Advances in Engineering Education

SUMMER 2014

Haptic Paddle Enhancements and a Formal Assessment of Student Learning in System Dynamics

JENNA L. GORLEWICZ
Southern Illinois University
Edwardsville, Edwardsville, IL

LOUIS B. KRATCHMAN

AND

ROBERT J. WEBSTER III
Vanderbilt University
Nashville, TN

ABSTRACT

The haptic paddle is a force-feedback joystick used at several universities in teaching System Dynamics, a core mechanical engineering undergraduate course where students learn to model dynamic systems in several domains. A second goal of the haptic paddle is to increase the accessibility of robotics and haptics by providing a low-cost device for middle and high school teachers and students, a goal which has been hindered to date by the lack of low-cost electronic solutions for motor control and computer interfacing. Prior assessments of the learning enabled by the paddles at the college level have been qualitative, consisting of anecdotal case studies illustrating student and educator belief that they enhance learning and increase student enthusiasm. In this paper, we describe haptic paddle design enhancements and provide a formal assessment of student learning when interacting with haptic paddles. Design enhancements seek to enable broad dissemination and improve the student experience by making the paddle less expensive (less than $100 for a complete system including electronics) and easier to use. Our formal assessment quantifies if and when learning occurs in a System Dynamics course featuring haptic paddle laboratories, via multiple choice quizzes presented at various time points to measure learning in the lecture, in the lab, during lab report writing, and at the end of the semester.

Key Words: Haptic paddle, systems dynamics, robotics

INTRODUCTION

The haptic paddle and an associated laboratory curriculum were developed in the late 1990s at Stanford University to provide a hands-on platform for students to physically interact with and "feel" simulated dynamic systems via force feedback (Okamura, Richard, and Cutkosky 2002). Since then, haptic paddles have been adopted at multiple universities (see (EduHaptics 2012) for an overview) including Johns Hopkins (Okamura, Richard, and Cutkosky 2002), Rice (Bowen and O'Malley 2006b; Bowen and O'Malley 2006a), Michigan (Gillespie, Hoffman, and Freudenberg 2003), Vanderbilt (Gorlewicz and Webster III 2012; VU Webpage 2012), ETH Zurich (Gassert, Megzger, Leuenberger, L., Tucker, Vigaru, Zimmermann, and Lambercy 2013), and Utah (EduHaptics 2012). Generally agreed upon engineering education objectives (Felder and Silverman 1988) have spurred adoption of haptic paddles, including the desires to engage students with a variety of learning styles, enable students to connect theoretical principles to practical applications, and to provide students with cooperative learning experiences.

The objectives of haptic paddles are in keeping with prior work incorporating hands-on demonstrations (Cox 2008; Dewoolkar, George, Hayden, and Neumann 2009; Abdulwahed and Nagy 2009), computer simulations (Wieman and Perkins 2005; Fraser, Pillay, Tjatindi, and Case 2007; Goeser, Johnson, Hamza-Lup, and Schaefer 2011), design projects (Terpenny and Goff 2006; Chen, Chase, Wang, Gaynor, and McInnes 2010), and laboratory experiences (Feisel and Rosa 2005), which have been found beneficial in the context of many different undergraduate courses. For System Dynamics, a core mechanical engineering undergraduate course required at most universities, haptic paddles provide a particularly good device upon which to build laboratory curricula (Okamura, Richard, and Cutkosky 2002; Grow, Verner, and Okamura 2007). They are one of the simplest possible robots that a student can build, having only one motor and one degree of freedom (DOF). Yet the modeling, mechatronics, and control work required to accomplish the haptic paddle laboratories is directly generalizable to more complex systems with more degrees of freedom.

The haptic paddle has the additional benefit that it is a haptic device, through which students can touch and feel dynamic system simulations. Each university that has adopted haptic paddles has contributed to the evolution of the haptic paddle in mechanical design, educational curricula, and software in various ways (see Figure 1 for pictures of the various haptic paddle designs, and the central web repository EduHaptics (EduHaptics 2012) for more information). However, none of these modifications have fundamentally altered what the haptic paddle is; it remains a one DOF haptic device that students can construct and/or program and use. Most hardware changes have been aimed at increasing robustness, reducing costs (though even the initial work at Stanford emphasized cost-conscious mechanical design), and using readily available materials and components. The initial curriculum proposed at Stanford consisted of sequential laboratory exercises focused on constructing,

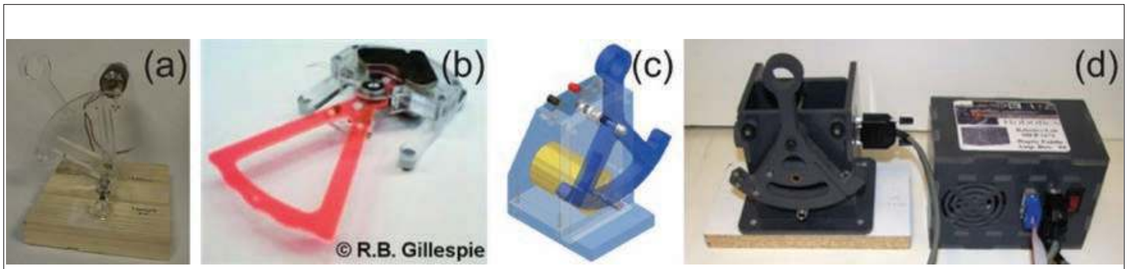

Figure 1. (a) The Stanford and Johns Hopkins Haptic Paddle. (b) The University of Michigan Haptic Paddle. (c) The Rice University Haptic Paddle. (d) The University of Utah Haptic Paddle.

calibrating, modeling, and controlling the paddle, before using it to interact with simulated dynamic systems (Okamura, Richard, and Cutkosky 2002). While some curricular adaptations have been made at various universities to suit the learning objectives of their respective courses, the originally proposed curriculum has been used with only minor modifications at Stanford, Johns Hopkins, and Vanderbilt, and is the subject of the formal assessment described in this paper. Prior assessments of the haptic paddle as a learning tool have largely been qualitative and anecdotal in nature, illustrating that students respond enthusiastically to the haptic paddle and that many students appeared to laboratory instructors to be developing a true understanding of core course concepts for the first time as they interacted with the haptic paddles (Okamura, Richard, and Cutkosky 2002). There has been one preliminary assessment measuring learning outcomes using the haptic paddle laboratory series compared to more traditional modular labs (Bowen and O'Malley 2006b). In this study, the outcomes were measured via a subjective scaling rubric, with student answers being graded as 0 (not at all correct) to 3 (fully correct). The results of the study show significant gains in student understanding of concepts associated with the course after completing the haptic paddle labs compared to disjointed labs.

There have also been a number of instances over the years where high school students and teachers have requested information and assistance from faculty and graduate students in building their own haptic paddles, having found information about them on the Internet. To address this, one of the main objectives in the initial work at Stanford was to enable broad dissemination at multiple educational levels. They sought to facilitate this adoption by making the haptic paddle mechanical components as low cost as possible. However, a major hurdle in the process of making haptic paddles accessible to the at-home and high-school settings has been the fact that the initial system at Stanford used an expensive D/A card, a desktop computer to host it, and a benchtop power supply. Thus, despite low-cost mechanical components, someone developing a haptic paddle setup from scratch needed to invest quite a bit of money in computer/electronics resources. They would also need to be sufficiently computer-savvy to be comfortable opening their computer case to install the card and then learning how to write a program to interface with it

To address these challenges, in this paper, we contribute enhancements to the haptic paddle infrastructure, as well as a formal assessment of the learning that is facilitated by haptic paddle laboratories during a semester of System Dynamics. More specifically, we present (1) a new friction drive design, which is more robust than the original capstan drive (enhancing learning by mitigating student frustration with re-stringing the paddles whenever they make them go unstable), (2) a new electronics implementation featuring a low-cost Arduino microcontroller and amplifier (≈$55), which connects to a computer using a universal serial bus (USB) interface, making it possible to operate the paddle from a laptop, and (3) a new Matlab/Simulink (The MathWorks Inc.) software framework, which is consistent with and reinforces Matlab use throughout the course. Further, we complement prior assessments that evaluated student perception of the valueof the haptic paddle laboratories and the enhancements that the haptic paddle lab series offers compared to traditional disjointed labs, with a formal assessment. The objective is to determine if and when students learn key course concepts: in lecture, in the lab activities themselves, or after reflecting on the lab activities while writing lab reports. We note that some material in this paper was presented in preliminary form in (Gorlewicz and Webster III 2012). Here, we present a new haptic paddle system and its accompanying hardware and software, a validation of the new friction drive design, a third year of assessment data, and a discussion of our dissemination efforts.

HAPTIC PADDLE HARDWARE AND SOFTWARE ENHANCEMENTS

The haptic paddle is similar in functionality to commercially available haptic devices (such as the PHANToM Omni by SensAble Technologies) in that it emulates interaction forces that occur when a user contacts an object, but it is simpler in design and construction since it is has just one DOF. As the user moves the paddle handle, the drive wheel attached to the motor rotates. The position of the drive wheel is sensed using a magnetic angle sensor and the Arduino (see Section 2.2). The Arduino is used for bidirectional communication between the motor and Simulink. In Simulink, the position and velocity of the paddle handle are calculated and desired forces are computed. Then, the motor generates these desired forces, which are felt by the user holding the paddle handle.

Mechanical Design Enhancements

The basic haptic paddle design (Figure 1) consists of an acrylic handle coupled to a single motor through a capstan drive. As with prior haptic paddles, ours (Figure 2) is designed to be low-cost

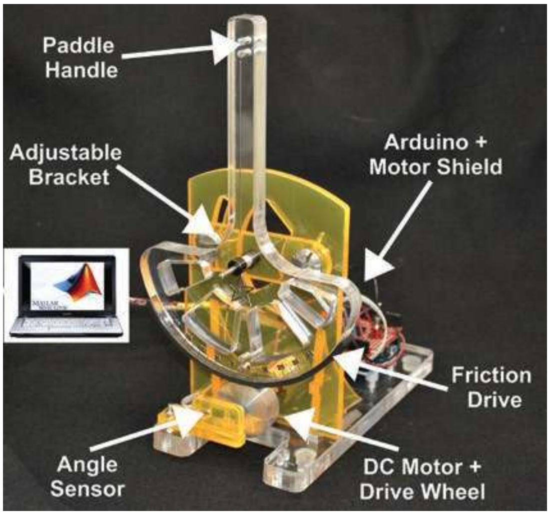

Figure 2. A schematic of the components of our Haptic Paddle, which relies on a friction drive design, runs in Matlab/Simulink, and uses a low-cost Arduino microcontroller for communication.

and easy to manufacture, consisting of laser-cut acrylic. All prior haptic paddle designs have used capstan drives, with the exception of the Michigan haptic paddle (iTouch Motor), which uses a direct drive device without a transmission (Gillespie, Hoffman, and Freudenberg 2003). While there is nothing intrinsically disadvantageous with capstan drives (indeed, they are preferred in many commercial haptic devices for their low friction and smoothness), several years of experience in the laboratory have illustrated that they can be a source of significant frustration for students and teaching assistants (TAs) as implemented on haptic paddles. For example, when students cause paddle instability (which they often do when learning about control, and occasionally at other times in the lab) the string will pop off of the motor drive wheel. It then requires several hands working in a small space to re-wrap and tension the string, while tightening screws and nuts to fix both ends of the string to the capstan. This process takes anywhere from 2-10 minutes, depending on student experience and frustration level. If done incorrectly, the string may be too loose and slip around the motor spool.

To address this, we have replaced the capstan drive with a friction drive. The friction drive consists of a strip of neoprene rubber adhered to the bottom of the paddle handle that rolls in contact with an aluminum drive wheel fastened to the motor shaft, as shown in Figure 3. This new design is much easier to assemble. If the paddle goes unstable, the neoprene strip simply rolls out of contact

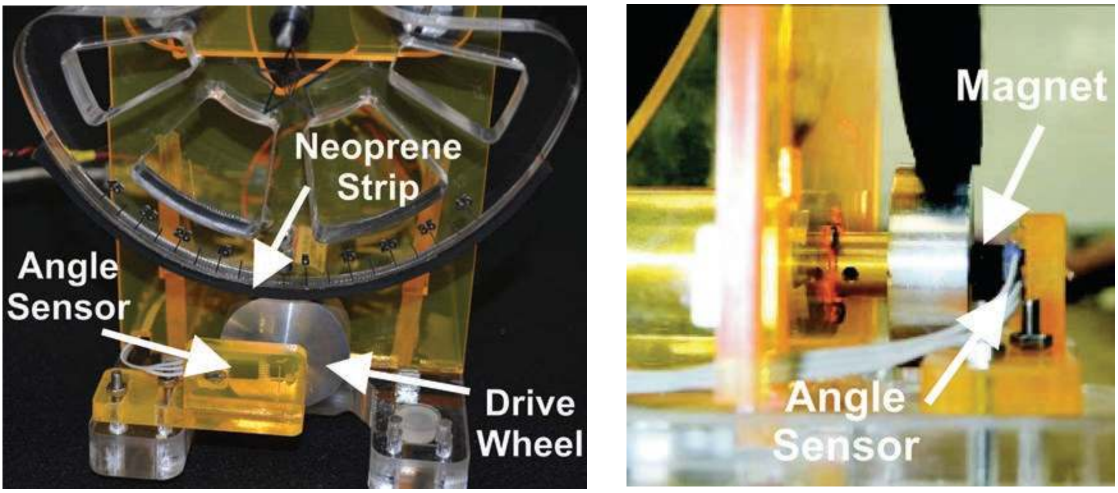

Figure 3. The new friction drive design of the haptic paddle consisting of a rubber strip at the bottom of the paddle handle which directly contacts the drive wheel on the motor. An inexpensive magnetoresistive angle sensor ($6, KMA199E, NXP Semiconductors), which measures the angle of the nearby rotating magnet on the drive wheel, provides the motor position.

with the drive wheel. To reset the paddle, all one needs to do is move the handle back into its normal vertical position. To ensure that the amount of contact force between the drive wheel and the rubber strip can be optimally adjusted (note that this only needs to be done once), we included an adjustable bracket that enables the entire paddle handle to move up and down, as shown in Figure 2.

While this friction drive trades off some haptic fidelity in exchange for robustness, we have observed no practical reduction in learning benefit with this new design. Indeed, it is qualitatively difficult to perceive a noticeable difference between the friction drive and capstan drive. To quantitatively compare the friction and cable drive designs, we measured the friction and inertia of our new paddle compared to the Stanford haptic paddle using an experimental setup similar to that described in (Abbott 2005), which estimated these parameters for the capstan drive haptic paddle. To do this, we attached a load cell (Entran ELFM-T2E-25L) with a small acrylic square at the handle of both paddles. A user lightly grasped the acrylic square and moved the paddle randomly using a variety of velocities. Force and position values were recorded throughout paddle motion. We modeled the haptic paddle as a mass with Coulomb-plus-viscous friction and used a pseudoinverse technique to solve for the mass (m), viscous friction coefficient (b), and Coulomb friction (f_c), as in (Abbott 2005). Twenty trials were performed for each paddle design, and the results were averaged.

The resulting parameter estimates for the capstan drive paddle were $m = 0.0596$ kg, $b = 0.0926$ Ns/m, and $f_c = 0.1083$ N, and for the friction drive paddle were $m = 0.0466$ kg, $b = 0.1218$ Ns/m, and $f_c = 0.1146$ N. We observe that the equivalent mass, viscous friction, and Coulomb friction in both paddles are similar. While the viscous friction coefficient is higher in the friction drive paddle compared with the capstan drive paddle, both are low. Further, in (Abbott 2005), the viscous friction coefficient in the cable drive paddle was found to be between $b = 0.15 – 0.23$ Ns/m, showing that this parameter can vary depending upon the construction of the paddle and the components used. We also note that our equivalent mass was lower in the friction drive paddle compared with the capstan drive paddle. Our calculated effective masses at the paddle handle, including the 0.0075 kg load cell and acrylic square, were $m = 0.053$ kg for the friction drive paddle and $m = 0.052$ kg for the capstan drive paddle. These calculated values were comparable to the estimated values of mass from our model, but there were some small discrepancies between the two values for each paddle. These discrepancies were likely due to slight variations in how hard the user's finger pressed onto the force sensor when holding the paddle in each trial. We mitigated this variation as much as possible by reminding the user not to squeeze the force sensor before each trial; however, the compliance of the fingertip pad likely contributed to small variations in the grasp force. Compared to commercial haptic devices such as the Phantom Omni, which has an apparent mass at the tip of 0.045 kg and backdrive friction of 0.26 N (SensAble Technologies), both haptic paddle designs perform well.

A minor additional mechanical change to the haptic paddle design was the incorporation of the larger aluminum drive wheel, shown in Figure 3, onto the motor shaft. This adds some inertia and changes the gear ratio slightly of our new paddle compared to the original haptic paddle. However, it enables the motor spin down test (a Lab 1 exercise, see (Okamura, Richard, and Cutkosky 2002)) to be performed with no disassembly of the paddle, as was required in prior capstan drive versions. Now, all that must be done to perform the motor spin down test is to rotate the handle until the neoprene strip is out of contact with the aluminum drive wheel.

Low-Cost Electronics and Computer Interfacing

Toward lowering the overall cost and thus the bar for entry to new users of the haptic paddle, and in keeping with the general goals of the original Stanford project which sought a widely disseminable device, we have developed a new low-cost and easy-to-use electronics solution based on the Arduino microcontroller. While the original haptic paddle could be built for $30 in mechanical components, it was assumed that a D/A (Digital to Analog) solution was already available to the person implementing the paddle (Okamura, Richard, and Cutkosky 2002). Initial instantiations of the design at Stanford and Johns Hopkins used Measurement Computing PCI cards which, at the time, retailed for between $1000 and $2000 and required a desktop computer. The recent introduction of the Arduino microcontroller has provided a low-cost microcontroller capable of D/A and (with the associated motor amplifier) motor control, that has catalyzed a large hobbyist community and has been introduced into the classroom at many universities over the past few years. For us, an ancillary benefit of Arduino use is that it reinforces the experience that students obtain in our undergraduate Mechatronics class, in which Arduino programming and interfacing are central topics. The Arduino is a USB-connected device (meaning the haptic paddle can now be run from a laptop) and is inexpensive, with the Arduino UNO retailing for $30 and the Motor Driver Shield (amplifier) retailing for $25 from SparkFun Electronics, as of the time of this writing. The microcontrollers are easy to program, with extensive online documentation and examples. Using the Arduino language, which is simply a set of C/C++ functions, and the Arduino programming environment (Arduino 2012), we developed code to read the haptic paddle's angle sensor and control the motor using pulse width modulation (PWM).

We also upgraded the hall effect sensors used in the original design to a new $6 magnetic angle sensor (KMA199E, NXP Semiconductors) as shown in Figure 3. These analog sensors are much more reliable and robust to misalignment and exact distance to the magnet, and provide a larger voltage output range compared with the hall effect sensors used in the initial Stanford design. They are also linear with respect to angle, obviating the need for a 3rd order model fit for calibration. We note that this calibration process, while perhaps useful educationally, was a significant source of frustration for students and TAs because the need for recalibration was frequent, as the sensors did not work well for imperfectly assembled paddles (i.e. those with distance variation between the magnet and sensor over the paddle sweep). To retain the educational aspects of calibration, we now include a calibration verification experiment in the lab, where students verify the linear relationship between handle angle and sensor output.

These electronics improvements have reduced the cost of the complete haptic paddle system to just under $90, ($55 D/A and motor control electronics + $6 angle sensor + $5 surplus motor + $20 Acrylic raw material). This lower cost may make it easier for universities, K-12 students and teachers, and hobbyists to adopt and use haptic paddles. This cost assumes that the user has a laptop or desktop computer and a power supply. Many users will already have access to a power supply, but if not, low-cost options such as a 12V Regulated Power Adapter, rated for 5A, can be easily found on several online retailers (e.g. Replacement AC Adapter, 12V, 5A Power Supply from Stiger) for less than $9. To connect this to the haptic paddle, an appropriate connector (e.g. CP-024B-ND, DigiKey Corporation, $3) will be needed, or one could simply cut off the plug and use the individual wires.

Matlab/Simulink Control Software

We have also modified the software interface that controls the haptic paddles, moving from C++ to Matlab/Simulink (The MathWorks, Inc.). Simulink's Real Time Windows Target and the 3D animation packages enable us to control the haptic paddle in real time and to create realistic visualizations and convenient user-interfaces for students. All of the code to do this is freely available at (VU Webpage 2012).

The move to Matlab/Simulink was made for two reasons. First, we use Matlab in the lecture portion of the class for model evaluation and dynamic simulation, and it is preferable to keep a consistent software language throughout the class. Note that many Mechanical Engineering students have only superficial knowledge of programming and little comfort with it, despite having a required programming course as freshmen or sophomores. Thus, often, one must re-teach many basic programming concepts in System Dynamics, and switching languages can cause confusion. Second, the original paddles were programmed using C++, and students were provided only with executables, which limited their ability to develop a deep understanding of what is going on inside the haptic paddle system "black box". Simulink's graphical interface enables students to build block diagrams, connecting what they have learned in class directly to hardware, and makes it easier for students to understand how the computer program works. Since students are now able to program the paddle themselves, they also have much more accountability during lab activities. Rather than blaming the TAs, the computer in the lab, the course instructor, or the university for any bugs they encounter, they are automatically inclined to begin debugging themselves, rather than relying on the TA to "fix" the system for them. We have qualitatively observed students to be more engaged and more enthusiastic in the lab with the interactive software environment provided by Simulink, than when simply double clicking in an executable.

Updates to the Original Stanford Laboratory Curriculum

The mechanical, electronic, and software changes described above 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. While these changes have required some curricular changes (notably the simpler design does not require most of one lab for paddle construction, and the improved sensor design has saved about half of another lab by eliminating sensor calibration), the learning objectives of the lab exercises remain comparable to the original Stanford and Johns Hopkins labs. These time savings have enabled us to devote more lab time to teaching the students Simulink and how to interface simulation and hardware.

Similar to the original Stanford curriculum, each of the five sequential lab assignments focuses on a different aspect of the haptic paddle, which relates to concepts covered in lecture. The following are descriptions of the current lab assignments (also posted at (VU Webpage 2012)), including minor modifications enabled by the hardware/electronics/software changes described above. The summaries and main educational goals of each lab are listed below, such that readers can easily refer back to them when interpreting the results of the formal assessment described in the Assessment Sections.

- **Lab 1:** In the first lab, the main educational objectives focus on first order systems and the components of dynamic systems. In lab, students are introduced to Simulink by creating simple models of virtual springs and dampers and then use the paddle to feel these virtual objects while changing their properties. Then, they conduct a motor spin down test as an example of a first order system and compare their experimental results with predicted results they obtain from their Simulink simulation of their motor. Students then include Coulomb friction in their simulation and compare the differences between it and their simulation with only viscous damping. Finally, they use their simulation results to estimate damping in their motor and compare it with the best fit damping constant they obtain from their analytical solution.
- **Lab 2:** In the second lab, the primary educational objectives include a more in-depth exploration of friction, damping, and other external factors that effect system performance. In lab, students experimentally measure the torque constant and Coulomb friction in the motor and analyze the paddle handle by measuring its moment of inertia through a bifilar pendulum experiment.
- **Lab 3:** In the third lab, the main educational objectives focus on the behavior of second order systems. In lab, students first perform a calibration exercise with the magnetic sensor. Then, they use Simulink to make their haptic paddle behave as a second order underdamped system in order to determine the equivalent mass, stiffness, and damping of the system. Finally, students build a pure simulation of a second order system and compare their predicted and experimental results.
- **Lab 4:** In the fourth lab, the main educational goals include stability and feedback control. In lab, students investigate PID (proportional-integral-derivative) control by developing a Simulink model to command step and sinusoidal inputs (as shown in Figure 4 (Left)), altering PID gains, and observing the paddle's response. Then, students add weights to the top of their paddle handle to make it an unstable system (inverted pendulum) and use feedback control to stabilize the paddle.
- **Lab 5:** In the fifth lab, the primary educational objectives focus on forced responses and vibrations. In lab, students interact with a multiple DOF mass-spring-damper system and explore its modes of vibration using the paddle and a real-time 3D visualization of the system, shown in Figure 4 (Right).

We note that majority of the content associated with the first two labs is material that students have been exposed to in prior engineering courses, while the content associated with the remaining three labs is likely new material for students. Thus, in principle, the first two labs may contain concepts that students can grasp more easily due to prior exposure as opposed to the latter three labs where students are trying to grasp concepts for the first time.

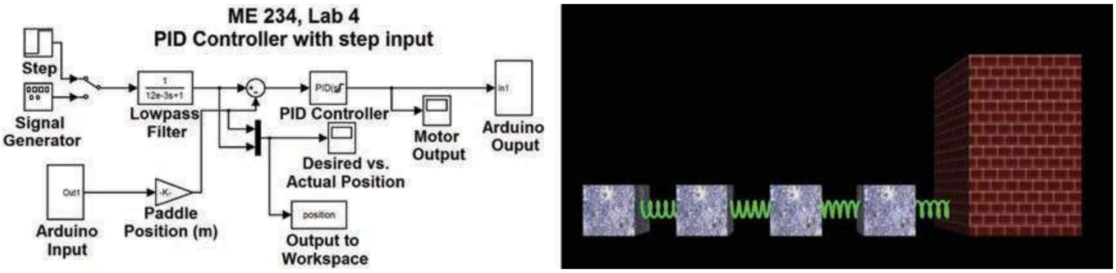

Figure 4. (Left) The Simulink model students build to investigate feedback control in Lab 4. (Right) The 3D visualization of a mass-spring-damper system students interact with using the haptic paddle in Lab 5.

A FORMAL ASSESSMENT OF STUDENT LEARNING: METHODS

This section addresses the formal assessment of the learning that takes place in haptic paddle laboratories in the context of a System Dynamics course. This assessment seeks to determine if students are learning key educational objectives for each lab and to determine when that learning took place. In conducting this assessment, we had the distinct advantage of having a large class (approximately 70 students, varying slightly from year to year) that was subdivided into four lab sections. Each lab section met for three hours on a different day of the week, five times throughout the semester. Within each lab section, teams of 2-3 students work together, as shown in Figure 5. These teams are self-selected by the students, and they remain in roughly the same groups for the duration of the semester.

To assess student learning, we constructed a 25-question multiple choice quiz (5 questions/lab × 5 labs, see (VU Webpage 2012)) covering the core concepts of the lecture and the lab exercises. The goal in our multiple choice quiz was to test for conceptual understanding. Because there does not exist a current concept inventory for system dynamics and the inventories available on online repositories (e.g. CIHub.org) for dynamics do not cover the scope of material for this course, we sought to generate our own System Dynamics Concept Inventory. To do this, we held iterative brainstorming sessions between the instructor and two teaching assistants who had been heavily involved in the course in the past. Questions were generated based on covering the core concepts of the course and then answers were generated to include common misconceptions or wrong answers that students would likely choose if they had only partial or incorrect understanding of the concept. The concept inventory is available online at (VU Webpage 2012). Each question had 4 possible answers, with one being correct. 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

Figure 5. Small teams of 2-3 students interacting with the haptic paddle during a lab activity.

during the first week of the semester to all students in order to assess their initial understanding of the course material and to provide a baseline measurement for statistical analyses and again at the end of the semester as a final evaluation. This 25-question assessment was then broken down into 5 quizzes, each containing 5 questions. Each quiz corresponded to key concepts from one lab.

To explore *when* student learning was occurring, we randomized the presentation of the 5-question lab quiz among the four student 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. The first time point enabled us to assess the value of the in-class lecture alone. The second time point provided insight on student learning after both an in-class lecture and a concise, introductory pre-lab lecture on the lab objectives. The third time point sheds light on the cumulative benefits of the lectures and lab activities in enhancing student learning, and the fourth time point enabled us to explore the overall success of all learning experiences including lectures, lab, and a formal lab report. 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 lab quizzes were administered to each section was systematically rotated (see Table 1) to remove any potential bias in data collection.

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

Table 1. Quiz placement for each lab for each student section. S1-S4 represents each of the four student sections.

Students were given a 10% extra credit bonus on each of their lab report grades for completing the respective lab quiz. These points were given based on completion of the quiz, not based on correctness of student answers. Students were aware of this grading policy, and thus, it is possible that they may not have always tried their hardest in answering the questions correctly. We note, however, that lab TA's stressed the importance of the quizzes to the students and provided ample time in lab for students to complete them. For assessment purposes, we recorded 1 point for every correct answer and 0 points for every incorrect answer. Means and standard deviations were computed for each of the students' quizzes.

Below, we present three years of data collected from the assessments, with $N_1 = 63$ students, $N_2 = 71$ students, and $N_3 = 74$ students, where N_x represents the total number of students in the class for each of the three years. We note that appropriate IRB approval was obtained for this study. In year 1 (Y1), we used the original Stanford version of the haptic paddle (subsequently used at Johns Hopkins University) and its C-executables, in year 2 (Y2), we used the inverted paddle design with a cable drive (similar to the Rice University design) and Matlab and Simulink software, and in year 3 (Y3) we used the friction drive paddle and Matlab and Simulink.

We note that the same instructor taught the course over the period of this study. There were no fundamental changes to the course structure or basic course content over the 3 years of this study, with only normal updates to individual lectures and minor modifications to homework problems being made. The one somewhat significant change to the course made during the time period of the study was that the final exam was replaced with a final take-home project in Year 2, but we do not expect this to have influenced the laboratories since all labs were completed prior to the final project assignment. It is possible that this take-home project may have influenced the end of the semester results, but another study would be required to explore this. We also note that the same 2 TAs taught all of the sections of the lab within a given year of the course. While there were different TAs in different course years, each TA was prepared for the laboratory exercises in the same manner, through discussions with the instructor. The presence of different TAs between years may be a positive feature of our study, since it is representative of how the lab would be conducted at a different university. We explore the effects of the different TA between years in Section 5.7.

Research Questions

The research questions we sought to answer analyzing quiz data are as follows:

1. Overall, did the students learn the core course concepts at some time point 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. 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. While our study architecture does not enable us to pinpoint the enhancements from the lab activities themselves, we were specifically interested in exploring the timepoint from the beginning of the semester to after completion of the lab activities, since this case encompasses the lab exercises presented in this work.

For each year of data collection, we conducted assessments at the beginning of the semester (pre-test), after the in-class lecture, after a pre-lab lecture, after the lab activity, after the lab report, and at the end of the semester (post-test). The one exception is in Year 1, where the final post-test data was not collected. As such, only Year 2 and Year 3 data are used in the course benefit analyses (Research Question 1), but all three years of data are used in the individual learning opportunities analyses (Research Question 2).

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 address the second research question and 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 throughout the semester. 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. For further insight into the magnitude of the difference in mean quiz scores in each analyses, we also computed the effect size between the two means of interest using Cohen's d with a pooled standard deviation. A positive value of d suggests an increase in student performance on the quiz at the specified time compared to the pre-test, and a negative value of d indicates a decrease in student performance on the quiz at the specified time compared to the pre-test.

Note that in all discussions, figures, and tables presented, significance at the 95% confidence level ($\alpha = 0.05$) was determined from the paired t-test analyses, and the interpretations made on effect size were based upon the Cohen's *d* computation. We note that these two statistical analyses are complementary to one another, with the t-tests providing insight on whether or not quiz means were significantly different from one another, and the effect sizes providing insight on the magnitude of this difference. In our discussions of effect size, we follow the standard interpretation that $d = 0.2$ is a small effect, $d = 0.5$ is a medium effect, and $d = 0.8$ is a large effect, where the value of *d* indicates the difference between two means as a fraction of the pooled standard deviation. All statistical analyses were performed in R 2.11.1, and the results are presented in the Assessments Sections.

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 for each year, and (3) No significant difference between student sections' initial pre-test scores *for each lab*, for each year. We assessed the normality of each student section's data for each year by creating quantile-quantile plots that included both pre-test scores and lab quiz scores for each student section. All 12 (4 student sections × 3 years) plots suggest a linear trend, and thus we can infer that our data is approximately normally distributed and that parametric statistical tests, such as the t-test, are applicable in our subsequent analyses. Second, we ensured that student sections within each year were comparable in their initial cumulative understanding of the course material by comparing 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. Finally, we ensured that student sections within each year were comparable in their initial understanding of the course material *for each lab* by separating the 25-question pre-test up into 5 parts, corresponding with the 5 lab quizzes, and comparing 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, but found a significant difference between section 1 and section 4 on the Lab 5 portion of the pre-test (p-value = 0.04), with section 1 having a significantly higher average on this portion of the material. Because of this, student section 1's data was omitted in the Lab 5 analyses for Y2.

Only section 1's data was omitted because there were no significant differences between any combination of sections 2, 3, and 4's scores on the Lab 5 portion of the pre-test. From the Y3 data, we observed no significant differences at the 95% confidence level ($\alpha = 0.05$) between any student sections'*cumulative* pretest score, but found a significant difference between section 1 and section 4 on the Lab 2 portion of the pre-test (p-value = 0.02), with section 1 having a significantly higher average on this portion of the material. For this reason, we omit student section 1's data in the Lab 2 analyses for Y3. Again, only section 1's data was omitted because there were no significant differences between any combination of sections 2, 3, and 4's scores on the Lab 2 portion of the pre-test

FORMAL ASSESSMENT OF COURSE: RESULTS AND DISCUSSION

We first sought to answer whether or not the students learned and retained the course concepts after completing the entire course. This enabled 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. Because we found a significant difference in cumulative scores from this comparison, we then 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. In both analyses, we also computed the effect size, d, of the difference in means. The results of this study are presented in Figures 6 and 7 and can be found in tabular form in (VU Webpage 2012). Note that post-test data was only available for Y2 and Y3, and thus no data is presented from Y1.

Discussion of Educational Benefit from Course

The results presented from the first study suggest that students learned and retained majority of the core course concepts throughout the semester. From Figure 6, we observe that students achieved a significantly higher cumulative score on the post-test compared to the pre-test in both years. Large effect sizes ($d > 0.8$) were also observed. This suggests that the learning opportunities provided to the students throughout the semester were successful in enhancing

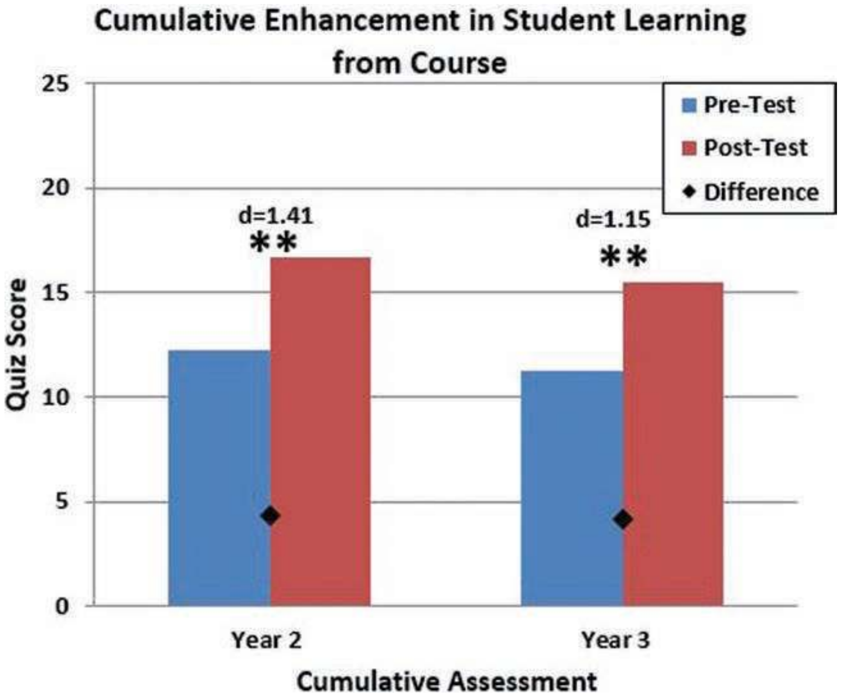

Figure 6. The cumulative mean (out of 25) of all students' pre-test score compared with their post- test score for years 2 and 3. Significant differences in quiz scores from the paired t-test are denoted with a ** at $\alpha = 0.05$. The effect size, d, between the two means is shown above the bars, and the difference between means is represented by the black diamond data points.

students' overall understanding of the course material. After looking at the pre-test and post-test scores separated by lab content (Figure 7), we observe that students did significantly better on the quizzes focusing on concepts from Labs 2, 3, 4, and 5 in at least one of the two years presented. Moderate to large effect sizes ($d > 0.5$) were also observed in these same labs. This suggests that students learned and retained these concepts throughout the duration of the course. Though quiz score increases are observed for Lab 1 in both years and Lab 2 in Y2, there were no significant differences between the pre-test and post-test scores in these cases, and the observed effect sizes were small ($d < 0.5$). This suggests that the material associated with these labs were particularly challenging for students to grasp, and could be sources of improvement in future labs. For further insight into our findings, we look to the next analysis focusing on when student learning was occurring.

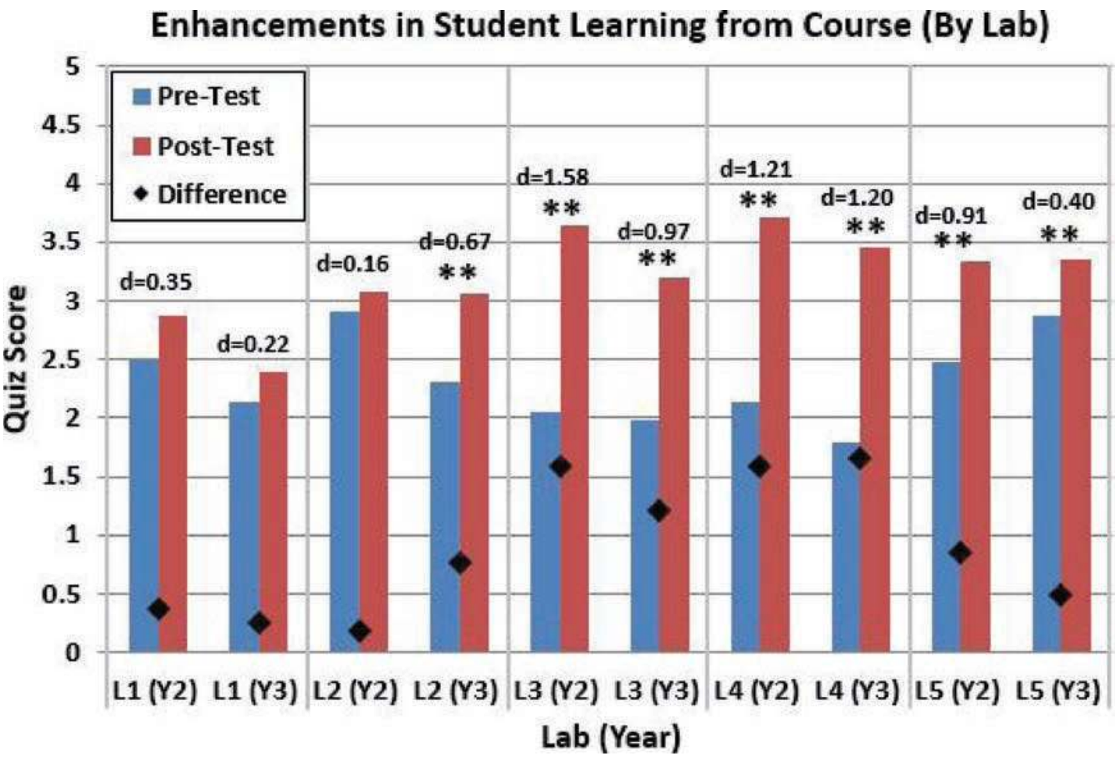

*Figure 7. The means (out of 5) of all students' pre-test score compared with their post-test score for years 2 and 3. Significant differences in quiz scores from the paired t-test are denoted with a** at $\alpha = 0.05$. The effect size, d, between the two means is shown above the bars, and the difference between means is represented by the black diamond data points.*

FORMAL ASSESSMENT OF LEARNING OPPORTUNITIES: RESULTS AND DISCUSSION

The second question we addressed provides a more in-depth analysis of when student learning was occurring throughout the semester and explores the cumulative learning enhancements for each of the course components. Namely, we explore the effects of the in-class lecture, the in-class and pre-lab lectures, the lectures and the lab activities, and the lectures, lab exercises, and lab reports.

Educational Benefit of Lectures

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

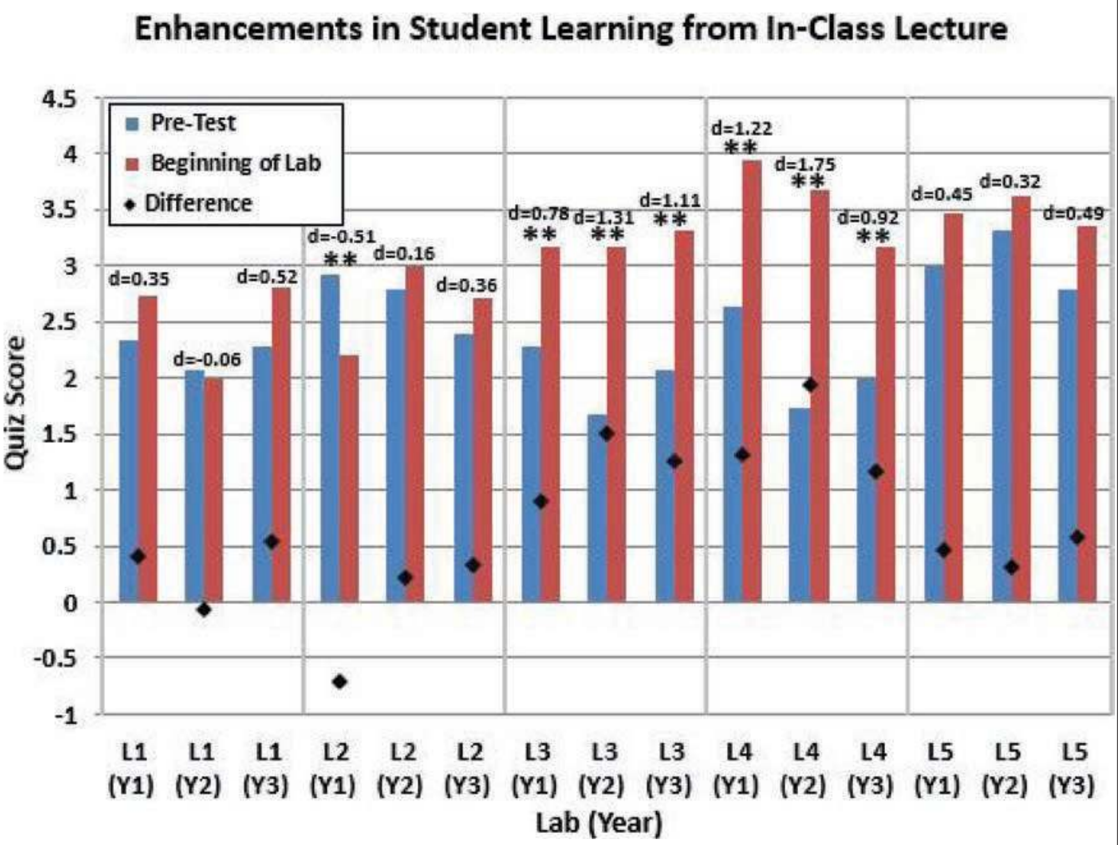

Figure 8. The means of the appropriate student section's quiz score on the pre-test compared with the quiz score taken at the beginning of lab for year 1 (Y1), year 2 (Y2), and year 3 (Y3). Significant differences in quiz scores from the paired t-test are denoted with a ** at $\alpha = 0.05$. The effect size, d, between the two means is shown above the bars, and the difference between means is represented by the black diamond data points.

at the very beginning of the lab. We also calculated the effect size, *d*, between the difference in means from pre-test to the very beginning of lab. 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 5.2, are shown in Figure 8.

Before students performed a lab exercise, they were given a short pre-lab lecture, which provided an introduction to the lab exercises and the main concepts associated with it. While our study design does not enable us to assess the value of this pre-lab lecture independently, we 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 by computing the effect size between the two appropriate means. 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 Figure 9 and are discussed below in Section 5.2 with results from the in-class lecture analysis.

Discussion of Educational Benefit from Lectures

From Figure 8, which shows the value of the in-class lecture alone, we observe that students scored significantly higher on quizzes associated with content from Labs 3 and 4 even before participating in the lab experience. Strong effect sizes ($d > 0.8$) were also observed in these two labs. 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 (and a moderate effect size ($d > 0.5$)) from pre-test to the beginning of lab for content associated with Lab 2. This suggests that students may have become confused by the in-class lecture on this material. In Y2 and Y3, however, we observe increases in quiz scores for this material, though they are not significant and the effects are small. Taken together, these results suggest that the lecture itself, while beneficial, was simply not enough in enhancing student understanding of majority of the material and reiterate the need for additional learning opportunities outside of the lecture.

From Figure 9, we observe that students appeared to listen and benefit from the pre-lab lectures, as reflected in the significantly higher quiz scores after hearing the pre-lab lecture in all of the labs except Lab 1, in at least one of the three years of data collected. Moderate to large effect sizes ($d > 0.5$) were also observed in these same labs. Though we cannot directly decouple the inclass 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 (Figure 8) and the results including both the in-class and pre-lab lectures (Figure 9). From these two figures, we see that students performed significantly better on more of the quizzes after the pre-lab lecture than after the in-class lecture. We suspect that the discrepancies in quiz performance between years is due in part to different TAs providing the lectures, in addition to external factors such as student population, student interest in a given lab activity, student listening and recall skills, among others. We note, however, that we explored the effects of different TAs in our linear regression models and did not find this to be a confounding factor (see Section 5.7). Overall, these results suggest that a pre-lab introduction may also be a useful component to the lab experience and may offer concise, repetition of main concepts.

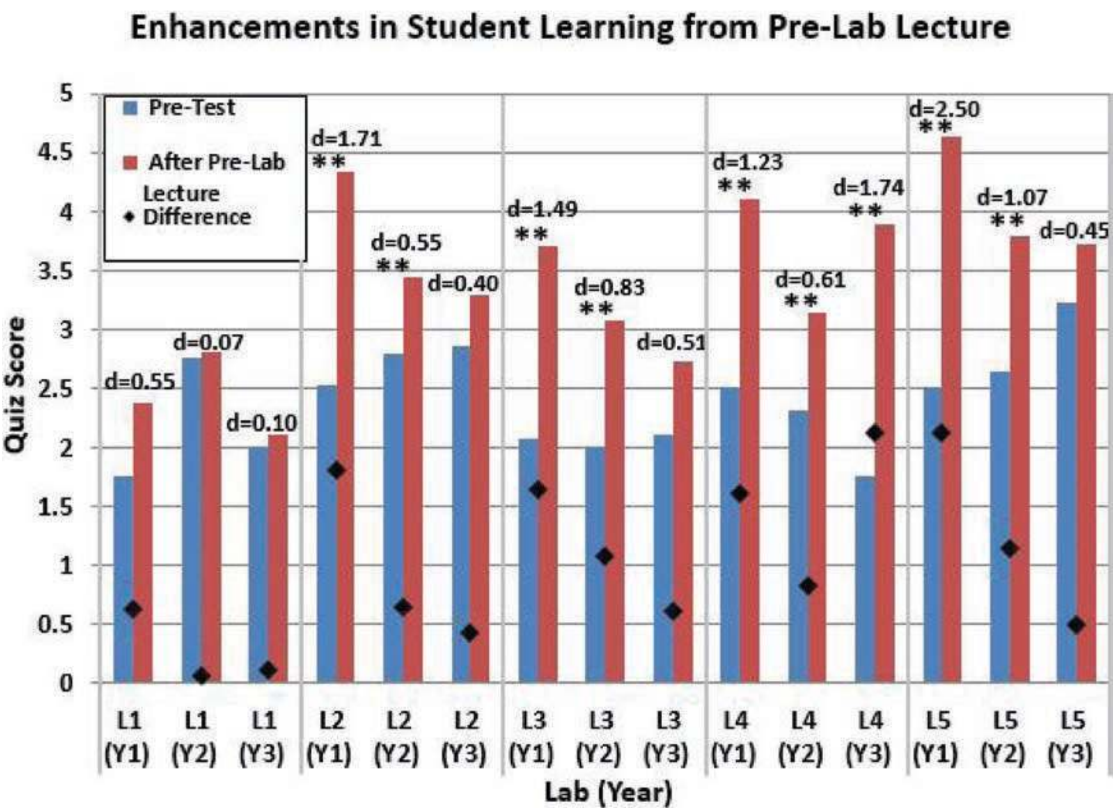

***Figure 9.** The means of the appropriate student section's quiz score on the pre-test compared with the quiz score taken after the pre-lab lecture for year 1 (Y1), year 2 (Y2), year 3 (Y3). Significant differences in quiz scores from the paired t-test are denoted with a ** at $\alpha = 0.05$. The effect size, d, between the two means is shown above the bars, and the difference between means is represented by the black diamond data points.*

Educational Benefit from Lectures and Lab

We were particularly interested in this timepoint immediately after the lab activity, as it encompassed the haptic paddle lab exercises. While our study design does not enable us to assess the lab activities themselves, this time point does enable us to explore the cumulative learning enhancements from the lectures and the lab activities. To do 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. We also computed the effect size, *d*, for the difference in means from pre-test to after lab. For both analyses, we compared student section 3's 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 Figure 10.

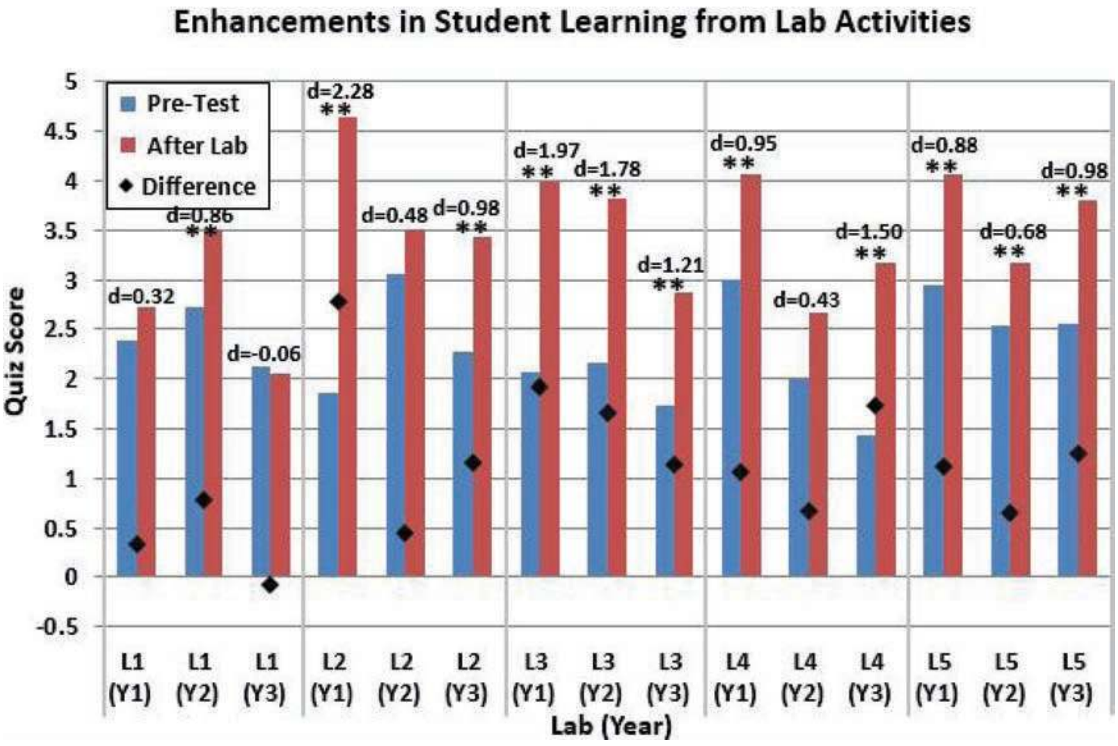

Figure 10. The means of the appropriate student section's quiz score (out of 5) on the pre-test compared with the quiz score taken after completing the lab for year 1 (Y1), year 2 (Y2), and year 3 (Y3). Significant differences in quiz scores from the paired t-test are denoted with a ** at $\alpha = 0.05$. The effect size, d, between the two means is shown above the bars, and the difference between means is represented by the black diamond data points.

Discussion of Educational Benefit from Lectures and Lab

The results from this second study shed light on the cumulative enhancements of the in-class lectures, pre-lab lectures, and lab activities. In comparing the pre-test scores with the quiz scores after completing the lab activities, we observe that students achieved significantly higher scores, for material in all of the labs, in at least one of the three years of data collected (see Figure 10). Large effect sizes ($d > 0.8$) were also observed for the content of each lab in at least one of the three years of data collected. Looking at each year individually, we observe that students achieved significantly higher quiz scores on material associated with 4 of the 5 labs for Y1 and Y3 and 3 of the 5 labs for Y2. A similar trend was observed in looking at effect sizes, with moderate to large effects ($d > 0.5$) being observed for content associated with 4 of the 5 labs for Y1 and Y3 and 3 of the 5 labs for Y2. For further insight into these results, we look at the content associated with each of the labs separately.

We begin with the material associated with Lab 1, which appears to be the most challenging for students to understand, as it was the only lab that did not have a significant increase in quiz score immediately after completing the lab exercise in Y1 or Y3. A significant increase was observed in Y2, however no significant increase was observed when comparing the pre-test scores to the post-test scores for Lab 1 in Y2 (see Figure 7). A similar trend was observed after the in-class lecture and the pre-lab lecture timepoints. Because we observe no change in observations after the Lab 1 timepoint, these results suggest that Lab 1 would benefit from further improvements to enhance student understanding and retention of the material. Because this was the students first lab experience, it is also possible that no measureable enhancements were observed due to external factors such as students becoming acclimated to the lab setup and becoming familiar with the course, instructor, and the TA.

For content associated with Lab 2, we observe that there was a large significant increase from pre-test to after lab for Y1 and Y3, but there was not a significant increase in quiz score from pre-test to after lab for Y2 (see Figure 10). Significance in Lab 2 in only some of the years was also observed after the in-class lecture timepoint and the pre-lab timepoint, though we note that the years were different between the two and between this after lab timepoint. Some changes were made between the three 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 an unpaired, two-sided t-test at the 95% confidence level comparing the mean pre-test scores for Lab 2 between all three years revealed that the Y2 Lab 2 pre-test score was significantly higher than the Y1 Lab 2 pre-test score (p-value = 0.02) and significantly higher than the Y3 Lab 2 pre-test score (p-value = 0.05). This suggests that the students from Y2 had a better understanding of the Lab 2 material at the beginning of the course compared to the students in Y1 and Y3. A similar trend was observed in the assessment of the educational benefit from pre-test to post-test for the content associated with Lab 2 (see Figure 7), as the pre-test score for Y2 was significantly higher than the pre-test score for Y3 (p-value = 0.01, from an unpaired, two-sided t-test at the 95% confidence level). Thus, part of the reason we may not observe a significant increase in Y2 Lab 2 after the in-class lecture, pre-lab lecture, and lab timepoints, may be due to the fact that students already knew a large portion of the material initially. These results, however, suggest that Lab 2 could also be a focus for future improvements to promote further learning enhancements supplementing the in-class and pre-lab lectures.

The results for content associated with Labs 3, 4, and 5 show significant cumulative learning enhancements from pre-test to after lab. For the Lab 3 content, we observe significant increases in quiz scores and large effect sizes ($d > 0.8$) from pre-test to after lab (see Figure 10) in all

3 years and from pre-test to post-test (see Figure 7). These positive results are also observed at earlier timepoints, both after the in-class lecture and after the pre-lab lecture, though we note that the effect sizes are larger after the lab timepoint. For the Lab 4 content, we observe a significant increase in quiz score and a large effect size from pre-test to post-test in Y2 and Y3 (see Figure 7) and from pre-test to after lab in Y1 and Y3. These findings are consistent with the observations made after the in-class lecture and pre-lab lecture timepoints. 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, and has a small to moderate effect size ($d > 0.2$). This result may be due in part to the fact that the sample size for this particular lab was relatively small due to several students switching lab sessions or not completing the quiz. The significant increase in Lab 4 in Y2 from pre-test to post-test, however, suggests that majority of the students learned the Lab 4 material, perhaps benefiting especially from the lab report and lecture discussions following the lab. For the content associated with Lab 5, we observe a significant increase in quiz score and moderate to large effect sizes ($d > 0.5$) from pre-test to after lab in all three years and from pre-test to post-test. Comparing these findings with the results from previous timepoints, we observe no significant increases in student learning after the in-class lecture timepoint, and only 2 of the 3 years had significant increases in student learning after the pre-lab lecture timepoint. While we cannot decouple the value of the Lab 5 activity itself, this additional enhancement after the lab activity is encouraging.

Educational Benefit of Lectures, Labs, and Lab Reports

Finally, to assess the value of all learning components (lectures, labs, and lab reports), we conducted a paired t-test with unequal variances to compare the mean quiz score from the pre-test to after the lab report. We again computed the effect size, d, of the difference between the two means. From Table 1, the pertinent data used in this analysis was student section 4's scores for Lab 1, student section 3's scores for Lab 2, and so on. The results, also discussed in more detail in Section 5.6, are shown in Figure 11.

Discussion of Educational Benefit from All Learning Opportunities

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 and computed corresponding effect sizes. From Figure 11, we again see that students scored significantly higher on quizzes for all of

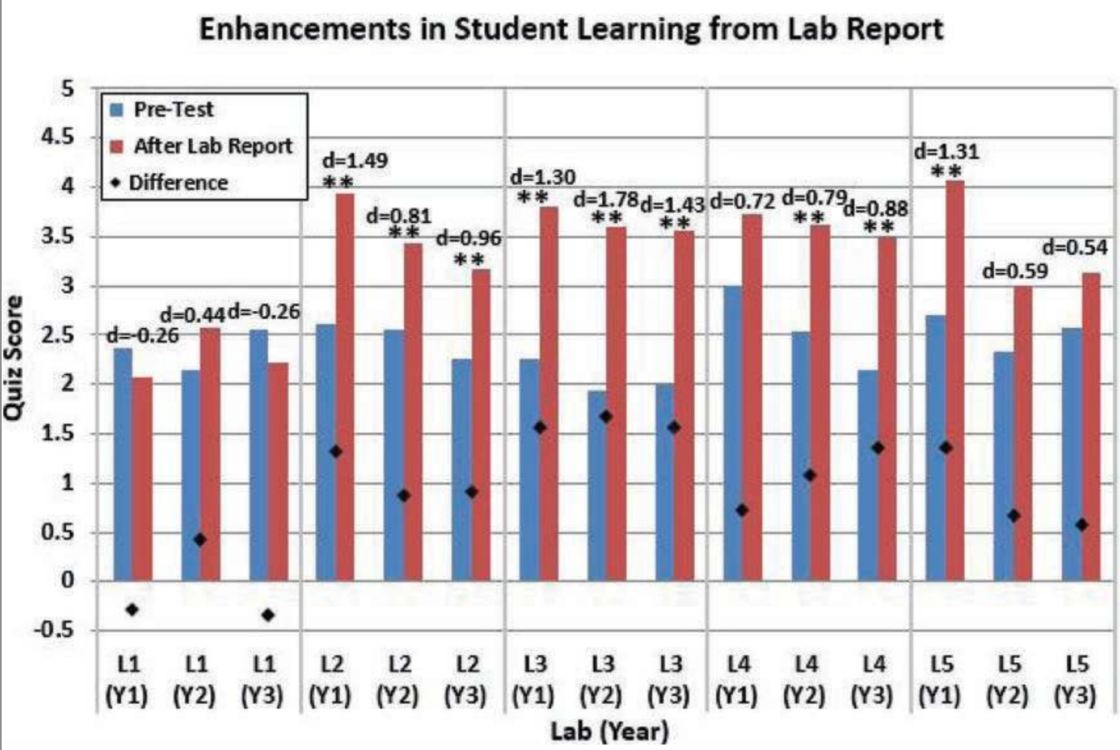

Figure 11. The means of the appropriate student section's quiz score on the pre-test compared with the quiz score taken after completing the lab report for year 1 (Y1), year 2 (Y2), year 3 (Y3). Significant differences in quiz scores from the paired t-test are denoted with a ** at $\alpha = 0.05$. The effect size, d, between the two means is shown above the bars, and the difference between means is represented by the black diamond data points.

the labs in at least one of the three years of data collected, except for content associated with Lab 1. Large effect sizes ($d > 0.8$) are also observed in at least one of the three years of data collected in all of the labs except for the content associated with Lab 1. Our study architecture limits us from being able to decouple the enhancements attributed to the lab report individually, however, we note that the sustainability of significant learning enhancements associated with content of Labs 2-5 is a positive outcome, and the fact that Lab 1 still does not result in significant learning enhancements reiterates the need for improvements in the learning opportunities provided with the Lab 1 content.

Effects of Teaching Assistants

Though care was taken to make the classroom and lab environments as controlled as possible, some of the data collected exhibits large variances from year to year, particularly the data from the after lab timepoint, shown in Figure 10. These variances may be due to several external factors and/or the contributions of each summed together. These factors could include the dynamics of student teams, the lab environment, the individual students and their interest and motivation in any given lab exercise or course content, the TA, and many others. As with any classroom hands-on activity, the authors did the best they could to minimize the effects of external variables, but the classroom is a dynamic environment. The one factor the authors could control was the presence of a different lab TA from year to year. While it is certainly possible that the TA may have contributed to the variances observed, we note that the TA did not vary within sections in any given year, but only between years. A regression analysis was done to ensure that the difference in TAs from year to year was not a confounder in our above data analysis.

To do this, we created univariate regression models exploring the relationship between quiz score and lab exposure (model 1) and quiz score and TA exposure (model 2). We then created a multivariate model exploring the relationship between quiz score and both lab exposure and TA exposure as factors (model 3). We did this for two cases: (1) comparing the after in-class lecture/very beginning of lab timepoint with the after lab timepoint and (2) comparing the after pre-lab lecture timepoint with the after lab timepoint. We chose to compare these two cases, as they both encompass when the students would have interacted with the TA in the lab activities. From this analysis, we observe that the coefficient estimates for the factors in both the univariate and multivariate regression models (model 1 vs. model 3 and model 2 vs. model 3) are nearly the same and within the confidence intervals. This was true for both of the cases explored. Additionally, we note that p-values for all factors are nearly the same in both the univariate and multivariate models for each case. For these reasons, we conclude that there is very little confounding occurring for lab exposure and TA exposure in our data, and that perhaps, the variances can be attributed to external factors beyond the control of the study design. This variance, however, isn’t necessarily a negative outcome of the study, as the purpose of our analyses was to shed light on student learning throughout the semester and sequence of learning opportunities, and it may be representative of the variability that will exist in the implementation of such hands-on activities at other institutions.

Summary of Results

We now summarize our findings and the key take-aways from the formal assessment.

- Overall, we found that the series of lectures and haptic paddle labs (including the pre-lab introduction, the lab activity, and the lab report) were 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 in one of the three years of data collected and in 3 of the 5 labs in the other two years. A similar trend was observed when looking at effect sizes of the differences in quiz scores after completing the lab activity. Beginning to end of semester comparisons also support this observation, resulting in significance on the cumulative post-test compared with the pre-test in both years.

- We found that the in-class lecture alone, though beneficial, is not sufficient for enhancing student understanding of the material, as reflected by students only scoring significantly higher on quizzes relating to 2 of the 5 labs, (Labs 3 and 4). Large effect sizes for the in-class lecture comparison were also only observed for Labs 3 and 4.
- While our current study setup cannot tease out the effects of the contributions from the haptic device itself or the individual lab activities, our results emphasize the importance of well-designed active, hands-on learning activities. An interesting future study would be to make a comparison between this set of lab activities with similar well-designed, non-haptic lab activities to shed light on the values of the haptic component explicitly.

CONCLUSION

In this paper, we have introduced a new, robust, inexpensive design of the haptic paddle, a force feedback device which has been adopted by several universities in teaching System Dynamics. Our haptic paddle relies on a friction drive, which we have experimentally shown is comparable in performance to the original, widely accepted, capstan drive, but is much more robust to classroom use. Further, by using the low-cost Arduino microcontroller for communication, our complete haptic paddle kit can be constructed for less than $100 including all electronics except a computer, and can be operated from a laptop, making it more portable than prior haptic paddle systems. We also transitioned the software from its original C-executable files over to Matlab and Simulink, software that enables students to take on a much more independent role in programming their haptic paddle and provides a convenient, engaging user interface.

We have also formally assessed the benefits of the haptic paddle laboratories combined with a traditional lecture-style course, probing both what material students are learning and when they are learning it. Our formal assessments, using 3 years of student data, suggests that the haptic paddle laboratories, including the pre-lab lecture, the lab activity, and the lab report are successful in enhancing student understanding of core concepts in this course. The results of our study demonstrate that well-designed, supplementary hands-on activities like the haptic paddle laboratories enhance the in-class lecture and significantly increase student performance on conceptual quizzes. These results, combined with prior assessments of the haptic paddles (Okamura, Richard, and Cutkosky

2002; Bowen and O'Malley 2006b), 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.

In order to encourage the adoption of the haptic paddle by other educators and interested university or K-12 students, we have developed a comprehensive website containing all of the information one needs to build the haptic paddle and conduct the lab exercises (VU Webpage 2012). This website contains all of the part files required to manufacture the paddle, a complete bill of materials and assembly guide for constructing the paddle, all of the lab handouts and lab report questions, all of the Arduino and Simulink files needed to complete the lab exercises, and all of our assessments. In addition, with support from The MathWorks, Inc., we have made an introductory video to the haptic paddle labs, which provides a discussion of the hardware and software of the paddle, examples of using Real Time Workshop in Simulink in combination with external hardware, an overview of the lab exercises, and our "lessons learned" on using the haptic paddle laboratories. We are also working with collaborators at California State University Long Beach to implement the haptic paddle in a freshman introduction to engineering course and in a graduate level course on teleoperation. The material developed for these courses will be made freely available on our website in the near future.

Our analyses also enable us to pinpoint areas for future improvement for the course and the haptic paddle lab exercises. In subsequent years, our primary focus will be on revising content associated with 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. Another area of future work is to take advantage of the flexibility and functionality of Simulink to provide simulations of additional dynamic systems beyond the mass-spring-damper system already used in the labs. We also plan to explore how the haptic paddle can be used in teaching other subjects at both the university and K-12 level, and will work with educators to develop lab modules around these ideas. One idea is to incorporate the haptic paddle in a physics lesson and compare its effectiveness using the standardized Force Concept Inventory as our assessment. Finally, we intend to develop new assessments of the haptic paddle that will enable exploration of the lab activities themselves, as well as the "haptic" interaction explicitly. 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.

ACKNOWLEDGMENTS

The authors wish to thank several people who have contributed to the development of the haptic paddle laboratories and to the collection and analysis of assessment data, especially Dr. Thomas Withrow, Dr. Caleb Rucker, 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

Abbott, J. J. (2005). Virtual fixtures for bilateral telemanipulation. *Ph.D. Dissertation submitted to The Johns Hopkins University*, 195–199.

Abdulwahed, M. and Z. Nagy (2009). Applying Kolb's experiential learning cycle for laboratory education. *Journal of Engineering Education 98*, 283–294.

Arduino (2012). Arduino Development Environment. *http://arduino.cc/en/Guide/Environment Accessed June 2012*.

Bowen, K. and M. K. O'Malley (2006a). 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)*, 147–152.

Bowen, K. and M. K. O'Malley (2006b). Haptic interfaces for a labview-based system dynamics course. *American Society of Engineering Education Conference*, 2627.

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

Cox, D. (2008). Hands-on experiments in dynamic systems and control for applied education in robotics and automation. *World Automation Congress*, 1–6.

Dewoolkar, M. M., L. George, N. J. Hayden, and M. Neumann (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*, 161–175.

EduHaptics (2012). EduHaptics Webpage. *http://eduhaptics.org/index.php/HapticDevices/HapticPaddles Accessed June 2012*.

Feisel, L. and A. Rosa (2005). The role of the laboratory in undergraduate engineering education. *Journal of Engineering Education 94 (1)*, 121–130.

Felder, R. M. and L. K. Silverman (1988). Learning and teaching styles in engineering education. *Engineering Education 78(7)*, 674–681.

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

Gassert, R., J.-C. Megzger, K. Leuenberger, P. W. L., M. R. Tucker, B. Vigaru, R. Zimmermann, and O. Lambercy (2013). Physical student-robot interaction with the ethz haptic paddle. *IEEE Transactions on Education 56*, 9–17.

Gillespie, R. B., M. B. Hoffman, and J. Freudenberg (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)*, 410–415.

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

Gorlewicz, J. L. and R. J. Webster III (2012). A formal assessment of the haptic paddle laboratories in teaching system dynamics. *American Society of Engineering Education Conference*.

Grow, D., L. N. Verner, and A. M. Okamura (2007). Educational haptics. *AAAI Spring Symposia - Robots and Robot Venues: Resources for AI Education*.

Okamura, A., C. Richard, and M. Cutkosky (2002). Feeling is believing: Using a force-feedback joystick to teach dynamic systems. *Journal of Engineering Education*, 345–349.

SensAble Technologies, I. http://www.sensable.com/documents/documents/STI Jan2009 DesktopOmniComparison print.pdf. Accessed February 2013.

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

VU Webpage (2012). *Vanderbilt University Haptic Paddle Webpage Accessed February 2014*. Wieman, C. and K. Perkins (November 2005). Transforming physics education. *Physics Today*, 36–41.

AUTHORS

Jenna L. Gorlewicz received her B.S. in mechanical engineering from Southern Illinois University Edwardsville in 2008 and her PhD in mechanical engineering from Vanderbilt University in 2013. At Vanderbilt, she was a National Science Foundation Fellow and a Vanderbilt Educational Research fellow. In Fall 2013, Jenna returned to Southern Illinois University Edwardsville as a faculty member in the Mechanical and Industrial Engineering Department. She is currently the director of the Intelligent Mechatronic, Haptic, and Robotic Systems (IMeHRS) Lab. Her research interests are in medical robotics, haptic devices, human-machine interaction, educational robotics, and novel learning technologies. Corresponding author: jgorlew@siue.edu

Louis B. Kratchman received the B.A. degree in psychology and B.S. degree in mechanical engineering from the University of Michigan, Ann Arbor, in 1998 and 2009, respectively. He is currently working toward the Ph.D. degree in mechanical engineering from Vanderbilt University, Nashville, TN, where he has taught robotics and system dynamics. His research interests include medical robotics, haptics, and image-guided surgery.

Robert J. Webster III received his B.S. degree in electrical engineering from Clemson University in 2002 and his M.S. and Ph.D. degrees from the Johns Hopkins University in 2004 and 2007, respectively. In 2008, he joined the faculty of Vanderbilt University where he is now an Associate Professor and directs the Medical & Electromechanical Design Laboratory. Dr. Webster is a co-founder and steering committee member of the Vanderbilt Initiative in Surgery and Engineering (ViSE), is a recipient of the IEEE Volz Award for PhD thesis impact in robotics, a National Science Foundation CAREER award winner, and a winner of the Excellence in Teaching award from Vanderbilt University. He is an Associate Editor for IEEE Transactions on Robotics and Chair of the SPIE Image-Guided Procedures, Robotic Interventions, and Modeling conference. His current research interests include medical robotics, image-guided surgery, continuum robotics, and educational robotics.