

Characterization of time scale for detecting impacts of reforms in an undergraduate physics program

Eleanor W. Close, Jean-Michel Mailloux-Huberdeau, Hunter G. Close, and David Donnelly
Department of Physics, Texas State University, 601 University Drive, San Marcos, TX, 78666

Over the past five years, we in the physics department at Texas State University have reformed our calculus-based introductory sequence by implementing a Learning Assistant program, introducing research-based instructional methods, and increasing faculty professional development and instructional mentoring. Students in all sections of these courses now spend a significant portion of class time in small groups, often working through materials from the curricular supplement “Tutorials in Introductory Physics.” Over this time, students’ normalized gains on the Force Concept Inventory (FCI) have increased dramatically and the DFW rate (the percentage of students receiving a grade of D or F, or Withdrawing from the course) has been cut in half. In addition, we have seen that FCI gains for individual faculty increase gradually over several semesters, for both novice and experienced instructors. These data suggest the need for patience with programmatic reform efforts, and for stability in instructor-course assignments.

I. INTRODUCTION

The calculus-based introductory physics course sequence at Texas State University (TXST) has been restructured over the past five years to include research-based curriculum and instructional methods with support from trained undergraduate Learning Assistants (LAs). The TXST LA program was created in 2012 based on the model developed at the University of Colorado (CU) Boulder [1] and the successful LA program at Seattle Pacific University [2]. We have previously reported on the impact of these changes on students who serve as LAs [3]; in this study, we analyze data gathered over 13 semesters documenting the impact on students in the LA-supported courses.

Our data are consistent with the well-established positive impact of research-based curriculum and interactive instructional methods (e.g., [4,5]). In addition, as previously shown with data from large-scale implementation of course reforms at CU-Boulder [6], our data show instructor improvement over time.

II. CONTEXT

A. Institutional growth

TXST is a Hispanic Serving Institution (HSI) with an enrollment of approximately 39,000 students in Fall 2016, of whom 34,000 were undergraduate students. The university has grown rapidly in recent years, increasing enrollment from approximately 28,000 in 2007, (24,000 undergraduates). During the decade from 2007 to 2016, the number of students in the TXST College of Science and Engineering has more than doubled (from 3,000 to 6,200), with especially fast growth in the recently-established Ingram School of Engineering (from 200 majors in 2007 to 850 in 2016). The demographics of the university have also changed: TXST was granted federal HSI status in 2011 with 26% Hispanic enrollment; in 2016 Hispanic enrollment was 35%.

B. Physics department instructional reforms

The calculus-based introductory physics sequence at TXST primarily serves students majoring in physics, chemistry, and engineering. Physics faculty have been using the Force Concept Inventory (FCI) [7] to assess student conceptual learning in the introductory calculus-based mechanics course (hereafter referred to as “mechanics”) since Spring 2011. The current study is based these data.

In Spring 2012 we piloted an LA program in one section of mechanics. The program expanded to serve all sections of mechanics in Fall 2012, and over the following three semesters expanded to serve all sections of all three courses in the calculus-based introductory sequence. All faculty who teach in this course sequence receive LA support.

Students enrolled in mechanics attend regular class sessions (“lecture”) for three hours each week, and attend a weekly three-hour laboratory. Most laboratory sections are taught by undergraduate students. Prior to 2012 both lab and lecture were traditionally taught, with primarily verification experiments in the lab and instructor-delivered lecture during class sessions. The course does not have an associated recitation section. In the interests of space we will not describe the demographics of the course here; this information can be found in ref. 3.

1. LA program implementation

Because the mechanics course structure does not include a recitation section, LA support for interactive reforms was implemented primarily in the lecture sections, mostly through the introduction of *Tutorials in Introductory Physics* [8] and other similar materials. From three to five LAs assist in each section, depending on student enrollment. In addition to assisting in class, many LAs staff the Physics Help Center, a walk-in tutoring center located near physics faculty offices that was created shortly after the introduction of Tutorials. The lab curriculum was also

modified in 2012 to include tutorials in selected weeks, and the physics department preferentially hires LAs as lab instructors; all mechanics labs are now taught by current or former LAs.

The TXST LA program structure includes collaboration across sections and courses: all physics LAs and LA-supported faculty meet for weekly content preparation sessions at the same time, and begin together in the same room, before splitting up by course to prepare for instruction. Some LAs assist in multiple sections of the same course, working with different instructors. In addition to supporting a sense of community among LAs, this creates a structure for faculty collaboration and mentoring, described below.

2. Instructor collaboration and mentoring

The LA program structure creates a constraint on faculty pacing: all LAs prepare for the same instruction at the same time. While instructors have significant freedom in shaping their use of class time, all instructors are expected to complete the same set of tutorials and to follow approximately the same pacing. They are additionally encouraged to use research-based instructional strategies beyond tutorials.

Mechanics faculty hold a brief (15-30 min.) weekly meeting before the preparation session to discuss progress; this meeting provides a regular context for faculty to discuss classroom issues, and for less experienced instructors to receive mentoring from colleagues. This has been an important mechanism for supporting novice instructors and faculty who are unfamiliar with research-based instructional methods and curricular materials.

III. METHODS

A. Conceptual learning gains

The FCI was administered as a pretest very close to the start of the semester, and as a posttest at some time after completion of instruction on Newton's Laws. The specific timing and format of the posttest administration was left up to individual instructors. For most semesters, the assessment was completed using paper and pencil either in class or in lab. Starting in Spring 2016, faculty had the option of administering the FCI online using the LASSO tool provided by the Learning Assistant Alliance. Incentives given for completing the FCI varied by semester and by instructor.

FCI data were filtered so that only data from students who had both a pretest and a posttest score were used in the analysis. The normalized gain ($\langle g \rangle$) was calculated using the formula [4]:

$$\langle g \rangle = \frac{\bar{x}_{post} - \bar{x}_{pre}}{100\% - \bar{x}_{pre}}$$

where \bar{x} represents the mean score on either the pre- or posttest. (For analysis of these data comparing $\langle g \rangle$ to the Cohen's d effect size, see [9].)

B. Student failure rates

DFW rates (the fraction of students receiving a grade of D or F, or Withdrawing from the course) were calculated for each section and by semester using institutional research data. Grades of I (Incomplete) were excluded from these calculations. While a grade of D counts toward graduation, it does not fulfill prerequisite requirements, therefore most students who receive a D grade in mechanics (or other major courses) must re-take the course. The overall DFW rate for each semester was calculated using a weighted average of course section DFW rates.

IV. RESULTS & DISCUSSION

A. Enrolled student population

Table I shows the number of students enrolled in mechanics each semester over the six years of FCI data collection, and the number who completed each FCI administration. Note that the total number of students per semester has more than doubled since 2011.

TABLE I. Enrollment and FCI participation 2011-2016. Columns show number of course sections, number of students enrolled in intro calculus-based mechanics, and number who completed FCI pre-test, post-test, and both (matched). Semesters are noted by calendar year and order, e.g., 2011.1 is the Spring 2011 semester.

Semester	Course Sections	Number of students			
		Enrolled	Pre	Post	Matched
2011.1	3	119	113	80	74
2011.2	3	144	129	92	77
2012.1	3	160	140	111	91
2012.2	4	160	133	135	108
2013.1	4	171	154	108	91
2013.2	4	203	190	166	153
2014.1	4	215	201	185	171
2014.2	4	232	216	198	182
2015.1	5	239	221	204	186
2015.2	5	229	210	199	180
2016.1	5	235	181	151	143
2016.2	5	261	209	192	170
2017.1	5	230	207	196	173

In order to determine whether the average preparation of the students enrolled in mechanics had changed over this time, we plotted FCI pretest scores by semester, averaged over all matched students (see Figure 1). Based on this plot, we approximate student preparation for the mechanics course as stable over the time span of our data collection.

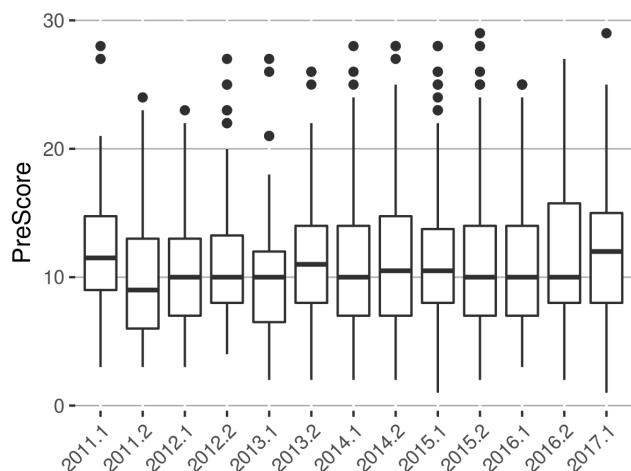


FIG 1. Box plot of mechanics FCI pretest scores (out of 30 points) by semester. Bold lines represent median scores; rectangles span from first to third quartile. Whiskers describe values up to a maximum of 1.5 times the inner quartile range (the box), points represent outliers. This plot shows stable levels of student preparation during program growth. Semesters are noted by calendar year and order, e.g., 2011.1 is the Spring 2011 semester.

B. Impact by semester

We examined the overall trend in student conceptual learning, as measured by $\langle g \rangle$. As shown in Figure 2, average student learning gains in 2011, prior to LA program implementation and associated course reforms, were below the $\langle g \rangle = 0.22$ average for lecture-based instruction found by Von Korff et al. in their meta-analysis of data from 50K students [5]; in all semesters subsequent to LA program implementation, average $\langle g \rangle$ has exceeded this figure, and in several semesters has exceeded the $\langle g \rangle = 0.39$ average for interactive engagement instruction [5]. Error bars in Figure 2 represent the 95% Confidence Interval for each semester, calculated by finding the pooled standard error of the average normalized gain $\langle g \rangle$ (see [10]) and multiplying by 1.96. The increase is not monotonic; we will discuss the variation in the post-reform data in the next section.

Figure 3 shows the mechanics DFW rate by semester, beginning in Spring 2011. (Note that the high DFW rate in 2011 is not representative of the long-term trend before LA program implementation—these numbers were unusually high for the course. From 2003 to 2010, typical DFW rates were between 25% and 40%.) The overall downward trend in DFW rate, in combination with the upward trend in $\langle g \rangle$, provides evidence that the course reform effort has succeeded in improving students' experience of mechanics.

One notable feature of Figure 2 is the timescale: the average $\langle g \rangle$ continued to increase over several semesters after course reforms were first implemented in Spring 2012. Also, as mentioned above, the trend is not uniform over time: $\langle g \rangle$ decreased for three consecutive semesters beginning in Spring 2015, then increased again. In order to

examine the variation in $\langle g \rangle$ more closely, we analyze the student data grouped by instructor and course section below.

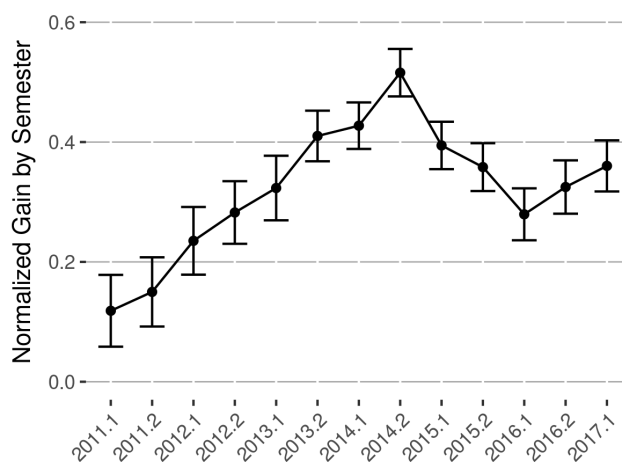


FIG 2. Graph of FCI normalized gain ($\langle g \rangle$) by semester, with 95% Confidence Intervals, shows increased student conceptual learning after course reforms implemented in Spring 2012 (labeled 2012.1).

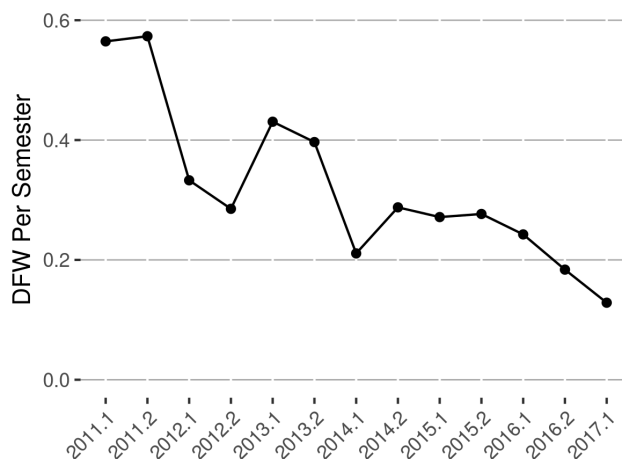


FIG 3. DFW rate (fraction of students who withdrew from course or receive grade of D or F) by semester—shows increased student success after course reforms implemented in Spring 2012 (labeled 2012.1).

C. Analysis by course section and instructor

Over the six years since we began FCI data collection, a total of 12 instructors have taught at least one section of mechanics, including lecturers and tenured or tenure-track professors. Some instructors had little or no prior teaching experience, while others had been teaching for many years; instructors' initial knowledge of research-based instructional strategies ranged from none to extensive experience in this area. As can be seen by the trend lines in

Figure 4, nearly all instructors' average $\langle g \rangle$ increased over multiple semesters, in some cases quite dramatically. This is true both for instructors who began teaching the course before instructional reforms were implemented, and for those who began when the LA program was well-established.

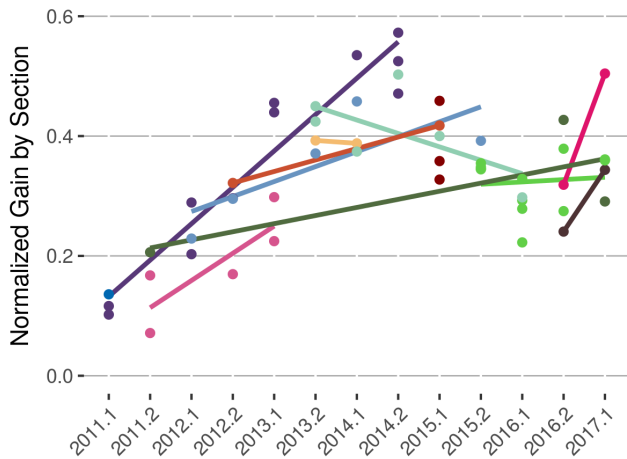


FIG 4. FCI $\langle g \rangle$ by instructor and section. Each color represents an instructor (N=12); the trend line for each instructor who taught two or more semesters (N=10) is shown in the same color. Semesters are noted by calendar year and order, e.g., 2011.1 is the Spring 2011 semester.

Similarly, Figure 5 shows that nearly all instructors experienced a downward trend in DFW rate over time. This strongly suggests that instructors become more effective users of research-based curriculum and methods as they gain experience, which has implications both for instructional assignments and for expectations of timescale for reform efforts.

The impact of instructor experience also provides insight into a possible cause of the dip in average $\langle g \rangle$ between 2014 and 2016: over this time span, the instructors for the course changed almost entirely. As the new instructors gained experience, the learning gains began to increase again.

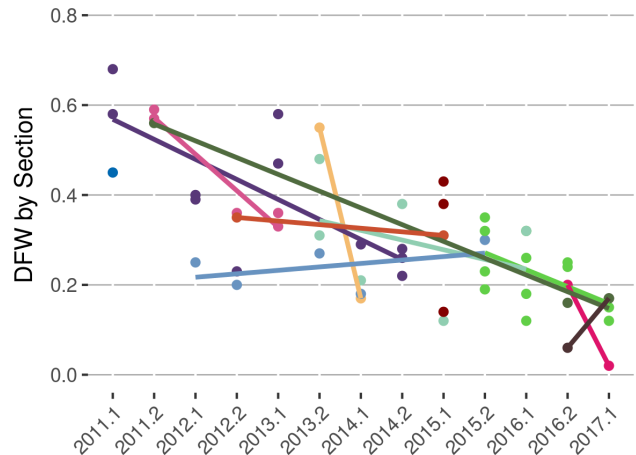


FIG 5. Fraction of students who withdrew from course or receive grades of D or F (DFW), by semester and section. Each color represents an instructor (N=12); the trend line for each instructor who taught two or more semesters (N=10) is shown in the same color. Semesters are noted by calendar year and order, e.g., 2011.1 is the Spring 2011 semester.

V. CONCLUSIONS

Because these data were not experimental, it is not possible to identify causal mechanisms for the observed positive changes. A number of factors changed during the time interval studied in parallel with the instructional reforms; in addition, students do not choose course sections at random, as suggested by the variation between sections shown in Figures 4 and 5. In spite of these limitations, the relationship between instructor experience (in the context of reformed courses) and student success is relevant for departments implementing course reforms. Our findings strongly suggest that persistence with reforms is rewarding, and also that instructor experience matters. These findings have implications for staffing practices, and demonstrate the value of long-term systematic data collection for programmatic improvements.

ACKNOWLEDGEMENTS

This work and the Texas State University LA Program are supported in part by NSF grants 1240036, 1557405, 0808790, and 1431578, and by the Halliburton Foundation.

[1] V. Otero, S. Pollock, & N. Finkelstein, *Am. J. Phys.* **78** (11), 2010.
 [2] E. W. Close et al., in *Effective Practices in Preservice Phys. Teacher Ed.: Recruitment, Retention, and Prep.* edited by E. Brewe and C. Sandifer. APS, 2015
 [3] E. W. Close, J. Conn, & H. G. Close, *Phys. Rev. Phys. Educ. Res.* **12**, 010109, 2016
 [4] R. Hake, *Am. J. Phys.* **66**(1), 1998.

[5] J. Von Korff et al., *Am. J. Phys.* **84**, 2015.
 [6] S. Pollock & N. Finkelstein, *2012 PERC Proceedings*.
 [7] Hestenes, Wells, & Swackhammer. *The Phys. Teach.*, **30**, 1992.
 [8] L. McDermott et al., *Tutorials in Introductory Physics*
 [9] D. Donnelly et al., *2017 PERC Proceedings* (in press).
 [10] T. Smith, M. Wittmann, & T. Carter, *Phys. Rev. ST Physics Educ. Res.*, **10**, 020102, 2014.