

SLIS L503
Summer Session II, 2001

Project Authors:

Jenn Riley
Marianne Kruppa
Marisa Hudspeth
Abbie Anderson

www.startrek.com

Assessment of Usability

Introduction

This study examines the overall usability of the official Star Trek web site, www.startrek.com. Tasks and post-test questions posed to study participants focus primarily on issues of navigation and organization, since these issues were determined to be the main usability concerns for a site of this nature.

Startrek.com is the official web site for all things Star Trek, designed and copyrighted for Paramount Studios. Its “Library” section covers every iteration of the franchise on television as well as on film, with plot synopses, biographical sketches both of characters and of the actors who play them, and encyclopedias of technological and cultural details featured in the Star Trek universe. A “News” section provides the latest information on Star Trek productions, publicity events, and the projects of Star Trek stars, while “Information” offers a gateway to practical things such as a calendar of significant Star Trek facts (both actual and fictional) and the air dates for Star Trek episodes. Separate sections access “Recreation” (online games and fan interaction) and “Gaming” (support and news for commercial game releases), while a final “Utilities” section offers free downloads of Star Trek images and cursor sets. A link to the Star Trek store takes the user to a separate site for purchasing Star Trek memorabilia and costumes. With a potentially overwhelming range of information and activities to make accessible, and a diversity of intentions to support for those approaching the site, it is particularly critical that the designers of startrek.com ensure that the content of the site is clearly discernible and easily navigable.

The major questions asked by this study are, “Does the design of www.startrek.com provide an enjoyable and effective information-seeking experience for casual Star Trek fans?” and “What usability problems do users encounter?” The user should be able to find desired information in an entertaining manner on the website with an acceptable level of ease and favorable degree of satisfaction. The site will be considered usable if all users complete at least 4 of 5 tasks, with a performance score of at least 50% of an expert’s, and with an average satisfaction rating of 4.5 on a 7-point scale. As benchmarks for measuring this success, we used our own “expert” times in completing tasks assigned to our testers (see Appendix 1 for task list and benchmark specifications).

Test design

Subjects for this study were chosen who are casual Star Trek fans, and have rarely or never been to www.startrek.com. In this way we aimed to test the site using people who were sufficiently familiar with the Star Trek universe to readily grasp the content of the site, but who were not overly familiar with the site itself. Prior to the commencement of formal testing, two volunteers evaluated the structure and content of the script, tasks and questionnaire in a “pre-test” phase of the project.

Five subjects participated in the formal testing of www.startrek.com. All subjects were tested in the CITED computer cluster in the School of Library and Information Science library on the Indiana University-Bloomington campus. Two testers were present for each test, one to time task performance and the other to administer the tasks and record steps taken by the subject in pursuing task completion. A script was followed verbatim for each test, to ensure consistent administration (see Appendix 1 for copy of script). For each subject, the nature of the study was outlined (stressing the subjects' complete anonymity) and the website was introduced. Subjects were then allowed approximately two minutes to explore the site before testing began.

After this initial exploration, the study was explained in more detail, and the subjects were given a series of five tasks to perform (see Appendix 1 for task list with screenshots). Tasks were administered in a different order for each subject, to avoid testing the learning curve of the site through the progression of tasks, rather than the tasks themselves. The tasks were presented on paper for the subjects to read. The subjects were instructed to say, "Ready" when they were about to begin, at which point timing of the task began. The subjects were encouraged to think aloud while pursuing the task, to give reactions to the pages as well as explain strategies. They were instructed to say "Done" when they felt the task was satisfactorily completed. The subjects were informed that they had the option to end a particular task prior to completion if desired. Between each task, one of the testers set the web site back to the [startrek.com](http://www.startrek.com) home page.

After the tasks were completed, the subjects were given a six-item questionnaire to fill out by hand (see Appendix 2 for questionnaire). The first five items presented statements to which the subjects responded on a seven-point scale of agreement (strongly disagree as 1, strongly agree as 7), circling the number felt most appropriate. The sixth item provided space for comments, asking “What would you change about this site, or what suggestions do you have for the designers of this site?” Subjects were thanked sincerely for their time and for their contribution to this project.

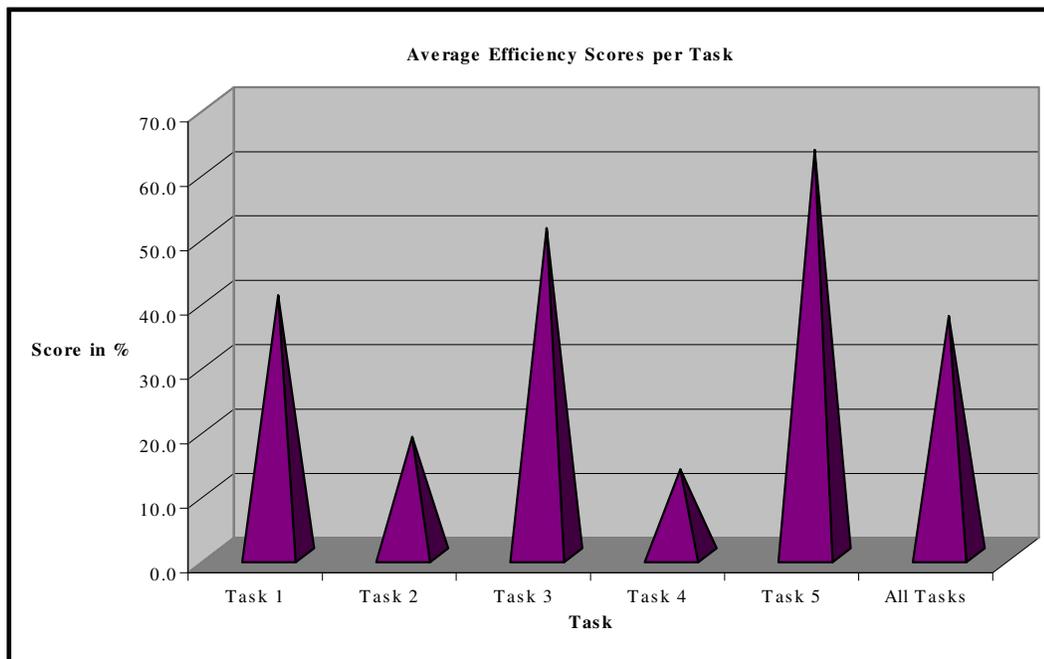
Following completion of all tests, results were collated and analyzed. Effectiveness for each task was measured by the number of actions the subject took toward fulfillment of the task, divided by the minimum number of actions required to complete the task (as established in our expert benchmarks). Efficiency for each test was measured by user completion rate multiplied by the sum of expert time divided by user time (see Appendix 3 for scoring tables). Responses to the questionnaire and comments made by subjects during task completion were analyzed for patterns both in assessment of the site and in individual behavior (see sections 3.1. and 3.2. for more details).

Findings

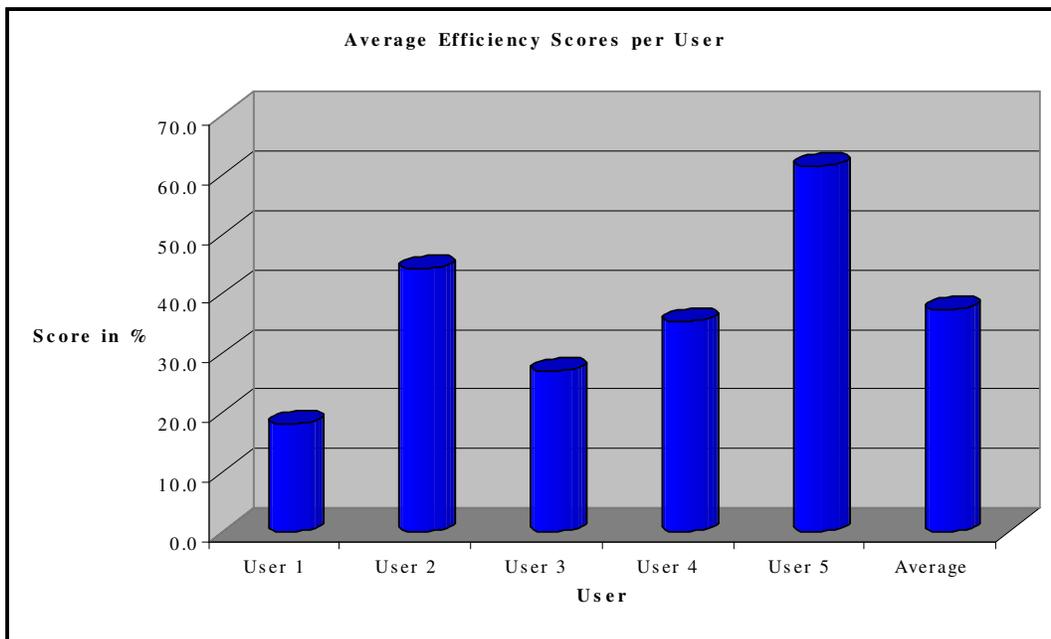
The users in this study exceeded the effectiveness benchmark we set. Four of five users completed all five tasks, and the fifth user completed four of five

tasks. The single uncompleted task in the study had a 0% effectiveness rating. The user in this case completed neither of the two actions required to complete the task.

Despite the high effectiveness ratings in this study, users' efficiency ratings were quite low. The chart below shows the average efficiency rating of all subjects over all tasks to be 37.2%, far below our benchmark of 50%.



No single subject was able to complete all of the tasks as fast as the experts, but there was one exceptional subject that did complete 3 of the 5 tasks at 50% or above. As shown in the chart below, this subject, User 5, was the only individual to perform above our benchmark of 50% efficiency.



In the post-test questionnaire, users gave an average satisfaction rating with the site of 5.8 on a 7-point scale, well above our benchmark of 4.5. However, the qualitative comments on the post-test questionnaire and users' verbal comments during the test seemed to indicate a lower satisfaction rating.

The most common errors observed during the test were clicking on Gaming instead of Recreation to find the online game "Odo's Investigation" in Task 1 (5 occurrences) and clicking on Canada instead of the United States on a world map to get to the TV listings in Task 3 (4 occurrences). The first error was caused by the fact that "gaming," on this site, refers to the commercially released games by Paramount and not those games that are available to play on the site itself. But this intellectual distinction was not meaningful to the users we tested.

The second error occurred because the map that users must click on to find TV listings for a geographic area does not provide boundary lines for countries or continents. It is too small for fine selection of areas. Most of the subjects clicked on the map near the southern most tip of Lake Michigan, but the site brought them to listings for Canada. The site does offer an alternative method for finding listings, based on a series of combo boxes, but only after the user passes the map page.

Some problems in the testing that we encountered were due to semantic ambiguity in two tasks and not foreseeing an alternate path to completion for one task. One of our subjects did not understand what was meant by “cursor” in Task 4. The biggest semantic ambiguity was our use of the word “milestones” to represent birthdays in Task 2. All users were eventually able to complete the task, but their efficiency ratings suffered. Two of our subjects found an alternate path for completing Task 5 by using the Media Archive search to find a bat'leth picture instead of finding it via the Alien Cultures page. While these users were given full completion scores, our efficiency results were skewed by this alternate route we had not taken into account.

Discussion

Startrek.com did not pass our original definition of usability for this project. Criteria for effectiveness and satisfaction were met: all users completed

at least four of five tasks, and users gave the site an average satisfaction rating of 4.5 on a 7-point scale. Efficiency criteria, however, were not met, since the average performance score of each subject did not attain at least 50% of our expert benchmarks. It is interesting to note that despite frustrations encountered on the site, users expressed a favorable level of satisfaction. While our subjects were generally impressed with the content of the site, obstacles to efficiency are built into its design. The organization, nomenclature, text color and size, and changing color schemes of the menu often led our subjects astray.

Organization of the menu bar posed problems for our subjects. For instance, the link to return to the Home page of the site was located at the bottom of the screen instead of the top, breaking with web convention and thereby confusing our subjects. The most significant problem encountered was the web site's definition of Gaming versus the subject's definition—more specifically, the division of Star Trek games between the Gaming section and the Recreation section. In relation to task 1 ("What Star Fleet rank will you be assigned in the online game 'Odo's Investigation'?"), our subjects keyed into the word "game" rather than the phrase "online game." This logically led them to begin their search under the Gaming section, rather than Recreation. What is found under Gaming on startrek.com are news and product support for commercial games, not games available to play online. Another menu item that subjects found unclear in its intent was Utilities. Three of the subjects expressed their uncertainty of what could be found in Utilities, and credited luck as the reason

they were able to complete task 4 (“Find a list of free Star Trek cursors for download on this site.”).

Text color and size were a problem for most of the subjects, although only one subject expressed this verbally. Through observation the testers could see that some were struggling to see the text. This was characterized by the subject leaning close to the screen and often using a finger as a guide to read the text. For most of the pages, the text was a bright color against a dark background. As one example, the color scheme for the Production News page is a light blue text on a medium blue background. The similarity of the colors makes it very hard to read. We included task 5 (“Find a picture of a bat’leth, a Klingon weapon”) for this very reason, to see how subjects reacted to the different color scheme and variant layout of the Klingon section of the Library.

Unfortunately, we did not sufficiently explore task 5 before testing began, and were surprised when two of our subjects found an alternate means of locating a picture of a bat’leth that we had not anticipated. These subjects used the Media Archives Search function in the Library to locate images of a bat’leth taken from Star Trek episodes—this in spite of the fact that “bat’leth” (as spelled in the Library’s Klingon culture section, the intended destination) is spelled “bat’leth” in the Search for Media Archives. Subjects were able to use this route to complete the task because a drop-down menu in the Search category “Technology” includes “bat’leth” (sic) as an option. The subjects were conversant enough with Star Trek to recognize the bat’leth in the picture, even though it was

not labeled as an example of the weapon (but was merely an excerpt from an episode which happened to include a bat'leth).

Despite these difficulties, the quantitative satisfaction rates were high. This could be due to the subjects being fans of Star Trek and their appreciation of the information contained in the site. This could also account for the high completion rates and the willingness of the subjects to persist until they felt they had completed the task. Points of negative satisfaction were revealed qualitatively under item 6 of the post-test questionnaire, where subjects were encouraged to express their thoughts more freely. Their comments included complaints about the site's "visual barrage", and legibility problems resulting from the inconsistency color-scheme from section to section. The major difficulty of the distinction between Gaming and Recreation was also stressed both in the post-test comments and verbally during task completion.

Conclusions

Performance of this study has provided our group with some important lessons. Given the opportunity to conduct another usability project, we would take more time to pilot test our tasks. With more time for pre-testing, we would aim to discern all possible alternate paths to task completion, as well as any potential difficulties with the wording of task instructions. Another thing we learned from startrek.com is the importance of being aware of last-minute changes to the site that may affect our tasks. For instance, task 2 required the

subjects to find the Star Trek calendar, which lists birthdays of Star Trek celebrities. The day we began formal testing, startrek.com added a new item to the News section--announcing August birthdays. The only thing that prevented subjects from using this path was the semantic ambiguity of "milestones" (an unfortunate choice in our wording of the task which in this case ended up helping subjects complete the task as we had anticipated). To bridge the gap between the quantitative and qualitative satisfaction rates we would include more quantitative questions in the post-test questionnaire and possibly extend the pre-test exploration to 3-5 minutes.

In spite of difficulties encountered, our testing of this site provided useful and interesting results. We achieved what we set out to do with this project. In our study plan, we defined usability for this site as the ability of the user to find desired information in an entertaining manner and with an acceptable level of ease and a favorable degree of satisfaction. While "acceptable level of ease" remains in question for startrek.com, the level of entertainment and wealth of material provided by the site appeared to offset frustrations and discomfort experienced by users in navigating its menus. We conclude that high completion and satisfaction rates can outweigh obvious problems with efficiency in determining the user's overall response to the site.

Appendix 1: Study Protocol

Major Question to be Addressed by the Test

Does the design of www.startrek.com provide an enjoyable and effective information-seeking experience for casual Star Trek fans? What usability problems do users encounter?

Working Definition of Usability for this Site

The user should be able to find desired information in an entertaining manner on the website with an acceptable level of ease and favorable degree of satisfaction. The site will be considered acceptable if all users will complete at least 4 of 5 tasks, with a performance score of at least 50% of an expert's, and with an average satisfaction rating of 4.5 on a 7-point scale.

Criteria for Selecting Subjects

Subjects will be chosen who are casual Star Trek fans, and have rarely or never been to www.startrek.com.

Tasks

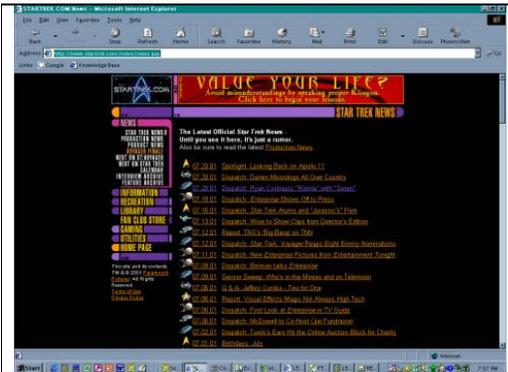
Task 1: What role will you be assigned in the online Game "Odo's Investigation?"

Expert time to completion: 24.05 seconds Number of actions required: 2.

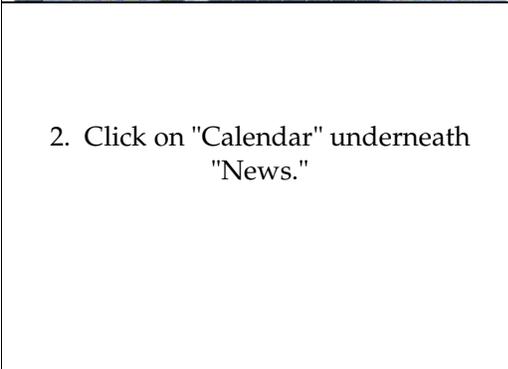
	<p>1. From home page, click on "Recreation" in the navigation bar</p>
<p>2. Click on the "Odo's Investigation" picture in the middle of the page. Answer: "Star Fleet Academy Cadet" is listed as the player's role in the first sentence.</p>	

Formatted: Bullets and Numbering

Task 2: What important Star Trek events are happening on August 19, 2001?
 Expert time to completion: 12.77 seconds Number of actions required: 3

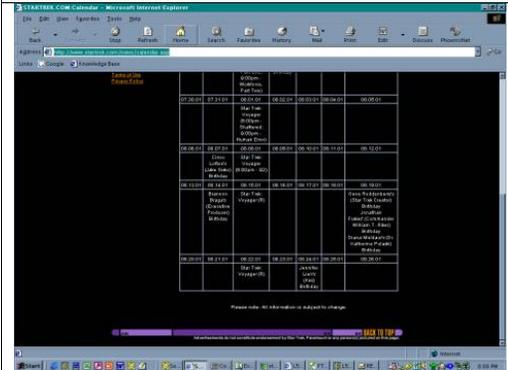


1. From home page, click on "News" in the navigation bar



2. Click on "Calendar" underneath "News."

Formatted: Bullets and Numbering



3. Scroll down the page to August 19.
 Answer: Gene Roddenberry's (Star Trek Creator) Birthday,
 Jonathan Frakes' (Commander William T. Riker) Birthday,
 Diana Muldaur's (Dr. Katherine Pulaski) Birthday.

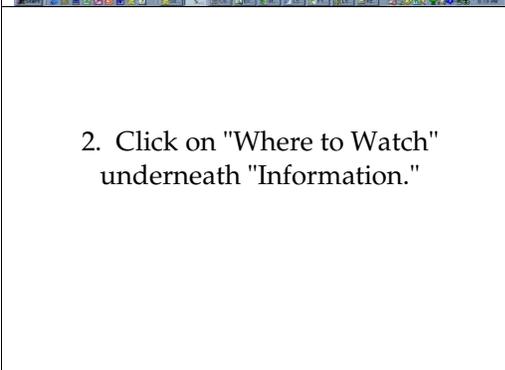
Formatted: Bullets and Numbering

Task 3: Find a list of Star Trek (all series) episodes showing on Indianapolis TV stations this month.

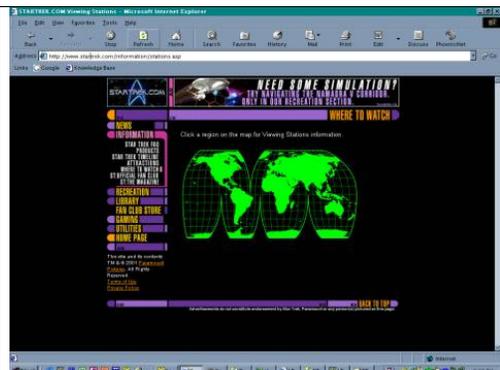
Expert time to completion: 26.86 seconds Number of actions required: 6



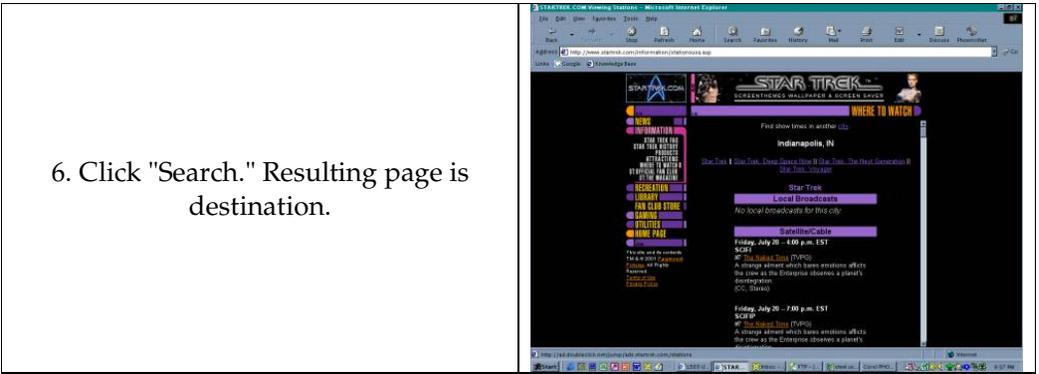
1. From home page, click on "Information" on the navigation bar.



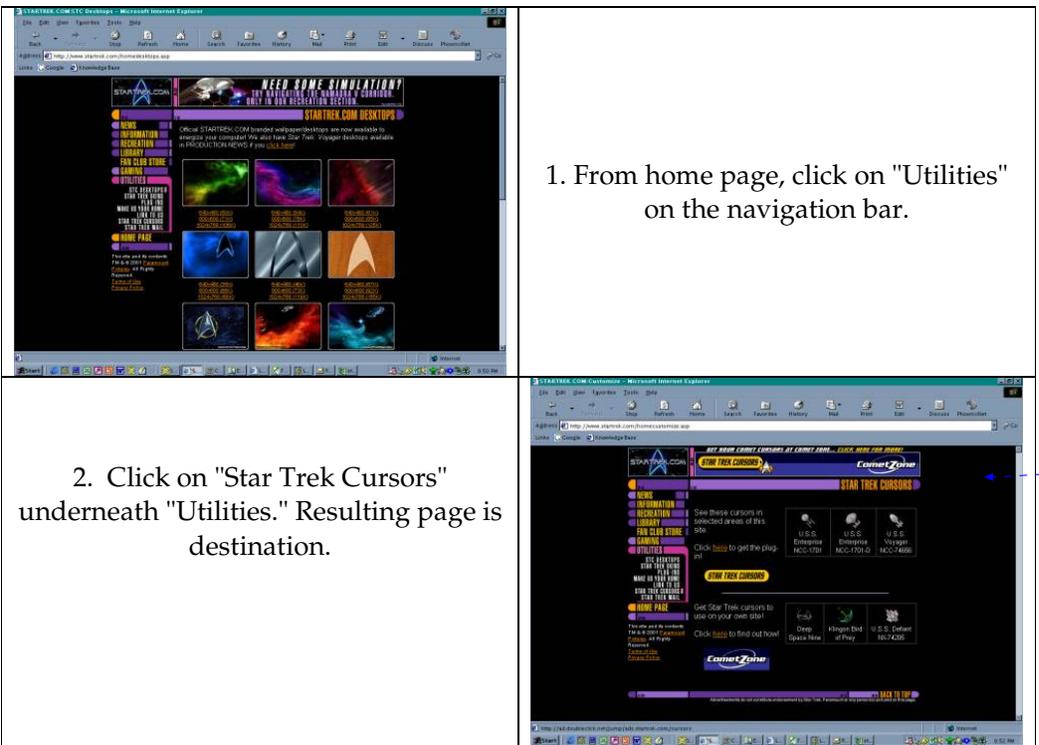
2. Click on "Where to Watch" underneath "Information."



3. Click on the map around Indiana.
4. Enter "IN" in state box.
5. Choose "Indianapolis" from the combo box.



Task 4: Find a list of free Star Trek cursors for download on this site.
 Expert time to completion: 10.33 seconds Number of actions required: 2



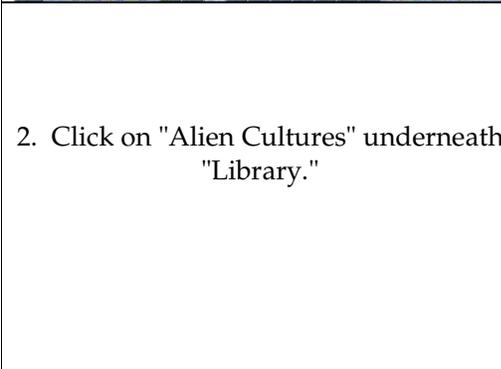
Task 5: Find a picture of a bat'leth, a Klingon weapon.

Expert time to completion: 28.18 seconds Number of actions required: 5



1. From home page, click on "Library" on the navigation bar.

Formatted: Bullets and Numbering



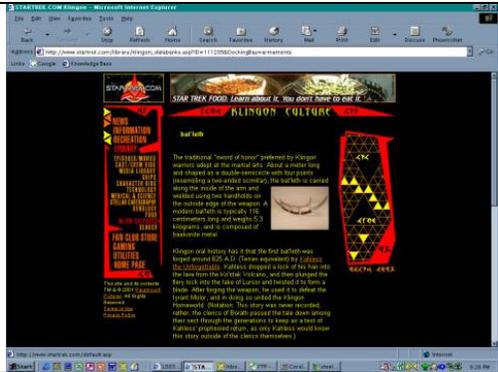
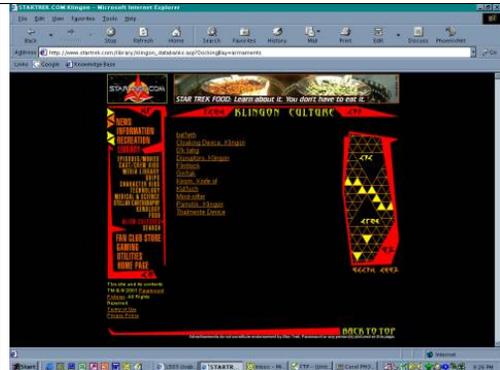
2. Click on "Alien Cultures" underneath "Library."



3. Click on "Klingon Culture" picture on page.



4. Click on "Weapons" from the list near the bottom of the page.



5. Click on "Bat'leth" link. Resulting page has picture.

Script

Before starting

Before the volunteer arrives, set up the computer to the startrek.com home page, with no other windows open.

When the volunteer arrives, greet them. Get them comfortable in front of the computer. Tell them that you'll be reading from a script, to ensure that all of our study volunteers hear exactly the same information.

Read this script

"Thank you for volunteering for this study. All information we gain from this study will be kept strictly confidential; even our professor will not know the identities of the volunteers we observe using this site. You may at any time stop and withdraw from the study, simply by telling me you'd like to do so.

"As part of a class assignment for SLIS L503, we're evaluating the usefulness and design of the web site startrek.com. We'd like to start out by giving you a few minutes to explore the site on your own and look at things that interest you. I'll give you about two minutes."

[Give the volunteer two minutes to play.]

"Let's move on to our test of the site. For the class assignment, we're observing people using the site and asking them to give us their opinions about it. Please remember that we're evaluating the web site, and not your performance. All of your interactions with the site are important to us, as are your observations of the site as you use it. I'll be taking notes about what you say and do as you work.

"We're going to give you a series of five tasks to perform on startrek.com. When you receive the task, take your time reading and understanding it. When you're ready to begin the task, say "ready." While you're performing each task, please talk about what you are doing and why, and give us your reactions to the pages you see. Your reasoning and opinions are important to us in evaluating this site. When you believe you have completed the task, please tell me that you're done. I can't tell you if you are "right" or not, you decide. After you complete each task, we'll set your web browser back to the startrek.com home page and give you the next task to do. If at any time you feel you can't complete the task, simply tell me you'd like to move on to the next task. After you say you are done with the fifth task, I'll ask you to fill out a short questionnaire about your experience on the site."

[Give the volunteer the first task. Remind them to say “ready” and “done.” Start timing when they say they’re ready. If they’re not talking very much while working, remind them to do so. When they say they’re done or want to give up, stop timing. Take the slip of paper with the task from them, they say:]

“Thank you. I’m now going to put your web browser back to the startrek.com home page.”

[Put the browser back to the startrek.com home page. Give them the next task, and say:]

“Here is the next task.”

*[Start timing when they say ready. Repeat until all 5 tasks done. When the volunteer says they’re finished with the last task, ask them questions about places they got stuck or got off-track in completing a task. When that interaction is complete, **hand them a pen**, and say:]*

“The last thing we’d like you to do is to fill out a short questionnaire about your experiences at startrek.com today.”

[When they’re done with the questionnaire, tell them they’re done and thank them profusely!]

After the test

Put the volunteer’s ID number on each piece of paper used for their test, including:

Each task
Your notes for each task
Timings list
Volunteer’s completed questionnaire

Put all of these materials in a single envelope marked with the ID number for that volunteer.

What Will be Measured and How

- Usefulness of main navigation menus on site. Measured by all tasks’ effectiveness and efficiency, and post-test questions 2-3.
- Usability of subnavigation menus on site. Measured by tasks 2, 3, 4, and 5 effectiveness and efficiency, and post-test questions 2-4.
- Organization of information in “Library” section. Measured by tasks 1 and 5 effectiveness and efficiency, and post-test question 4.

- Organization of information in “Calendar” section. Measured by task 2 effectiveness and efficiency, and post-test question 4.
- Usability of the “click on map, then enter state and city” interface for TV listings. Measured by task 3 effectiveness and efficiency, and post-test question 4.
- Overall satisfaction with the site. Measured by post-test questions 1, 5, and 6.

Effectiveness will be measured by the following formula:

Number of actions user took toward fulfilling the task / Minimum number of actions required to complete the task.

Efficiency will be measured by the following formula:

User completion \times (expert time / user time)

Appendix 2: Post-Test Questionnaire

Test ID:

Usability Questionnaire for www.startrek.com

Please circle one number for each question on the scale of 1 to 7. Feel free to ask your testers for clarification if any of the questions seem unclear.

1. I enjoyed using startrek.com today.

Strongly Disagree 1 2 3 4 5 6 7 **Strongly Agree**

2. The design of the site was comfortable to look at.

Strongly Disagree 1 2 3 4 5 6 7 **Strongly Agree**

3. I thought the startrek.com menus were confusing.

Strongly Disagree 1 2 3 4 5 6 7 **Strongly Agree**

4. The individual pages I encountered on startrek.com today were clearly laid out.

Strongly Disagree 1 2 3 4 5 6 7 **Strongly Agree**

5. I would recommend startrek.com to other Star Trek fans like me.

Strongly Disagree 1 2 3 4 5 6 7 **Strongly Agree**

6. What would you change about this site, or what suggestions do you have for the designers of this site?

Appendix 3: Raw Data

- Task 1** Which Star fleet rank...
- Task 2** What important star trek milestones...
- Task 3** Find a list of Star Trek episodes
- Task 4** Find a list of free star trek cursors...
- Task 5** Find a picture of a bat'leth...

Effectiveness on a scale of 1-100%

	Task 1	Task 2	Task 3	Task 4	Task 5
User 1	100	100	100	100	100
User 2	100	100	100	100	100
User 3	0	100	100	100	100
User 4	100	100	100	100	100
User 5	100	100	100	100	100
Average	80	100	100	100	100
Median	100	100	100	100	100
Expert	100	100	100	100	100
Range	100	0	0	0	0

Efficiency Scores

	Task 1	Task 2	Task 3	Task 4	Task 5	All Tasks
User 1	22.2	13.5	29.7	12.0	13.4	18.2
User 2	41.4	19.7	122.7	8.9	28.0	44.1
User 3	0.0	28.9	37.5	16.7	51.9	27.0
User 4	23.8	8.7	9.0	12.9	121.7	35.2
User 5	114.3	21.0	55.1	16.2	100.0	61.3
Average	40.3	18.4	50.8	13.3	63.0	37.2
Median	23.8	19.7	37.5	12.9	51.9	
Expert	100.0	100.0	100.0	100.0	100.0	
Range	114.3	20.2	113.7	7.7	108.3	

decimal time = minutes + (seconds/60)
 efficiency = effectiveness x (expert time/user time)

Usability Questionnaire Results

	Question 1	Question 2	Question 3	Question 4	Question 5
User 1	7	2	3	5	6
User 2	5	4	5	5	7
User 3	4	2	6	2	3
User 4	6	5	4	4	6
User 5	7	7	6	6	7
Average	5.8	4	4.8	4.4	5.8

Recurring Task Errors

Task	Error	Frequency
1	Clicking on Gaming option in menu instead of Recreation	5
3	Clicking on Canada instead of US on map	4