AC 2012-4170: A FORMAL ASSESSMENT OF THE HAPTIC PADDLE LABORATORIES IN TEACHING SYSTEM DYNAMICS

Jenna L. Gorlewicz, Vanderbilt University

Jenna L. Gorlewicz received her B.S. in mechanical engineering from Southern Illinois University, Edwardsville, in 2008. She is currently in the forth year of her Ph.D. work in mechanical engineering at Vanderbilt University, and she is a National Science Foundation Graduate Research Fellow. Her current research interests are in developing novel devices and methods for engineering education, including haptic touch screen interfaces to help teach graphical mathematics concepts to blind children.

Dr. Robert James Webster III, Vanderbilt University

Robert J. Webster III received his B.S. degree in electrical engineering from Clemson University in 2002 and the M.S. and Ph.D. degrees in mechanical engineering from the Johns Hopkins University in 2004 and 2007, respectively. In 2008, he joined the faculty of Vanderbilt University as an Assistant Professor of mechanical engineering, where he currently directs the Medical & Electromechanical Design Laboratory. His current research interests include medical robotics, image-guided surgery, continuum robotics, and engineering education. Webster received the NSF CAREER Award in 2011, and the IEEE Volz award for Ph.D. thesis impact in 2011.

A Formal Assessment of the Haptic Paddle Laboratories in Teaching System Dynamics

Abstract

In this paper, we present a formal assessment of the haptic paddle, a one degree-of-freedom force-feedback joystick that has been adopted at several universities as a teaching tool in System Dynamics. System Dynamics is a core mechanical engineering undergraduate course that teaches students to combine knowledge from physics and differential equations to model dynamic systems in several domains. Haptic paddles are useful for hands-on modeling and control laboratories and enables students to physically interact with simulated dynamic systems via their sense of touch. Prior qualitative assessments of haptic paddles have shown that both students and educators believe them to be beneficial in conveying the core concepts of the course and that they increase student enthusiasm for learning the material. Toward quantifying these perceived benefits, we have conducted a systematic assessment of learning objectives associated with haptic paddle laboratories. To do this, we developed a multiple choice quiz consisting of five questions for each of the five sequential laboratories during the semester. To establish if and when students learned the concepts associated with our learning objectives for each lab, we randomized presentation time of the lab quiz to each of our four sections, with the following options: (1) at the beginning of a lab session, (2) after a pre-lab lecture, (3) immediately after completion of the lab, and (4) after completion of the lab report. This assessment architecture enables us to determine whether learning happened in lecture, in the lab itself, or during subsequent reflection on laboratory results during the process of writing the lab report.

1 Introduction

A common challenge in engineering education is to develop students' intuitive understanding of how physical systems behave, despite the fact that many students have never physically observed or interacted with the systems they are learning about. A variety of approaches have previously been developed to address this, including implementation of hands-on demonstrations[1–3], computer simulations[4–6], and design projects[7,8]. In this paper, we are particularly interested in System Dynamics, a core mechanical engineering course where students learn to model and control systems in multiple domains. Prior results for this course, aimed at developing student intuition through hands-on examples, have centered around laboratories that feature the haptic paddle, a low-cost device that enables students to "feel" dynamic systems by interacting with computer simulations via force feedback. The haptic paddle is a one degree-of-freedom haptic interface developed in the late 1990s at Stanford University[9]. In the sequential laboratory format developed with the haptic paddles, students analyze, construct, calibrate, model, and interact with the paddle. In addition to enabling direct use of many of the concepts taught in the course, the haptic paddle engages students through several sensory modalities, appealing to those who are audio, visual, and kinesthetic learners[9]. Qualitative assessments of the laboratories have shown that students respond enthusiastically to the haptic paddle, and it has been observed by laboratory instructors that many were developing a true understanding of core course concepts for the first time[9].

In view of the perceived benefits of the haptic paddle, it has rapidly been adopted as an educational tool in System Dynamics courses[10] at universities including Johns Hopkins[9], Rice[11], Michigan[12], Vanderbilt[13], and Utah (see[14] for more information). Since its inception, the paddle has evolved in both hardware and software, as well as mechanical design, as shown in Figure 1. None of these modifications have fundamentally altered what the haptic paddle is, or what it is able to accomplish in the classroom, and most have been aimed at increasing robustness, reducing costs, using readily available hardware, etc. The laboratory curricula implemented at Stanford, Johns Hopkins, and Vanderbilt in particular, have been extremely similar, with the curricula at other universities sharing some aspects of the original Stanford labs while also incorporating new concepts. Assessment of the learning effectiveness of the haptic paddle and associated labs has, to date, been qualitative in nature. Thus, the purpose of this paper is to complement these previous qualitative results with a formal (that is, data-driven) assessment of the haptic paddle laboratories as implemented at Vanderbilt. The objective is to determine if and when students learn key concepts, whether in lecture, in the lab activities themselves, or after reflecting on the lab activities while writing their lab reports.

This paper is organized as follows. Section 2 briefly describes the Vanderbilt Haptic Paddle and the laboratories associated with it. The assessment methods and statistical analyses used are discussed in Section 3, and the results are presented in Section 4. Section 5 contains a discussion of the results as well as future planned improvements for the laboratories based on the results, with conclusions in Section 6.

2 Overview of the Vanderbilt Haptic Paddle Hardware and Laboratory Assignments

At Vanderbilt, we have contributed to the evolution of the haptic paddle through mechanical, electrical, and software changes. Initially, we used a paddle similar to that used at Stanford and Johns Hopkins, which was a cable-driven capstan that ran on C-executable files[9]. The most recent version of our setup is shown in Figure 2. The mechanical design of our paddle is similar to that of Rice[11] and Utah[14], in that the motor is at the bottom, with the capstan and handle consisting of a single piece, unlike the original Stanford design[9]. A key difference in our paddle is that we have replaced the capstan drive with a friction drive, in view of enhancing student usability (we

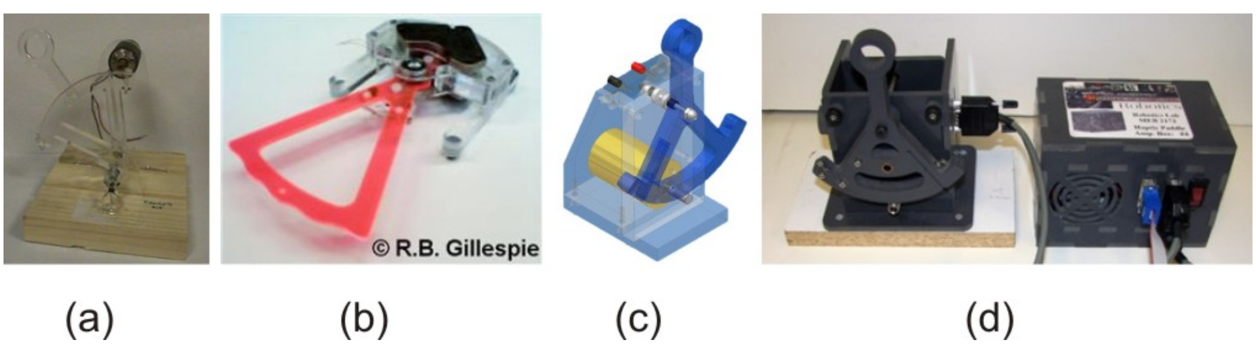

Figure 1: (a) The Stanford and Johns Hopkins Haptic Paddle. (b) The University of Michigan Haptic Paddle. (c) The Rice University Haptic Paddle. (d) The University of Utah Haptic Paddle.

found that the capstan drives were difficult for undergraduates to wind effectively and that the cable can be prone to falling off the spool, leading to student frustration and lab delays). The friction drive trades off some haptic fidelity for robustness and usability, but we have qualitatively observed no reduction in learning benefit with the friction drive design. A second important change we have made is replacing internal PCI cards for D/A with the low-cost, USB-connected Arduino microcontroller, which reduces the cost of setting up a haptic paddle by hundreds of dollars (in previous work, it was assumed that a D/A solution was already owned by the person implementing the haptic paddle). Lastly, we have developed software to run the paddles based on Matlab and Simulink (The MathWorks, Inc.). Together, the above changes have reduced the complexity and cost of the entire system while also providing students with a flexible software interface through which they can quickly develop real-time models and interface them with their haptic paddle.

The focus of this paper is a formal assessment of the haptic paddle laboratories that employ these devices. The lab is associated with our System Dynamics course, a required mechanical engineering course typically taken in the first semester of a student's 3rd year. The course consists of a standard lecture component and five laboratory sessions roughly evenly distributed over the semester. In the laboratory sessions, the class is divided into four sections, each containing approximately 15-20 students. Each lab section meets for three hours on a different day of the week. Within each section, teams of 2-3 students are formed to work together on the labs, as shown in Figure 3. These teams are chosen by the students, and they remain in roughly the same groups for the duration of the semester.

Each of the five lab assignments (which follow the basic Stanford-Johns Hopkins format with some modifications) focuses on a different aspect of the haptic paddle, which relates to concepts covered in lecture. In the first lab, students become acquainted with Simulink by building a simple model of a first order system. Then, students perform a motor spin down test and compare their

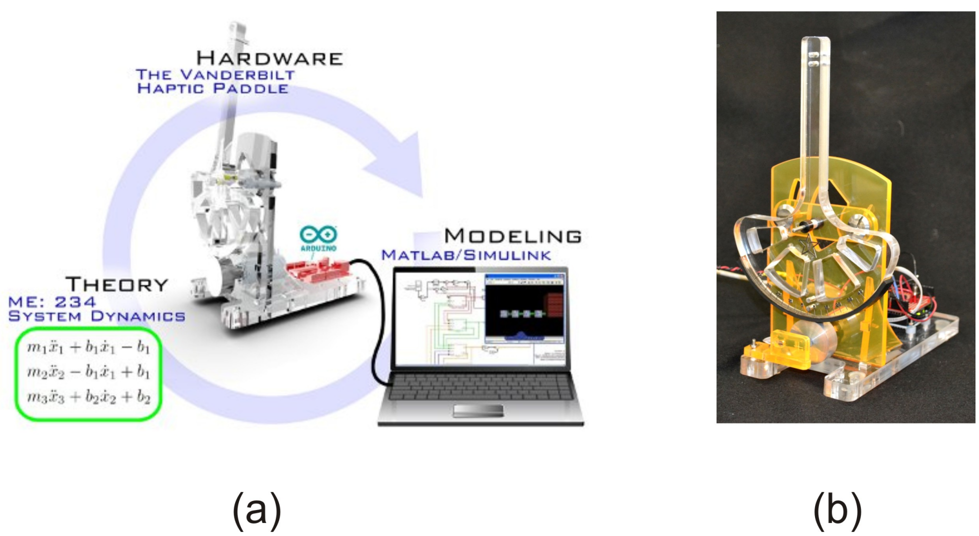

Figure 2: (a) A schematic of the components of the Vanderbilt Haptic Paddle. (b) The assembled Vanderbilt Haptic Paddle. Our paddle relies on a friction drive design, runs in Matlab/Simulink, and uses a low-cost Arduino microcontroller for communication.

experimental results with predicted results they obtain from their Simulink simulation of their motor. In the second lab, students analyze the paddle handle by measuring its moment of inertia through a bifilar pendulum experiment, and they experimentally measure the torque constant and Coulomb friction in the motor. In the third lab, students interface the Arduino microcontroller with Simulink to verify encoder readings of the motor and obtain a relationship between encoder counts and angle of rotation of the paddle. Then, they set up an experiment to measure the equivalent mass and damping of the assembled paddle by attaching two stiff springs to the paddle, making it behave as a second order underdamped system. In the fourth lab, students investigate feedback control by observing the paddle's step response while adjusting PD control gains. In the fifth lab, students interact with a multiple degree of freedom mass-spring-damper system and explore its modes of vibration using the paddle and a real-time 3D visualization of the system.

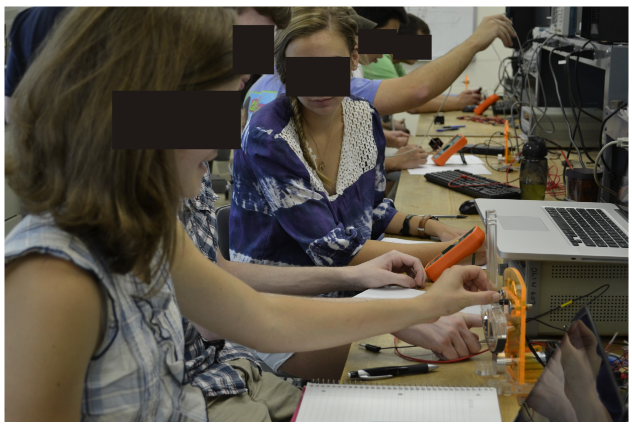

Figure 3: Small teams of 2-3 students interacting with the haptic paddle during a lab activity.

3 Formal Assessment Methods

The core concepts of the lecture and the lab exercises were used to construct a 25-question multiple choice quiz (5 questions/lab) which was used to assess student learning. Each question had 4 possible answers, one of which was correct. The three remaining incorrect choices were chosen to include common wrong answers or misconceptions that students often have. In a few questions, we asked students to "Choose all of the answers that apply," instead of selecting just one.

This 25-question assessment was administered at the beginning of the semester to all students in order to assess their initial understanding of the material and to provide a baseline measurement for statistical analyses. The assessment was then broken down into 5 quizzes, each containing 5 questions pertinent to one of the labs. We administered the appropriate 5-question quiz to each of our four student lab sections at different time points throughout the lab session in order to assess if and when the students were learning the material. We randomized the presentation of the quiz among the four sections at (1) the beginning of the lab session, (2) after a pre-lab lecture, (3) after completing the lab, or (4) after completing the lab report (typically 1-2 weeks after completing the lab), as shown in Table 1. Though the timing of the lab quiz differed between student sections, the same lab quiz was administered to each. Using this approach, each student section took one quiz for each lab, varying only by the time point at which they took it. The time at which the quizzes were administered to each section was systematically rotated to remove any potential bias in data collection. Students were given 5 points toward each of their lab report grades for completing the lab quizzes, and thus, were not given a score based on their performance on the quiz.

Table 1: Quiz placement for each lab for each student section (SX).

Placement	**Lab 1**	**Lab 2**	**Lab 3**	**Lab 4**	**Lab 5**
Beginning	S1	S4	S3	S2	S1
After Pre-Lecture	S2	S1	S4	S3	S2
After Lab	S3	S2	S1	S4	S3
After Lab Report	S4	S3	S2	S1	S4

Below, we will present two years of data collected from the assessments, with $N_1 = 63$ students and $N_2 = 71$ students. We note that appropriate IRB approval was obtained for this study. In year 1 (Y1), we used the original version of the haptic paddle (similar to Johns Hopkins University) and its C-executables, and in year 2 (Y2), we used the inverted paddle design with a cable drive (similar to Rice University) and Matlab and Simulink software. The friction drive is a new innovation that is being used during the current academic year, and quiz data is not yet available for it. The lab content was the same for both Y1 and Y2. Students were given 1 point for every correct answer and 0 points for every incorrect answer. Means and standard deviations were determined for each of the students' quizzes.

3.1 Research Questions

The research questions we were interested in answering in analyzing our assessments are as follows:

1. Overall, did the students learn the core course concepts during the semester? Statistically, we were interested in determining if there was a significant increase in mean quiz score from the beginning to the end of the semester.
2. Did the labs increase student understanding of the course material? Statistically, we were interested in determining if there was a significant increase in mean quiz score from the beginning of the semester to after completing the lab.
3. When did the students learn the material? Statistically, we were interested in determining if there were any significant differences between mean quiz scores from the beginning of the semester to any of the time points at which the quizzes were administered.

In order to address the first question, paired t-tests were performed to compare the mean quiz score on the pre-test with the mean quiz score of the post-test. To assess the value of the labs, we performed paired t-tests to compare the mean quiz score on the pre-test with the mean quiz score of the appropriate student section after completing the lab. Finally, to assess *when* student learning was occurring, we performed paired t-tests comparing the mean quiz score of the pre-test to the mean quiz score of the appropriate student section at various time points for each lab. Note that all analyses consist of pairwise comparisons, in order to not compare across student sections and implicitly assume that each student section is equivalent at every time point throughout the lab. This was done to ensure valid interpretations of our results. All statistical analyses were performed in R 2.11.1, and the results are presented in Section 4.

3.2 Verification of Normality and Comparable Student Sections

Before performing the above statistical analyses, we sought to verify three things: (1) Normality of our data, (2) No significant difference between student sections' initial cumulative pre-test scores, and (3) No significant difference between student sections' initial pre-test scores *for each lab*. First, we assessed the normality of each student section's data for each year. To do this, we created a quantile-quantile (Q-Q) plot for each student section, including both pre-test scores and lab quiz scores, resulting in a total of 8 plots. For simplicity, only one of the plots is shown in Figure 4, however, it is representative of the other plots obtained. Because this plot suggests a linear trend, we can infer that our data is approximately normally distributed, and thus, parametric statistical tests, such as the t-test, are applicable in our subsequent analyses.

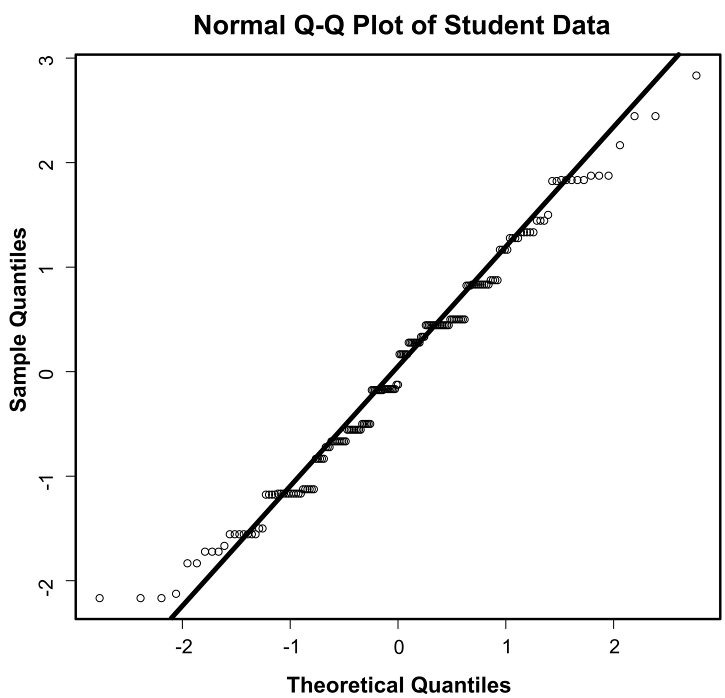

Figure 4: A Q-Q plot from one student section assessing the normality of our data. The linearity of this plot suggests that the data follows an approximate normal distribution. Q-Q plots were created for each student section for each year, for a total of 8 plots. The plot shown is representative of all 8 plots, and thus we can infer that each student section followed am approximate normal distribution.

Because we use multiple student sections' quiz scores within our analyses, we also need to ensure that all student sections were comparable in their initial cumulative understanding of the course material and in their initial understanding of the material *for each lab*. To do this, we first compared the mean cumulative pre-test score (all 25 questions) of each student section with the other 3 student sections using a two-sample t-test with unequal variances. Then, we separated the 25-question pre-test up into 5 parts, corresponding with the 5 lab quizzes, and compared the mean quiz scores on each part between each student section using a two-sample t-test with unequal variances. The null hypothesis for all tests was that no difference in mean pre-test score existed between any two sections. From the Y1 data, we observed a significant difference between student section 1 and student section 2 (p-value = 0.04) in their *cumulative* pre-test score, but found no significant differences at the 95% confidence level ($\alpha = 0.05$) between student sections on individual parts of the pre-test. For this reason, we only omit the cumulative pre-test scores of student sections 1 and 2 in appropriate subsequent analyses. From the Y2 data, we observed no significant differences at the 95% confidence level ($\alpha = 0.05$) between any student sections' *cumulative* pre-test score, and thus, we include this data in the appropriate analyses in Section 4. When looking at individual parts of the pre-test, however, we did find a significant difference between section 1 and section 4 on the Lab 5 portion of the pre-test (p-value = 0.04), with section 4 having a significantly lower average on this portion of the material. Because of this, student section 4's data was omitted in the Lab 5 analyses for Y2.

4 Results

4.1 Educational Benefit from Course

The first question we sought to answer was whether or not the students learned and retained the course concepts after completing the course. This question allows us to generally assess if the combination of learning opportunities we are providing (lectures, homework assignments, labs, lab reports) is beneficial for students. To address this question, we performed a paired t-test comparing all students' cumulative mean score on the pre-test with their cumulative mean score on the post-test. We also separated the pre-test and the post-test into 5 parts (corresponding with the lab quizzes), and performed paired t-tests comparing all students' mean quiz score on one part of the pre-test with their mean quiz score on that same part of the post-test. This latter analysis allowed us to observe the cumulative learning of portions of the course material, in order to pinpoint which areas appear more difficult for students to grasp and may benefit from more emphasis in the future. Note that post-test data was only available for Y2, and thus no data is presented from Y1. Further, because the Lab 5 pre-test data of section 1 and section 4 from Y2 showed a significant difference, we also omit section 4's data in the Lab 5 analysis of the separated pre-test and post-test. Only section 4's data was omitted because there were no significant differences between any combination of sections 1, 2 and 3's scores on the Lab 5 portion of the pre-test. The results of the above two analyses are shown in Table 2 below. These results, which will be discussed in more detail in Section 5, suggest that students learned and retained majority of the course material by performing significantly better on their cumulative quiz score and on 3 of the 5 lab quizzes at the end of the semester.

4.2 Educational Benefit from Lab

The second question we addressed was if students increased their conceptual understanding after having participated in the lab. To assess this, we performed a paired t-test comparing the mean quiz score obtained after completing the lab with the mean quiz score obtained on the corresponding section of the pre-test for each lab. For this analysis, we compared student section 3's

Table 2: The means (standard deviations) of all students' pre-test score compared with their post-test score for year 2. The corresponding p-value from the paired t-test is shown, denoting significance at $\alpha = 0.05$ with a ** and at $\alpha = 0.1$ with a *. The sample size for each test, N, is shown in the last column.

	Pre-Test	**Post-Test**	**p-value**	**N**
Lab 1	2.49 (1.14)	2.87 (1.06)	0.1212	39
Lab 2	2.90 (1.07)	3.08 (1.20)	0.4132	39
Lab 3	2.05 (1.00)	3.64 (1.01)	1.393e-10**	39
Lab 4	2.13 (1.30)	3.72 (1.32)	9.138e-07**	39
Lab 5	2.80 (1.06)	3.43 (0.77)	0.0140**	30
Cumulative	12.26 (2.94)	16.64 (3.27)	1.913e-11**	39

scores from pre-test to after lab for Lab 1, student section 2's scores from pre-test to after lab for Lab 2, and so on, as shown in Table 1. The results are shown in Table 3. These results, which will be discussed in more depth in Section 5, suggest that all of the lab activities were beneficial, as students achieved significantly higher quiz scores after lab, for each of the labs, in at least one of the two years of data collected.

Table 3: The means (standard deviations) of the appropriate student section's quiz score on the pre-test compared with the quiz score taken after completing the lab for year 1 (Y1) and year 2 (Y2). The corresponding p-value from the paired t-test is shown, denoting significance at $\alpha = 0.05$ with a ** and at $\alpha = 0.1$ with a *. The sample size for each test, N, is shown in the last column.

	Pre-Test	**After Lab**	**p-value**	**N**
Lab 1 (Y1)	2.39 (1.04)	2.72 (1.02)	0.3808	18
Lab 1 (Y2)	2.72 (1.02)	3.5 (0.79)	0.0116**	18
Lab 2 (Y1)	1.86 (1.61)	4.64 (0.63)	1.223e-5**	14
Lab 2 (Y2)	3.06 (0.93)	3.50 (0.89)	0.1862	16
Lab 3 (Y1)	2.07 (1.14)	4.00 (0.78)	0.0011**	14
Lab 3 (Y2)	2.17 (0.94)	3.83 (0.94)	6.603e-4**	12
Lab 4 (Y1)	3.00 (1.30)	4.07 (0.92)	0.0295**	14
Lab 4 (Y2)	2.00 (1.32)	2.67 (1.73)	0.3856	9
Lab 5 (Y1)	2.94 (1.64)	4.06 (0.75)	0.02302**	17
Lab 5 (Y2)	2.53 (0.94)	3.18 (0.95)	0.0686*	17

4.3 Educational Benefit of Other Learning Opportunities

The last question we sought to answer in this study addressed if and when students were learning the material. In other words, we wished to pinpoint at what stage(s) learning was occurring. In order to assess this, we analyzed the components of the learning process in a separate, but cumulative fashion.

To assess the value of the in-class lecture, we conducted a paired t-test with unequal variances

Table 4: The means (standard deviations) of the appropriate student section's quiz score on the pre-test compared with the quiz score taken at the beginning of lab for year 1 (Y1) and year 2 (Y2). The corresponding p-value from the paired t-test is shown, denoting significance at $\alpha = 0.05$ with a ** and at $\alpha = 0.1$ with a *. The sample size for each test, N, is shown in the last column.

	Pre-Test	**Beginning of Lab**	**p-value**	**N**
Lab 1 (Y1)	2.33 (1.29)	2.73 (0.96)	0.3200	15
Lab 1 (Y2)	2.06 (1.16)	2.00 (0.93)	0.7921	15
Lab 2 (Y1)	2.92 (1.49)	2.21 (1.31)	0.0354**	14
Lab 2 (Y2)	2.79 (1.19)	3.0 (1.47)	0.6198	14
Lab 3 (Y1)	2.28 (1.02)	3.17 (1.25)	0.0054**	18
Lab 3 (Y2)	1.67 (0.97)	3.17 (1.29)	0.0013**	18
Lab 4 (Y1)	2.63 (1.15)	3.94 (1.00)	0.0031**	16
Lab 4 (Y2)	1.73 (1.33)	3.67 (0.82)	1.815e-4**	15
Lab 5 (Y1)	3.00 (1.00)	3.47 (1.06)	0.1689	15
Lab 5 (Y2)	3.31 (1.11)	3.62 (0.77)	0.3925	13

to compare the mean quiz score from the appropriate part of the pre-test to the mean quiz score from the student section who took the quiz at the very beginning of the lab. The null hypothesis was that no difference in means existed. From Table 1, the pertinent data used in this analysis was student section 1's scores for Lab 1, student section 4's score for Lab 2, and so on. The results, which will be discussed in detail in Section 5, are shown in Table 4, and suggest that the lecture was beneficial for the concepts covered in Labs 3 and 4.

We then assessed the value of the in-class lecture and the pre-lab lecture combined by comparing the mean quiz score on the appropriate part of the pre-test with the mean quiz score from the student section who took the quiz after the pre-lab lecture, but before the lab activity. We did this using a paired t-test with unequal variances, and the null hypothesis was that no difference in means existed. From Table 1, the pertinent data used in this analysis was student section 2's scores for Lab 1, student section 1's score for Lab 2, and so on. The results are shown in Table 5 and will be discussed in more detail in Section 5. They suggest that the pre-lab lecture is a beneficial component of the lab, enhancing students immediate recall of the material in all of the labs.

Because the value of the lectures and the lab combined are presented above in Table 3, we omit them here. Finally, to assess the values of all learning components (lectures, lab, and lab report), we conducted a paired t-test with unequal variances to compare the mean quiz score from the pre-test to after the lab report. The null hypothesis was that no difference in means existed. From Table 1, the pertinent data used in this analysis was student section 4's scores for Lab 1, student section 3's score for Lab 2, and so on. The results, also discussed in more detail in Section 5, are shown in Table 6 and suggest that the lab report is also a necessary component of the lab activity, as it enabled students to apply and retain the material they learned in four out of the five labs.

Table 5: The means (standard deviations) of the appropriate student section's quiz score on the pre-test compared with the quiz score taken after the pre-lab lecture for year 1 (Y1) and year 2 (Y2). The corresponding p-value from the paired t-test is shown, denoting significance at $\alpha = 0.05$ with a ** and at $\alpha = 0.1$ with a *. The sample size for each test, N, is shown in the last column.

	Pre-Test	After Pre-Lab Lecture	p-value	N
Lab 1 (Y1)	1.75 (1.18)	2.38 (1.09)	0.0859*	16
Lab 1 (Y2)	2.75 (0.86)	2.81 (1.05)	0.8489	16
Lab 2 (Y1)	2.53 (0.74)	4.33 (1.29)	1.688e-4**	15
Lab 2 (Y2)	2.79 (1.19)	3.43 (1.16)	0.0445**	14
Lab 3 (Y1)	2.07 (1.00)	3.71 (1.20)	3.395e-4**	14
Lab 3 (Y2)	2.00 (1.18)	3.07 (1.39)	7.552e-4**	14
Lab 4 (Y1)	2.50 (1.62)	4.11 (0.90)	0.0024**	18
Lab 4 (Y2)	2.31 (1.35)	3.13 (1.31)	0.0431**	16
Lab 5 (Y1)	2.5 (1.03)	4.63 (0.62)	1.149e-6**	16
Lab 5 (Y2)	2.64 (1.34)	3.79 (0.70)	0.0041**	14

5 Discussion

The data presented above suggests that the students learned many of the core course concepts throughout the semester. From Table 2, we observe that students achieved a significantly higher cumulative score on the post-test compared to the pre-test. This suggests that the learning opportunities provided to the students throughout the semester were successful in enhancing students' overall understanding of the course material. After looking at the pre-test and post-test scores separated by the material covered in each lab (see Table 2), we observe that students did significantly better on the quizzes focusing on concepts from Labs 3, 4, and 5. This suggests that students learned and retained these concepts throughout the duration of the course. Though quiz score increases are observed for Labs 1 and 2, there were no significant differences between the pre-test and post-test scores in these two cases.

For further insight into this result, we look at Table 3, which shows student performance immediately after completing the lab. We focus on the Y2 data since Table 2 contained only Y2 data. For Lab 1, it appears that students did learn the material after completing the lab, as there was a significant increase in quiz score from pre-test to after lab. This suggests that students learned the Lab 1 material in the short-term but did not retain it at the end of the semester. Note that some improvements were made to Lab 1 between Y1 and Y2, which may be the reason a significant increase in quiz score was observed in Y2 but not in Y1. These results, however, suggest that Lab 1 would benefit from further improvements to enhance student retention of the material. For Lab 2, we observe from Table 3, that there was not a significant increase in quiz score from pre-test to after lab for Y2, but there was a large significant increase for Y1. Some changes were made between the two years in the Lab 2 curriculum which may have contributed to this discrepancy, though the changes were primarily hardware and software rather than lab content. We note, however, that the pre-test score for Lab 2 (Y1) was much lower than the pre-test score for Lab 2 (Y2), suggesting that the students from Y2 had a better understanding of the Lab 2 material at the begin-

Table 6: The means (standard deviations) of the appropriate student section's quiz score on the pre-test compared with the quiz score taken after completing the lab report for year 1 (Y1) and year 2 (Y2). The corresponding p-value from the paired t-test is shown, denoting significance at $\alpha = 0.05$ with a ** and at $\alpha = 0.1$ with a *. The sample size for each test, N, is shown in the last column.

	Pre-Test	After Lab Report	p-value	N
Lab 1 (Y1)	2.36 (1.28)	2.07 (0.92)	0.4533	14
Lab 1 (Y2)	2.14 (1.03)	2.57 (0.94)	0.2896	14
Lab 2 (Y1)	2.61 (0.92)	3.94 (0.87)	1.812e-4**	18
Lab 2 (Y2)	2.56 (1.25)	3.44 (0.92)	0.0311**	18
Lab 3 (Y1)	2.25 (1.12)	3.81 (1.26)	1.754e-4**	16
Lab 3 (Y2)	1.93 (1.03)	3.60 (0.83)	2.443e-5**	15
Lab 4 (Y1)	3.00 (1.20)	3.73 (0.80)	0.0853*	15
Lab 4 (Y2)	2.54 (1.45)	3.62 (1.26)	7.230e-4**	13
Lab 5 (Y1)	2.71 (1.27)	4.07 (0.73)	0.0073**	14
Lab 5 (Y2)	2.33 (1.50)	3.00 (0.71)	0.1950	9

ning of the course compared to the students in Y1. Thus, part of the reason we may not observe a significant increase in Lab 2 (Y2), may be due to the fact that students already knew a large portion of the material initially. Nonetheless, the results suggest that Lab 2 could also be a focus for future improvements.

The most promising result we obtained from our assessments was that after completing the lab exercises, students achieved significantly higher quiz scores for all of the labs in at least one of the two years of data collected (see Table 3). Further, from Tables 2 and 3 combined, we observe that students achieved significantly higher quiz scores for Labs 3, 4, and 5 both after lab and on the post-test compared with the pre-test. This suggests that the lectures and labs are beneficial in enhancing student understanding **and retention** of the concepts associated with these labs. The one exception is the Lab 4 data from Y2, which does not show a significant increase in quiz score from pre-test to after lab. This result may be due in part to the fact that the initial pre-test score was quite low and that the sample size for this particular lab was relatively small due to several students switching lab sessions or not completing the quiz. Looking at Table 2, however, we do see a significant increase in Lab 4 in Y2 from pre-test to post-test, suggesting that majority of the students learned the Lab 4 material, perhaps benefiting especially from the lab report and lecture discussions following lab. Even with this one exception, however, the results presented both after lab and at the end of the semester suggest that the haptic paddle labs significantly enhance student understanding and retention of the core concepts in the course.

Finally, we sought to gain a better understanding of when student learning was occurring, by looking at the individual learning opportunities. From Table 4, which shows the value of the in-class lecture, we observe that students scored significantly higher on quizzes for Labs 3 and 4 even before participating in the lab itself. This suggests that the in-class lecture is particularly beneficial for the concepts associated with Labs 3 and 4. We also note that the Y1 students had a significant decrease in quiz score from pre-test to the beginning of lab for Lab 2. This suggests that students may have become confused by the in-class lecture on this material. In Y2, however, we see an increase in quiz score for this material, though it is not significant. These results suggest that the lecture itself was simply not enough in enhancing student understanding of the Lab 2 material and reiterate the need for additional learning opportunities outside of the lecture.

Before students perform a lab exercise, they are given a short pre-lab lecture, specifically addressing the learning objectives of the lab. These objectives correspond directly with the concepts covered in the quiz, such that if students paid close attention during this introductory lecture, they should know every answer on the quiz. Thus, this quiz primarily tested students' listening and recall skills. In order to assess the value of the in-class lecture and the pre-lab lecture, we compared pre-test scores with quiz scores taken after the pre-lab lecture but before completing the lab. From Table 5, we observe that students appeared to listen and benefit from these pre-lab lectures, as reflected in the significantly higher quiz scores in all of the labs after hearing the pre-lab lecture. Though we cannot directly decouple the in-class lecture from the pre-lab lecture in this analysis, we speculate that the pre-lab lecture had a significant benefit on its own when comparing the results from the in-class lecture individually (Table 4) and the results including both the in-class and pre-lab lectures (Table 5). From these two tables, we see that students performed significantly better on more of the quizzes after the pre-lab lecture than after the in-class lecture. Overall, these results suggest that students' immediate recall of the material appears to be very good and that a pre-lab introduction is a useful component of the lab itself.

After finishing a lab exercise, we ask students to complete a lab report where they answer questions about the lab exercises and analyze and interpret the data they collected in lab. The purpose of these lab reports is to teach students how to be reflective learners, give them another opportunity to connect theoretical concepts to their lab activities, and to enhance their ability to write technical reports. To assess the value of the lectures, lab, and lab report together, we compared pre-test scores with quiz scores taken after completing the lab report. From Table 6, we again see that students scored significantly higher on quizzes for all of the labs in at least one of the two years of data collected, except for Lab 1. These results suggest that the lab report is beneficial in helping students translate and retain the concepts they learned in the lab.

Overall, we found that the in-class lecture alone, though beneficial, is not sufficient for enhancing student understanding of the material, as reflected in students only scoring significantly higher on quizzes in 2 of the 5 labs. Further, we found that the haptic paddle labs (including the pre-lab introduction, the lab activity, and the lab report) were very successful in increasing student understanding of the core concepts, as students scored significantly higher on quizzes in 4 of the 5 labs after completing all parts of the lab experience. Specifically looking at the pre-test scores compared to the after lab scores, we observed that students scored significantly higher on all 5 of the lab quizzes in one of the two years of data collection. Finally, our results suggest that student retention of the material is also good, with significantly higher scores attained in 3 of the 5 labs, and cumulatively, on the post-test compared with the pre-test.

6 Conclusion

In this paper, we have taken the first steps toward formally assessing the benefits of the haptic paddle laboratories, which have been adopted by several universities in System Dynamics. Our formal assessments suggest that the haptic paddle laboratories are successful in enhancing student understanding of core concepts in this course. The results of our study, which probe both what material students are learning and when they are learning it, show that the lab activities complement and enhance the in-class lecture and significantly increase student performance on conceptual quizzes. These results, combined with prior qualitative assessments of the haptic paddles[9], suggest that this set of laboratories engages students, provides an inexpensive, versatile platform for educators to use, and results in significantly higher scores on multiple-choice conceptual quizzes in System Dynamics.

These analyses also enable us to pinpoint areas for future improvement. In subsequent years, our primary focus will be on revising Lab 1, which was the lab that consistently appeared to be the most difficult for students in the analyses discussed in this paper. One possible thought in addressing this issue is to split Lab 1 up into two labs. The first "lab" session would simply be an introduction to the lab and the equipment, and the second lab session would be the actual first lab, with modifications from previous years. The motivation behind this is to allow students more time to get acquainted with the hardware and software of the haptic paddle before performing any in-depth analysis. As mentioned earlier, we have also recently redesigned the paddle and transitioned its software to a Matlab and Simulink implementation running on Arduino microcontroller boards and implemented a friction-based drive train. We are currently administering assessments with this new architecture and plan to compare data collected from it with the data presented here, in our future work. Finally, we will continue using assessments like the one presented here to evaluate and improve the haptic paddle laboratories. We believe that this type of assessment and reflective analysis has the potential to significantly improve the educational experience and performance of both teachers and students.

7 Acknowledgments

The authors wish to thank several people who have contributed to the development of the Vanderbilt haptic paddle laboratories and to the collection and analysis of assessment data, especially Dr. Thomas Withrow, Dr. Caleb Rucker, Louis Kratchman, Hunter Gilbert, Neal Dillon, and all of the former TAs of the system dynamics lab. Support for this work was from The MathWorks Inc. Education Curriculum Grant and the National Science Foundation under award #IIS-1054331 and a Graduate Research Fellowship.

References

[1] Cox, D., 2008, "Hands-on experiments in dynamic systems and control for applied education in robotics and automation," World Automation Congress, pp. 1–6.

[2] Dewoolkar, M. M., George, L., Hayden, N. J., and Neumann, M., 2009, "Hands-on undergraduate geotechnical engineering modules in the context of effective learning pedagogies, ABET outcomes, and our curricular reform," Journal of Professional Issues in Engineering Education and Practice, **135**, pp. 161–175.

[3] Abdulwahed, M. and Nagy, Z., 2009, "Applying Kolb's experiential learning cycle for laboratory education," Journal of Engineering Education, **98**, pp. 283–294.

[4] Wieman, C. and Perkins, K., November 2005, "Transforming physics education," Physics Today, pp. 36–41.

[5] Fraser, D., Pillay, R., Tjatindi, L., and Case, J., 2007, "Enhancing the learning of fluid mechanics using computer simulations," Journal of Engineering Education, **No.4**, pp. 381–388.

[6] Goeser, P., Johnson, W. M., Hamza-Lup, F. G., and Schaefer, D., 2011, "VIEW - a virtual interactive web-based learning environment for engineers," Advances in Engineering Education.

[7] Terpenny, J. and Goff, R., 2006, "Utilizing assistive technology design projects and interdisplinary teams to foster inquiry and learning in engineering design," International Journal of Engineering Education, **22**, pp. 609–616.

[8] Chen, X., Chase, J. G., Wang, W., Gaynor, P., and McInnes, A., 2010, "Embedding design projects into multidisplinary engineering education," International Conference on Educational and Information Technology, **3**, pp. 398–402.

[9] Okamura, A., Richard, C., and Cutkosky, M., 2002, "Feeling is believing: Using a force-feedback joystick to teach dynamic systems," Journal of Engineering Education, pp. 345–349.

[10] Grow, D., Verner, L. N., and Okamura, A. M., 2007, "Educational haptics," AAAI Spring Symposia - Robots and Robot Venues: Resources for AI Education.

[11] Bowen, K. and O'Malley, M. K., 2006, "Adaptation of haptic interfaces for a labview-based system dynamics course," Proceedings of the 14th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (HAPTICS 06), pp. 147–152.

[12] Gillespie, R. B., Hoffman, M. B., and Freudenberg, J., 2003, "Haptic interface for hands-on instruction in system dynamics and embedded control," Proceedings of the 11th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (HAPTICS 03), pp. 410–415.

[13] Vanderbilt University Haptic Paddle Webpage. http://research.vuse.vanderbilt.edu/MEDLab/haptic_paddle.html.

[14] EduHaptics. http://eduhaptics.org/index.php/HapticDevices/HapticPaddles.