2011

Usage of and Satisfaction with Online Help vs. Search Engines for Aid in Software Use

Charles J. Welty Ph.D.

University of Southern Maine

Follow this and additional works at: http://digitalcommons.usm.maine.edu/computer-science-papers

Part of the Databases and Information Systems Commons

Recommended Citation

Welty, C. Usage of and Satisfaction with Online Help vs. Search Engines for Aid in Software Use. Submitted to ACM sigDOC 2011 Conference (special interest group on Design of Documentation). October 3-5, 2011, Tirrenia, Pisa, Italy

This Conference Proceeding is brought to you for free and open access by the Computer Science at USM Digital Commons. It has been accepted for inclusion in Faculty Publications by an authorized administrator of USM Digital Commons. For more information, please contact jessica.c.hovey@maine.edu.
Usage of and Satisfaction with Online Help vs. Search Engines for Aid in Software Use

Charles J. Welty
University of Southern Maine
Department of Computer Science
96 Falmouth Street
Portland, ME 04104-9300
Tel: 1.207.780.4240
welty@usm.maine.edu

ABSTRACT
Computer users have long been frustrated by software problems. It is unusual that the Help menu actually helps with the software problems they have. At the same time, computer science students and professionals have been using search engines to get help with the complex software they use. The use of search engines to get help with software by both computer scientists and students in other disciplines is investigated. Students from all disciplines tested were found to use and be more satisfied by search engines than Help. Further investigation showed that, generally, students went to other people more than Help or search engines but found search engines and people to be the most satisfactory sources of help. Recommendations are made to improve Help systems by incorporating aspects of search engines.

Categories and Subject Descriptors
H.1.2 [Information Systems]: User/Machine Systems—Human factors

General Terms
Documentation, Experimentation, Human Factors

Keywords
Help systems; search engines; statistical study.

1. INTRODUCTION
In the early 1970s a school in California had an annual faculty-student sports day. Neither the students nor the faculty were very good athletes. One student, a talented systems programmer, always made a dismal showing at cricket, lacrosse, baseball or any sport involving a bat or stick. A faculty member said, “Give him a manual, he can do anything with a manual.” Everyone seemed to get the picture of him swinging a big heavy systems manual and winning the game. Laughter ensued.

This student had the ability to read, understand and use manuals to find solutions to difficult programming problems. Manual use is rare now [11, 12, 18, 16, 5].

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

SIGDOC’11, October 3–5, 2011, Pisa, Italy.
Copyright 2011 ACM 978-1-4503-0936-3/11/10...$10.00.

With the advent of graphical user interfaces, software houses no longer routinely generated and distributed these massive manuals1. Instead, they now include the Help menu item and put the documentation directly at hand. In many ways this is an improvement over paper documentation. It is searchable and it automatically comes with any new version of the software. On the other hand, various aspects of Help dialogs are frustratingly poorly conceived, implemented [14, 1] and accepted [15, 2] by users. Also, the Help system usually does not contain nearly as much information as the manuals it replaces. This is not an argument for returning to paper manuals but for rethinking Help. All the above cited papers deal with approaches to alleviate computer-related user frustration but others have addressed this frustration more directly [3, 4, 6, 13, 9]: Help systems have not made significant progress since the early and embarrassing efforts with Microsoft Bob and Clippy [8]. Help systems were problematic then and continue to be problematic.

Anecdotal evidence implies that computer scientists are now using search engines for software help almost exclusively in place of Help systems. Has the use of search engines supplanted Help not only for computer scientists but for people in other disciplines? What long term direction will Help take? Also, due to its problems, could Help be changed so that Help in its current form ceases to exist?

The basic hypotheses that will be investigated here are:

HA. Computer science students use search engines for software help more than students from other disciplines.
HB. Computer science students are more satisfied with search engines than students in other disciplines.
HC. Students in diverse disciplines use search engines for help more than they use Help.
HD. Students from diverse disciplines are more satisfied with search engines than Help.
HE. Students from diverse disciplines are more satisfied with search engines than with any other source of help.

The following conjecture, although not statistically analyzed, is of interest:
CA. Students in diverse disciplines use search engines for help more than they use any other source of help.

1 Of course, many bookstores have large sections devoted to texts on how to use software. This will be discussed briefly later.
Confirming or rejecting these hypotheses will lead to recommendations for improving help systems.

2. THE CASE FOR SEARCH ENGINES
When finding the answers to a multitude of questions, people no longer go to physical libraries or encyclopedias; they now go to the search engine, especially the Google™ search engine. The popularity of search engines is due to the excellence of their responses to search queries. At the same time many more users with diverse backgrounds, and few computer skills, have become involved with office software such as word processing, spreadsheets, presentation builders and many others. Each of these software products contains a Help menu attempting to supply the user with much needed help. Help systems have limited search capabilities and their navigation often conforms to the structure of the associated software. They also bring only the support documentation supplied by the software house to bear in finding answers to software questions. There are many other available resources of help they do not reveal.

Faculty and student experience points to computer science students and professionals using search engines to obtain help in using software such as .NET and Java nearly to the exclusion of Help. Search engines allow the use of multiple terms, including terms that may not be part of the Help vocabulary but make sense to the user. If these terms are also used by others, a search engine will find the relevant sites. The search engine will find forums, blogs and individual web sites that Help would never find. The search engine also usually does a better job of indexing into the product’s online Help than the software product itself. Thus, even if a search uses the same terms that the Help builders used, search will index into their documentation better than Help. The user will also benefit from other sources of information. In addition, much open source software is only detailed in forums and blogs, which lend themselves to search engines.

Several of the above points result in it being much easier to navigate using a search engine than Help (Novick and Ward, 2006b). The hierarchical organization of the Help system is often problematic. The flatness of the Web eliminates this problem.

Another advantage that a search engine brings is that it presents a single, often familiar, interface. If several pieces of software are used, each will have its own Help structure. The search engine’s interface is always the same. Recreational experience with a search engine can transfer to searching for help with a much reduced learning curve. This is an aid to new users of a software system.

3. POPULATION DEMOGRAPHICS AND REPRODUCIBILITY OF RESULTS
The University is an urban, comprehensive university. Its urban setting results in a large number of commuter students although the majority are full-time students. The average age is 25. The term “comprehensive” means that it is mainly an undergraduate institution with several master’s degree programs but few doctoral programs. Participants in this questionnaire were not randomly chosen; they were whatever students chose to take the surveyed classes.

Reproducibility would depend on the population demographics of a future study as well as any changes to software and web usage that could affect the results.

4. THE SURVEY
The questionnaire went through several pilot tests with faculty and 56 students. The resulting one-page questionnaire was distributed to students in different majors: arts and humanities (Arts, 94 students), business (Bus, 92 students), computer science (CS, 62 students) and sciences and engineering (Sci, 85 students). Questionnaires were distributed in first year through fourth year classes in a variety of subjects including physics, chemistry, introduction to business, market research, political science, criminology, philosophy, English, and the entire range of computer science courses offered in that semester. The students were given approximately 10 minutes to complete the questionnaire at the beginning of each surveyed classes.

The first question determined the major, year in school, years of computer use and favorite search engine of participants. Question two served to introduce participants to what sorts of software we were discussing, from social networking through programming. To encouraging their thinking fairly deeply about their use of each software category, participants were asked to rate their expertise in each area. This data was not found to be useful for analysis.

Question 3., asking how frequently they used Help or search engines and how satisfied they were with the results, is shown in Fig. 1.; it uses a Likert scale.

The final question (Q.4., also in Fig. 1.) asked what resources students used to get help with software and how satisfied they were with those resources. In both the following question statement and the brief oral introduction to the questionnaire, participants were told that the use of human resources, text and any other sources were of interest, not just search engines and Help.

---

2 Microsoft eliminated the Help menu in Office 2007, using a question mark icon instead. Other manufacturers use function keys. Both the question mark and function keys have the same functionality as Help. Thus, in the following, the term Help refers to any of these mechanisms.
5. OVERALL SURVEY RESULTS

The survey was run using the above questions and the four groups of participants. Questions Q1 and Q2 were not statistically analyzed. Q1 did show that 86% of participants preferred Google: 79% of arts students preferred Google, 86% of business students, 95% of computer science students and 89% of science and engineering students.

Detailed analysis of Q3 and Q4 is given in the following section. Questionnaire question Q3 gathered data on participants’ use and satisfaction with Search and Help. The participant groups were statistically compared (see below) to each other to see if there were any differences between them in their use and satisfaction with Search. While there were no statistically significant differences between the non-CS groups, computer science participants used and were satisfied by Search significantly more than all other groups. Thus hypotheses HA and HB were confirmed.

Also in questionnaire question Q3, the paired responses of participants evaluating their use and satisfaction with Help and Search were statistically tested, see below. The overall group and the individual groups all showed that they used and were satisfied by search significantly more than Help. Thus, hypotheses HC and HD were confirmed. The results are graphed as percentages in Figures 2 – 5. (If these figures are seen in black and white, the bars are shown in the order Arts, Bus, CS, Sci, starting from the left.)

WhenFig.’s 2 and 3 (Help) are compared to Fig.’s 4 and 5 (Search), the higher usage and satisfaction with Search over Help is immediately noticeable. This seems to be true for all groups, especially CS. These differences will be statistically analyzed in the next section.

Fig. 1 Questions 3. and 4. from the survey.

Figure 2. Frequency that participants used Help. Survey question Q3.A.

Figure 3. Frequency that participants were satisfied with Help. Survey question Q3.B.
Questionnaire question Q4.A. was concerned with sources of help that participants used in addition to search engines and Help, see Fig. 6. Responses to Q4.A. were categorized as being from Person (friend, family, teacher, help desk, etc), Search (Google, Yahoo, Blackle, etc.), Help, and Text (books, manuals, user guides, etc.). Figure 6 summarizes the result of Q4.A. These results were not statistically analyzed as discussed below. Overall, Person was the most cited source of help (37%). Thus, conjecture CA is not supported.

In question Q4.A., participants were not asked to state their help preferences in any order. Therefore choice order is not analyzed. Asking for choice order caused problems in the pilot studies, see below.

Questionnaire question Q4.B. assessed satisfaction with the help sources found in part Q4.A. These results are not as amenable to graphic representation and will be presented thoroughly in the next section. Overall, there were significant differences between the levels of satisfaction with the four help sources. Further investigation showed that Arts, Business and CS each showed statistically significant differences in satisfaction between the help sources.

6. DETAILED SURVEY RESULTS
Each of the following sections will address pairs of the 5 hypotheses and one conjecture formulated above.

6.1 Hypothesis HA. Computer science students use search engines for software help more than students from other disciplines.

Hypothesis HB. Computer science students are more satisfied with search engines than students in other disciplines.

The results for Science were not significant. Drilling down one more level shows that Arts, Business, CS and Sci were significantly more satisfied with Search than Help. Overall, Person was favored over Search. Only CS participants were statistically more satisfied with Search than Person. Also, only the CS group found Help significantly less satisfying when compared to any other help source. See below for the analysis. HE. is not supported.

Details of the statistical analysis are given in the next section. A test was significant if the p value was less than 0.01. Testing with 3 or 4 groups was done using the Kruskal-Wallis test. Tests of two separate groups used Mann-Whitney. Paired tests used the Wilcoxon-Test [7]. All are non-parametric tests, as discussed below.
Are these differences statistically significant? As stated above, participants in this survey were not randomly chosen. Also, the values on the Likert scale (1 to 7) are ordinal, not cardinal, values and do not allow the arithmetic operations required to find means, standard deviation, etc. For both of these reasons, the standard parametric statistical tests (e.g., analysis of variance) cannot be used. Instead, the non-parametric Kruskal-Wallis test is used.

Table 2. summarizes the results of the Kruskal-Wallis test. Given the 4 participant groupings used, the test had 3 degrees of freedom. The null hypothesis (that the distribution of the various population medians were equal) was rejected if the test statistic, p, was less than 0.01. This will be the form of the null hypothesis and its rejection whenever Kruskal-Wallis is used.

Table 2. Statistical significance of the dependent variables on the combined discipline groups using the Kruskal Wallis Test. Cells in bold, italic face type show significant differences, those in plain face are not significant.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median (mean)</td>
<td>Median (mean)</td>
<td>Median (mean)</td>
<td>Median (mean)</td>
</tr>
<tr>
<td>Arts</td>
<td>3 (3.20)</td>
<td>4 (3.43)</td>
<td>6 (5.25)</td>
<td>6 (5.57)</td>
</tr>
<tr>
<td>Bus</td>
<td>4 (3.82)</td>
<td>4 (3.77)</td>
<td>6 (4.81)</td>
<td>6 (5.20)</td>
</tr>
<tr>
<td>CS</td>
<td>2.5 (3.02)</td>
<td>3.5 (3.27)</td>
<td>6 (6.08)</td>
<td>6 (6.13)</td>
</tr>
<tr>
<td>Sci</td>
<td>4 (3.86)</td>
<td>3 (3.55)</td>
<td>5 (4.96)</td>
<td>6 (5.53)</td>
</tr>
<tr>
<td>All</td>
<td>3 (3.52)</td>
<td>3 (3.53)</td>
<td>6 (5.23)</td>
<td>6 (5.47)</td>
</tr>
</tbody>
</table>

The Mann-Whitney test, another non-parametric test comparing medians of two groups at a time, was used to drill down into the data. These tests were run on dependent variables that showed significant differences in the Kruskal-Wallis tests, see Table 2. This pursuit of the source of statistical significance is generally deprecated due to the dependent variables probably not being linearly independent of each other. These results are reported anyway but the increased probability of error must be noted. The null hypotheses were that the CS group’s median distributions are less than or equal to that of the compared group and were rejected with p < .01. These results are in Table 3.

The only significant results between groups of two were for CS versus each of the other groups; none of the non-CS groups (Arts/Business, Arts/Science and Science/Business) showed significant differences. This shows that the statistically significant overall results (Table 2.) were entirely due to CS students and, thus, CS students used and were satisfied with search significantly more than students in other majors. Hypotheses HA and HB are accepted.

Table 3. The first 3 rows show the statistical significance of the comparisons of the CS group with each of the non-CS groups using the Mann-Whitney test. A Kruskal Wallis test (last row of Table 3.) summarizes the 3 Mann-Whitney tests that showed the differences between the non-CS groups were not significant. Cells in bold, italic face type show significant differences, those in plain face are not significant.

<table>
<thead>
<tr>
<th></th>
<th>Q3.C. Search Used</th>
<th>Q3.D. Search Satisfied</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>p</td>
<td></td>
</tr>
<tr>
<td>CS/Sci</td>
<td>&lt; 0.0001</td>
<td>0.0069</td>
</tr>
<tr>
<td>CS/Arts</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>CS/Bus</td>
<td>&lt; 0.0001</td>
<td>0.0003</td>
</tr>
<tr>
<td>Sci/Art/Bus</td>
<td>0.4488</td>
<td>0.1948</td>
</tr>
</tbody>
</table>

6.2 Hypothesis HC. Students in diverse disciplines use search engines for help more than they use Help.

Hypothesis HD. Students from diverse disciplines are more satisfied with search engines than Help.

Further investigating Question Q3 (level of use and satisfaction with search and Help); the pairs of responses for each participant on their use and satisfaction are tested using the non-parametric Wilcoxon test, with p < .01. This investigates whether or not each individual participant used and/or was satisfied more with Search or Help. The following results were found.

Table 4 shows the paired use and satisfaction data for all (column All) participants were statistically tested first and the results were significant. Drilling deeper, the tests showed a significant difference in each group’s use and satisfaction with Search and Help. Thus, these overall results are not due to just a subset of the groups as was found earlier (see Tables 2. and 3.) with computer science participants skewing the results for all the groups. Students from all disciplines used and were satisfied with Search significantly more than they were with Help, confirming hypotheses HC and HD.

6.3 Conjecture CA. Students in diverse disciplines use search engines for help more than they use any other source of help.

Hypothesis HE. Students from diverse disciplines are more satisfied with search engines than with any other source of help.
Table 6. Median and mean satisfaction of participants with the various help sources in question Q4.B.

<table>
<thead>
<tr>
<th>Source</th>
<th>Arts (Median)</th>
<th>Bus (Median)</th>
<th>CS (Median)</th>
<th>Sci (Median)</th>
<th>Overall (Median)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person</td>
<td>5 (5.54)</td>
<td>6 (5.53)</td>
<td>4 (4.93)</td>
<td>6 (5.27)</td>
<td>6 (5.36)</td>
</tr>
<tr>
<td>Search</td>
<td>6 (5.55)</td>
<td>6 (5.58)</td>
<td>6 (5.82)</td>
<td>6 (5.44)</td>
<td>6 (5.59)</td>
</tr>
<tr>
<td>Help</td>
<td>6 (4.30)</td>
<td>6 (5.60)</td>
<td>5 (5.24)</td>
<td>5 (5.21)</td>
<td>5 (5.09)</td>
</tr>
<tr>
<td>Text</td>
<td>7 (6.4)</td>
<td>6.5 (6.5)</td>
<td>6 (6.20)</td>
<td>6 (5.80)</td>
<td>6 (6.18)</td>
</tr>
<tr>
<td>Other</td>
<td>7.5 (6.4)</td>
<td>6.5 (6.5)</td>
<td>6 (6.20)</td>
<td>6 (5.80)</td>
<td>6 (6.18)</td>
</tr>
</tbody>
</table>

Question Q4.A. asked each participant to state three actual sources from which they get help using software. The responses were varied, but were classified as being from Person (e.g., teacher, tutors, friend, family, help desk, help line), Search (e.g., Google, Yahoo, and other software with a search feature, such as YouTube), Help, Text (books\(^3\), user manuals) or Other (not easily classifiable sources). The results are shown in Table 5.

The data show that in all groups, except CS, the Person count was higher than the Search count. Help was third for the non-CS groups but was fourth for CS. This shows that three of the four groups preferred Person to Search. No appropriate statistical tests were found that fit this situation without stretching the limits of statistical credulity. The percentages have to speak for themselves. Conjecture CA is apparently, but not statistically, contradicted because Person was the most used help source, not Search. (Note: The original test design had the students rate their three help sources in order of preference. This caused a variety of problems. It was decided just to have them list the top three, not in order of preference.)

The final question, Q4.B., asks participants to rate their satisfaction with each of the help sources listed in their response to Q4.A. Both medians and means are reported in Table 6 as discussed above.

Although Table 5 shows that participants listed Person most often as the source of help, the medians in Table 6 show Search and Person to be very close in level of satisfaction. Both Person and Search were more satisfactory than Help or Text. Help is the least satisfactory for all groups; it is the only row with 4 or 4.5 as the medians. Computer scientists were most satisfied with Search and the least satisfied with Help of all groups.

\(^3\) The primary text source was an assigned class text. Students did not seem to buy books specifically to aid in the use of software.
help Person/Tex
help but results were mixed for other help pairings. Thus, help systems.
help
help, People
participants significantly preferred all other sources of help to show that Search significantly dominates H
see where any differences occurred. The results, in Table 8,
and Sci were further investigated using the paired Wilcoxon test
and Text.
other mechanisms that participants reported using: H
other help sources (People and Text) used by the
participants differed in their satisfaction with the four sources (Person, Search, Help, and Text).
This was followed by the tests within each major. Cells in bold, italic face type show significant differences, those in plain face are not significant.

<table>
<thead>
<tr>
<th>Satisfaction with sources</th>
<th>Overall</th>
<th>Arts</th>
<th>Bus</th>
<th>CS</th>
<th>Sci</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>0.0009</td>
<td>&lt; 0.0001</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Table 8. Statistical results using the paired Wilcoxon test of satisfaction within majors and between help sources. Cells in bold, italic face type show significant differences, those in plain face are not significant.

<table>
<thead>
<tr>
<th></th>
<th>Person/Search</th>
<th>Person/Help</th>
<th>Person/Text</th>
<th>Search/Help</th>
<th>Search/Text</th>
<th>Text/Help</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arts</td>
<td>0.9210</td>
<td>0.0362</td>
<td>0.4223</td>
<td><strong>0.00706</strong></td>
<td><strong>0.0041</strong></td>
<td>0.7983</td>
</tr>
<tr>
<td>Bus</td>
<td>0.2040</td>
<td><strong>0.0012</strong></td>
<td>0.7226</td>
<td><strong>0.000914</strong></td>
<td>0.8454</td>
<td>0.4237</td>
</tr>
<tr>
<td>CS</td>
<td>&lt; 0.0001</td>
<td><strong>0.0005</strong></td>
<td>0.9192</td>
<td><strong>0.000132</strong></td>
<td>0.0186</td>
<td><strong>0.0003</strong></td>
</tr>
<tr>
<td>Sci</td>
<td>0.5153</td>
<td>0.1925</td>
<td>0.8597</td>
<td><strong>0.0051</strong></td>
<td>0.1557</td>
<td>0.7276</td>
</tr>
</tbody>
</table>

Table 7. First shows the results of statistical testing of the overall participant satisfaction with the help sources. For this, Other was eliminated due to the few respondents and the variety and ambiguity of the sources classified as Other. The overall test was significant so the component majors were then compared. Again, the Kruskal-Wallis test is used with significance at p < 0.01.

All majors except Science showed significance differences in their overall satisfaction with the help sources. Arts, Business, CS and Sci were further investigated using the paired Wilcoxon test to see where any differences occurred. The results, in Table 8, show that Search significantly dominates Help for all groups. CS participants significantly preferred all other sources of help to Help. No group other than Computer Science significantly preferred Search to Person. All pairings other than Search/Help are inconclusive for some groups of participants. These results do not reject the null hypothesis. Thus, hypothesis HE is rejected. Search is not significantly more satisfactory than any other source of help.

7. CONCLUSIONS AND RECOMMENDATIONS

Three hundred and thirty three undergraduate students were surveyed about their sources of help in using software. The participants were placed in 4 groups, based on their major: arts, business, computer science, and science (with engineering). When use and satisfaction with a search engine was compared directly with the use of the standard software help mechanism, every group used and was satisfied significantly more with a search engine. Computer science students used and were significantly more satisfied with a search engine than any other group. Thus, Search was perceived by the study participants as superior to Help for aid in software use.

The results are much less conclusive when comparing Search and Help with other help sources (People and Text) used by the participants. Each tested group still significantly preferred Search to Help but results were mixed for other help pairings. Thus, Search is not the preferred help mechanism when compared to the other mechanisms that participants reported using: Help, People and Text.

Given the above, how might software companies provide better support to users? The obvious recommendation is to augment the Help mechanism with a tailored search engine. In order to implement this recommendation, all applications, whether web-based or not, would have searchable online help documentation supported by more in-depth, manual-like, online, searchable documentation. The search should be tailored to allow users to also access trusted forums, blogs and other relevant sites. The included sites could be expanded as more are found relevant. The user could also choose to search additional specific sites or the entire web.

Adobe Creative Suite 5 uses Google Site Search to power its Adobe Community Help Centre system [10]. Google Site Search allows Adobe to index into its own content as well as specific community-recommended sites. These sites include blogs, forums, etc. This application almost fully embodies the above recommendation.

It is interesting that, while participants went to other people for help more than to any other resource (Table 5), they were not consistently more satisfied with people than any other source (Table 8). In a campus setting, people are usually available and often helpful. As computing has become more ubiquitous, people have the web available wherever they go. This could drastically affect where people go most often for aid.

In conclusion, more complex software has made its use more and more difficult. People need tools to easily find solutions to their software problems. User satisfaction is the key to successful software. People are not satisfied with current Help systems. Therefore, Help must be actively reviewed and changed by both new and established software enterprises.

7. ACKNOWLEDGEMENTS

Many thanks are due to the students and faculty that participated in this study.
8. REFERENCES


