do these patterns relate to individual differences and learning styles? (Fenley, 1998; Jonassen, 1992; Wild and Quinn, 1998)

• What types of navigational tools or guidance should be provided to the learners? (Jonassen, 1992).

Some of these research questions have been explored through built-in tracking features that record the paths learners take while interacting with the programs. For example, Lawless and Brown (1997) reviewed studies that used computerized tracking systems to examine navigational patterns and identified three user profiles: (a) knowledge seekers (rely on systematic strategies in covering program contents; (b) feature explorers (mainly focus on audio and visual interactive features); and (c) apathetic users (follow navigational paths linearly with limited exploration of the overall program contents). In studies focusing on learners with different levels of knowledge, students with higher content knowledge selected more efficient navigational techniques (Dillon, 1991; Shyu and Brown, 1995) and viewed more screens. In comparison, students with low levels of prior knowledge exited the lessons more quickly (Schroeder and Grabowski, 1995).

Usability can be an important precursor to conducting research in navigation in multimedia programs. Ideally, usability studies should be performed as part of the software development process. This will minimize the adverse effect of poor navigational system on learning and produce more meaningful results in research studies of computer-based learning (Clark, 1992).

Purpose of the study

The purpose of this study was to explore the principal interface issues in the GramStain-Tutor™ (GST), particularly the navigation of the program and the use of embedded design features. An initial review of the program by the authors pointed to several problems regarding the navigational structure of the GST. The study examined whether users experienced these problems while performing various tasks.

The following questions were examined in this study:

1. Navigational patterns among users.
   1.(a) Do users discover that there are three ways to navigate through the program?
   1.(b) What difficulty do users experience while moving through the program?
   1.(c) Does previous knowledge of the program’s content influence the way users interact with the program? For example, is there a difference in the number of steps taken and errors made in locating visual structures between content novices and experts?

2. Usability of embedded features.
   2.(a) Can users identify and use animation buttons without problems?
   2.(b) Can users follow instructions in the Image Atlas? More specifically, can users display a single image, two images from the same morphologic category, and two images from different morphologic categories?
Before discussing the methods of the study, the navigational structure and embedded features of the GST are briefly described in the following section.

**GramStain-Tutor™ (GST)**

The GramStain-Tutor™ (GST) is a computer program on CD-ROM that was developed by the Department of Laboratory Medicine at the University of Washington. The program teaches the performance and interpretation of the direct Gram stain of body fluids (Cookson et al., 1994; for more information about GST, see http://www.labmed.washington.edu/tutor/products). The direct Gram stain is used to initially diagnose bacterial and yeast infections and it is the most common test performed in a clinical microbiology laboratory. GST is used in three learning contexts. First, laboratory technologists in clinical microbiology use GST for initial training and continuing medical education. Second, students enrolled in medical technology programs are required to use GST during their rotation in microbiology laboratory. In many cases, these students learn about Gram staining techniques for the first time through GST. Lastly, medical students are required to use GST during the first year of their medical training as part of the course, Microbiology and Infectious Diseases.

GST uses digitized microscope images of Gram-stained clinical samples, illustrations, animations, videos, text, and labels. Approximately 100 concepts are accompanied by example images and text describing the visual characteristics and clinical implications of the organisms, cells, or artifacts. To support the content structure, three complementary modes to navigate the program are available as illustrated in Figure 1: (a) a horizontal navigational bar contains concepts; (b) a vertical navigational bar lists examples; (c) arrows may be used to see additional examples.

**Figure 1: Interface design of the GramStain Tutor™ (GST)**

examples of concepts; and (c) arrow buttons lead to more images of each concept. The three levels of navigation were designed to make the content structure as maximally visible to the users as possible; and to provide multiple ways to access the content depending on user preferences.

To stimulate learning, a number of design features have been embedded in the program. First, video clips are available to illustrate the Gram stain technique. Second, animations are inserted to show structures and shapes of organisms from various angles (see http://128.95.186.167/tutor/products/prod1/animation/theory/default.asp). Third, features that highlight and focus on specific organisms are available to accentuate structures that are hard to distinguish (see http://128.95.186.167/tutor/products/prod1/screenshots/screen1C.asp). Fourth, case studies and exam questions enable students to select their responses and view immediate feedback to their responses (see http://128.95.186.167/tutor/products/prod1/screenshots/screen1E.asp). Lastly, an Image Atlas feature enables the users to select and compare two images simultaneously (see http://128.95.186.167/tutor/products/prod1/screenshots/screen1F.asp). The educational effectiveness of the Image Atlas has been reported in a previous study (Kim and Aston, 2000).

Methods

Subjects
A total of 22 users were recruited for the study. Seventeen users were students enrolled in various health sciences departments at the University of Washington (3 medical technology students, 6 physician assistant students, and 8 nursing students). The remaining 5 users were laboratory technologists, who were employees in the Department of Laboratory Medicine at the University of Washington Medical Center. These technologists had on average 16.2 years of experience (SD = 8.2; range: 8–27 years) in the field. Administrators of the physician assistant program and nursing school, as well as a course director of the medical technology program recruited students through email postings. Laboratory technologists were recruited through supervisors of the laboratory, who permitted the technologists to participate in the study. Only the students were paid $25 for their participation.

Based on the pre-session questionnaire, the following user characteristics emerged. Eleven users (50%) reported that they had experience with the Gram stain procedure and 16 users (73%) mentioned that they had general experiences in the preparation and analysis of blood and urine samples. More than 60% of the users reported that they used computers including email and the World Wide Web at least 2 hours a day.

Tasks
A total of 16 tasks were developed for users to complete. Of the 16 tasks, one task asked the users to explore the introduction section, 8 tasks specifically dealt with navigation within the program and the use of animation buttons, 5 tasks covered the use of the
Image Atlas feature, and 2 dealt with case studies and exam questions. An example of a simple task stated:

"Go to the specimen sites and select Wound. Display a Proteus mirabilis image on the screen. When finished, go back to the main menu."

A more complex task instructed the users to do the following:

"Go to the Image Atlas screen. Select the split image mode. In the top image box, display Streptococcus group B and in the bottom panel, display Candida albicans. When finished, go back to the main screen of Image Atlas."

The users performed each task until they accomplished fully the goals of the tasks.

Data collection procedure
The usability sessions were conducted using a laptop with GST installed. The following qualitative and quantitative data were collected during the sessions.

1. Users filled out pre- and post-session questionnaires. The pre-session questionnaire included questions related to user training backgrounds, prior experiences with laboratory procedures, and their computer skills. The post-session questionnaire asked for feedback on navigation and design features.
2. Users were encouraged to "think aloud" as they performed each task by describing what they were thinking and what problems they were experiencing. These comments were audio taped.
3. Debriefing sessions were conducted by the investigator (SK) to discuss difficulties each user encountered while completing the tasks.
4. Time on task was measured during each session.
5. Screen shots captured by the video recorder were analyzed to determine major errors committed by the users. A digital video recorder connected to a scan converter (Focus, TView gold card) recorded every screen viewed by the users while they performed the tasks.

Following the completion of the usability sessions, the investigator met with the designers of the program to review the interface design problems identified by the users.

Results
The results of this study are presented in the following order: (a) time spent on tasks; (b) navigational patterns; (c) the use of the image atlas; and (d) the use of video/animation/simulation functions. No significant usability issues were found in the use of the case studies and exams.

Time spent on tasks
The 22 users spent on average 33.3 minutes (SD = 6.3; range 21–45 minutes) completing 16 tasks. Of the total time, users spent 13.9 minutes (SD = 3.4) on tasks dealing with navigation of the program and 7.7 minutes (SD = 2.7) on using the Image

Atlas feature. The rest of the time was spent exploring the introduction section of the program and performing tasks related to case studies and exam questions.

Laboratory technologists (n = 5) spent an average of 11.4 minutes covering the tasks on the navigation of program content and an average of 10.2 minutes using the Image Atlas feature. In contrast, students (n = 17) spent an average of 14.6 minutes navigating the program and an average of 7.0 minutes using the Image Atlas feature. The difference in time spent on the image atlas was statistically significant (p < 0.05; Mann-Whitney test), but the difference in navigation time was not statistically significant.

Navigational patterns
Overall exploration: The first task of the usability session instructed the users to explore the “Introduction” section. The purpose of this task was to familiarize the users with the interface of the program and to observe whether users relied on a particular navigational mode (horizontal bar, vertical bar, or arrow keys) or used a combination of navigational modes. Of the 22 users, 6 users (27%) explored the section using more than one navigational mode and the remaining users confined their navigation to one mode. All 3 users who indicated that they had experience with designing their own web pages used a combination of all available navigational modes.

Navigational Barriers: Two major barriers to navigating the program emerged based on user comments during debriefing sessions. First, users did not understand the relationship between the horizontal and vertical bars with the former representing concepts and the latter, examples of those concepts. Second, users did not see the arrow key pointing to the text, “1 of 3”, for viewing embedded learning materials. Of the 22 users, only 3 users (14%) noticed the arrow line and viewed the embedded image. During debriefing sessions, users were surprised that they did not notice the line while performing the task.

Content Knowledge: One task asked users to display an example (Proteus mirabilis) of “Gram-Negative Bacteria” in the “Wound” section, which could be accessed from the main menu in two steps. Only two users (less than 10% of total users) took no extra steps in displaying the image. In contrast, about 50% of the users took an extra step and approximately 40% of the users took more than one step as they explored the program to look for the example. During the debriefing session, the two users who made no extra steps reported that they knew the prompted example belonged to the category of Gram Negative Bacteria whereas the rest of the users remarked that they did not know the morphological category of the example.

The use of image atlas
Data analysis on the use of the Image Atlas feature mainly focused on the number of steps the users took to complete tasks, such as displaying a single image, displaying two images from the same morphologic category, and displaying two images from different morphologic categories. Only 6 users (27%) completed all tasks related to the use of the Image Atlas feature without taking any extra steps. The number of extra steps taken by
the rest of the 16 users (73%) ranged from 1 (27% or 6 users) to 8 (5% or 1 user), with an average of 2 extra steps. The users took extra steps by selecting the wrong item, attempting to recover from errors, and backtracking to the screen that included the instruction. Users reported two barriers to effective use of the feature: (a) the instructions on how to use the feature only appeared on the main page of the Image Atlas section and users at times had to backtrack to view the instructions; and (b) the icons denoting single versus split image mode were not self-explanatory.

The use of video/animation/simulation buttons
All users noticed and used buttons that played videos and animations as well as the buttons that simulated microscope features. Users reported that having visual cues that showed the duration of animations would be useful.

Discussion
This usability study focused on users' navigational behavior and their ease of use of various interactive features including the Image Atlas. The main results of the study and future design implications are summarized below:

1. The majority of the users did not use all available navigational options supported by the program.

Although three complementary navigational options were available in the program (horizontal and vertical bars, and arrow keys), only 6 of 22 users used more than one option. Most users did not discover the horizontal and vertical navigation bars, which represented a hierarchical knowledge structure, until they were prompted to select an item appearing on the navigation bars. Users commented that a description of the content structure and instructions on navigating the content would be useful. The need to provide a guided structure to the learners has been highlighted in a number of studies (Koneman, 1994; Litchfield, 1993; Persico, 1997; Wild and Quinn, 1998). Without the prior knowledge of how to access content materials, learners become disoriented while interacting with the program and explore the program inefficiently.

2. Navigational patterns differed as a function of training background and content knowledge of the users.

When measuring the number of steps it took the users to locate a specific concept, users' content knowledge affected navigational behavior. For example, users who were familiar with a conceptual category took no extra steps in displaying a concept within that category. Those who were unfamiliar with the concept clicked on all available categories until the correct concept appeared on the screen. This result implies an association between navigational patterns and the level of prior knowledge, and this association has been suggested in other studies (Dillon, 1991; Fenley, 1998; Schroeder and Grabowski, 1995; Shyu and Brown, 1995). The findings are consistent with the study by Brickell (1993), who emphasizes the need to reduce a learner's frustration associated with cognitive overload. Learners with low levels of prior knowledge should
be given the means to efficiently navigate through an interface aided by adequate instruction and intuitive navigational structure. A content map in the program, which presents a structure of conceptual categories, may help learners understand the relations between and among concepts (West, Farmer, and Wolff, 1991).

3. Several visual cues critical to program use were not discovered by the majority of the users. In one task that required the users to display embedded example images as prompted by an arrow line, only 3 of 22 users noticed the arrow line. This suggests that subtle visual cues are likely to go undetected when an interface demands that a user recognize embedded text, animations, navigational features, and images. Increasing the physical attributes of visual cues or using unique color may help draw a user's attention (Fleming, 1998; Park and Hannafin, 1993).

4. Icons representing certain commands were not intuitive

Visual icons associated with executable functions, such as icons denoting single or split image screens in the Image Atlas section, were not intuitive. When designing new icons, designers should test the usability of the icons with users or with designers who are not involved in the product design. In addition, this problem may be solved with simple instructions that explain the functions of the icons. Lastly, users asked that the instructions describing the Image Atlas feature remain visible at all times to avoid backtracking to the screen containing the instructions. This is consistent with an interface design guideline that suggests that all relevant information be kept in view for easy access (Marshall et al., 1987).

Conclusion

This study focused on key usability problems from users' perspectives based on multiple data collection methods. The analyses of video and audio-tapes along with user feedback collected from questionnaires and debriefing sessions helped us modify the interface design. First, to address the problem of users failing to notice visual cues for viewing embedded content, we have redesigned the vertical menu bar so that the content is no longer embedded. The formerly embedded material is now available directly from the vertical menu bar, and therefore, we have improved the likelihood of users accessing this content. Second, we have clarified the relationship between the horizontal “concept” bars and vertical “example” bars by using a matched color scheme (Fleming, 1998) that strongly suggests to the user that the colored examples on the horizontal bar are related to the concept on the vertical bar that is highlighted in the same color. Users who have tested this revised interface reported that the matching colors between the two navigational bars helped establish the conceptual relationship represented by the two bars. The recent literature on navigation, particularly in the area of Web interface design, continue to provide suggestions to us for best representing layers of conceptual hierarchy on a two dimensional screen. These suggestions include use of pull-down menus, tables of contents, indexes, and site maps (Rosenfeld and Morville, 1998).
We demonstrated how a usability study of a relatively small number of users could identify specific problems in interface design. We recommend these interface problems be addressed before conducting educational studies that examine how educational software programs affect student learning. Our results suggest specific design improvements that should be made to software that shares similar content organization, navigational structure, and interactivity to GST. An example of such software is the “TUTOR” family of computer tutorials that has been developed by the University of Washington Department of Laboratory Medicine since 1992. Improvement in this series is expected to benefit approximately 23,000 users in 1,500 laboratories in the US and abroad.

**Acknowledgement**

We would like to thank all participants in the usability sessions for their time and insightful comments regarding the quality of the program. Also, we thank the following individuals who helped recruit participants for this study: Dr Mary Lampe (Department of Laboratory Medicine), Patty Morrison (Physician Assistant Program), Sandra Krueger-Nielson (Department of Laboratory Medicine), and Cliff Solomon (School of Nursing).

**References**


Virzi R (1992) Refining the test phase of usability evaluations; how many subjects is enough? Human Factor 34, 457–468.
