Assessing the Efficacy of Mobile Computing Platforms in Mathematics Education via a Mobile Learning Usability Scale (MLUS)

John E. Ehrke

Abilene Christian University

Abstract

Over the next decade, it is anticipated that mobile learning technologies will significantly impact the future of the graphing calculator platform. The impact of integrated devices (devices which blend productivity, social media, and computing) on educational design in mathematics remains largely unexplored. In this study, we analyze the results of a fall 2010 focused comparison of two sections of a first-year, general education mathematics course. Student performance data and student perceptions of usability are compared across two platforms: the SpaceTime™ mobile computing app and the Texas Instruments TI-84® series of graphing calculators. Pedagogical implications of the case study results are assessed and discussed within the mathematics TPACK (technology pedagogy and content knowledge) framework. (*Keywords: graphing calculators, technology pedagogy and content knowledge, mobile learning, mathematics education, mobile computing, reliability and validity, usability scale*)

Introduction

The use of graphing calculators in educational settings has grown rapidly in the last twenty years. Even though traditional graphing calculator platforms are now more than two decades olds, their impact on educational design and philosophy is still a heavily researched question (Ellington 2003, Clark-Wilson 2010). Just as advances in electronics and display improvements led to the birth of today's graphing calculators, similar technological advances in mobility and multi-touch interfaces are poised to take the reins in driving the graphing calculator to the mobile platform.

The integration of technology—for technology's sake—risks impeding any performance gains to be realized by the use of graphing calculators unless careful consideration is given to the effect of such platforms, not only on student performance, but also on instructional practices, content knowledge, and teacher knowledge. Therefore, it is a principal aim of this study to understand how mobile computing platforms fit within the mathematics TPACK standards (Niess 2009). This report details the findings of a fall 2010 case study which looked at a comparison of student perceptions of usability and performance across two sections of a first-year undergraduate mathematics course. A comparison of the experimental and control platforms is also included.

Theoretical Framework

This report draws heavily upon the TPACK framework's assumptions about knowledge and learning as they relate to mathematics education. Illustrations of the model's use in assessing emerging technologies, development of standards for mathematics education, and integration of the model can be found in (Lee, Hollebrands 2009, Niess 2009, Groth, 2008). As called for in the TPACK standards and development model, embracing a new technology requires we need to reassess the technology based on three criteria: (1) its effect on the teaching and learning of mathematics, (2) changes in instructional pedagogies required to facilitate the technology, and (3) effect on content knowledge building. Our discussion centers on these criteria and provides a lens through which to view any empirical results.

Based on the work of Mishra and Koehler (Mishra 2006) and adapted by the AMTE (Association of Mathematics Teacher Educators), specific guidelines for planning, improving, and evaluating technology-based mathematics instruction at all levels are organized around four major areas:

1. **Design and develop technology-enhanced mathematics learning environments and experiences.** In designing the course careful consideration was given to the opportunities provided by a mobile computing platform in creating a more learner friendly environment. Students in the mobile learning group had access to screencasting resources via their mobile

devices which helped them accelerate the learning curve associated with the use of a new technology. Future resources are being devoted to expanding the pool of screencasts available to students via their mobile devices as a way of delivering a “technology-enhanced” learning environment. While we acknowledge such efforts are platform independent, we believe the prevalence of student use of mobile devices outside the classroom (as opposed to traditional graphing calculator handhelds) makes students more likely to integrate mathematics and mobility.

2. **Facilitate mathematics instruction with technology as an integrated tool.** Students in the mobile learning group frequently utilized their iPhones® not only as a graphing calculator, but also as a clicker using ResponseWare™ from Turning Technologies, and as a collaborative learning tool using the HeadsUp™ app developed for on-campus use. In this way, the new technology is very much an “integrated tool.”
3. **Assess and evaluate technology-enriched mathematics teaching and learning.** The TPACK framework provides the overarching assessment tool for not only evaluating student learning, but educator learning as a result of technology implementation. As such, not only do we blend qualitative and quantitative assessments, but we provide a pedagogical framework within which to view such assessments.
4. **Engage in ongoing professional development to enhance technological pedagogical content knowledge.** Clearly this study is an outgrowth of this guideline and seeks to inform future research on mobile learning in mathematics as it relates to enhancing the knowledge, productivity, and research practices for other educators.

Together these considerations provide the underpinnings of the research presented in this study. We now consider the specific results in the literature that pertain to this study.

Review of Literature

Mobile Learning

Mobile learning is defined by (Quinn 2000) as, "the intersection of mobile computing and eLearning; accessible resources wherever you are, strong search capabilities, rich interaction, powerful support for effective learning, and performance based assessment." In some sense, this study is compelling in that we are comparing two different mobile learning platforms: traditional handheld graphing calculators and iPhone® based mobile computing applications. In mathematics and science instruction, there is widespread belief that technology is pedagogically effective and can lead to improvements in learning.

Technology in the classroom has been mandated by curriculum standards for over a decade. For example, the NCTM standards for grades 9-12 state that all students should *develop fluency in operations with real numbers, vectors, and matrices using mental computation or paper and pencil calculations for simple cases and* ***technology for more complicated cases***. This awareness of the need for technology in computation is not exactly new either as similarly worded NCTM standards date back to 1989. (NCTM 9-12) It makes perfect sense therefore that the use of graphing calculators have had a far reaching effect on math and science instruction since they represented a more affordable, more portable, and more user friendly alternative to the laptop or desktop computer. This study wonders if the mobile computing app is the next step in the evolution of such standards.

Unfortunately much of the current mobile learning research is restricted to usages of mobile devices as communication or multimedia tools. This means there is little precedent for the empirical or qualitative study of the mobile platform as a computing tool. The few examples that exist in the literature utilize empirical results with small sample sizes to claim there is no significant statistical difference in student performance using the devices. (Mayrath 2011, Schou 2007) In many cases though, similar studies have shown an increase in motivation and student attitudes in a mobile environment. Much of the current mobile learning research focuses on Palm OS devices or laptops ranging in age from five to ten years old. While five to ten years may not seem very long, within the last three years we have seen the introduction of multi-touch interfaces, the incredible growth of the tablet market with products like the iPad®, MobileCAS® apps like Wolfram Alpha™ and SpaceTime™, and the integration of augmented reality tools. Given this exciting array of tools we feel much of the current mobile learning research is already outdated from a platform perspective.

Survey Design and Usability Scales

One of the primary sources of data in this study are student survey results. There have been numerous survey studies that concluded that students could improve their learning achievement and attitude, (Liu 2010) while others have questioned the lack of empirical support for these claims and called for the alignment of outcome measures with technological innovation. (Bebell 2010) Still others have warned that survey data is potentially biased by students having correct notions about "socially desirable answers." (Cheung 2009) When coupled with survey fatigue we feel that any assessment of mobile learning must rely on more than survey data to substantiate claims. In lieu of using only student reported data, case studies represent an effective means of performing qualitative research (Merriman 1998) and have been used to "explore the possible effects on teaching and learning of wireless and mobile technologies." (Liu 2010) Much of the current research on these effects is in the form of various case studies (Chen 2009, Wang 2009, Shen 2009) as they can form the basis for a more robust general theory to which this study hopes to add.

As we seek to understand the usability of these new devices we are reminded that there is a difference between usability of mobile devices and "good usability" which according to (Kukulska-Hulme 2007) means learning can proceed without obstacles and might even be enhanced by the availability of certain features (e.g. screencasting, collaborative learning apps, computing apps, web access, etc…). This study seeks to define "good usability" as it relates to a mobile graphing calculator platform. The idea of usability in general is considered by many authors to be a prominent deciding factor in the adoption of a new educational technology. (Forster 2006, Looi 2010, Kukulska-Hulme 2007)

At the heart of this research is a realization that mobile devices represent a pedagogical shift from didactic teacher-centered to participatory student-centered learning (Looi 2010) but there has been no acknowledgement that such a shift has long already occurred in mathematics education due to the prevalence of handheld graphing calculators historically.

Perhaps a more promising direction for inquiry in mathematics and the sciences is in the distinction between formal and informal learning. Formal learning as described in (Looi 2010) is learning based on fixed curricula enacted in classroom environments whereas informal learning takes place outside the confines of the classroom. The authors argue the two forms of learning are not in conflict and that mobile learning can help bridge the gap.

One of the challenges faced by researchers whose studies focus on survey and comparison of test results, is to quantify the extent to which learning takes place in an informal context. In (Looi 2010) the authors propose the use of the experience sampling method (due Csikszentmihalyi and Larson, 1987) with hopes that the method may “provide a better understanding and natural assessment of how students are engaged in informal learning every day with mobile devices as they are using it.” While this study does not pursue this direction of inquiry it acknowledges such considerations as an important goal of future research. In general it seems there is not an agreement among researchers what methodology appropriately determines user experiences and effects on student performance over time, which is why in addition to empirical results we analyze results within a larger pedagogical context.

Parameters of the Study

Context

In the fall of 2008, Abilene Christian University distributed Apple iPhone® and iPod Touches® to 964 incoming first year students. In the last three years the campus wireless infrastructure and institutional practices have reflected the success, and at times the challenges, of institutionally embracing mobile learning. As of fall 2010, the campus reached full saturation of mobile devices with many faculty using the devices in class. This unique mobile friendly environment is the catalyst for exploring research of the type presented in this study. For more information on the ACU Mobile Learning Initiative visit http://www.acu.edu/connected.

This study was conducted in two first-year general education mathematics courses (MATH 120) in the fall 2010 semester. One section was designated as an experimental section in which students used their iPhone® and the SpaceTime™ computing app for all calculations normally reserved for a handheld calculator. The other section used the traditional TI-83, TI-84 handheld graphing calculator. Course content—a mix of probability, statistics, and mathematical finance—was consistent across both sections, and all assessments were the same. In-class instruction differed only in the student use of their device.

Data Sources

There were a total of (n = 171) student respondents to the initial mobile learning attitudes survey conducted across all sections of the course (including those who were in neither of the comparison groups). The comparison groups consisted of (n = 16) students in the control group and (n = 24) 24 students in the experimental group. Among students in the experimental group the average SAT math score was 450. The average ACT math score was 17.9, and ACT composite score average was 19.6. Among the control group the average SAT math score was 442.5, the average ACT math score was 16.9 and the ACT composite score average was 17.8.

At the beginning of the semester students across all sections of the course were invited to participate in an anonymous online survey initiated by their own section's instructor. In fall of 2010 a total of 287 students were taught across all sections of MATH 120, of these a total of 171 students participated representing a 60% coverage. Additionally, within each comparison group, usability and perception surveys were conducted after the completion of each major unit in which the calculator was used heavily, namely statistics and finance. Student participation in these surveys was voluntary and of the 16 students in the control group 12 completed the post statistics survey while 18 of the 24 students in the experimental group completed the same survey. This represents a coverage of 75% in each of the comparison groups. For the post-finance survey 7 of 16 students completed the survey in the control group and 10 of 24 completed the survey in the experimental group for a coverage of 44% and 42%, respectively.

Research Questions

The purpose of this study was to determine the efficacy of a mobile computing platform by comparing student perceptions of usability, student performance on in-class exams, and the logistics of calculator use between the SpaceTime® iPhone® app and Texas Instruments TI-8x series of calculator.

To examine the issue of usability, we developed a mobile learning usability scale (MLUS) based on student responses. The questions asked on the scale fit into two basic constructs: general usability and content-specific areas of instruction.   For the statistics portion students were asked to rate the usability of their calculator on four sub-areas: (1) data entry and management, (2) sample statistics calculations, (3) linear regression and correlation, and (4) normal distributions. These topics represent four of the main statistics topics in the MATH 120 course. Similarly, for the finance portion, students were asked to rate the usability of their calculator on four more sub-areas: (1) solving equations, (2) TVM solvers, (3) interest calculations, and (4) effective rate or APY calculations. By comparing the student responses across these sub-areas we hope to determine if a mobile platform is well suited as a computational tool for use in MATH 120.

To assess the degree to which the goals of our research were met, we evaluated the following hypotheses:

1. There is no difference in mean score on the MLUS content sub-areas between the traditional calculator and the mobile platform.
2. There is no difference in mean score on either content exam (statistics, finance) between the traditional calculator and the mobile platform.

3. There is no difference in mean score on the comprehensive final exam between the traditional calculator and the mobile platform.

Results

Logistics of Calculator Use

One of the chief aims of this study was to identify the number and sources of calculators being brought into the course by students as well as the potential economic ramifications to students. Of the 171 students surveyed at the beginning of the semester 115 (67%) had a calculator from high school, 20 (12%) were going to borrow a friend's calculator, 28 (16%) were going to purchase a new calculator, 5 (3%) were going to rent a calculator, and 3 (2%) were going to obtain a calculator via other means. Assuming the average cost of the TI-84 calculator at retail is around $120.00 (prices range from $90.00 to $140.00 depending on model), the 28 students willing to buy a new calculator for the course represent an investment of $3360.00 at a per student cost of $19.65. If all the students surveyed were asked to pay $9.99 for the SpaceTime™ app this would represent an overall investment of approximately $1708.29. At these per student costs the potential savings in using a mobile app across a typical 300 student fall enrollment is around $2900.00.

Additionally, we tried to quantify the extent to which a mobile platform made a student more or less likely to bring their calculator. The results are categorized in the table below. Of the 42 respondents, students using the traditional handheld calculators were more than twice as likely to forget their calculator at least once as compared to those students using the iPhone®. When it comes to battery life however, 50% of students reported being unable to use their calculator on at least one occasion due to a low battery, compared with only 4.2% of TI-8x users. The incidence rate of battery issues among iPhone® users can partly be attributed to the number of students using the iPhone 3G or iPhone 3Gs. Among those students using the newer iPhone 4, there were far fewer incidents.

	Totals	How many times have you forgotten to bring your calculator to class with you this semester?					How many times have you brought your calculator to class and been unable to use it because of a low battery?			
		0	1-2	3-4	5-6	6+	0	1	2-3	4+
	42	27	12	3	0	0	32	4	4	2
Texas Instruments (TI-83, TI-84 series)	24 57.10%	13 48.10%	9 75.00%	2 66.70%	0 0.00%	0 0.00%	22 68.80%	1 25.00%	0 0.00%	0 0.00%
SpaceTime (iPhone platform)	18 42.90%	14 51.90%	3 25.00%	1 33.30%	0 0.00%	0 0.00%	10 31.30%	3 75.00%	4 100.00%	2 100.00%

Table 1: Logistics of Calculator Use

General Usability (MLUS)

The MLUS survey consisted of three constructs created based on the literature review which suggested that “good usability” is a primary factor in the adoption of a new educational technology. The first construct, general usability, consists of three items assessing students ease of use, ease of instruction, and perceived performance with the device. The second construct assess usability as it related to the statistical content sub-area, and the third construct assessed usability as it related to the mathematical finance content sub-area.

The internal reliability estimates for the MLUS constructs based on Cronbach’s alpha were 0.729 (General usability), 0.933 (Statistical usability), and 0.942 (Financial usability)—See Table 2. Each construct measured above the generally accepted level of 0.70, and recorded very high levels within each content sub-area.

Construct	Number of Items	Range	Internal reliability
General usability	3	1 to 4	r = 0.729
Statistical usability	8	1 to 5	r = 0.933
Finance usability	6	1 to 5	r = 0.942

Table 2: MLUS Construct Internal Reliability Results

Statistics Usability

In order to determine the extent to which the means differ across the various items of our MLUS statistics sub-area we employed a 2-sample ttest across the four main areas (1) data entry, (2) sample statistics calculations, (3) linear regression, and (4) normal distributions. We also include 95% and 99% confidence intervals, F-test for equality of variances, and the t-test statistics. The top table below (Table 5) contains the descriptive statistics, while the bottom table (Table 6) contains the results of the test.

Upon examination of the results the null hypothesis is rejected ($\alpha = 0.05$) for the data entry sub-area. This suggests a difference in the mean usability between platforms for this area. This was not completely unexpected as a number of students remarked the difficulty the smaller buttons provided in entering and manipulating data with the SpaceTime® calculators. Such size limitations are unfortunate consequences of the platform, but it should be noted that an iPad platform might prove more reliable in this regard. It should also be noted that of the four content sub-areas only data entry failed the equality of variances assumption ($\alpha = 0.05$).

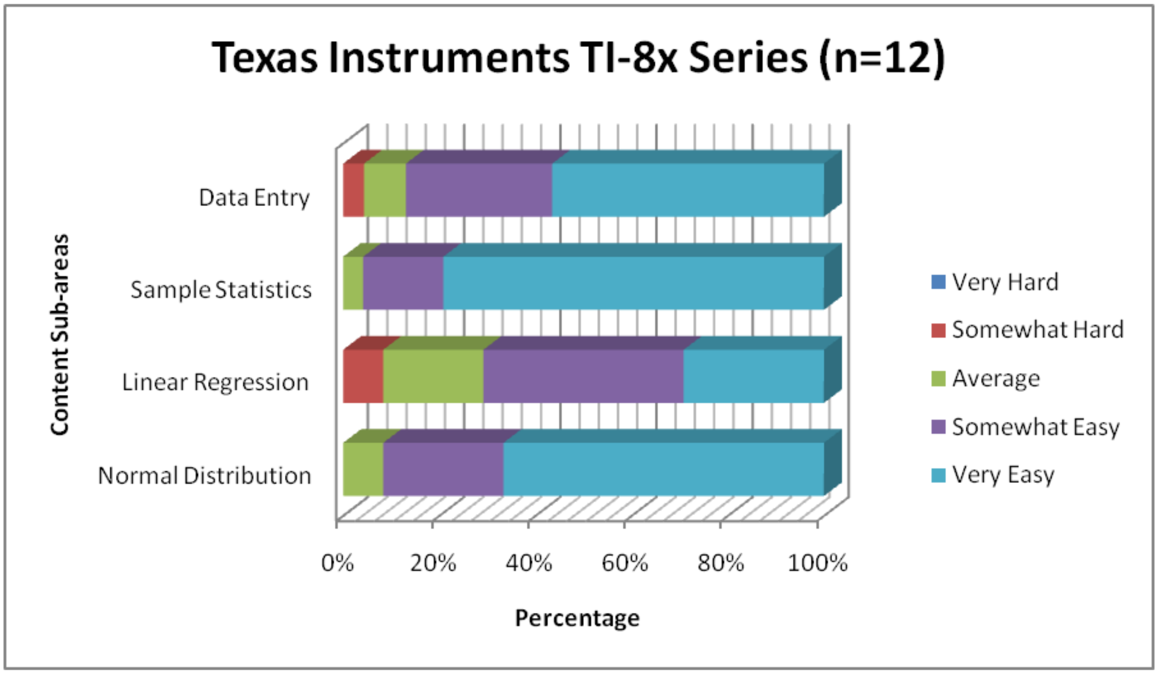
On the pages which follow the correlation matrix for each of the sub-area constructs is provided (see Tables 3-4). For the statistics sub-area, there was a strong correlation observed between ease of use and perceived personal performance, but very weak to no correlation between content areas and student reported ease of use, personal interaction with the platform, or perceived performance within the platform. This suggests that the general usability construct tests a different student-platform interaction as compared to the statistics construct and is further evidence of the inter-reliability of the scale.

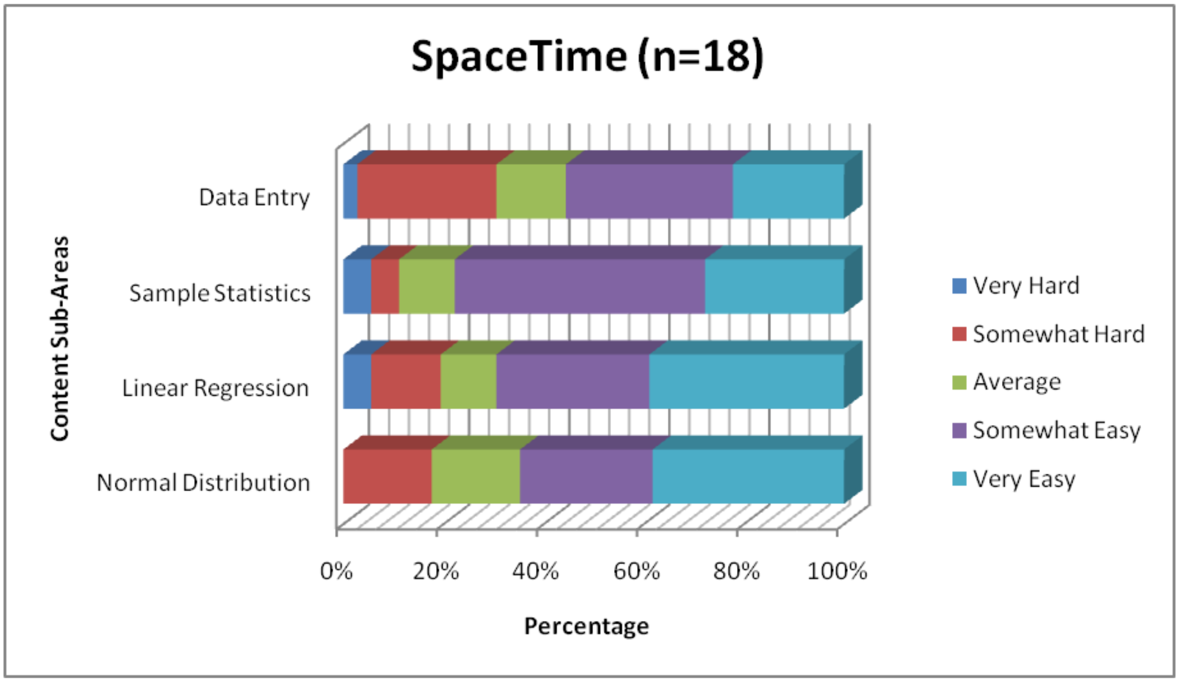
	Data Entry	*Sorting Data*	*Sample Statistics*	*Percentiles*	*Correlation*	*Line of Best Fit*	*Normal CDF()*	*InverseNorm()*	*Ease of Use*	*Personal Interaction*	*Personal Performance*
Data Entry	1.000										
Sorting Data	0.675	1.000									
Sample Statistics	0.527	0.510	1.000								
Percentiles	0.390	0.548	0.788	1.000							
Correlation	0.431	0.502	0.696	0.772	1.000						
Line of Best Fit	0.481	0.535	0.531	0.601	0.729	1.000					
NormalCDF()	0.549	0.587	0.707	0.814	0.805	0.771	1.000				
InverseNorm()	0.495	0.573	0.699	0.785	0.761	0.731	0.913	1.000			
Ease of Use	0.046	-0.028	0.084	0.052	0.065	-0.015	0.091	0.115	1.000		
Personal Interaction	0.241	0.172	0.303	0.298	0.270	0.135	0.293	0.299	0.665	1.000	
Personal Performance	0.177	0.063	0.149	-0.074	0.031	-0.046	0.053	0.062	0.403	0.519	1.000

Table 3: Correlation Matrix For Statistics Usability

	Solving Eqns	*TVM Solver*	*Loan Payments*	*Remaining Balance*	*Interest Earned*	*Calculating APY*	*Ease of Use*	*Personal Interaction*	*Personal Performance*
Solving Eqns	1.000								
TVM Solver	0.638	1.000							
Loan Payments	0.837	0.760	1.000						
Remaining Balance	0.804	0.723	0.968	1.000					
Interest Earned	0.853	0.691	0.968	0.936	1.000				
Calculating APY	0.874	0.718	0.941	0.899	0.915	1.000			
Ease of Use	0.615	0.460	0.616	0.620	0.658	0.518	1.000		
Personal Interaction	0.573	0.346	0.464	0.503	0.561	0.315	0.673	1.000	
Personal Performance	0.375	0.267	0.475	0.480	0.506	0.453	0.604	0.469	1.000

Table 4: Correlation Matrix for Finance Construct





Finance Usability

In order to determine the extent to which the means differ across the various items of our MLUS statistics sub-area we employed a 2-sample ttest across the four main areas (1) solving equations, (2) TVM solver, (3) interest calculations, and (4) APY calculations. We also include 95% confidence intervals, F-test for

Statistics		N	Mean	Std. Deviation	Std. Error Mean
Data Entry	TI-84	24	4.54	0.658	0.134
	SpaceTime	18	3.5	1.249	0.294
Sample Stats	TI-84	24	4.42	0.717	0.146
	SpaceTime	18	4.06	0.998	0.235
Linear Regression	TI-84	24	3.67	1.167	0.238
	SpaceTime	18	3.94	1.259	0.296
Normal Distributions	TI-84	24	4.13	1.154	0.236
	SpaceTime	17	3.88	1.166	0.283

Table 5: Descriptive Statistics (Statistics Sub-area)

		F-test for equality of variances		t-test for Equality of Means						
		F*	p	t	Df	p	Mean Difference	Std. Error Difference	95% CI	
									Lower	Upper
Data Entry	Equal Variances	3.603	0.036	3.4992	40	0.001	1.0417	0.298	0.438	1.642
	Unequal Variances	0.911		3.2202	24.05	0.004	1.0417	0.323	0.374	1.709
Sample Stats	Equal Variances	1.937	0.157	1.365	40	0.180	0.3611	0.265	-0.174	0.894
	Unequal Variances	0.719		1.3030	29.45	0.203	0.3611	0.277	-0.205	0.928
Normal Distribution	Equal Variances	1.164	0.323	0.6604	39	0.513	0.243	0.368	-0.481	0.980
	Unequal Variances	1.343		0.6592	35.18	0.514	0.243	0.369	-0.484	0.984
Linear Regression	Equal Variances	1.261	0.295	-0.7381	40	0.465	-0.2778	0.376	-1.031	0.491
	Unequal Variances	1.457		-0.7299	32	0.470	-0.2778	0.381	-1.050	0.495

*F-test performed for equal variances, standard pooled variance reported for unequal case

Table 6: Inferential Statistics (Statistics Sub-area)

Statistics		N	Mean	Std. Deviation	Std. Error Mean
Solving Equations	TI-84	7	4.143	0.899	0.340
	SpaceTime	10	4.100	0.994	0.314
TVM Solver	TI-84	7	4.571	0.535	0.202
	SpaceTime	10	4.000	0.943	0.298
Interest Earned	TI-84	7	4.429	0.787	0.297
	SpaceTime	10	3.900	0.994	0.314
Calculating APY	TI-84	7	4.286	0.756	0.285
	SpaceTime	10	3.700	1.059	0.335

Table 7: Descriptive Statistics (Finance Sub-area)

		F-test for equality of variances		t-test for Equality of Means						
		F*	p	t	Df	p	Mean Difference	Std. Error Difference	95% CI	
									Lower	Upper
Solving Equations	Equal Variances	1.223	0.322	0.792	15	0.440	0.043	0.054	-0.962	1.048
	Unequal Variances	0.708		0.8791	13.88	0.394	0.043	0.049	-0.951	1.037
TVM Solver	Equal Variances	3.107	0.074	1.441	15	0.170	0.571	0.396	-0.275	1.417
	Unequal Variances	0.648		1.5866	14.56	0.134	0.571	0.359	-0.198	1.341
Interest Earned	Equal Variances	1.595	0.235	1.170	15	0.26	0.529	0.452	-0.434	1.492
	Unequal Variances	0.841		1.221	14.68	0.241	0.529	0.433	-0.396	1.453
Calculating APY	Equal Variances	1.962	0.175	1.251	15	0.230	0.586	0.468	-0.411	1.583
	Unequal Variances	0.902		1.330	14.97	0.203	0.586	0.441	-0.350	1.524

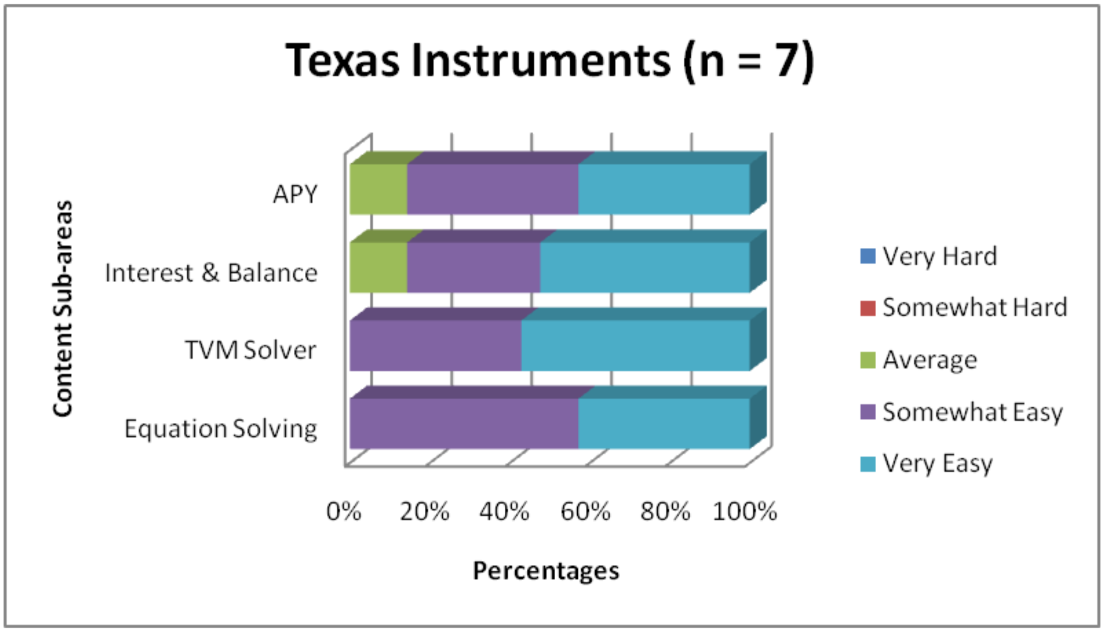
*F-test performed for equal variances, standard pooled variance reported for unequal case

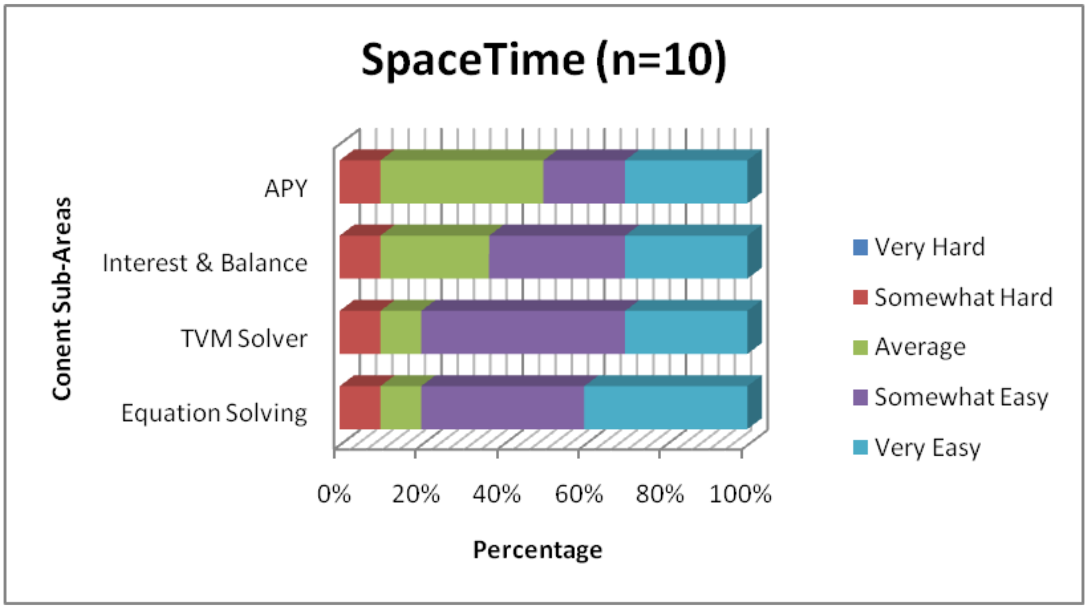
Table 8: Inferential Statistics (Finance Sub-area)

equality of variances, and the t-test statistics. The top table on the previous page (Table 7) contains the descriptive statistics, while the bottom table (Table 8) contains the results of the test.

Upon examination of the results, we see that all the null hypotheses are accepted, but due to small sample sizes (n = 7, n = 10) confidence interval results do not tell us much especially in the case of solving equations usability. The TVM solver (finance command) usability almost failed the equal variances assumption. It would not have been entirely unexpected for the TI-84 to demonstrate a statistically significant mean difference in this area given how robust and user friendly their TVM solver is as compared to the SpaceTime® Finance() command which is argument driven.

Examining the correlation matrix comparing the general usability construct with the mathematical finance construct (see Table 4), we note an even stronger correlation between personal performance and ease of use as compared to the statistics result. Based on these results observed r = .615 (ease of use, solving equations) and r = .561 (personal interaction, interest earned) suggest moderate correlations between these construct items. This suggests we should revisit some of the elements of the finance construct and try to differentiate them from general usability items to improve our results.





Effects on Student Performance

The final set of tests conducted concern student performance data on two content exams (post-statistics and post-finance) as well as the comprehensive final exam. Once again a 2-sample t-test was employed to gauge the likelihood of any statistically significant difference in mean student performance. Student performance data is summarized in the table below (Table 7).

In every case the null hypothesis was accepted suggesting there is no statistically significant difference in mean student performance between the two platforms. Based on our sample sizes it is clear we can only be confident of differences in the mean on average of size 20, or two letter grades. While this is comforting in the sense there is not a substantial difference between the two platforms, further study is warranted to reduce the size of the confidence interval to at most a single letter grade. To detect significant differences ($\alpha = 0.05$) in mean size within a letter grade ($\Delta\mu = 10$) would require approximately

$$n = \frac{16\cdot(507.75)}{10} \approx 813 \text{ or } n = \frac{16\cdot(544.41)}{10} \approx 688$$

813 samples between the two groups for the finance result and 688 samples between the two for the statistics result.

Student Performance Data		N	Mean	Std. Deviation	Std. Error Mean
Post – Statistics	TI-84	16	78.28	15.905	3.976
	SpaceTime	24	71.21	23.333	4.763
Post – Finance	TI-84	16	73.43	19.890	4.973
	SpaceTime	24	63.60	23.242	4.744
Final Exam	TI-84	16	73.50	12.972	3.243
	SpaceTime	24	66.33	9.993	2.039

Table 9: Descriptive Statistics (Student Performance)

		F-test for equality of variances		t-test for Equality of Means						
		F*	p	t	Df	p	Mean Difference	Std. Error Difference	95% CI	
									Lower	Upper
Post – Statistics	Equal Variances	2.152	0.151	1.0576	38	0.297	7.073	6.688	-5.488	19.633
	Unequal Variances	429.367		1.139	37.96	0.261	7.073	6.210	-6.466	20.612
Post – Finance	Equal Variances	1.365	0.249	1.385	38	0.173	9.831	7.098	-4.530	24.192
	Unequal Variances	507.752		1.430	35.53	0.161	9.831	6.875	-4.114	23.776
Final Exam	Equal Variances	1.685	0.202	1.977	38	0.055	7.167	3.625	-0.171	14.505
	Unequal Variances	133.165		1.874	26.39	0.072	7.167	3.824	-0.689	15.02

*F-test performed for equal variances, pooled estimated variance reported for unequal case

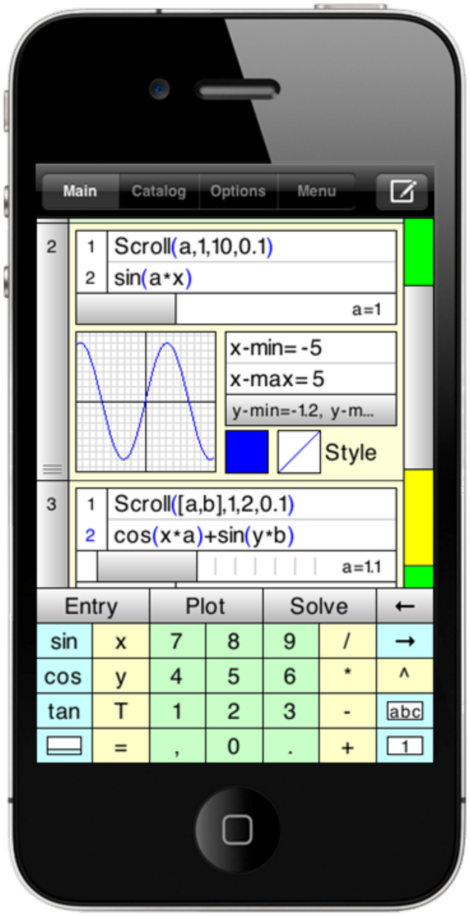
Table 10: Inferential Statistics (Student Performance)

Discussion of Results

What do the results of this study tell us? From a purely empirical standpoint the results give us hope that the mobile platform for computing is a viable alternative to the traditional handheld calculator. In only one of the tests did we see evidence of a potential difference in mean performance or usability—that belonging to data entry. In this section we will consider potential reasons for this observation as well as discuss possible interventions classroom teachers can make within the mathematics TPACK framework.

Data Entry Discrepancy

One of the major limitations of designing a user interface for a calculator on the mobile platform is screen real estate. For example, the iPhone 4® supports 960 x 640 resolution. When considering the SpaceTime® interface (see Figure 1), only a fraction of the screen is devoted to a keypad. The responsiveness of the multi-touch display and the ease with which keystroke errors can be made was reported as "a significant frustration" among some students. We believe the difference in student responses observed for the data entry item in the statistics usability construct is directly attributable to this issue. Several student were unaware of the setting within the SpaceTime® options menu which allowed them to increase the "button pad size". Utilizing this option or switching to landscape mode proved a suitable solution for many students.



1 Scroll(a,1,10,0.1)
2 sin(a*x)
a=1
x-min= -5
x-max= 5
y-min=-1.2, y-m...
1 Scroll([a,b],1,2,0.1)
2 cos(x*a)+sin(y*b)
a=1.1

Figure 1: SpaceTime iPhone® Interface

From a teaching perspective, utilizing the mobile platform to its fullest potential actually provides a way to alleviate the data entry problem altogether, while also allowing students to manipulate larger, more meaningful data sets. The SpaceTime® app allows educators to make SpaceTime state files, preloaded with data sets already saved to variables. These files can be downloaded wirelessly from the university file server, or downloaded as an attachment sent through email. The ability to quickly and dynamically change between large data sets without manually entering the data is a nice functionality and a potential intervention for the "data entry problem." It should be noted that similar avenues exist to load data sets into the TI-84 series of calculator, but more often than not these require linking calculators in class (time consuming), or necessitate that students load the sets before coming to class.

Pedagogical Interventions

In applying an integrated approach to technology use, Lee and Hollebrands (Lee, 2008) describe technology tools as being in one of two categories: *amplifiers* and *reorganizers*. A technology which amplifies student performance allows students to implement techniques and make connections more quickly—in other words, accelerate the learning curve. Reorganizers force students to reassess the structure of a concept and how it is linked to other problems or concepts. We will organize our discussion of the pedagogical interventions we plan on making after observing student work during the case study under these two categories.

Amplifiers

One of the most significant "amplifier" properties of the mobile platform is the multi-touch display. One of the more difficult things to teach a student to do on traditional handheld calculators is manipulating the window of a graph to see all of the important behaviors. With a multi-touch display a simple pinch can zoom the graph out, or a swipe re-center the graph, while spreading the fingers apart can allow a student to zoom in. Couple this with color displays and higher resolutions and graphing on the mobile platform is a potentially more intuitive and rewarding experience for students.

One of the few areas tested where the mean score for the SpaceTime® platform exceeded the traditional platform was in the application of linear regression techniques. I believe this is primarily a function of expedited learning due to the ability to interact with scatter plots and lines of best fit via multi-touch inputs. Plotting a scatter plot and related linear regression equation in the TI-84 involves several different menus located across multiple distinct and somewhat unrelated areas. In the SpaceTime® app all concepts related to regression (correlation, statistical plots, regression type) or accessible within a single command.

Students frequently reported the SpaceTime® catalog as being their favorite feature of the app. The SpaceTime® catalog is a list (not unlike the TI-84 catalog) of all available commands that can be issued, but unlike the TI-84 calculator each command has an associated entry describing the arguments of the command and examples of their use. In most cases, students were able to quickly pick up commands allowing them to move beyond basic applications at an accelerated pace.

Not every feature of the SpaceTime® app was well received though and some posed obstacles, or *de-amplifiers* if you will. One example of this were the limitations of the tabular environment associated with graphing multiple functions. With the TI-84 you can graph multiple functions and compare their results on discrete intervals within the table menu. The SpaceTime® app did not support table output for multiple tables (only a single function was viewable at a time). This severely limits the student ability to compare function values at important domain points.

Omissions of this type highlight the distinction between functional design and pedagogical design, the latter of which is a primary reason for the widespread adoption of the TI-84 among educators today. If the mobile computing platform is to improve upon the current generation of handheld calculators the design must echo both functional concerns as well as pedagogical ones.

Reorganizers

One of the main ways in which the mobile platform serves as a reorganizer is in its ability to integrate a wide range of resources including educational screen casts, student response systems, or just simple storage of assets (pictures, documents, video, etc…) related to a specific concept. These assets can be called on anytime, anywhere to reinforce a concept.

It's not entirely clear based on our results that students are more likely to use their iPhone® as a calculator outside of class than their handheld, but we did observe a larger percentage of students bringing their mobile devices to class. We believe this translates to an increase in calculator use "outside" class, but future research needs to be done to try and quantify this difference.

Within the classroom setting the role of technology as a reorganizer is highly teacher dependent. The MobileCAS® capabilities and scripting interface allow a teacher great flexibility in designing alternative investigations within the flow of a discussion. For example, within a group setting two students could work a problem by hand (algebraically with associated diagrams, models etc…) and film their process using their iPhone®. The other two students in the group could work the problem clearly documenting each step within a separate SpaceTime® entry. When both groups are finished they can share their product with the other group and comment on the similarities and differences within each group allowing students to make connections they might have not otherwise made.

Conclusion

The purpose of this study was to outline findings of any potential impact a mobile computing app might have on student performance and perceptions of usability in a general education mathematics course. We developed and evaluated the mobile learning usability scale consisting of three constructs: (1) general usability, (2) statistics usability, and (3) finance usability as a response to the literature review's call for a characterization of “good usability”. All items showed very good to acceptable internal reliability and sufficient evidence of weak inter-correlation to assume the constructs assess disparate concepts.

Mean student perceptions of usability were compared across computing platforms and found to not be statistically different in all but one case, data entry. We have given observed reasons for this discrepancy and provided pedagogical interventions an educator might employ to offset this deficiency. Finally, student performance data was collected from content area post-tests and a comprehensive final exam. No statistically significant difference in student performance was observed, though we acknowledge the need to conduct further research to narrow the interval over which we could detect potential differences.

The design of this study appears to be reliable and holds potential to produce exciting results. Nonetheless, there are several limitations of the study which should be addressed in future research. First, the research should be conducted across multiple classes to produce a combined sample size of near 700 to obtain better statistical results. Second, future research needs to clarify the extent to which informal learning (learning outside the class) takes place as a result of platform choice. Third, the mobile learning usability scale should be refined to further separate the finance construct from the general usability construct, and items within each construct should be narrowed in focus. The potential to integrate new constructs for other content sub-areas should be explored and construct validity assessed. Finally, methodological issues associated with survey design, and student reported data should be addressed.

As higher education institutions continue to integrate mobile devices the need to assess the usability of these platforms in various contexts is a necessary component of future research. Identifying the extent to which the mobile computing platform enhances student performance and use is a critical component to improving the design and quality of instruction experienced by students not only in mathematics but in the larger university setting. Additionally investigations such as this provide the impetus for mobile design to reflect pedagogical practice giving educators the best of both worlds.

References

1. Bebell, D., O’Dwyer, L. M., Russell, M., & Hoffmann, T. (2010). “Concerns, considerations, and new ideas for data collection and research in educational technology studies.” Journal of Research on Technology in Education, 43(1), 29-52.
2. Chen-Chung, L., Chen-Wei, C., & Nian-Shing, C., (2009). “Analysis of peer interaction in learning activities with personal handhelds and shared displays.” Journal of Educational Technology and Society, 12(3), 127-142.
3. Cheung & Hew (2009). “A review of research methodologies used in studies on mobile handheld devices in K-12 and higher education settings.” Australasian Journal of Educational Technology, 25(2), p. 153-183.
4. Clark-Wilson A. (2010). “Connecting mathematics in a connected classroom: Teachers emergent practices within a collaborative learning environment.” The Mathematics Centre-University of Chichester.
5. Ellington, A.J. (2003). “A meta-analysis of the effects of calculators on students’ achievement and attitude levels in precollege mathematics classes.” Journal for Research in Mathematics Education 34(5): 433-463.
6. Forster, P. A., (2006). “Assessing technology based approaches for teaching and learning mathematics.” International Journal of Mathematical Education in Science and Technology, 37(2), p. 145-164.
7. Groth, R., Spickler, D., Bergner, J. & Bardzell, M. (2009). “A qualitative approach to assessing technological pedagogical content knowledge.” Contemporary Issues in Technology and Teacher Education, 9(4), 392-411.
8. Kukulska-Hulme, A. (2007). “Mobile Usability in Educational Contexts: What have we learned?” The International Review of Reasearch in Open and Distance Learning, 8(2).
9. Lauricella, S. & Kay, R. (2010), “Assessing laptop use in higher education classrooms: The laptop effectiveness scale (LES).” Australasian Journal of Educational Technology, 26(2), 151-163.
10. Lee, H. and Hollebrands, K. (2008). “Preparing to teach mathematics with technology: an integrated approach to developing technological pedagogical content knowledge.” Contemporary Issues in Technology and Teacher Education, 8(4), 326-341.
11. Liu, G., Hwang, G. (2010). “A key step to understanding paradigm shifts in eLearning: Towards context aware ubiquitous learning.” British Journal of Educational Technology, 41(2), p. E1-E9.
12. Looi, C., Seow, P., Zhang, B., So, H., Chen, W., Wong, L., (2010). “Leveraging mobile technology for sustainable seamless learning: a research agenda.” British Journal of Technology, 41(2), p. 154-169.
13. Mayrath, M., Nihalani, P., & Perkins, S. (2011). “Digital texts and the future of education: Why books?” Educause Quarterly, 34(1).
14. Merriam, S. (1998). “Qualitative research and case study applications in education.” Jossey-Bass, San Francisco.
15. Mishra, P., and Koehler, M. J. (2006). “Technological and pedagogical content knowledge: A new framework for teacher knowledge.” Teachers College Record. 108(6), 1017-1054.
16. National Council of Teachers of Mathematics, Curriculum Standards for Grades 9-12, http://standards.nctm.org/Previous/CurrEvStds/currstand9-12.htm.

17. Niess, M. L., Ronau, R. N., Shafer, K.G., Driskell, S.O., Harper, S. R., Johnston, C., Browning, C., Ozgun-Koca, S.A., & Kersaint, G. (2009). “Mathematics teacher TPACK standards and development model.” Contemporary Issues in Technology and Teacher Education, 9(1), 4-24.
18. Quinn, C. (2000). “Mobile, wireless, in your pocket learning.” Linezine, Fall 2000.
19. Schou, S. B. (2007). “A study of student attitudes and performance in an online introductory business statistics course.” Electronic Journal for the Integration of Technology in Education, 6(5), 71-78.
20. Shen, R., Minjuan, W., Wanping, G., Novak, D., & Tang, L., (2009). “Mobile learning in a large blended computer science classroom: System function, pedagogies, and their impact on learning.” IEEE Transactions on Education, 52(4), 538-546.
21. Wang, M., Shen, R., Novak, & D., Pan, X., (2009). “The impact of mobile learning on students’ learning behaviors and performance: Report from a large blended classroom.” British Journal of Educational Technology, 40(4), 673-695.