Using On-line Interactive Statistics

for Evaluating Web-based Instruction

Submitted to Journal of Educational Media International

Chong-ho Yu, Ph.D., Angel Jannasch-Pennell, Ph.D.,

Samuel A. DiGangi, Ph. D., Barnaby Wasson

Arizona State University

Jan 17, 1998

(Draft. Please do not quote)

Correspondence:
Chong-ho Yu, Ph.D.
Instruction and Research Support
Information Technology 3S70
Arizona State University
Tempe AZ 85287-0101
Phone: (602)965-7402
Fax: (602)965-6317
Email: alex.yu@asu.edu
Http://seamonkey.ed.asu.edu/~alex

Abstarct

This article describes the use of online interactive statistical analysis to evaluate and enhance university level web-based courses. In this study, SAS/IntrNet was employed within a web-based instructional environment to enable instructors to continuously evaluate and dynamically refine courses through analysis of the test scores and the user logs with regression model, and calculation of test item reliability. This approach allowed for real-time formative evaluation and assessment of treatment effectiveness and provided for on-going refinement of instruction.

Using On-line Interactive Statistics for Evaluating Web-based Instruction

Chong-ho Yu, Ph.D., Angel Jannasch-Pennell, Ph.D.,

Samuel A. DiGangi, Ph. D., Barnaby Wasson

Arizona State University

This article describes the use of SAS/IntrNet (SAS Instittute, 1997) to evaluate and enhance university level web-based courses. Learning environments, which incorporate continuous monitoring of learner performance and allow for data-based instructional decisions, provide the educator and the student with a means for ensuring mastery of content. Computer-based instruction provides a mechanism for immediate feedback, allowing students to monitor their own learning progress. In the present study, SAS/IntrNet was employed within a web-based instructional environment to enable instructors to continuously evaluate and dynamically refine courses through analysis of the user logs, and calculation of test item reliability. This approach allowed for real-time evaluation and assessment of treatment effectiveness and provided for on-going refinement of instruction.

SAS/IntrNet

For several decades SAS has dominated the fields of statistical analysis and to a lesser degree, data warehousing. Since the release of SAS/IntrNet in October, 1997, a number of major corporations and government agencies such as Ford Credit, National Semiconductor, and Danish Institute of Agricultural Sciences have implemented SAS/IntrNet for their intranet and internet servers (SAS Institute, 1997). However, at the time of this writing, no web-based courses directly utilizing the analysis capabilities of SAS/IntrNet were found.

SAS/IntrNet provides the ability to create both static and dynamic webpages. The product uses Web Publishing Tools (SAS Institute, 1998) to convert SAS output to hypertext markup language (html) for the creation of static webpages. For dynamic pages SAS/IntrNet provides internet-database tools for users to retrieve real time data. A feature entitled 'Application Broker' provides the ability to connect and utilize data sets spanning across several servers and platforms (SAS Institute, 1998). Further, SAS/IntrNet functions as a web *service* rather than a web *server*. Some web server implementations prohibit multiple web server packages from coexisting in the same physical machine. In contrast, web services are not mutually exclusive. A web service runs on top of an existing web server and thus it does not conflict with other services provided all services are properly configured. For example, Cold Fusion (Allaire, 1998) and Active Server Page (Homer, 1997), which are web-database services, can be used in the same server concurrently.

Background

An on-line, web-based course was developed for use by graduate-level, k-12 in-service teachers. The course objectives move participants toward mastery of basic networking and computing applications, including the use of email, telnet and ftp services, a web browser, and synchronous and asynchronous discussion environments. A goal of the course is to efficiently and effectively move the students from beginner and novice levels of competence and confidence with computer networking and utilities to mastery of these basic tool skills.

The course then focuses on the incorporation of the use of these tools and resources within the content/curriculum areas of each educators' expertise.

The participants are 180 teachers who have little or no experience with networked computing. Following the instructional process, participants use Internet resources to build, contribute to, and maintain a close, effective electronic community of educators. The project is structured to provide ongoing support tailored to the educator's specific needs and interests. Educators receive the necessary networking tools (computer, modem, access to the Internet), on-line support throughout the project, and in-person and on-line training in the integration of technology and telecommunications into an educational curriculum (http://eruditio.asu.edu).

Use of SAS/IntrNet

The content of the course are stored in Lotus Domino/Learning Space and Netscape Web Server with SAS/IntrNet as a supporting tool. This approach enhances the web-based course by:

1. Providing descriptive statistics such as test score distribution and page usage distribution.
2. Calculating Cronbach Coefficient Alpha to estimate the test item reliability and evaluate construct validity (Cronbach & Meehl, 1955).
3. Performing dependent t-tests of pretest and posttest scores to evaluate the effectiveness of the web-centric instruction.
4. Implementing regression analysis with the user activity log data and the test scores to determine the best strategies of web-based learning.

The following are examples of the webpage input (see figure 1) and output (see figure 2) for instructor analysis:

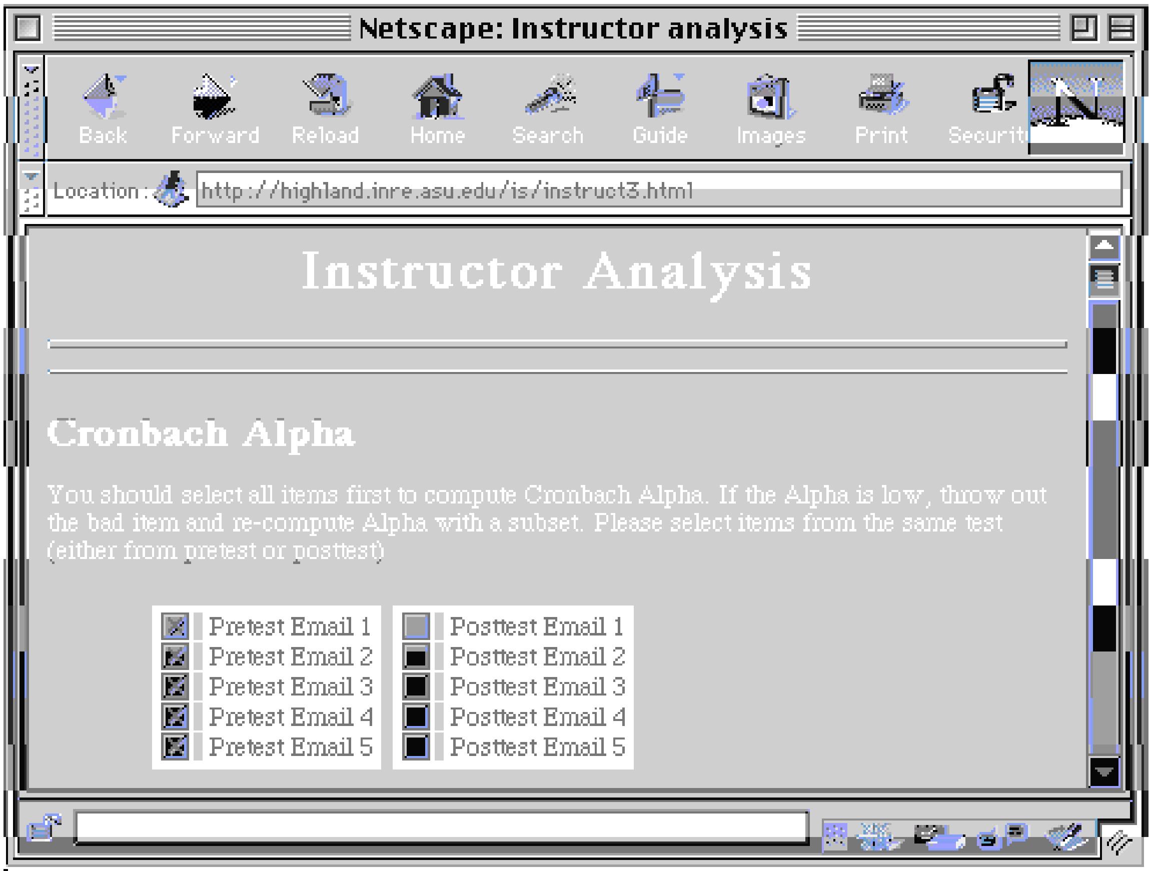

Figure 1. Input page of Instructor Analysis

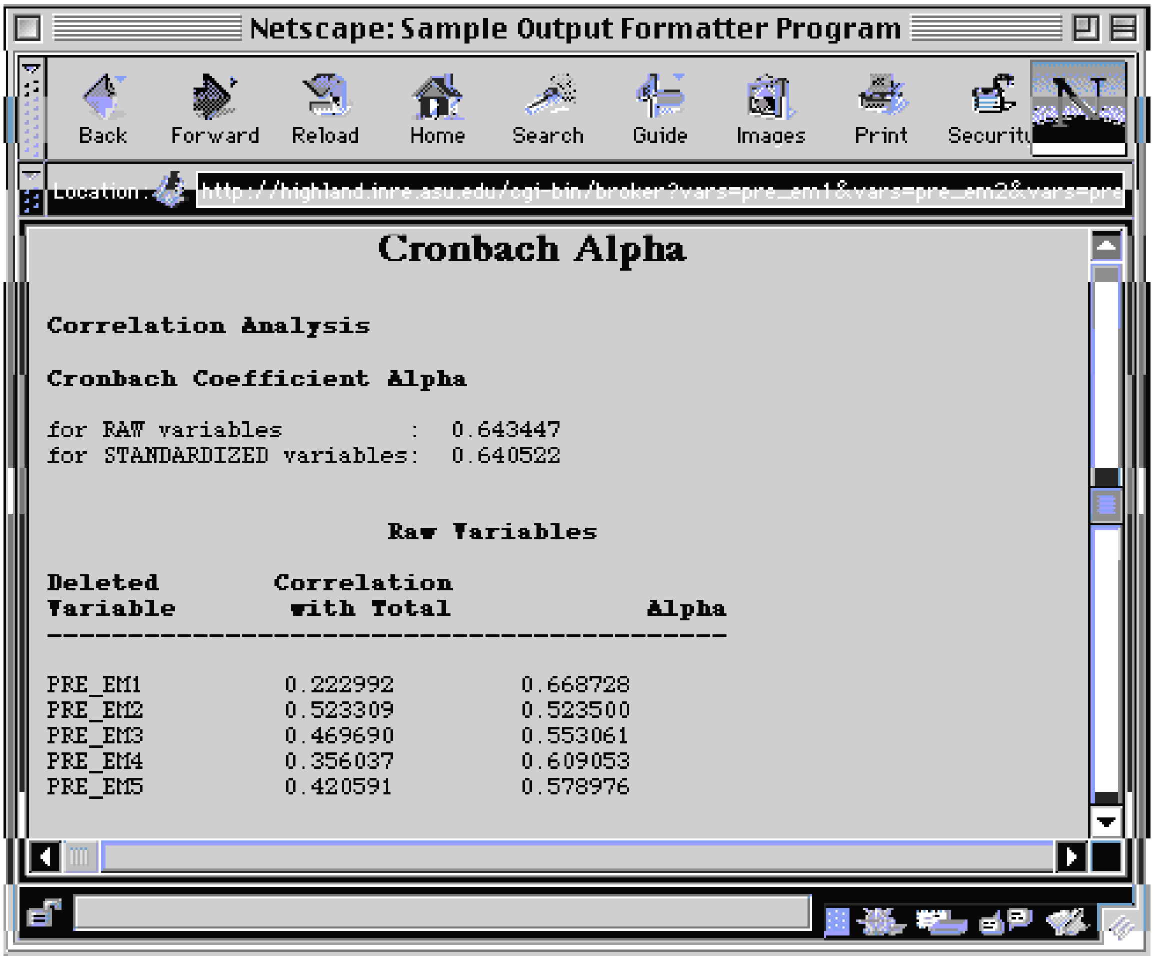

Figure 2. Output page of Instructor Analysis

User log Analysis

The Lotus Domino Server and Netscape Web Server were configured to record the user activity log such as the number of page access, the number of hits, and the time spent on each page. Although both Lotus and Netscape administrative tools can provide an overview of the user activities, they do not have a built-in statistical tool to analyze these data. Thus, we use SAS/IntrNet to access the Lotus user log for further analysis.

A typical user activity log gives the number of page accesses, hits, and unique hosts as shown in Figure 3.

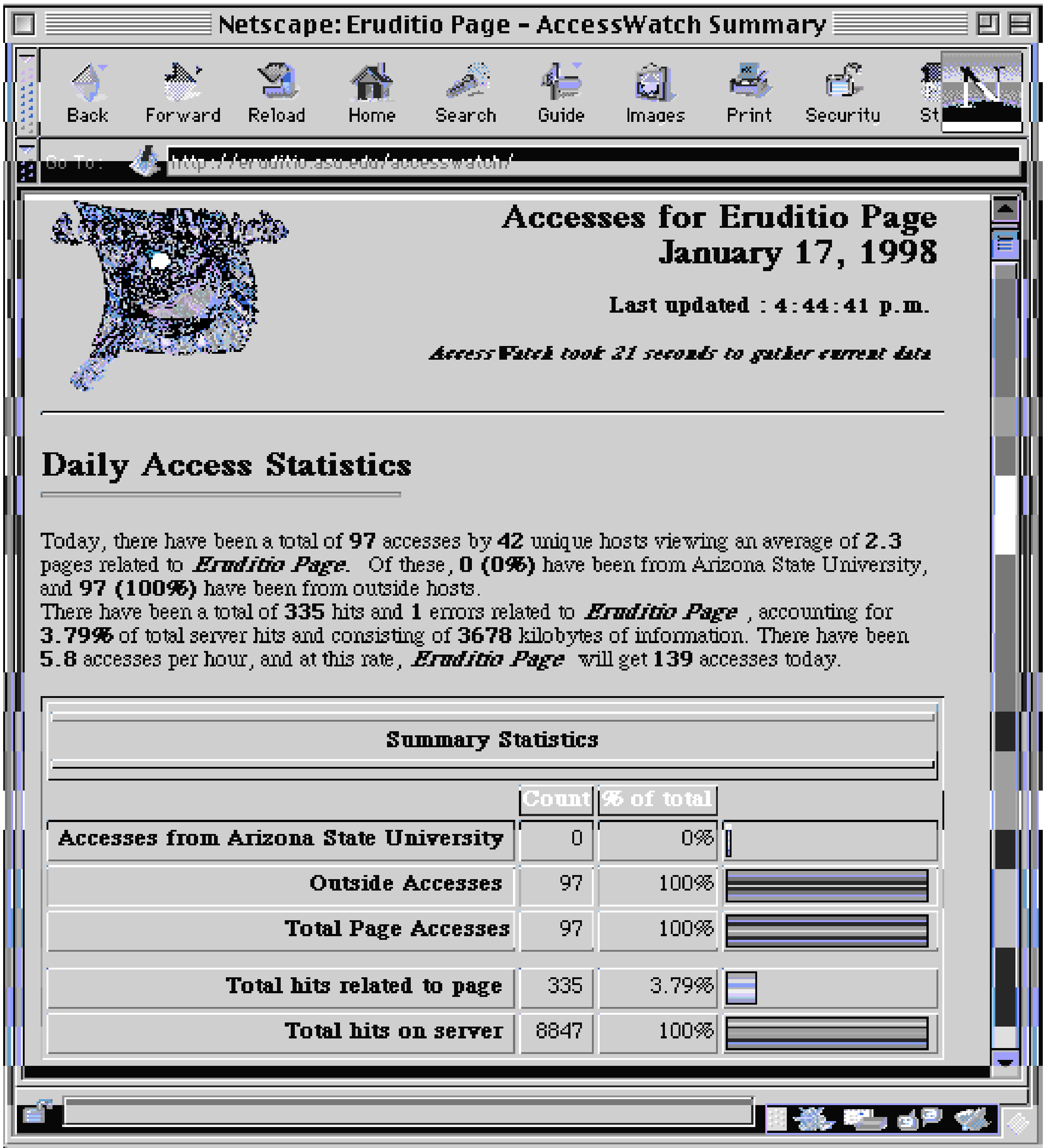

Figure 3. User activity log

Hit is defined as a downloaded object. For example, if a page carries multiple graphics or images; four GIF images, one JPEG image, and one Java Applet, a hit count would reveal seven ‘hits’--although only one physical page has been accessed. To the contrary, *page access* records access to only the html file itself. When a webpage with several graphics and applets is viewed, it is recorded as one *page access*. Hit counts may best convey the system resource load on a particular server, but potentially artificially inflate the individual user activities. In the present study page access served as the focus of analysis. Figure four details the configuration of the user log program to filter hits. In the variable *$excludeURL*, graphics such as GIF and JPEG images, Java-related files such as files with class and zip extensions, and sound clips such as WAV and AU files are excluded and thus are not counted as page accesses.

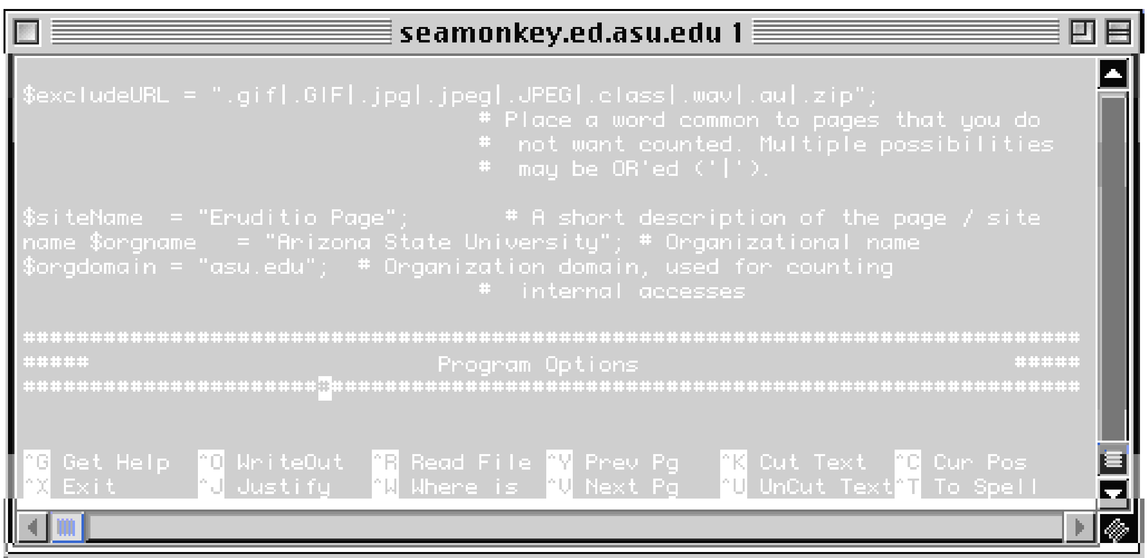

Figure 4. User log program configuration

A unique host is the hostname or IP address connected to the web server, which can be interpreted to identify individual users. However, in the present course students use a point-to-point protocol (PPP) server to obtain internet access via telephone dial-up. Hostnames and IP numbers change each time the user obtains access to the internet and thereby providing an inaccurate means of tracking user activity within the website. In order to identify each student's activity the web environment was configured to require logon with a unique ID and password. The amount of time that the user's browser displays each specific webpage is recorded, as well as the total session time. The following (Figure 5) is an example of a user activity log.

```
129.219.254.22
       00:23:27 /~eruditio/alexyu.html
       00:23:27 /~eruditio/computer/windows/email.html
       00:23:58 /~eruditio/navigation.html
       00:31:25 /~eruditio/computer/windows/email.html
       00:31:34 /~eruditio/computer/windows/listserv.html
       00:31:55 /~eruditio/computer/windows/websearch.html
       00:32:25 /~eruditio/computer/windows/quiz.html
```

Figure 5. User activity log

Performance Evaluation

Each learning module contains a pretest and a posttest, which users are required to complete immediately prior to and immediately following viewing the instructional module. Students who score perfect on the pretest are asked not to take the posttest. The rationale is as the following: As it is likely that students who score high or perfect on the pretest will also achieve the same result on the posttest, this "ceiling effect" will cause the t-test to show virtually no improvement and also bias our interpretation of the regression analysis. To allow for a more accurate analysis of the instructional impact of the learning modules, these scores are coded as “outliers” and are excluded from computation of the t-test and regression.

Regression model

After both the user log and the test scores have been obtained, SAS/IntrNet is employed to merge the two files by user ID. User access data can then be used to predict performance in terms of test scores. The regression analysis is based upon the following hypotheses:

Performance as measured by change in scores between the pretest and posttest is a function of how much time the user spent accessing the web, how many pages the user read, and the interaction between the two. Time engaged or page access alone is not a good predictor to performance. A student may open many pages but stay only a few seconds on each page. On the other hand, it may appear that a student spends a significant amount of time accessing a few pages, however s/he may have logon, go away, and then come back an hour later to go to the next page. Therefore, we included the product of time engaged and page access in the regression model as shown below:

Variables:

Y = Change in scores

X_1 = Time engaged

X_2 = Page access

X_3 = Time engaged X Page access

Model:

$$Y = A + B_1X_1 + B_2X_2 + B_3X_3 + E$$

The interaction of time engaged and page access is strongly correlated with either time engaged or page access. This multicollinearity could lead to variation inflation and threatens the validity of the regression model. In order to avoid multicollinearity, centered scores and residuals of interaction are used because their vectors are orthogonal (Aiken & West, 1991; Burrill, 1998; Yu, 1998). The original variables are transformed and the model is reconstructed as the following:

Variables:

Y = Change scores

C_X1 = Time engaged - Mean of time engaged

C_X2 = Page access - Mean of page access

C_X3 = C_X1 * C_X2

R_X3 = Residuals of regressing C_X3 with C_X1 and C_X2

Model:

Y = A + C_X1 + C_X2 + R_X3

SAS source code for computing the above model is hidden from course instructors. Instructors use the web interface rather than SAS interface to select variables and the range of data.

Conclusion

Historically, analysis of this depth has been employed primarily by well-trained statisticians rather than by the content-area expert or the instructional designer. Although item and performance data is used in commercial test development, the ability of classroom instructional designers and instructors to effectively incorporate this type of analysis into their daily routine has been limited. With the advancement of web-based technologies, general instructors can now analyze and interpret their student user data and employ the results dynamically within the instructional process. This approach effectively extends the utility of formative evaluation in the instructional design and delivery process. In contrast to summative evaluation (Herman, Carol, & Fitz-Gibbon, 1987), which is conducted for the purpose of obtain information for summary statements and judgments about a program and its values, formative evaluation (Hopkins, Stanley, & Hopkins, 1990) is conducted for the purpose of bringing about improvement in practice. However, in conventional courses it is difficult to put formative evaluation into practice due to the difficulty of collecting and interpreting data. In addition, data such as test scores collected from conventional classes are not sufficient for formative evaluation. Formative evaluation is not merely reporting pretest-posttest difference. The learning process--how students use the instructional materials--is also important (Cognition and Technology Group at Vanderbilt, 1993; National Council of Teachers of Mathematics, 1991).

In the design of our website we provided space that can be only accessed by the instructors. On these pages, instructors have the ability to assemble any variables to construct their own analysis of student data. Besides quantitative analysis, user activity data can also be analyzed in a qualitative manner such as looking at the pattern of page access. Through this multi-method approach toward data collection, evaluation, and dynamic instructional revision, we move toward the improvement of instructional courseware. Design decisions can be grounded in instructional theory, while driven by continuous assessment of user performance. As researchers we look forward to the extension, customization and use of instructional learning environments which effectively enable instructors to evaluate and enhance their online materials.

REFERENCES

Allaire Corp. (1998). Cold Fusion 3.1. [On-line]. Available: http://www/allaire.com

Aiken, S. & West, G. (1991). Multiple regression: Testing and interpreting interactions. Newbury Park, CA: Sage.

Burrill, D. (1998). Modelling and interpreting interactions in multiple regression. [On-line]. Available: http://www.minitab.com/burril2.htm.

Cognition and Technology Group at Vanderbilt. (1993). The Jasper experiment: Using video to furnish real-world problem solving contexts. Arithmetic Teacher, 40(8), 474-478.

Cronbach, L. J. & Meehl, P. E. (1955). Construct validity in psychological tests. Psychological Bulletin, 52, 281-302.

Herman, J. , Carol, L. & Fitz-Gibbon, T. (1987). Evaluator's handbook. Newbury Park, CA : Sage Publications.

Homer, A. (1997). Professional Active Server Pages. Birmingham, UK: Wrox Press.

Hopkins, D., Stanley, C. & Hopkins, R. (1990). Educational and psychological measurement and evaluation.. Englewood Cliffs, N.J.: Prentice Hall

National Council of Teachers of Mathematics. (1991). Professional standards for teaching mathematics. Reston, VA: Author.

SAS Institute. (1998). SAS Institute Web Tools. [On-line] Available: http://www/sas.com/rnd/

SAS Institute. (1997) Serving up SAS applications on the World Wide Web. SAS Communications, 3Q, 4-7.

Yu, C. H. (1998). Multicollinearity, variation inflation, and orthogonalization. [On-line] Available: http://seamonkey.ed.asu.edu/~alex/computer/sas/collinear.html