1-1-2002

Evaluating Java Applets for Teaching on the Internet

Michael R. Healy '04
Claremont Graduate University

Dale E. Berger
Claremont Graduate University

Victoria L. Romero '07
Claremont Graduate University

Christopher L. Aberson '99
Claremont Graduate University

Amanda Saw '11
Claremont Graduate University

Recommended Citation

Evaluating Java Applets for Teaching on the Internet

Healy, Michael R., Berger, Dale E., Romero, Victoria L., Aberson, Christopher L. and Saw, Amanda

Abstract— **The Web Interface for Statistics Education (http://wise.cgu.edu) is a website built around interactive tutorials designed to teach introductory and advanced statistical concepts. The tutorials use Java applets that dynamically illustrate the statistical concepts being taught. By using Java applets, we teach statistics in a manner not possible in a traditional classroom environment. In this paper, we provide examples of the applets, illustrate how students use them, and we report the outcome of a study that examined tutorial effectiveness as a learning tool.**

Index Terms—**Java, Applets, Usability, Education**

I. INTRODUCTION

One of the most attractive features about teaching on the Internet is the interactivity that the medium affords. For the Web Interface for Statistics Education's (WISE; [1]) website, we have designed tutorials on topics in Psychology and Statistics that capitalize on this interactivity by using Java applets. In this paper, we present an assessment of the usability of one of our tutorials and its Java applet. The materials discussed are available to the public and can be accessed at http://wise.cgu.edu.

A. Benefits of Java

Visualization of statistical concepts can be difficult for students. Statistics professors with a chalkboard as their only teaching tool often resort to various forms of hand-waving when trying to illustrate the intricacies of statistical theory. With Java applets, however, we can create simulation programs that graphically illustrate statistical concepts and show how these concepts function together as a system. Java applets are computer applications designed for the Internet. Applets are platform-independent, meaning that they can run on any operating system that has a Java Virtual Machine to translate applet bytecodes into appropriate platform-dependent instructions. Applets are among the most powerful and sophisticated ways of obtaining online interactivity.

Teaching with applets has many cognitive, social, and technical benefits. One of our primary motivations to use applets was to facilitate the visualization of statistical principles, making them easier for students to understand. As such, our applets are designed to make abstract principles concrete. Among the cognitive benefits of this teaching approach is that the graphic and interactive presentation format enhances semantic elaboration, leading to better long-term retention of the material [2]. Teaching in this manner also allows students to take control of their learning process, which is the hallmark of constructivist learning theory [3]. In addition, applets allow students to go beyond simply learning about a theory to actually seeing its applications, which improves both procedural and declarative forms of knowledge [4-6]. From a technical viewpoint, given the wide-scale accessibility to computers with Java-enabled web browsers (but see [7]), applets are an inexpensive solution to obtaining interactivity and a good way of communicating to the large audience of the Internet. Furthermore, instructors can use these applets as a lecture aid, and researcher's can use them to analyze and visualize the results of their studies.

B. The Signal Detection Theory Applet

In this paper, we present an evaluation of an applet and its accompanying tutorial that teach an introduction to Signal Detection Theory (SDT) [8], [9]. SDT is a model of decision making analogous to statistical null hypothesis significance testing [10]. SDT was developed as a means to measure an observer's sensitivity to the difference between two classes of items (Old/New, Diseased/Not Diseased, Present/Not Present, and so forth) while accounting for response biases. SDT can be applied to almost any decision-making task in which there is uncertainty, such as memory [11], radiology [12], and bioinformatics [13]. To illustrate SDT, imagine a situation in which the observer is presented with an ambiguous stimulus item, and is asked to respond 'Yes' if a signal is present and 'No' if a signal is not present. Further suppose that the observer is given a large reward for responding 'Yes' when the signal is present (Hit) and is given a small punishment for responding 'Yes' when the signal is not present (False Alarm). The observer's best strategy is to say 'Yes' when there is even a small chance that the signal is present, and to say 'No' only when the signal is almost certainly absent. If the reward contingencies are altered so that the observer is punished severely for a False Alarm, but rewarded only minimally for a Hit, then the observer's best strategy is to say 'Yes' only when the signal is almost certainly present and to say 'No' on every other trial. The bias of the observer is toward 'Yes' when using the first strategy, and toward 'No' when using the second strategy. The ratio of Hits to False Alarms would be very different under these two strategies, and the proportion of correct responses would be different, even though the observer's sensitivity to the difference between the presence and absence of the signal is the same. SDT provides a way of measuring this sensitivity and measuring the 'bias' in the

response strategy.

The WISE SDT applet (see Figure 1) illustrates the key components of an equal-variance signal detection theory model. It displays the theoretical Gaussian distributions underlying the model and the receiver-operating characteristic (ROC), calculates indices of sensitivity and bias, and dynamically shows the relationships between these three ways of representing the model. The applet is controlled through mouse pointing-and-clicking, and it also accepts keyboard input. The applet appears within a tutorial that leads students through exercises and gives illustrative examples of the concepts being discussed. The tutorial concludes with a follow-up section that can be used as a homework assignment.

C. Usability Testing

When we first examined the effectiveness of our website, we had a simple, focused research question: Are we able to teach statistics on the internet as well as we can in the classroom? To address this question, we conducted two outcome evaluations in which we compared how much students learned from an online tutorial versus a lecture. In the first evaluation [14], no significant differences favoring either approach were found, although both were beneficial, suggesting that the choice of which method to use depends on factors such as time, cost, and technical constraints. However, this evaluation identified a small group of people who did not like the online approach. One possible explanation for their dislike of online education is they had lower technical skills compared to their peers, a possibility suggested by an observed positive relationship between computer experience and ratings of online education [15]. The finding that about 5% to 10% of students do not like online education has been observed by others as well [16]. When viewed in terms of the potential size of the Internet audience, these numbers are substantial and should not be dismissed. In a second study [17], we conducted an outcome evaluation on another applet-based tutorial whose design was based upon input from the previous evaluation. The results of this study indicated that students who used the tutorial learned more than their classroom counterparts. Therefore, we gain increasing support for the utility of applet-based online tutorials, and see evidence that evaluation is important for improving online teaching efforts. For both studies, the applets received high ratings of educational utility and students reported that they enjoyed using them, a finding also reported by others [18].

Many online educational sites provide static text information. By analogy, this is like lecturing to a classroom of students that one cannot see or hear. Without student feedback, the instructor cannot modify the lecture to improve the fit to students' abilities, whether this be changing the pace of the lecture, answering questions, clarifying confusing materials, or addressing student curiosities. By performing usability testing, educators can determine which aspects of their on-line approach are successful and which are problematic, and modify their teaching strategy accordingly.

The research issue addressed in this paper was how students used the SDT Java applets and what, if any, problems they encountered. We did not use a standard experimental design, but rather we gathered information on the individual experiences of students who used the SDT tutorial.

II. METHOD

The students who participated in the evaluation were upper-division and graduate level Psychology students who received extra credit in a course for their cooperation. Twelve students were included in the study. Although, the sample size was small, there is a diminishing return from conducting usability tests with a large number of individuals, as most problems become apparent after observing only a few people [19]. A testing session consisted of an individual student completing the tutorial in a teacher's office under the teacher's supervision. The teacher was a member of the WISE project who did not participate in the development of the SDT tutorial. Students were asked to report verbally what they were thinking and to complete the tutorial as if they were at home doing a homework assignment. Students' questions were answered directly. The evaluation consisted of both observational and survey assessment methods. Detailed observational notes included student questions and comments, teacher observations and comments, and time spent viewing each web page. At the end of a testing session, students completed a survey in which they rated (on a 5-point scale) the tutorial's clarity of exposition, the site's navigational design, and the educational utility of the applet. The survey also included open-ended questions in which students were asked what they liked best and least about the tutorial. Data from user evaluations of this nature can be time consuming to collect and more difficult to analyze than data collected in standard experimental designs. Nevertheless, the information is richer and more informative than survey data; thus we believe that collection of qualitative usability data is an essential part of any website evaluation.

III. RESULTS AND DISCUSSION

Our evaluation focused not only on the applet, but also on its accompanying tutorial, as the two are inextricable in terms of our teaching method. Overall, the tutorial and applet were rated highly and comments about the applet included that it was both 'helpful' and 'fairly easy to use.' Students were comfortable manipulating the applet as instructed and thought that it added to their educational experience. The computing skill level of students who completed the tutorial ranged from beginner to advanced. All of the students had the requisite skills to use the applet, and there were no observed difficulties in using the mouse, keyboard, or web browser. Ratings of how easy the site was to navigate ranged from 'adequate' to 'too slow,' with more skillful users giving the latter rating. These students wanted to navigate the site quickly and have easier access to previous material, while the novice users were content with the tutorial's linear progression. Differential computing skills resulted in variable success using the applet. The advanced students remarked that they had used applets before and were comfortable using this one. Some novices,

however, did have problems at first as they did not understand what an applet was. One student, when asked to use the applet, merely stared at it, and when asked what she was doing, stated that she thought the applet was a picture, and she was confused as to how she was supposed to 'use' it. Variable skill levels are potentially an important consideration in the development of online tutorials, because in addition to designing a good GUI, one must provide adequate instruction and support so that students know how to use it in ways beneficial to their learning.

As is often discovered in usability tests [e.g., 20], we found that students did not use our materials as we anticipated. We found that some students did not use the applet unless explicitly instructed to do so. Throughout the tutorial, students were asked to perform calculations that either require statistical tables and hand computations or a simple application of the applet. Even though students had used the applet for these calculations previously, a few still opted to do the calculations by hand. In addition, it was observed that students did not always use the applet to verify the accuracy of their answers when asked to do so. One student indicated that she avoided the applet because she didn't want to spend time learning to use something she would never use again. This observation is in accord with other evaluations which have found that less technically skilled people tend to be avoidant of technology [15] and less willing to spend time learning how to use it [21].

Based upon these observations, the applet and its accompanying tutorial were modified in the following ways. The first issue addressed was a tendency for students not to use the information available to them [cf. 22]. We concluded that a likely reason for this was that the original applet contained too much information, making it difficult for students to isolate features that they needed. To alleviate this problem, a new GUI was developed that restricted the information being displayed to match the needs of the current tutorial topic. This was accomplished by removing advanced features and modifying the applet so that information not pertinent to the current discussion is deleted or appears visually dimmed. A critical goal in the design of the applet is to explicitly show the relationship between the theoretical distributions, receiver-operating characteristic, and detection measures. The GUI was adapted, both literally and figuratively, to do this by adding lines which show these connections. In addition, an online-help system was developed that is sensitive to mouse movement. When the mouse enters a defined target area, definitional text appears immediately in the Help window. The help system is designed to always be visible, in response to the observation that students often were not aware of resources available to them. We also modified various elements of the GUI, such as font sizes, for text that was difficult for students to read.

Java offers platform-independence, but many details of how GUIs are implemented are left to the virtual machine, meaning that although an applet will operate correctly, its appearance can vary depending upon what type of computer it is running on. Developers of online materials are advised to view their applets on as many systems as possible to ensure that the GUI is consistent on all platforms. Finally, in the original tutorial, students were given an option to read instructions on how-to use the applet before beginning, but they invariably chose not to use this resource. We thus modified the tutorial so that students are required to read critical instructions before starting.

IV. CONCLUSIONS

Without evaluation, there can be no assurance that a new approach to teaching is effective. From our experience, observation of student responses provides important information to guide modifications of online materials. The purpose of this usability assessment was to improve our application of Java technology in the design of an educational tool. We found that the signal detection tutorial and applet were well liked by students. The applet's interface was important, but the primary factors in usability were its context, supporting materials, and the technical skills of the students. A primary finding was that the usability studies provided surprising, but critically useful, information into how students were actually using the materials. Heterogeneity among students was apparent in their computing skills, learning styles, and relevant knowledge; this variability is important to consider in the design of online educational material.

Throughout the assessment, we found that students' self-reports often differed from our observations. Some students who reported that the tutorial was 'easy' and 'helpful' in fact did not answer any of the follow-up questions correctly. People quite often overestimate how well they have learned something. This overestimation of how much one has learned is referred to as an 'illusion of knowing' [23-25]. Providing performance feedback, whether it be through multiple-choice questions with answers in HTML or via an applet, is an important way of helping students to objectively assess their performance.

As part of our testing procedure, we answered student questions, which increased the tutorial's usability, and potentially inflated ratings of the applet. With help from the instructor, students were able to use the materials quite well. Conceivably, without instructor support, students might have had difficulty using the tutorial and applet, and would not have completed it on their own. In fact, online education students tend to rely upon the support services of their instructor or a teaching assistant [22]. Nevertheless, our research has shown that these tutorials with interactive Jave applets are an effective means of teaching. Perhaps the most effective way to use these materials, however, is to use the online materials in conjunction with more traditional lecture approaches. Java applets allow a teacher to supplement lecture with dynamic illustrations of relationships among concepts.

REFERENCES

[1] Aberson, C. L., Berger, D. E., Healy, M. R., and Romero, V. L. (2001, February). Teaching Statistics with Web Technology: The WISE Project. *Syllabus*, pp. 43-45.

[2]   Paivio, A. and Csapo, K. (1969). Concrete Image and Verbal Memory Codes. *Journal of Experimental Psychology* (2, Pt. 1), pp. 279-285.

[3]   Piaget J. (1960). The Child's Conception of the World. Rowman and Allenheld, New York.

[4]   Adams, G. (1992). Why Interactive? *Multimedia and Videodisc Monitor, 10 (3)*, pp. 20-24.

[5]   Geisman, J. (1988). Beyond CBT Interactive video. *Computers and Personnel*, pp. 35-38.

[6]   Malloy, T. E. (2001). Difference to Inference: Teaching logical and statistical reasoning through on-line interactivity. *Behavior Research Methods, Instruments, & Computers*, 33 (2), pp. 270-273.

[7]   Microsoft Corporation (2001, August). *An Open Letter Regarding Windows XP and Java Support*. [Online] Available: http://www.microsoft.com/java/issues/openletter.htm.

[8]   Wickens, T. D. (2001). *Elementary Signal Detection Theory*. Oxford University Press, Inc.: New York.

[9]   Tanner, W. P., Jr., and Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review*, *61*, pp. 401-409.

[10] Nickerson, R. S. (2000). Null hypothesis significance testing: A review of an old and continuing controversy. Psychological Methods, 5, 241-301.

[11] Banks, W. P. (1970). Signal detection theory of human memory. *Psychological Bulletin*, *74*, pp. 81-99.

[12] Metz C.E., Goodenough D. J., and Rossmann K. (1973). Evaluation of receiver operating characteristic curve data in terms of information theory, with applications in radiography. *Radiology*, *109*, pp. 297-303.

[13] Baldi, P., Brunak, S., Chauvin, Y., Andersen, C. A. F., and Nielsen. (2000, May). Assessing the accuracy of predication algorithms for classification: an overview. Bioinformatics, 16, 5, pp. 412-424.

[14] Aberson, C. L., Berger, D. E., Healy, M. R., Kyle, D., & Romero, V. L. (2000). Evaluation of an interactive tutorial for teaching the Central Limit Theorem. *Teaching of Psychology, 27,* 289-291.

[15] Pear, J. J. and Novak, M. (1996). Computer-aided personalized system of instruction: A program evaluation. *Teaching of Psychology*, *23*, pp. 119-123.

[16] Anido-Rif, L., Llamas-Nistal, M., and Fernandez-Iglesias, M. J. (2001) *A Component Model for Stardardized Web-based Education. Procedings of the ACM.*

[17] Aberson, C. L., Berger, D. E., Healy, and Romero, V. L. (in press). An Interactive Tutorial for Teaching Statistical Power. *Journal of Statistics Education.*

[18] Street, S. and Goodman, A. (1998). *Some experimental evidence on the educational value of interactive Java applets in Web-based tutorials. Proceedings of the ACM.*

[19] Nielsen, J., and Landauer, T. K. (1993). A mathematical model of the finding of usability problems. *Proceedings ACM/IFIP INTERCHI'93 Conference* (Amsterdam, The Netherlands, April 24-29), 206-213.

[20] Jolicoeur, K. & Berger, D.E. (1986). Do we really know what makes educational software effective? A call for empirical research on effectiveness. *Educational Technology, 26,* 7-11.

[21] Johnson, J. (2000). GUI Bloopers: Don'ts and Do's for Software Developers and Web Designers. Academic Press: San Diego, CA.

[22] Varnhagen, C. K. Drake, S. M., and Finley, G. (1997). Teaching Statistics With the Internet. *Teaching of Psychology*, pp. 275-278.

[23] Bjork, R. A. (1996). Memory and metamemory considerations in the training of human beings. In J. Metcalf & A. P. Shimamura (Eds.), Metacognition: Knowing about knowing (pp. 185 - 205). Cambridge, Massachusetts: MIT Press.

[24] Garhart, C., and Hannafin, M. (1986). The accuracy of cognitive monitoring during computer-based instruction. *Journal of Computer-Based Instruction.* 13, pp. 88-93.

[25] Hofer, B. K., Yu, S. L., & Pintrich, P. R. (1998). Teaching college students to be self-regulated learners. In D. H. Schunk, & B. J. Zimmerman (Eds.), *Self-regulated learning: From teaching to self-reflective practice* (pp. 57 - 85). New York: Guilford Press.

Figure 1

Normal Distributions

Criterion = 0.63
<<drag>>

d' = 2.97
<<drag>>

Signal Absent
Signal Present

ROC

Hit Rate

1.00
.75
.50
.25
.00

.00   .25   .50   .75   1.00
False Alarm Rate

d' is a measure of how sensitive an observer is to the difference between stimuli coming from the Signal Present or the Signal Absent distributions.

False Alarms 0.265   Hits 0.990   Set Hits + FAs   d' 2.968   c -0.856