Quantitatively Analyzing the Use and Usefulness of the Design Learning Simulator

Jennifer Turns, Farrokh Mistree, Janet K. Allen
Georgia Institute of Technology

Abstract: Current trends in engineering design education, which promote teams of students solving open ended problems, often result in classes which create a wide variety of logistical, cognitive, and motivational problems for students. Software resources can help students but only if students perceive them to be useful and make use of them. Our Design-Learning Simulator at the Georgia Institute of Technology contains a variety of resources that support the different problems students have in the doing of and learning about design through experience. Through the Design Learning Simulator research, we have been exploring issues about what resources to provide to students and how to make them available. In the Spring, 1996 quarter, the Design Learning Simulator was implemented in a Web-based platform and included model design reports, electronic versions of class documents, resources to support team formation, resources to negotiate project requirements, and an on-line parts catalog. During this period, we received generally positive feedback from the students through survey questions, exit interviews, and testimonials. In this paper, we explore a quantitative approach for understanding the use and usefulness of the software, the analysis of the log files of student activity. Web log files were analyzed to address questions about how and where the resources were being used in general, as well as over time and in relation to critical periods in the course. We found that all resources, with the exception of the parts catalog, were used effectively and that the web-based implementation, promoting platform independent and universal access, was important. In the paper, we report on the analysis and conclude with recommendations for the continued development of the software and for the next steps in the research.

1. ME3110 AND THE DESIGN LEARNING SIMULATOR

A recent survey conducted by the National Society of Professional Engineers found that while industry places a very high value on design and teamwork skills, the preparedness of the engineering graduates is very low [1]. Findings such as these are used to argue that engineering students need more and qualitatively different design experiences than currently exist within the curricula. Such experiences are supposed to provide students with the opportunity to solve open ended problems, to work in teams, and to treat design in a more formal manner [2-4]. In practice, such experience-based engineering design education can be difficult to create and challenging to sustain. Strategies and resources, including *software resources*, are needed to make the teaching with such experiences more feasible.

ME3110: Creative Decisions and Design is a junior level engineering design course which we have been teaching for over 10 years. In this course, we strive to help students learn about design as an intellectual cognitive activity through their participation in extended, team-based design experiences with open ended problems [5]. Self formed teams of students tackle open ended design problems, moving through a term long, structured design process - the Decision Support Process Technique [5] - to arrive at a working complex mechanical artifact by the end of the term. The process consists of ten phases that distribute the workload across the eleven weeks of the term (see Table 1) and takes the student through the stages of conceptual design and meta-design (planning), through detailed design, to construction, demonstration, and even a little marketing.

Table 1. ME3110 Class Periods

Period ID	Weeks	Description
Bid	1.5	**Bid**: Newly formed student teams must submit a report detailing their team's work ethic and their statement of the problem, resembling a contractor bid.
DR1	1.0	**System Conceptualization** : The first step for the team is to identify and allocated subsystems and identify possible concepts for each subsystem.
DR2	1.0	**Meta-Design** : The team now creates a plan of their activities for the remainder of the quarter, using a variety of planning representations.
DR3a	1.0	**Preliminary Selection** : Individual team members return to their subsystem ideas from DR1 and use mathematics to determine the most promising ideas.
DR3b	1.0	**Selection** : Individual team members use harder information to determine the best candidate solution from the set identified in DR3a.
DR4	1.0	**System Configuration** : The team now comes together in full force to determine how to go about interfacing the subsystems and building the final system.
DR5	1.0	**Parameterization** : Team members use optimization methods to determine appropriate values for subsystem parameters like gear size and batter voltage.
Demo	1.5	**Demonstration** : The demonstration of a working device is required for the team to complete the course.
Report	1.0	**Final Report** : The final report includes both a sales presentation of the device, a sales brochure, the device itself, and the collected work over the term.
Exam	1.0	**Final Exam** : The final exam focuses on general design concepts such as meta design as well as the solving of design problems with mathematics.
Total	11.0	

Because teaching such a course presents serious logistical, cognitive, and motivational challenges for all parties involved, we have been developing different *software resources* to support the students with some of the various difficulties that they encounter. We call our collection of software resources the *Design Learning Simulator* [6]. We have been developing the Design Learning Simulator for over two years. Throughout this period, students have used the Design Learning Simulator and given general qualitative feedback suggesting that it is useful. In this paper we report on our use log file analysis to quantify the use of the Design Learning Simulator in order to address questions we have about its design. We anchor our study in the use of the Design Learning Simulator during the Spring 1996 quarter.

2. THE DESIGN LEARNING SIMULATOR

The Design Learning Simulator (D-LS)[1] in a computer-based environment to support students in their efforts to learn about design and learn through the process of designing. Through the Design Learning Simulator research, we have been exploring questions about what resources to provide to students, how to categorize and organize the resources, and how to distribute the resources.

Previously, we have suggested that the D-LS be organized around a taxonomy of resources including process resources and knowledge-base resources. Process resources are those which help students through class and design specific processes such as forming teams. Knowledge-base resources are those which are collections of information such as a catalog of data about mechanical parts [6]. The D-LS is meant to be available in addition to general purpose tools such as CAD, word processing, spreadsheet, and presentation software.

The components of the D-LS change frequently as we test the usefulness of different resources. For example, past versions of the D-LS included a resource for electronic collaboration, CaMILE. CaMILE's role in the D-LS was phased out as we learned that the benefits *reported* by students did not outweigh the time and effort required to support the software [7]. Current versions of the D-LS include a resource for the articulation of lessons learned (the Reflective Learner [8]) and resources for solving several kinds of design decision problems.

[1] The Design Learning Simulator is currently instantiated in the ME3110 Web, located at the following address: http://srl.marc.gatech.edu/education/ME3110/ME3110-Web.html.

During the Spring 1996 quarter, the D-LS included the five resources distributed across the two categories. The process resources included ones to support team formation and rule clarification while the knowledge-base resources included electronic versions of class documents, example design deliverables, and a catalog of common design components. For the most part, these resources have been developed in response to specific problems that students in ME3110 experience as they work through the design process over the term, as discussed below:

- **Class Documents -** Referencing public class documents: Over the years, the organizers of the course have developed many different paper resources for students. Traditionally, students have purchased this collection of information from the copy center. Increasingly, this information is being put on-line. The on-line information in the Spring 96 quarter included the forms (guidelines on each activity the students are asked to perform), memos (documents which are to be filled out and turned in), and copies of old final exams.

- **Team Information -** Self Organizing into Teams: Within only two weeks of the beginning of the quarter, the students in the class must find a set of teammates with whom their aspirations, work habits, and interests are compatible. By the Spring, 96 quarter, the D-LS had come to contain an interactive facility which allows students to 1) enter and edit profile information about themselves and 2) search the database for other possible teammates matching specified criteria.

- **Design Reports -** Making Sense of the Report Requirements: The students in the class must submit seven design reports concurrent with the construction, and prior to the demonstration, of the mechanical device. The specifications of the design reports seem vague to the students who are often accustomed to having all details of class assignments well defined. The D-LS provides on-line versions of a complete set on submitted design reports from the Spring 1995 term. Students can reference these reports to resolve ambiguities.

- **Rules and Clarifications** - Negotiating Project Requirements: The class project consists of a cover story describing the goals the artifact is to be designed to fulfill and the rules which govern the set of acceptable solutions. As the term progresses and students develop their design ideas, they begin to petition for clarifications of the rules. During the Spring 1996 quarter, petitions for rule clarifications could be submitted either on paper or electronically through the D-LS. Clarifications were made available solely through the D-LS. At any point in time, the D-LS served as the repository of the currently negotiated set of rules.

- **Parts Catalog -** Searching for Design Components: While the students in the class do perform extensive paper-based prototyping and submission of reports, ultimately they must build and demonstrate a working complex mechanical artifact. Many of these students have never had to locate and purchase the types of parts required for their artifact, such as gears and power sources. The students are often uncertain about the types of components available (e.g. what size gears are standard) and the suppliers of such parts. The parts catalog is a resource where students can explore these issues prior to going out and making the purchases.

In addition to being concerned with what types of and which specific resources to provide, we have been exploring different mechanisms for making software available to students. Our prior experience strongly suggests that the choices one makes in how the software is distributed will affect how much students use it and subsequently, how useful they find it. Many of the initial resources that we initially developed and/or made available to students (e.g. CaMILE - the electronic collaboration software, an electronic version of the textbook, and software to support the solving of various design decision problems) were made available solely on Macintosh computers and only in one or two lab facilities on the school's campus. Students reported low levels of use and suggested that the potential usefulness of the resources did not surpass the combined effort required to 1) learn a new computer platform (i.e. most students reported being PC users) and 2) go to the resource's location. As a result, the students continued to perform their activities in the old manner and the resources *apparently* went largely unused. Based on these experiences, we have moved the D-LS to a Web-based platform to provide the students with almost universal access and platform independence.

3. EVALUATION QUESTIONS

Throughout its life, the D-LS has been publicly available to students and to the engineering education community. During this time, we have relied on comments from these groups, ad-hoc surveys, and our perceptions of classroom experiences to judge the productiveness of the Design Learning Simulator research. In this paper, we start to formalize our evaluation of the D-LS. We begin by exploring the answers to several questions suggested by our current approach to designing the D-LS. In particular, we have created a set of resources that we believe address some students problems over the quarter and have made them available through a Web implementation in order to provide platform independence and universal access. We believe that, as a result, students will find it easy to access the collection of resources and upon use, will find them useful and continue to access and use them. Thus, we expect students from many diverse locations to be accessing the software repeatedly over the quarter. In addition, because different resources are targeted at different class activities (e.g. forming a team, verifying the project rules), we expect the use of the resources to change over the quarter reflecting the activities over the quarter.

In particular, we explore the following specific question: <u>To what extent are the D-LS resources being used over the term and from where?</u> This overall question can be broken down into several sub questions which represent our expectations and are covered in the analysis:

- *Resource Use*: Were all of the resources in the D-LS used? Did any resources go unused? Were the resources accessed repeatedly? Does the resource use constitute effective use by the students and teams? Does the resource use reflect the changing demands of the course (i.e. team formation through conceptual design to detailed design, construction, and demonstration)?
- *Resource Accessibility*: Did having the D-LS available through the Web seem to support making it "universally" accessible (to a variety of users)? Did the users of the D-LS access it from diverse locations?

4. THE METHOD: LOG FILE ANALYSIS OF SPRING, 1996 USE OF THE D-LS

We have explored the answers to these questions through log file analysis. Log files are files produced by software which record a "trace of user events", typically a sequence of time stamped user actions, during interactions with the software [9]. Log file analysis is attractive for at least two reasons: the data is cheap since data gathering can be totally automated and the data is ecologically valid since it can be collected outside of the laboratory while the user is engaged in real tasks [10]. Raw log file data can be manipulated to provide information on unique events, sequences (orderings) of events, and the duration of events [9]. The grain size of the analysis that can be conducted with log files is constrained by the size and types of events which are being recorded. In our case, we used Web log files generated automatically by the web server.

Web log files contain a single line for each *hit* (i.e., a request for a file containing information) to the Web server. The recording of the hit includes the IP (i.e., internet protocol) address from which the request originated, the file that was requested, and the time stamp of the request. A section of the log file analyzed in the paper is included in Figure 1. Web servers automatically record all hits to all files accessed through the server (e.g., html files, ftp sites, graphic images, movies, etc.). Perhaps the two main attractions of analyzing Web server log files is the ease of data collection (i.e. it happens automatically) and the volume that is available.

Web log file analysis is not without a few issues.

- First, log files generated by a Web server software contains a record of hits to ALL information accessed through the server over time, which may include a lot of information unnecessary for a particular analysis. The files associated with the D-LS represent only a subset of those available through the server. If the software of interest represents only a portion of the information accessible through the server, then it will be necessary to extract that information from the larger log file.
- Second, the units of recording used in Web log files, hits, represent requests to the server for all types of files. A request for a single *page* as viewed by a user can result in several hits if the page references files such as images and movies. Thus, these hits may not necessarily the best unit of analysis for the log file analysis.
- Third, Web log files identify accesses based on the IP address of the machine which made the request, not the identity of the actual user. While it is sometimes appropriate to infer that different IP addresses represents different users (e.g. IP addresses associated with private computers), this inference is often not appropriate (e.g. IP addresses associated with publicly available computers such as those available to students in a library).
- Fourth and finally, hits are not recorded for all activity performed by a user. In particular, when a user, during an interaction with the Web, returns to a page that he/she has already viewed, there is a good chance that the information for that page is still available in the local memory of his/her browser (e.g., Netscape). If the information is still available, the browser will redisplay the page from the cached information, never accessing the server and thus not causing a hit to be recorded in the log file. Thus, not all of a user's activity will be recorded in the log file.

The first two of these issues simply make the analysis of Web log files more challenging and are

```
srlbrass.marc.gatech.edu - - [26/Mar/1996:11:03:13 -0500] "GET /cgi-bin/rules/3110RulesDisplay HTTP/1.0" 500 0
fr_ibm11.gatech.edu - - [27/Mar/1996:12:47:28 -0500] "GET /cgi-bin/rules/3110RulesDisplay HTTP/1.0" 500 0
apts132.residence.gatech.edu - - [27/Mar/1996:21:15:46 -0500] "GET /cgi-bin/rules/3110RulesDisplay HTTP/1.0" 500 0
```

Figure 1. Excerpt from D-LS Log file

the basis for the first two steps in the procedure below; the third and fourth issues represent limitations of the analysis and must be embedded in the interpretation of the data. Despite these challenges and limitations, log files contain extensive (and free) information about how the software was used, as will be demonstrated in the following analysis.

4.1 The Procedure

Performing a log file analysis using the automatically generated Web server log file consists of three steps: 1) data extraction, 2) analysis concept definitions, and 3) exploration and hypothesis testing. The first two steps are necessary to respond to the first two log file analysis issues presented above. Data extraction is the process of identifying and collecting from the overall log file, the entries that are relevant for the given analysis, and relocating them into a separate workspace. Since Web log files identify hits based on the name (and path within the file system) of the file accessed, this extraction process is based on the features of the names and paths of the files that compose the resources of interest. In our case, we used two heuristics to extract the D-LS hits:

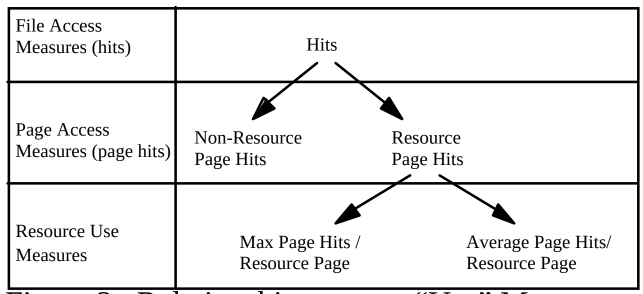

Figure 2. Relationships among "Use" Measures

- Extract all files with "ME3110" in their path (i.e., all files anywhere within the ME3110 directory). This provided the hits to all resources except the Team Information resource.
- Extract all files with "info_sheet" in their path. The term "info_sheet" is the name of the directory containing all of the files associated with the Team Information Resource, which was located outside of the ME3110 directory.

Combining the heuristics with the beginning and ending dates of the time period for the analysis (see Section 4.2) made it possible to identify and extract the portion of the overall Web log file relevant for the analysis.

Analysis concepts are the units of the analysis to be read or inferred from the log file data. In our analysis, we identified three sets of concepts to be used in addressing the analysis questions: resource use measures, class period, and address. Below we describe the nature of each concept and the basis for its inference (if necessary).

- *Resource Use Measures* - Hits are commonly used to gauge the level of use of Web resources, but as we have described earlier, hits are inflated by accesses to graphic images, movies, and other types of media stored in files. In response to this issue, we developed several additional concepts to use for our analysis. In Figure 2, we show these concepts and provide an indication of their relationships to each other. *Page hits* represent the subset of all hits which are to pages of information. Page hits can be classified as hits to a resource (*resource page hits)* or hits to parts of the D-LS which are not associated with any particular resource (*non-resource page hits)*. For example, a hits to the current listing of rule clarifications would be a resource page hit. A hit to the first page of the D-LS would not be a resource page hit since the first page of the D-LS is not directly associated with any particular resource. Page hits provide an indication of the level of use of a resource, but are in turn, biased by the number of pages associated with any particular resource. The *maximum number of pages hits for a resource page* (resource access measure #1) and the *average number of pages hits per resource page* (resource access measure #2) take this variable number of resource pages into account and thus provide estimates of the level of use which are comparable across resources

- *Class Period* - The date and time of each hit is recorded in the log file. Based on the due dates of the different assignments in the class, each hit can be assigned as corresponding to one of 10 class periods, as defined in Table 1.
- *Address* - While the actual identity of the user who invoked the hit is not available, the address of the machine, from which the request was made, is available.

In the exploration and hypothesis testing stage, the data is manipulated to answer the question and hypotheses of interest. In our case, we inserted our extracted data into a database and used structured query language (SQL) to pose the questions (e.g. select count(IPaddress) from dls_use_table).

4.2 The Context

We focused our analysis on the use of the D-LS during the Spring, 1996 academic term. This term lasted for approximately 90 days. During this term, three sections of ME3110, constituting approximately 90 students and 18 teams, were being encouraged to use the D-LS. The students were not required to use any of the features of the D-LS although many features were discussed in class, demonstrated in lab sessions, and suggested to be useful by both the class instructors and the teaching assistants. The class workload was distributed across ten major sequential periods, as discussed earlier and presented in Table 1.

5. THE RESULTS

The extraction process resulted in over 27000 hits, over 10000 page hits, and approximately 7700 resource page hits. In the following sections, we describe how we use the information embedded in these page hits to explore our expectations about resource use and accessibility.

5.1 Resource Use

In Table 2 we have summarized descriptive statistics associated with the use of the five main resource and also the seven example design reports. From the page hits in column two of the table, we see that all of the resource experienced what seems to be non-trivial use (particularly if we recall that these values are not inflated by the over 15000 hits to image files). These page hit values suggest that the design reports were the most valuable resource with by far the most page hits, followed by the team information resource with less than half as many page hits, followed by the other three resources with total page hits similar to each other. Of the available example reports, the page hit values suggest that the DR4 and DR3b reports are the most used.

Table 2. The distribution of Hits over Pages in each of the Functional Areas

Column	1	2	3	4	5	6	7	8
Measure / **Functional Area**	Page Hits	Total Page	Max. Hits /Page	Avg. Hits /Page	Ind?	Mul?	Hits/ Page/ Person	Hits/ Page/ Team
Resources								
Class Documents	832	35	164	24			0.3	1.5
Team Information	1602	11	240	146	y	y	1.6	--
Design Reports	3577	32	837	112			1.2	7.0
Rules and Clarifications	743	6	378	124		y	1.4	7.7
Parts Catalog	1020	106	222	10			0.1	0.6
Main Page	1187	1	1187	1187	-	-	13.2	74.2
Other	1444	5	1148	289	-	-	3.2	18.1
Total	10405	196						
Specific Design Reports								
Bid	101	1	89	101			1.1	6.3
DR1	226	2	188	113			1.3	7.1
DR2	332	2	197	176			2.0	11.0
DR3a	293	1	286	293			3.3	18.3
DR3b	437	2	236	218			2.4	13.6
DR4	785	15	136	52			0.6	3.3
DR5	232	2	161	116			1.3	7.3
Main Page	837	1	837	837	-	-	9.3	52.3
Other	334	3	68	111	-	-	1.2	7.0
Total	3577	29						

A striking feature of the values in the table is the large numbers of page hits in the categories *Main Page* and Other . Web sites often contain index, table of contents, or organizational pages which are simply pages with list of links (i.e., like a table of contents). Both the overall Web site that is the D-LS and the set of pages that represents the Design Reports Resource have such index pages. The number of hits to these "main pages" represents a rough estimate of the number of accesses to the D-LS overall (i.e., roughly 1200 accesses) and to the Design Report Resource (i.e., roughly 837 accesses). This number most likely represents an upper bound on the number of sessions with the D-LS, since it may include accesses to the Main Page in which a user does not go any farther. The "Other" categories include features of the D-LS which did not get categorized as part of any specific resource. These include a problem statement page in the Design Report Resource and a set of pages which provide a tutorial of the overall D-LS. The page hit counts in these two categories provide a measure of the overhead associated with using the D-LS to access and use the resources.

As discussed in section 4.1, using the resource page hits as a measure of resource usage provides a biased viewpoint. This value can be inflated for resources with a greater number of component pages. For example, the *Parts Catalog* resource experienced over 1000 page hits, but this is certainly related to the over 100 pages which make up the catalog. Column three of Table 2 shows the number of pages associated with each of the resources. For example, there are over one hundred pages associated with the *Parts Catalog* resource while only six pages associated with the resource "Rules and Clarifications".

Two resource use measures which take the number of pages into account are the *maximum number of page hits to any page* in the resource and the *average number of page hits per resource page* for a resource (see Figure 2). The values for these two measures occupy columns four and five of the table, respectively. These values provide a different portrayal of resource use than the page hit values, but generally similar to each other.

- The maximum page hits measure, which provides a possible upper bound on the number of times someone started to use the resource, again suggesting that the Design Reports were the most useful. The next most useful resources, though, are the Rules and Clarifications and then the Team Information resources. While this measure still suggest that the Class Documents were little used, it also suggests that the Parts Catalog may not have experienced much use. The only significant difference in the portrayed use of the individual example design reports is the use of DR4, which appears to be much less than that suggested by the page hits.
- The average page hits measure represents the level of use under the assumption that all pages are used during each access to the resource. These values suggest that the Team Information, Rules and Clarifications, and Design Reports resources were used frequently, followed far behind by the Class Documents and the Parts Catalog. Considering the individual example design reports, the DR3a report was most used, followed by the reports for DR3b and DR2, followed by the reports for DR1, DR5, and the Bid. The example report for DR4, which had previously appeared to have been the most widely used based on the page hits, is found to be the least widely used when the number of pages is taken into account.

While there are slight discrepancies between the level of use as portrayed by the two measures; taken together, they make it possible for use to infer that the three resources, Design Reports (including most of the individual reports), Rules and Clarifications, and Team Information, were widely used while the Class Documents and Parts Catalog resources were not. Interpreting these levels of resource use, though, requires additional assumptions.

5.1.1 Interpreting the Resource Use Levels

One may wonder, do these levels of use support a hypothesis that students found the resources useful and thus used them effectively? In order to answer this question, we need to consider the strategies that the teams of students could and would follow when accessing these resources. Two particular features of such a strategy seem important - would the resource be accessed by teams or by individuals and would the resource be accessed repeatedly or only once.

With some resources, it makes sense to believe that each member of the team would reference the resource individually (e.g., the team information resource is for *individuals* to find teammates, making it an individual resource). On the other hand, an effective distribution-of-labor strategy in a team might include assigning one team member the responsibility of keeping track of the current rules. In such a scenario, the number of accesses to the rules and clarifications resource would be interpreted relative to the number of teams rather than the number of individuals. Since the Team Information resource is the only resource which can be considered to be truly individual, we have coded it as an individual resource in column six of Table 2, leaving the others as team resources.

A second relevant issue for interpreting the level of use is the number of times we would expect students to return to a particular resource. Accessing a class document or even a model design report once might be an effective use of the resource; on the other hand we would expect multiple accesses per student (or team) for dynamically changing resources like the Rules and Clarifications and the Team Information resources. In accordance with this, we have coded the Team Information and the Rules and Clarifications resources as multiple use resources in column seven of Table 2, leaving the others as single use resources.

Based on the number of students and teams using the D-LS during the Spring, 1996 quarter, we can calculate a resource use measure per student and per team. In our case, we used the measure, average page hits per resource page, and calculated it per student and per team (i.e., column 5 of the Table 2 divided by 90 and 18 respectively). The resulting values are shown in columns eight and nine of Table 2.

Where the value in the eighth column exceeds one, we can tentatively suggest that each student referenced the resource at least once. This was true for the team information, the design reports, and the rules and clarifications resources and specifically for all of the example design reports with the exception of DR4. The Team Information resource, the only resource coded as individual use, was also coded as multiple use. The values suggest that individual students did use the resource more than once, but not quite twice. It appears that some students may have used the Team Information resource only once and never referred to it again, suggesting that the use may not have been completely effective.

When the value in the ninth column exceeds one, we can tentatively suggest that each team referenced the resource at least once. This was true for all resources with the exception of the parts catalog. The values for the Class Documents resource and the DR4 design report are particularly interesting since the number suggest that while each student did not make at least one use of the resource, each team did. In the particular case of the Rules and Clarifications resource, we expected the value to greatly exceed one showing repeated use. The average number of accesses per team to this resource is 7.7, which does not contradict an effective use belief. Tentatively, these values suggest reasonable levels of use for the each resource with the exception of the Parts Catalog, since each team did not use it at least once, and the Team Information resource, since many individuals may have used it once or less.

5.1.2 Resource Use over the Team

We were also interested in whether the resource use conformed to the demands of the term, suggesting that students found the resources useful in addressing their emerging problems. In Figure 3, we show the resource use measure, average page hits / resource page, for each of the class periods laid out in Table 1. By looking at the values in this manner, we can verify whether the use of the resources across the class is consistent with our expectations.

From the graph in Figure 3, it is possible to make several observations.

- We can see that the team information resource (team formation) is clearly very useful to students in the beginning, while the use of it subsides as the team is formed, with almost all use terminating once the first design report is submitted (the teams *must* be formed by that time).
- We can also see that the pattern of use associated with the Rules and Clarifications resource use mimics student behavior in the class. The use of this resource peaks at DR2. Prior to and during DR2, the students are struggling to define the problem and understand the rules. The use of the rules drops during the next two class periods as students complete DR3a and DR3b, individual assignments. The use of the resource increases again as students enter into the period in which they are constructing the actual artifact. Finally, the use of the resource drops off after the demonstration of the device.
- We can also see that the individual design reports are being used throughout the term, although with some variability in the amount of use during each period.
- Finally, we can see the marginal use of the class documents and the almost non-existent use of the Parts Catalog over the term. For the Class Documents, which are convenient to have in the D-LS but are redundant with materials most students own, the low level of use is not surprising. The lack of use of the parts catalog is surprising and something which needs to be explained.

In Figure 4, we have shown the design report use measure, average page hits / design report page, again for each of the class periods laid out in Table 1. We can see that the peaks at each class period correspond to the on-line version of the report that is due (as we would expect).

- One surprising feature of the graph is the evidence of student's planning behavior in their accesses to the design reports over the quarter. The trends for several of the example design reports show that the use of the individual design report gradually increased through the periods prior to its due date, peaked during the period in which the design report was due, and then dropped to almost no use in the periods after the due date. This

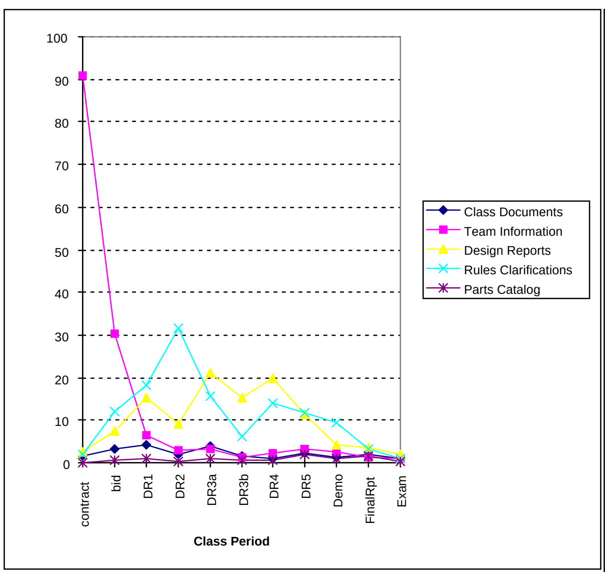

Figure 3. Average Page Hits / Resource Page within Class Periods

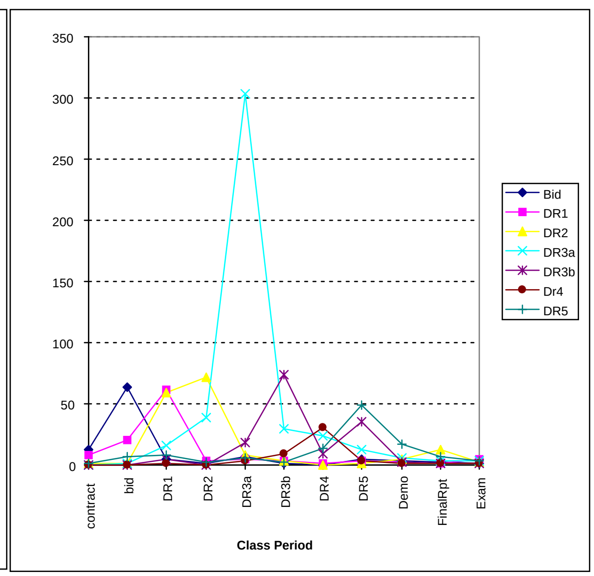

Figure 4. Average Page Hits / Design Report Page within Class Periods

behavior is most pronounced for the use of the example report, DR2. There was almost as much use of the DR2 model design report during the DR1 class period as during the DR2 class period.

- Of the individual model design reports, the one associated with class period DR3a appears to have been the most useful to the students. This report is the first truly individual activity as well as the first to require mathematics to be used for making design decisions. Students were repeatedly turning to the on-line design report for guidance.
- Another interesting feature is the existence of two peaks in the use of the DR3b report. The first period of high use occurred during the DR3b class period (as expected). The second period of high use corresponds to the DR5 class period. This feature of the use of DR3b is probably related to a combination of characteristics of the class. The course is difficult and demanding, and over time students get tired. Coupling this increasing tiredness with the class policy permitting students to resubmit any design report, results in many students submitting DR3b reports (an individual report) which need to be resubmitted. It seems that perhaps many of these DR3b reports were resubmitted during the DR5 class period, and the continued work on the DR3b reports during the DR5 periods explains the second peak in the use of that report.

5.2 Resource Accessibility

The second set of research issues concerns the accessibility of the system. We have been implementing the D-LS as a set of Web pages in order to provide nearly universal access and to provide platform independence. We would like to know whether these precautions are necessary. In our analysis, we explore distribution of the IP addresses (hereafter "addresses") in order to address these questions.

As we have stated earlier, addresses are associated with each recorded server hit. An address consists of a series of generally three to six identifiers, separated by periods (e.g. apts999.residence.gatech.edu). The identifiers, read from right to left, provide increasing specificity about the location of the machine making the hit. Although some addresses do not contain semantically meaningful units between the periods (e.g., 123.45.678.912 might be an address), most addresses do seem to make use of semantically meaningful units (e.g., apt999.residence.gatech.edu). Thus, from the address we can know not only know the number of different machines from which hits originate but also something about the location and owner of the machine.

In the top of Table 3, we have shown the distribution of addresses and page hits across five categories of addresses. The total number of hits provides a rough measure of the magnitude of activity from any given class of addresses. These five categories, educational at Georgia Tech, educational at other institutions, commercial, foreign, and unknown, are based primarily on the last two units of the address. Educational addresses end with the postfix "edu", commercial accesses with "com" and "net", and foreign end with two letters uniquely identifying the country. The second to last unit in the IP address of an educational institution usually identifies the institutions. In our case we could use the unit "gatech" to separate the Georgia Tech addresses from those of other educational institutions.

We see some interesting features in the top half of the table. First, over 600 different addresses were used to access the D-LS during the Spring 1996 quarter. Of these addresses, the majority are Georgia Tech addresses, as would be expected since the software is for the students of Georgia Tech. Additional accesses by Georgia Tech students are embedded in the categories for *other universities* and *commercial*. Some students accessed the D-LS from computers at the nearby Emory University. Many students accessed the D-LS through local commercial internet providers such as Mindspring. The majority of the addresses of unknown origin are those IP addresses which consist of only numbers. It is fascinating to see the large interest in the D-LS from the foreign audience, showing how the web can be used to distribute educational ideas across geographical borders.

Table 3. Distribution of Unique Addresses and Total Hits

Category	Distinct Addresses	Total Hits
All Accesses		
Educational-Georgia Tech	315	9374
Commercial	147	822
Unknown	55	677
Foreign	49	104
Educational-Other	39	216
Totals	*605*	*11193*
Georgia Tech Accesses Only		
Residence	40	2384
Systems Realization Lab	19	1295
Library Cluster	69	1136
Student Center	41	1058
Mechanical Engineering	15	793
Acme Cluster	3	512
Unknown	19	500
Manufacturing Building	8	364
French Cluster	27	319
Rich Building	21	285
Commons Cluster	17	253
Industrial Engineering	6	132
Matheson Cluster	8	108
Research Institute	9	82
Electrical Engineering	2	65
College of Computing	6	56
Physics Building	2	17
Management Building	3	15
Totals	*315*	*9374*

The 315 Georgia Tech addresses are particularly important since it seems a fair assumption that these accesses were made by individuals associated directly with the class. These addresses are very widely spread across campus locations, as we have also shown in the bottom half of Table 4. The campus locations were inferred based on the third-to-last unit in the IP address, the one before the unit “gatech”, as in apts999.*residence*.gatech.edu and lib99.*library*.gatech.edu. The largest collections of hits came from private residence internet connections, the second largest was from the laboratory housing the teaching assistants in ME3110 and the developers of the D-LS, and the next two from two highly used public campus clusters available to students. The number of distinct machines being used to access the information testifies to the importance of having the software widely available, not just installed on a few machines in a particular lab. With an average of 90 students taking the class, the over 300 unique addresses implies that each student used around of 3 different machines to access the software. The 40 different *residence* addresses used to access the D-LS suggests that just under half of the students were able to make use of the D-LS from the dormitory computers.

Another way of looking at the Georgia Tech specific addresses is by looking at the total number of hits for any given IP address. Two addresses each accumulated over 500 page hits. One, with 1033 page hits, was the address associated with the development machine for the Design Learning Simulator. The large use is not particularly surprising. The other, with 560 page hits, was a residential address. It would be interesting to uncover what this particularly student did with so much access. If, as seems likely, this one residential address represents a single user, it would also be interesting to relate the student’s use of the D-LS with his performance and learning in the class. This is not possible, though, since all we know about the student is the address of his/her machine. Overall, though, the total number of page hits for any given address tended to be rather small, with over half of the addresses having been used to access a total of 25 pages or less. This suggests that some users accessed the D-LS from a single place, but that most users moved around quite frequently.

Finally, in our prior experience with software in classes, we had noticed that students have strong platform preferences and that not having software available on both platforms can lead to students rejecting the software. Because many of the Georgia Tech addresses embed "MAC" and "IBM" directly in the address, it is possible to explore how much of the above use is on the different platforms. Just under half of the addresses were indecipherable with respect to platform but accounted for over two thirds of the page hits (139 addresses for 6332 page hits). Of the decipherable addresses, there is a tendency toward Macintosh use with both the unique addresses and the total page hits being about double the IBM values. There were 116 distinct Mac addresses accounting for 2161 page hits compared to 60 distinct IBM addresses for 881 page hits. Given that the platform of the remaining addresses is unknown, we cannot tell whether the students prefer Macintosh over IBM or vice versa. What we can tell is that both platforms are used quite frequently by students.

6. DISCUSSION AND RECOMMENDATIONS

In our analysis, we explored how the resources in the D-LS during the Spring, 1996 were used. We learned that 1) the Design Report resource was the most widely and consistently used resource (with all individual reports experienced non-trivial use), 2) the use of both the Rules and Clarifications resource and the Team Information resource peaked during the expected class periods, and 3) the Class Documents were not referenced frequently but did seem to be used. The only resource which did not experience use consistent with our expectations was the parts catalog, which was rarely used. We also learned that the accesses to the D-LS came from an exceptionally wide variety of locations both within the academic campus and beyond its borders, that the use for each address was varied but generally low, and that, as far as we know, both Macs and IBM's were used to the same degree.

Our findings suggest several conclusions. In general, student volume of use and the repetitions of use over the quarter suggest that the students found the resources useful. In particular, it appears that resources which specifically target student problem areas (e.g. finding group members) are more successful than general purpose functions that might be useful (e.g. the parts catalog). While this distinction is similar to our distinction of process resources and knowledge-base resources, the high use of the Design Reports resource suggests that a true distinction is not as clear as the one we have proposed. Second, having the resources available on the Web does seem to support nearly universal access and promote frequent accesses as of the D-LS. Because log files only identify the address of the user, not the individual user, it is difficult to make strong claims about the impact of the software on the individual students. We cannot relate their use to their performance or learning data from the class or any observational, survey, or interview data collected from individual students.

Based on this analysis, we make the following recommendations.

1. Explore further why the parts catalog did not experience higher use by students. Maybe it was hard to learn to use or hard to understand why they would want to use it. Perhaps a tutorial or better description of it could be designed. Maybe it does not have enough data to make it really useful and populating it with more data would make it appear more useful to the students.

2. Explore the possibility of automating this type of analysis so that it can be performed in real time during the term. With such an automated analysis, one could monitor the level

of use during the term, and in cases where resource use is not conforming to expectations (e.g., the use of the Parts Catalog), one could immediately start to explore why the resource is not being used and possibly intervene.

3. Find out if and how students assign responsibility for learning from what is in the D-LS. Does one person on the team read, monitor, and report back to the team. Alternately, is each individual responsible for learning what information resides within a resource. Without knowledge of the students'/teams' strategies for using the D-LS, the strength of the conclusions that we can draw about the level of use is quite limited.

4. Add a registration facility so that student use (not address use) can be tracked. If we ultimately would like to perform an analysis relating student learning outcomes to the use of the D-LS, we will need to know how each student use the software.

Overall, we have shown that these issues can be generally explored in a very quantitative manner based on data that is widely available - Web log file data. We are encouraged that the resources we are providing are being used and are being used repeatedly by a variety of users. These patterns of use suggest that students are finding them useful and that we are on the right track.

ACKNOWLEDGMENTS

This analysis was supported by the Edutech Institute through a grant from the Woodruff Foundation. The design and development of the Design Learning Simulator and has been supported by a number of research efforts including NSF grant DM1-9420405. As always, we would like to extend our sincere appreciation to the teaching team and all of the students of ME3110, without whom this research would not be possible.

BIBLIOGRAPHY

1. NSPE, *Engineering Education Issues: Report on Surveys of Opinions by Engineering Deans and Employers of Engineering Graduates on the First Professional Degree*. 1992, National Society of Professional Engineers: Washington, D.C.
2. ASEE, *Engineering Education for a Changing World: Project Report*. 1994, ASEE: Washington, D.C.
3. Eder, W.E., Engineering Design Education: Situation Report. *Design Studies*, 1991. **12**(4): p. 261-267.
4. Jones, R.C. and D.R. Reyes-Guerra, Directions for Engineering Education in the United States. *International Journal of Applied Engineering Education*, 1987. **3**(4): p. 349-356.
5. Mistree, F., and Muster, F., Engineering Design as It Moves from an Art toward a Science: Its Impact on the Education Process. *International Journal of Applied Education*, 1988. **5**(2): p. 239-246.
6. Turns, J., Mistree, F., Rosen, D., Allen, J.K., Guzdial, M., and Carlson, D. A Collaborative and Multimedia Design Learning Simulator. in *ED-Media Conference*. 1995. Austria.
7. Newstetter, W. and J. Kolodner. Learning to Change the World: A Case Study of a Mechanical Engineering Design Course. in *Frontiers in Education*. 1995. Atlanta, GA.
8. Turns, J., Newstetter, W., Allen, J.K., and Mistree, F. The Reflective Learner: Supporting the Writing of Learning Essays that Support the Learning of Engineering Design through Experience. in *ASEE*. 1997. Milwalkee, WI:
9. Guzdial, M. Analyzing Student Use of Educational Software with Event Recordings. *submitted to Interactive Learning Environments*, 1997. .
10. Guzdial, M., Santos, P., Badre, A., Hudson, S., and Gray, M., *Analyzing and Visualizing Log Files: A Computational Science of Usability*. 1994, Georgia Institute of Technology; Graphics, Visualization, and Usability Center:

JENNIFER TURNS

Jennifer Turns is a research assistant in the Center for Human-Machine Systems Research in Georgia Tech's Department of Industrial and Systems Engineering. She received her B.S. and M.S. in Systems Engineering from the University of Virginia. Her interests revolve around issues in engineering design education and include teaching engineering design through project--based approaches, studying the cognitive underpinnings of design activity, and developing technology to help teaching, learning, and assessment. She is a member of ASEE and AERA.

FARROKH MISTREE

Professor Farrokh Mistree's design experience spans mechanical, aeronautical, structural, industrial and software engineering. His research focus is on learning what happens when concurrent engineering principles are applied to the design, deployment, operation and support of open engineering systems. He is committed to developing a design pedagogy that is rooted in Decision-Based Design and adaptive action learning. It is in this context that he enjoys experimenting with ways in which design can be learned and taught. Professor Mistree is responsible for two books and over 180 technical publications. Since 1992 he is a Professor at Georgia Tech. He received his Bachelor of Technology with Honours from I.I.T. Kharagpur, India in 1968 and his M.S. and Ph.D. from the University of California, Berkeley in 1970 and 1974, respectively. He is a Fellow of ASME, a Senior Member of AIAA and a member of ASEE. URL http://www.srl.gatech.edu

JANET K. ALLEN

Janet K. Allen is a senior research scientist in the Woodruff School of Mechanical Engineering at the Georgia Institute of Technology. She received her S.B. degree in life sciences from the Massachusetts Institute of Technology and her Ph.D. degree in Biophysics from the University of California at Berkeley. Her research interests include modeling design problems, especially using constructs from living systems theory, and design education.