Developing Software for Teaching: An Exploratory Study

Juan Manuel Gómez Reynoso
Universidad Autónoma de Aguascalientes, Mexico
jmgr@correo.uaa.mx

Mónica R. Brizuela Sandoval
Universidad Autónoma de Aguascalientes, Mexico
mbrizuel@correo.uaa.mx

ABSTRACT

Global competition creates a need for organizations and individuals to implement an every-day learning activity. People learn using several approaches such as having experiences, studying, and practicing. Technology is a tool that can help people learn. Commercial software has been developed to help people learn, for example, how to use some productivity tools such as a spreadsheet, a word processor or email. However, people need to perform tasks that are complex and, sometimes, require previous knowledge. We believe that software that is intended to teach a special knowledge that requires previous understanding of the subject is a good option to help people learn. In order to study this issue, special software was developed and tested for usability. In addition, a pilot test was conducted so that the measuring instrument was tested. In order to probe that special software for teaching allows better performance compared to teaching with traditional means two studies were conducted: an experiment conducted with two groups and a quasi-experiment using three groups. Results obtained in our studies confirm this claim.

INTRODUCTION

Individuals and organizations often face the challenge of achieving better training/learning. It is important for industry and researchers to create new ways to train people (Desai, Richards, & Eddy, 1999). In addition, it has been found that people learn better when they are enjoying what they are doing (Csikszentmihalyi, 2000; Webster & Martocchio, 1995).

Past research in training/teaching through software made use of commercial software (Bowman, 2002; Bowman, Grupe, & Simkin, 1995; Desai et al., 1999; Desai, Richards, & Eddy, 2000). Particularly software that is intended to teach people how to use email, spreadsheets, and word processors mainly. However, there are tasks that are more complex to learn because they require previous knowledge. For example, learning statistics requires some previous knowledge in mathematics. Software that helps people to acquire complex knowledge might provide an aide for trainers. We believe that software that is developed to teach a specific complex subject (SDTCS) is a better option to train people. A study that develops its own software and it is intended to teach a complex knowledge was not conducted before, which is the purpose of the present study.

In order to study whether a SDTCS has better outcomes we developed special software through an incremental prototyping approach. We tested the final version for usability so that it complies with user requirements. In order to measure the effectiveness of teaching using a SDTCS as an approach, an exam was created and tested in a pilot study.

Two studies to measure the effectiveness of a SDTCS were conducted. The first study was an experiment with two groups; one group was taught using a SDTCS, and the other group using traditional means for training. The second study was a quasi-experiment using three groups, one group was taught using traditional means, a second group was taught using a reduced version of our developed tool, and a third group was taught using our SDTCS. Results of both studies confirm that SDTCSs has better outcomes in learning performance.

TRAINING/TEACHING

Individuals and organizations often face the challenge of delivering/achieving better training/learning. Industry and researchers are challenged to create new ways to train people (Desai et al., 1999). Learning through training can be achieved in a classroom but also by performing the tasks using stimulating materials (Oldfield, 2004). In addition, new developments in information technology (IT) allows organizations to deliver training free from time and/or place constraints (Hornik, Johnson, & Wu, 2007).

There are numerous studies in the training/teaching field using IT as means. For example, research has been conducted on learning performance tools (Bowman et al., 1995; Desai et al., 1999), low complexity tasks (Bowman et al., 1995), perspectives on online learning (Kazmer & Haythornthwaite, 2005), virtual learning initiatives (Hornik et al., 2007), and learning on the Web (Spaniol, Klamma, Springer, & Jarke, 2006). However, at the moment of conducting our study we did not identify research that developed its own software tool for delivering a special complex knowledge.

Past research has identified that cost of any training program is an important issue (Bowman et al., 1995). Huge amounts of resources have been invested in CBT and education (Palvia & Palvia, 2007). Organizations that intend to cut training cost through the use of commercial packages result in negative user attitudes and a low-term equilibrium (Lassila & Brancheau, 1999). Montesino (Montesino, 2002) states that training presents a low return on investment and training investments have experienced a dramatic growth over time. Olfman (Olfman & Pitsatorn, 2000) explains that money spent in training represents a significant amount of organizational budget. It has been identified that computer technology helps to reduce the cost of training (Brown, 2001; Buch & Bartley, 2002). Thus, by using the right training approach organizations may reduce costs (Bowman et al., 1995). Therefore, we believe that it is very important to increase teaching/training efficiency and to reduce its associated cost.

It is well known that using traditional means for training delivers good results. In addition, this model cannot be eliminated, especially when there is a need for high level of interaction or use expensive resources (Goode, Willis, Wolf, & Harris, 2007). However, we believe that improved approaches/tools have the ability to enhance outcomes. Teaching not always takes place as formal, pre-usage instruction (Boudreau & Seligman, 2005). Some users may benefit from training during ongoing use (Spitler, 2005). Further, learning performance influences attitudes (Bostrom, Olfman, & Sein, 1990).

Previous research identifies several advantages in using electronic means for learning. For example, Piskurich (Piskurich, 2006) states that one of the most important advantages is the savings in both travel time and costs, which is very important in any training/learning program. We believe that an available 24x7 all-year round training tool helps to lower the costs of training and helps to increase outcomes.

Training/Learning Using Technology

Technology has been incorporated into a variety of human activities. It has been reported that technology has added new options to deliver training (Buch & Bartley, 2002; Hornik et al., 2007) as well as a way to improve people's knowledge (Chin-Chih & Chien-Chung, 2008). In addition, IT helps to build new tools and methods that could help to share knowledge and learning easily (Kekäle, Pirolt, & Falter, 2002). Training that is characterized as play results in higher performance (Webster & Martocchio, 1995) and it has been reported that it is important that students engage in problem solving (Bates & Watson, 2008). The vehicle for delivering training is very important because negative reactions may be created when learning to use computer systems (Bozionelos, 1996). Past research questions the effectiveness of learning through computer based education (Palvia & Palvia, 2007). Thus, it is very important to provide effective training/learning tools to help people learn.

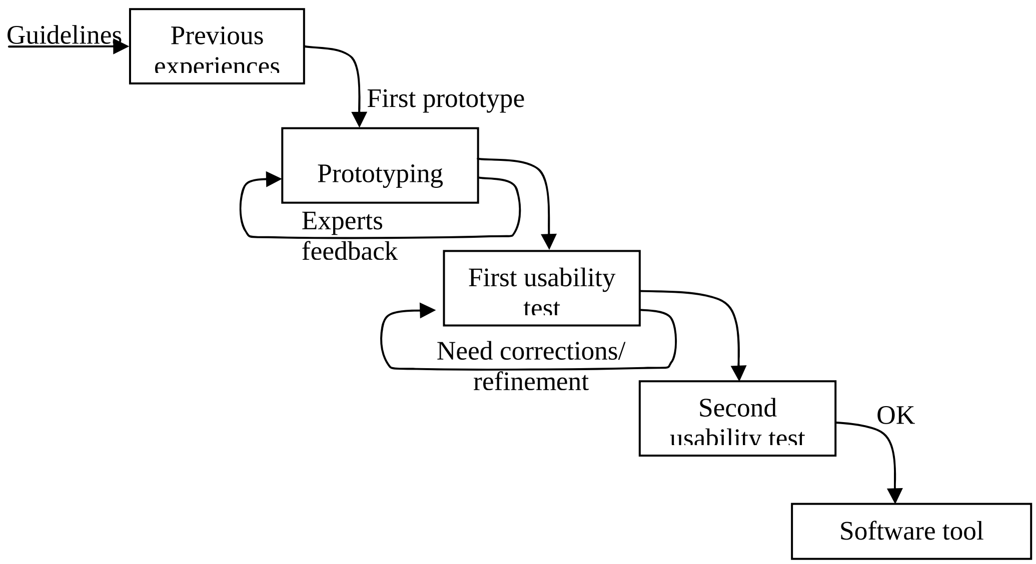
Developing Special Software for Training

Educators seek new ways for improving educational quality (Goode et al., 2007). Training/teaching effectiveness is measured based on performance outcomes. Yi and Davis (Yi & Davis, 2001) classifies performance outcomes into two categories: cognitive learning and skill based learning. The former tries to provide knowledge in a particular domain, while the latter refers to hands-on proficiency in using a product or service. Previous research on Computer Based Training (CBT) made use of commercial software (Bohlen, 1997; Bowman et al., 1995; Desai et al., 2000; Grainger, 2001) as a means to deliver training. For example, research has been done with web-based 3-D training systems, combining the use of an expert system to provide dynamic coaching advice and feedback to teach maintenance technicians how to perform shaft alignment tasks, based on the user's activities in a 3-D practice environment (Johns, 2000). Such study intended to develop mechanic abilities. Our study intends to develop mental abilities through special software.

A SDTCS has the opportunity of addressing all the requirements for a special area of knowledge because it contains the knowledge needed to be transferred and includes all requirements to maximize knowledge transfer. For the present study, we developed a SDTCS that performs pattern analysis using an Expert System approach in a visual way through a GUI to teach central tendency and dispersion concepts by using an object-oriented incremental prototyping approach. A well-designed interface can give the trainee a much clearer overall picture of the system's view (Hayes, 1999; Vokurka, Flores, & Pearce, 1996). A GUI can be tailored so that it matches users' abilities. In addition, GUIs can minimize the dialogue between the system and users, and can reduce the need to pre-train the users in the use of the system before conducting the study. It is not only important to developing such software, but also testing it so that potential users receive benefits from it. The first version of the software tool was developed using guidelines based on researchers' past experiences as well as recommendations from software engineering (Pressman, 2005; Sommerville, 2005). Four lecturers with experience in teaching statistics evaluated each prototype so that no important knowledge was omitted. In addition, we tested the final developed prototype through usability approach. In the case of the present study, only two usability tests were necessary.

People perform better when they enjoy what they are doing (Csikszentmihalyi, 2000; Webster & Martocchio, 1995) and learn at a different pace. Therefore, it is very important that learners have the freedom to explore knowledge in a play-like system at their own pace. It is important to design and deliver innovative, exciting and relevant learning experiences if we want to deliver good learning experiences (Goode et al., 2007). Rather than performing transactional activities, a vast majority of the work now involves identifying and solving problems using knowledge (Romanik, 2000). Thus, we believe that a training tool that helps trainees to learn more effectively would help to create knowledge at a more rapid pace than traditional training. Figure 1 shows the development approach process for our SDTCS.

Figure 1: Purposely-Development software tool development process.



USABILITY TEST

It is extremely important for the study that our SDTCS should be easy-to-use. In order to do that, we conducted two usability tests. Subjects in the usability test used the tool for three hours and completed a post-test questionnaire. The questionnaire was created using some items from an usability test previously conducted (U.of W., 2004) as well as suggestions from literature (Dumas & Redish, 1999). All questions were presented using a 5-point Likert scale (the lower the number, the higher the score is). In addition, the questionnaire included a section for comments and a set of three open-ended questions to collect evaluators' recommendations of new features and improvements.

We invited potential evaluators based on two criteria: first, they are working in software development-related jobs; and second, they have knowledge of the statistics included in the software tool. Twenty-seven students in the ninth semester of the Computer Science bachelor program that comply with the criteria were invited and nineteen of them agreed to participate. Each usability test was conducted at the same time in the same computer laboratory. In order to eliminate any problems related with hardware specifications all computers had exactly the same technological characteristics. Evaluators were free to drop out of the usability test. However, all of them completed the evaluation.

The mean and standard deviation was calculated for close-ended questions. Open-ended questions' answers were grouped together based on content. After finishing the first usability test, issues detected by testers were addressed by researchers.

Results obtained through open-ended questions and comments made for closed-ended questions were grouped together into two major categories: suggestions and errors found. Suggestions were qualified by researchers as: “definite”, “possible”, and “discard”. “Definite” suggestions were resolved because those would increase usability, minimize problems, and would increase system’s quality. One suggestion was qualified as “possible”. It was considered but discarded because target users are very likely not to be proficient in using computers. To solve this issue, it would require pre-training users in the use of the keyboard. Thus, adding the use of a computer keyboard would reduce system’s usability. Five suggestions were qualified as “Discard”. Those were ignored because they were not relevant or do not add value to the systems. In addition, usability testers detected seven problems during first session (shown in Table 1, including the number of times mentioned). Those problems were addressed and resolved.

Table 1: Problems detected during first usability evaluation.

Improvements	Mentions
Improve rendering of point in charts	5
Two different buttons trigger the same application	2
System fails while saving own examples created by users	4
Problems with GUI resolution	2
Some topics are duplicated in help documents	1
Plots are not cleaned before starting a new example	1
GUIs lost plots after reading online help	2

The second usability evaluation was performed in order to determine whether issues detected by the testers were resolved. There was not mortality of testers of those that participate in the first usability evaluation, which increased system evaluation reliability and helped to strengthen results. In the second usability evaluation no issues were detected by testers. Further, in question number 6 eighteen testers said “Yes” (sixteen in the first evaluation), zero said “No” (same number as first evaluation), and one said “Do not know” (three in the first evaluation). Hence, it can be argued that evaluators were more likely to recommend the system after the second evaluation. From both usability evaluations, results show that the system is very likely to have high usability. Thus, after completely addressing the issues and suggestions made by usability testers, the final version of the system was ready to be used.

Results from the second usability evaluation changed compared with the first evaluation and show less dispersion. A comparison of means and standard deviations of both usability evaluations is presented in Table 2. In all cases, a reduction in the means as well as the standard deviations can be observed.

Table 2: Comparison of results from usability evaluation.

	First test		Second test		Change	
Question #	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
1	2.053	0.780	1.737	0.653	0.316	0.126
2	1.947	0.848	1.579	0.507	0.368	0.341
3	1.684	0.582	1.526	0.513	0.158	0.069
4	2.263	0.653	1.842	0.602	0.421	0.051
5	2.474	0.905	1.947	0.621	0.526	0.284
6	1.316	0.749	1.105	0.315	0.211	0.434
7	1.684	0.749	1.421	0.507	0.263	0.242
10	2.211	0.535	1.684	0.478	0.478	0.058

PILOT TEST

In order to validate the final exam for the training courses, a pilot test was conducted. Twenty-four students in their eighth semester of a bachelor degree in Computer Sciences participated in the pilot test.

They were selected because they had already been taught several statistics courses. A day before the exam, each participant received notes for the training course that were the base to write the final exam. No other training or lecture sessions were given to participants. The score was calculated in a 0 to 10 scale. Time was recorded starting off at the beginning of the exam. The first person finished twenty-one minutes after the starting time, and the last person finished forty-one minutes after the starting time. Figure 2 shows the participants' grades distribution obtained.

Figure 2: Histogram results based on overall grade obtained in the pilot test.

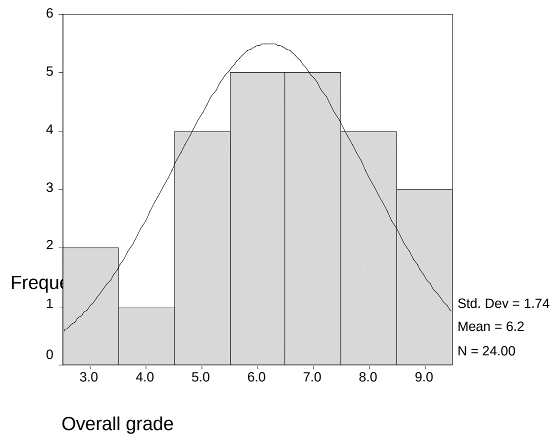


Table 3 shows descriptive statistics for the pilot test. Past research (Tabachnick, 1996) recommends that a sample's skewness value should not be beyond two times standard errors for skewness (SES). In our case, the calculated skewness value (-.489) for the score is below the SES (two times ±0.472). Thus, it is assumed that the skewness is within the expected range of chance fluctuations; which means that there is no significant skewness problem. In addition, a sample's kurtosis value should not exceed two times the standard errors of kurtosis (SEK). The calculated kurtosis value (-.599) is below SEK (two times ±0.918). Hence, it is assumed that the kurtosis is within the expected range of chance fluctuations. It can be said that the instrument results exhibit a normal distribution. Thus, the results of the pilot test indicate that the instrument is able to measure what is intended.

Table 3: Statistics obtained in the pilot test.

Overall grade

N	Valid	24
	Missing	0
Mean		6.1979
Median		6.6250
Mode		5.75[a]
Std. Deviation		1.7383
Skewness		-.489
Std. Error of Skewness		.472
Kurtosis		-.599
Std. Error of Kurtosis		.918
Minimum		2.50
Maximum		8.75

a. Multiple modes exist. The smallest value is shown

RESEARCH DESIGN

Two-Group Study

Learn to learn is very important. Frequently, organizations require people that have knowledge in a special area. Thus, organizations should either hire trained people or train their own people. The issue is not how to train people, but rather how to ensure they learn better. Thus, how people can learn to learn is an issue that needs to be addressed further by both academicians and practitioners.

For the first study, two groups had five sessions using different training approaches. In order to control the effects of teaching styles and to avoid compensatory coaching for any group, only one instructor taught both courses.

Participants were recruited through invitations sent to four different colleges. Potential participants should be registered in either their first or second semester of the program. Only two institutions provided a list of potential participants to researchers. One hundred and twenty potential participants were invited and fifty-six agreed to participate. Only fifty were selected because this was the number of computers available. Then, selected participants were assigned randomly to each group.

The first study measured the effectiveness of teaching on those two groups (the dependent variable) by applying an exam at the end of a course. Both courses were conducted as follows, on a daily basis, for a set of five consecutive days.

- The experimental group made use of the learning system (EXPG), which allowed participants to review stored examples, example creation, analyze examples (either stored or created), review electronic notes of the course, and store created examples.
- The control group (CONG) was taught through traditional means. That is, they did not use any type of system. The CONG group received a package with the notes for the course. During each session, participants were lectured on the topic of the day. In the next activity, the group solved a set of ten examples previously prepared by the researchers and some examples proposed by participants.

At the end of the experiment, an exam was given to participants in order to measure the outcomes of the training program. The exam was designed by independent sources (two lecturers that do not participate in any stage of the study)

Table 4 shows the descriptive statistics regarding the final exam scores for both groups. The distribution for EXPG was not normal; it was negatively skewed (with 9 of 25 subjects scoring a perfect 10). Following suggestions made in the literature (Tabachnick, 1996), EXPG skewness (-1.416) is beyond two times SES (±0.928), which means that the data is negatively skewed. On the other hand, data for CONG was not skewed. CONG skewness (-0.499) is between two times SES (±0.928), which means that the sample is normally distributed.

Table 4: Descriptive statistics for the final exam results.

		EXPG	CONG
N	Valid	25	25
	Missing	0	0
Mean		9.2700	5.6900
Median		9.5000	5.7500
Mode		10.00	2.50[a]
Std. Deviation		.8383	1.8851
Variance		.7027	3.5535
Skewness		-1.416	-.499
Std. Error of Skewness		.464	.464
Kurtosis		2.047	-.811
Std. Error of Kurtosis		.902	.902
Range		3.25	6.25
Minimum		6.75	2.00
Maximum		10.00	8.25

a. Multiple modes exist. The smallest value is shown

Time required for completing the exam was recorded for each subject as minutes eployed (see Table 5). Average time for EXPG was just below 22 minutes. There was no significant correlation ($r = -0.130$, $p = 0.37$) between time and score. Average time for CONG was just above 34 minutes. There was no significant correlation ($r = 0.305$, $p = 0.139$) between time and score. EXPG subjects, on average, required less time to complete the exam.

Table 5: Descriptive statistics: time required for answering the exam.

	N	Range	Minimum	Maximum	Mean	Std. Deviation
EXPG-TIME	25	16.00	15.00	31.00	21.8400	3.9017
CONG-TIME	25	19.00	22.00	41.00	34.0800	6.1774
Valid N (listwise)	25					

Data analysis shows that both groups performed differently. By comparing overall grade means and time required for answering the final exam means, overall, EXPG participants performed better. EXPG group mean (9.27) is 3.58 points higher than CONG group mean (5.69). Thus, overall, participants in the EXPG had the best performance using less time for answering the final exam.

Subjects' mean grades were tested using ANOVA, which shows that the means of the both groups are different (see Table 6). ANOVA test shows that the training method makes a difference in performance ($p<0.001$). Hence, the purposely-developed software for teaching helps to perform better comparing with teaching using traditional means.

Table 6: ANOVA table for comparison of groups.

Overall grade

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	160.205	1	160.205	75.280	.000
Within Groups	102.150	48	2.128		
Total	262.355	49			

Three-Group Study

The second study measured the effectiveness of teaching on the three groups (the dependent variable) by applying an exam at the end of a course. All courses were conducted as follows, on a daily basis, for a set of five consecutive days.

- One group (ESTG) made use of the full version of the developed training system. This version allows participants to review stored examples, example creation, analyze examples (either stored or created), review electronic notes of the course, and store created examples.
- A second group (ISTG) made use of the reduced version of the developed training system, which allowed participants to review stored examples, example creation, analyze examples (either stored or created), review electronic notes of the course, but cannot store created examples.
- The control group (NOSG) was taught through traditional means. That is, they did not use any type of system. The CONG group received a package with the notes for the course. During each session, participants were lectured on the topic of the day. In the next activity, the group solved a set of ten examples previously prepared by the researchers and some examples proposed by participants.

All participants completed all five sessions and answered the final exam. Table 7 shows the descriptive statistics regarding final exam scores for the three groups. For the ESTG, eleven of twenty-five subjects scored a perfect 10. Following the guidelines of Tabachnik and Fidell (1996), ESTG skewness (-1.328) is beyond two times the standard error of skewness (SES) (±0.968), which means that sample is not normally distributed. It was negatively skewed, which means that the data is skewed to the highest grades, meaning that people performed well above average (as expected).

In addition, the distribution for the ISTG group was not normal with twelve of twenty-five subjects scoring either a 9 or a 9.5. ISTG skewness (-1.050) is beyond two times SES (±0.968). For this group, sample is not normally distributed also. It was negatively skewed but not as skewed as the first group (ESTG).

Contrary to ESTG and ISTG groups, data for NOSG was not skewed. NOSG skewness (-0.546) is between two times SES (±0.968), which means that sample it is normally distributed . Thus, the data is not skewed to any end of the grades distribution meaning that people performed normally (as expected).

Table 7: Descriptive statistics for the final exam results.

Statistics

		Overall grade group 3	Points group 3	Overall grade group 2	Points Group 2	Overall grade Group 1	Points Group 1
N	Valid	25	25	25	25	25	25
	Missing	0	0	0	0	0	0
Mean		5.6900	22.84	8.2900	33.16	9.2900	37.16
Mode		5.75[a]	23[a]	9.00[a]	36[a]	10.00	40
Range		6.25	25	4.00	16	3.25	13
Minimum		2.00	8	5.50	22	6.75	27
Maximum		8.25	33	9.50	38	10.00	40

a. Multiple modes exist. The smallest value is shown

The number of minutes required to complete the exam was recorded for each subject. Descriptive statistics for time required for the three groups are shown in Table 8. The average time was just below 22 minutes. There was no significant correlation (r= -0.152, p=0.469) between time taken and score. For the

ISTG the average time was just above 22 minutes. There was no significant correlation (r=0.100, p=0.635) between time taken and score. The average time was just above 34 minutes for the NOSG. There was no significant correlation (r=0.317, p=0.123) between time taken and score.

Table 8: Descriptive statistics: time required for answering the exam.

Descriptive Statistics

	N	Range	Minimum	Maximum	Mean	Std. Deviation
Time expent solving the exam Group 3	25	19	22	41	34.08	6.18
Time expent solving the exam Group 2	25	16	15	31	22.40	5.00
Time expent solving the exam Group 1	25	16	15	31	21.84	3.90
Valid N (listwise)	25					

Data analysis shows that the three groups performed differently. We can argue that, overall, participants in the ESTG had the best performance. The ESTG mean (9.29) is one full point higher than the ISTG mean (8.29), and 3.7 points higher than the NOSG mean.

In addition, ESTG participants, on average, required less time to complete the final exam. Thus, overall, subjects in the ESTG had the best performance and required less time for answering the same exam. Participants' mean grades were tested using ANOVA. The test shows that the means of the three groups are different. Table 9 shows descriptive statistics.

Table 9: Summary of descriptive statistics.

Descriptives

Overall grade

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	25	9.2900	.8859	.1772	8.9243	9.6557	6.75	10.00
2	25	8.2900	1.0721	.2144	7.8475	8.7325	5.50	9.50
3	25	5.6900	1.8961	.3792	4.9073	6.4727	2.00	8.25
Total	75	7.7567	2.0314	.2346	7.2893	8.2241	2.00	10.00

A test of homogeneity of variances (Levene's Statistic) shows that the three groups' variances are not equal (p<0.001). While this violates an assumption of ANOVA, failing to meet this "assumption of homogeneity of variances is not fatal to ANOVA, which is relatively robust, particularly when groups are of equal sample size", which is the case of the study (Garson, 1998). Table 10 shows that the training method makes a difference in performance (p<0.001). The ESTG had better performance than the other training approaches, and the ISTG's performance was better than the NOSG.

Table 10: ANOVA table for comparison of training groups.

ANOVA

Overall grade

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	172.667	2	86.333	46.841	.000
Within Groups	132.705	72	1.843		
Total	305.372	74			

As a backup, nonparametric tests, Kruskal-Wallis, Median, and the Dunn-Sidak paired t-test were also performed and results are highly significant. The ESTG had better outcomes than the ISTG ($p<0.004$) and the NOSG ($p<0.001$). The ISTG had better outcomes than the NOSG ($p<0.001$).

CONCLUSIONS

The purposely-developed software is intended to help people learn statistics. Such type of system did not exist at the time when the present study was conducted. The system was developed using a prototyping approach. Several prototypes were constructed and evaluated using pre-solved examples and pre-designed errors. It was very useful to develop the software by prototyping because each version was tested so that errors and improvements were identified and addressed. This approach also helped in developing a system that was up to what users would require.

It was important for the present study to develop an easy-to-use system. In order to asses such easiness, a usability test was performed. Thus, after developing the system (through prototyping) and evaluating its usability, it could be concluded that the system may be used in the study.

Given the statistical test performed, it is clear that the grades of the groups that used the tool were significantly better than those of the group submitted to traditional, non-technological training techniques. Comparing the results, it can be concluded that the purposely-developed system really made a difference in learning outcomes. Hence, a purposely-developed training/teaching software does really makes a difference in performance outcomes.

Because the market cannot offer software that can be tailored, or that it is intended to teach some specific areas of knowledge, the researchers highly recommend developing software that can be used to train/teach people specific subjects that are complex in nature. We believe that purposely-developed software is a good option to help people learn specific knowledge or tasks. This recommendation is based in the results obtained in the present study.

Limitations of the Study

Experts in teaching statistics that also have knowledge in developing information systems were not available. Hence, testers could have ignored important issues during prototype evaluations and were not addressed. Training system might be enhanced and improved by using a different programming language, or a different AI approach, which could increase usability and quality. Using a non-prototyping software system development approach might result in a different training system.

Maybe results are only true for participants in the study. Since the demographics of the population were not known, then, sample may represent part of the population, but not the population itself. Hence, results might be different for a representative sample. Thus, results cannot be generalized.

There might be some issues that were unnoticed during the prototyping stages of the project or by usability testers. Such issues could have influence in the outcomes of the study.

Areas for Additional Research

The study can be conducted longitudinally so that several evaluations of participants can be performed. Such a study could evaluate not only whether participants learn during training courses but also how much they retain after training.

Findings cannot be generalized. Even tough hypothesis testing outcomes were very good, sample groups are not representative of the population and external validity is not strong. It is call for a future study with a representative sample.

REFERENCES

Bates, C., & Watson, M. (2008). Re-learning Teaching Techniques to be Effective in Hybrid and Online Courses. *Journal of American Academy of Business, Cambridge, 13*(1), 38-44.

Bohlen, G. A., & Ferrat, T. W. (1997). End User Training: An Experimental Comparison of Lecture versus Computer-Based Training. *Journal of End User Computing, 9*(3), 14-27.

Bostrom, R. P., Olfman, L., & Sein, M. K. (1990). The Importance of Learning Style in End-user Training. *MIS Quarterly*(March 1990), 101-119.

Boudreau, M.-C., & Seligman, L. (2005). Quality of Use of a Complex Technology: A Learning-Based Model1. *Journal of Organizational and End User Computing, 17*(4), 1, 22 pgs.

Bowman, B. J. (2002). Building Knowledge Management Systems. *Information Systems Management, 19*(3), 32-40.

Bowman, B. J., Grupe, F. G., & Simkin, M. G. (1995). Teaching End-User Applications With Computer-Based Training: Theory and an Empirical Investigation. *Journal of End-User Computing, 7*(2), 12-18.

Bozionelos, N. (1996). Psychology of Computer Use: XXXIX. Prevalence of Compuer Anxiety in British Managers and Profesionals. *Psychological Reports, 78*(3 pt.1), 995-1002.

Brown, K. G. (2001). Using Computers to Deliver Training: Which Employees Learn and Why? *Personnel Psychology, 54*, 271-296.

Buch, K., & Bartley, S. (2002). Learning Style and Training Delivery Mode Preference. *Journal of Workplace Learning, 14*(1), 5-10.

Csikszentmihalyi, M. (2000). *Beyond Boredom and Anxiety*. San Francisco, CA: Jossey-Bass Publishers.

Chin-Chih, L., & Chien-Chung, L. (2008). Instructional Strategies and Methods of e-Learning for Nurturing Appreciative, Expressive, and Creative Abilities. *Journal of American Academy of Business, Cambridge, 13*(1), 199-207.

Desai, M. S., Richards, T., & Eddy, J. P. (1999). End-User Training: A Meta Model. *Journal of Instructional Psychology, 26*(2), 1-8.

Desai, M. S., Richards, T., & Eddy, J. P. (2000). A Field Experiment: Instructor-Based Training vs. Computer-Based Training. *Journal of Instructional Psychology, 27*(4), 239-243.

Dumas, J. F., & Redish, J. C. (1999). *A Practical Guide to Usability Testing*. Bristol, England: Intellect Books.

Garson, D. A. (1998). *One Way Analysis of Variance Example*. Retrieved 10/19/2006, 2006, from http://www2.chass.ncsu.edu/garson/pa765/anova1.htm

Goode, S., Willis, R. A., Wolf, J. R., & Harris, A. L. (2007). Enhancing IS Education with Flexible Teaching and Learning. *Journal of Information Systems Education 18*(3), 297-302.

Grainger, C. (2001). *Just-for-me training: personalized options for training on demand.* Paper presented at the 29th annual ACM SIGUCCS conference on User services, Portland, Oregon, USA.

Hayes, P. J. (1999). *Providing Expert Systems with Integrated Natural Language and Graphical Interfaces.* Paper presented at the Proceedings of 1986 fall joint computer conference, Dallas, Tx.

Hornik, S., Johnson, R., & Wu, Y. (2007). When Technology Does Not Support Learning: Conflicts Between Epistemological Beliefs and Technology Support in Virtual Learning Environments. *Journal of Organizational and End User Computing, 19*(2), 23.

Johns, J. F. (2000). *Implementing Intelligent 3-D Web-Based-Training Solutions.* Paper presented at the American Society of Training and Development International Conference Proceedings, Alexandria, VA. USA.

Kazmer, M. M., & Haythornthwaite, C. (2005). Multiple perspectives on online learning. *ACM SIGGROUP Bulletin, 25*(1), 7-11.

Kekäle, T., Pirolt, K., & Falter, C. (2002). IT Tools in Personnel Training in Small Enterprises: Results of the Project "Apply". *Journal of Workplace Learning, 14*(7), 269-277.

Lassila , K. S., & Brancheau, J. C. (1999). Adoption and utilization of commercial software packages: exploring utilization equilibria, transitions, triggers, and tracks. *Journal of Management Information Systems, 16*(2), 63-90.

Montesino, M. U. (2002). Strategic Alignment of Training, Transfer-Enhancing Behaviors, and Training Usage: A Postraining Study. *Human Resource Development Quarterly, 13*(1), 89-108.

Oldfield, N., & Rose, C. (2004). Learning delivers the best return. *Industrial and Commercial Training, 36*(1), 25-28.

Olfman, L., & Pitsatorn, P. (2000). End-User Training Research: Status and Models for the Future. *Framing the Domains of IT Management. Projecting the Future Through the Past. Pinnaflex, Cincinnati, Ohio.*

Palvia, S. C. J., & Palvia, P. C. (2007). The Effectiveness of Using Computers for Software Training: An Exploratory Study. *Journal of Information Systems Education, 18*(4), 479-489.

Piskurich, G. M. (2006). E-Learning Fast, Cheap, and Good. *Performance Improvement Quarterly, 45*(4), 18, 18pgs.

Pressman, R. S. (2005). *Software Engineering: A Practitioner's Approach, 6/e*. New York, N.Y.: Mc Graw-Hill.

Romanik, K., & Snart, F. (2000). Enhancing Employability: The Role of Prior Learning Assessment and Portfolios. *Journal of Workplace Learning, 12*(1), 29-34.

Sommerville, I. (2005). *Ingeniería de Software*. Madrid, España: Pearson Education.

Spaniol, M., Klamma, R., Springer, L., & Jarke, M. (2006). Aphasic Communities of Learning on the Web. *International Journal of Distance Education Technologies, 4*(1), 31.

Spitler, V. K. (2005). Learning to Use IT in the Workplace: Mechanisms and Masters. *Journal of Organizational and End User Computing, 117*(2), 1, 25 pgs.

Tabachnick, B. G., & Fidell, L.S. (1996). *Using multivariate statistics (3rd ed.)*. New York, N.Y.: Harper Collins College Publishers.

U.of W. (2004, October 05, 2004). *Usability test for the University of Washington libraries' information gateway*. Retrieved December 2004, 2005, from http://www.lib.washington.edu/usability/redesign03-04/surveyResults.html

Vokurka, R. J., Flores, B. E., & Pearce, S. L. (1996). Automatic Feature Identification and Graphical Support in Rule-Based Forecasting: A Comparison. *International Journal of Forecasting, 12*(1), 495-512.

Webster, J., & Martocchio, J. J. (1995). The differential effects of software training previews of training outcomes. *Journal of Management, 21*(4), 757-787.

Yi, M. Y., & Davis, F. D. (2001). Improving Computer Training Effectiveness for Decision Technologies: Behavior Modeling and Retention Enhancement. *Decision Sciences, 32*(3), 521-544.