ASEE's 123rd Annual • Conference & Exposition • New Orleans, LA • June 26-29, 2016

New Orleans JAZZED about Engineering Education

Paper ID #15665

Probing the Flipped Classroom: Results of A Controlled Study of Teaching and Learning Outcomes in Undergraduate Engineering and Mathematics

Dr. Nancy K. Lape, Harvey Mudd College

Nancy K. Lape is an Associate Professor of Engineering at Harvey Mudd College.

Dr. Rachel Levy, Harvey Mudd College

Rachel Levy is an Associate Professor of Mathematics and the Associate Dean of Faculty Development at Harvey Mudd College. In addition to her work on fluid mechanics, she is an investigator on two NSF-funded education projects: one studying flipped classrooms and the other preparing teachers for mathematical modeling in the elementary grades. She is the Vice President for Education for SIAM, the Society for Industrial and Applied Mathematics and the founder of the blog Grandma got STEM.

Dr. Darryl H. Yong, Harvey Mudd College

Darryl Yong is currently a professor of mathematics and associate dean for diversity at Harvey Mudd College. His research interests relate to partial differential equations and the preparation and professional development of secondary school math teachers.

Ms. Nancy Hankel, Cobblestone Applied Research & Evaluation, Inc.

Ms. Hankel is a Research Associate II at Cobblestone and is currently pursuing an Ed.D. from UCLA in Educational Leadership. She manages multiple evaluation projects related to teacher training and professional development as well as various federally-funded STEM-focused programs at the post-secondary level. She has extensive experience in all phases of data collection (such as instrument development and administration, observations, focus group and individual interviews) as well as experience in site recruitment, developing logic models, quantitative and qualitative data analyses and reporting, and presenting results to a variety of audiences.

Dr. Rebecca Eddy, Cobblestone Applied Research & Evaluation, Inc.

Dr. Eddy received her doctorate in Applied Cognitive Psychology and has spent her career focused on applying the principles of learning and cognition to evaluation of educational programs. Her work includes published articles and client technical reports as President of Cobblestone Applied Research & Evaluation, Inc. and a faculty member at Claremont Graduate University. Work at Cobblestone focuses on advancing the numbers of underrepresented minority students in Science, Technology, Engineering and Mathematics (STEM) fields. Dr. Eddy has conducted evaluation or applied research studies on numerous university projects including clients programs funded by the National Science Foundation; U.S. Department of Education Title III and Title V; National Institutes of Health; Howard Hughes Medical Institute, among others. Dr. Eddy also trains professional evaluators from around the world as a faculty member at Claremont Graduate University in the Advanced Certificate in Evaluation Program.

Probing the Inverted Classroom: A Controlled Study of Teaching and Learning Outcomes in Undergraduate Engineering and Mathematics

Introduction

A flipped classroom reverses the paradigm of traditional lecture courses by delivering lectures outside of class – by means such as videos or screencasts – and using class meeting time for instructor-mediated active learning. This format has the potential to transform STEM education by increasing student time spent on what research has demonstrated to be the most effective teaching techniques (i.e. active learning) without sacrificing material coverage or educational scaffolding. Many educators are beginning to invert their classrooms, but there is limited (or no) data on learning gains currently available. We are rigorously examining the impact of three instructors inverting two STEM courses, in engineering (thermodynamics) and mathematics (differential equations), by measuring student learning gains and attitudes towards the course material. Our expected measureable outcomes are:

1. Higher learning gains;
2. Increased ability to apply material in new situations (transfer);
3. Increased interest in and positive attitudes towards STEM fields (affective gains); and
4. Increased awareness by students of how they learn and strategies that support their learning (metacognitive gains).

We hypothesize that increased student learning will arise primarily because of the additional time that students will have with instructors actively working on meaningful tasks in class. If our hypotheses prove true, that will have implications for institutions that are seeking to push more instruction online, where instructor-mediated learning is limited. In addition, because this study involves two different disciplines, the results may be applicable across STEM fields.

The inverted classroom model (and a traditional classroom model) was implemented at Harvey Mudd College over four years (AY 12-13 through AY 15-16) in two courses: Engineering 82 (a thermodynamics course) and Math 45 (introductory differential equations). The last three years of this research was supported by the National Science Foundation (NSF).

Using a variety of implementation and outcome measures (e.g., pre and post assessments of student surveys; content assessments; homework and course grades) the evaluation assessed the extent to which the inverted classroom model impacted students in three primary areas: Academic Learning Gains; Transfer of Knowledge; Metacognitive Gains. In addition, the evaluation also included student satisfaction and faculty experiences; these results are not included in this proceeding but are available upon request.

Method

Design

The quasi-experimental study design used during the first three years of the study was developed to compare students from inverted sections with those in control sections (i.e., traditional course model). Treatment and control students completed the same measures (e.g., content assessments and student attitude surveys) and faculty members, who taught in both conditions, also completed reflection papers related to their experiences. The guiding research questions for the study and an overview of the assessment measures are shown in Table 1 below (more details on assessment measures are included in a subsequent section of this paper). In the final year of the study, the researchers designed what they felt were "best practices" for the inverted model in all sections of their courses and the same outcome measures were used.

Table 1. Evaluation Questions and Outcome Measures

Evaluation Question	**Measure**	
	Engineering 82	**Math 45**
Implementation		
1. Do students in inverted classrooms spend additional time actively working with instructors on meaningful tasks in comparison to those students in control classrooms?	Student survey	
2. Do students in inverted classrooms actively participate and prepare for class through the videos and other materials?	Student survey	
Outcomes		
3. Do students in inverted classrooms, especially under-prepared students, show higher learning gains as compared to students in traditional classrooms?	Thermal Concept Inventory (TCI); Chemical and Thermal Process Assessment (CTP)	Pre/post math content assessment
4. Do students in inverted classrooms demonstrate an increased ability to apply material in new situations as compared to students in traditional classrooms?	Thermal Inquiry Projects (TIP)	Selected questions from the pre/post math content assessment
5. Do students in inverted classrooms demonstrate increased metacognitive gains as compared to students in traditional classrooms?	Selected items from Motivated Strategies for Learning Questionnaire (MSLQ) and Metacognitive Awareness Scale (MAS) on student surveys	
6. What are faculty experiences when teaching inverted course sections?	Faculty Reflection Papers	

Course Format

Engineering 82 met twice a week in 75-minute sessions. The control section was composed of 10-15 minute mini-lectures punctuated by conceptual and long form (calculation required) iClicker questions. Most students worked on the longer iClicker questions in informal, self-selected groups of 2-3. The inverted section meetings began with a 5-10 minute review of the video materials and 5-10 minutes answering questions asked in minute papers from the previous class meeting. The students then worked in self-selected groups of 3-5 on one problem extracted from the control section's homework assignment, while the instructor circulated to answer questions and intervene when students were reinforcing each other's misconceptions. After completing the problem, each team explained their solutions and reasoning to the instructor, and the instructor attempted to clear up any remaining misconceptions.

Math 45 met three times a week in 50-minute sessions. The control section was mainly a traditional lecture format, with many pauses, example problems, and "check-in" problems to check on student understanding. In the flipped class, the first five minutes were usually spent answering questions about the video that was watched. Then, the instructors would ask students to work on homework questions that were directly related to the videos. Sometimes students worked in groups; sometimes they worked individually. The instructors walked around the room to check on student understanding and ask and answer questions.

For both Engineering 82 and Math 45, all PowerPoint slides and tablet writing shown in the control section were contained in the video watched by the inverted section. For both courses, all students completed the same problems that students in the control section completed as homework. In Engineering 82, students in the inverted section completed specified problems during class meeting time (and turned them in at the end of class) and turned others in as homework. In Math 45, students in the inverted section used in-class time to work on any problems from the homework assignment and turned in all of their work as homework. As a final note, students in both sections of Math 45 had access to the videos; only students in the inverted section of Engineering 82 were allowed access to the videos.

In the final year of the study, Engineering 82 was taught using the Team-Based Learning (Michaelsen et al.) and was entirely inverted (both sections). Math 45 will be taught in Spring 2016 using a combination of flipped sessions and traditional sessions (new materials delivered during class meeting time).

Measures

Students in both sections of each course were administered a pretest and posttest attitude survey. The pretest survey contained a total of 28 selected items from established instruments including from the Research on the Integrated Science Curriculum (RISC), Motivated Strategies for Learning Questionnaire (MSLQ), Metacognitive Awareness Scale (Schraw & Dennison), and the STEM Questionnaires developed by the STEM team at the Higher Education Research Institute (HERI). A factor analysis was conducted on the pretest survey questions to determine which questions were most appropriate to represent the various constructs of interest including self-efficacy for learning, metacognitive self-regulation, peer learning, and help seeking behavior. Based on these data, a truncated scale was administered to students at posttest. Items used as part of the posttest include 14 items from the MSLQ and 4 items from the Metacognitive Awareness Scale (MAS). The posttest also included additional items from the HERI questionnaire as well as course-specific questions. Data from specific survey constructs were used to answer evaluation questions related to metacognition and attitudes. In addition to the surveys, students completed content assessments (described below) related to the subject area. These content assessments were used as indicators of learning gains, as there is no one clear measure of learning gains.

Engineering 82 Achievement Measures

- The **Thermal Concept Inventory** (TCI) is an online assessment created "to identify fundamental misconceptions about … thermodynamics in engineering students" (http://www.thermalinventory.com/). The TCI has a total of 24 points possible and contains five sub-measures including: Entropy and Second Law (8 points possible), Internal Energy vs. Enthalpy (4 points possible), Steady State vs. Equilibrium (4 points possible), Ideal Gas Law (4 points possible), and Conservation of Mass (4 points possible). The TCI was used to assess learning gains from pretest to posttest (Evaluation Question 3). In years 3 and 4 of the student, the TCI was replaced by the the Concept Inventory for Engineering Thermodynamics (CIET), developed by researchers at Bucknell University from the original TCI. The CIET consists of 35 multiple choice questions related to five separate sub-measures: Entropy, Reversibility, Steady State vs. Equilibrium, Internal Energy vs. Enthalpy, and Reaction Equilibrium vs. Reaction Rate. For the purposes of this course, questions related to Reaction Equilibrium vs. Reaction Rate were removed from the instrument since those topics are not covered in the course. The CIET was used to assess learning gains from pretest to posttest.

- The **Chemical and Thermal Process Assessment** (CTP) contains two complex problems for students. Each problem is graded in two areas: *Identify and Formulate Problem* and *Apply Knowledge and Solve Problem.* Each of the two areas had a total of five points possible. The CTP was used to assess learning gains from pretest to posttest (Evaluation Question 3).

- For the **Thermal Inquiry Project** (TIP), students were given the assignment to investigate two "inquiries" of their choice over the course of the semester. For each inquiry, students generated a pretest report and mini-poster and a posttest report and mini-poster. The main purpose of the projects was to provide students with a project to get them "thinking about thermodynamics beyond the textbook" (TIP student handout). Each project was done with a partner and projects had a total of five points possible for each of five domains: Ability to Communicate Effectively (Paper), Ability to Communicate Effectively (Poster); Ability to Identify and Formulate Engineering Problems in Thermodynamics; Ability to Apply Knowledge and Solve Engineering Problems in Thermodynamics; and Demonstration of an Understanding of the Impact of Inquiry in a Global, Economic, Environmental, and Societal Context. A total weighted score was also calculated (i.e., Ability to Identify and Formulate Engineering Problems in Thermodynamics: weighted x 3; Ability to Apply Knowledge and Solve Engineering Problems in Thermodynamics: weighted x 5) for a total of 55 points possible. TIPs were used to assess if students could apply material to new situations (Evaluation Question 4).

Math 45 Achievement Measures

- The **Math 45 pretest and posttest assessments** were created by the Mathematics Department. The pretest assessment consisted of five problems worth 10 points each for a total of 50 points and was not factored into students' final grades in the course. The posttest assessment used the same five problems from the pretest assessment plus an additional four new problems and was used as the final assessment for the course. For the purposes of the evaluation, only the five problems that were used for the pretest and posttest assessments were used to compare the growth from the beginning to the end of the course for the inverted and traditional sections (Evaluation Question 3). In addition, the faculty identified a subset of questions from the pretest and posttest that could be used to assess if students could apply material to new situations. We created a composite score to address this for Evaluation Question 4.

- There were five quizzes that were administered throughout the course. We analyzed the course's **quiz composite score** which was the average of all the quiz scores with the lowest score dropped. The composite score is reported as percent correct (i.e., 0% to 100%). The quiz composite score was used to assess learning gains from pretest to posttest (Evaluation Question 3).

- The **homework composite score** was calculated in the same manner as the quiz composite score. There were nine homework assignments and a final homework project that made up this composite score. The composite score was calculated by taking the average of the homework and project scores with the lowest homework score dropped. The composite score is reported as percent correct (i.e., 0% to 100%). The homework composite score was used to assess learning gains from pretest to posttest (Evaluation Question 3).

Participants

Across the first three study years, a total of 593 students completed at least one assessment or survey and were included in the overall sample. However, not all students completed all assessments, so the number of students included in each analysis differs slightly (see Table 2). For Engineering, there were a total of 83 students in inverted sections and 66 students in control sections. For Math, there were a total of 221 students in inverted sections and 234 students in control sections.

Table 2. Number of Students Completing Study Measures.

Course	Survey	Content Assessments			
Engineering	Student Survey (pre & post)	Chemical and Thermal Processes (pre & post)	Thermal Concept Inventory (pre & post)	Thermodynamic Inquiry Projects (post only)	
	127	131	114	139	
Math	Student Survey (pre & post)	Course Assessment (pre and post)	Homework Composite Score	Quiz Composite Score	Transfer Questions
	361	450	447	451	447

Students were primarily Caucasian or Asian with a nearly equal gender distribution across classes. See Table 3 for demographic characteristics. Tests of equivalence showed that, while not randomly assigned, students in the inverted section and the control section were well-matched at the time of pretest.

Table 3. Student Demographic Characteristics.

Demographic Information		% of Students
Gender	Female	46%
	Male	54%
Ethnicity	Asian	26%
	Black/Hispanic/Multiple	16%
	Caucasian	43%
	Unknown	6%
	International Student	9%

Results

The following provides results for the second year of the study except where noted (for results from the first year, see Lape et. al. 2014), organized according to each research question. Given the differences between the Engineering and Mathematics content, data are provided for each discipline separately.

Research Question 1: Do students in inverted classrooms spend additional time actively working with instructors on meaningful tasks in comparison to those students in control classrooms?

Feedback from students remained fairly consistent across study years. Students in both formats indicated that they had access to their instructors either in or out of the classroom and, generally, were actively engaged in activities and tasks.
Across years, many students indicated via surveys they did not perceive any difference between class formats. For those who did perceive a difference, students have expressed polarized opinions regarding the structure and benefits of the inverted classroom format. Students in the traditional format classrooms tended to perceive inverted format classrooms as easier or less work for students. Students in the inverted classrooms indicated having the videos to review was helpful, but these did not necessarily provide a noticeable advantage over the traditional students.

Student Participation and Preparation

Research Question 2: Do students in inverted classrooms actively participate and prepare for class through the videos and other materials?

Generally, those in inverted classrooms indicated via surveys that they typically watched the videos most or all of the time and generally watched the videos with full attention or almost full attention with minor distractions.

Answers could only be combined for Years 2 and 3 given that Math students responded to different questions regarding video viewing in Year 1 of the study (Engineering students did not answer video viewing questions in Year 1). Over 90% of Math students indicated they engaged mainly in *some*, *very little* or *no* multitasking while watching videos in Year 1 of the study.

Beginning in Year 2, both Engineering and Math students responded to the same question regarding video viewing experiences on the posttest surveys. Across both study years, response patterns remained similar, with students indicating they mostly watched the videos with full or almost full attention (see Figure 1).

Figure 1. Frequency of Video Viewing Experiences for Students in Inverted Sections, Years 2 and 3 (n = 159)

Students mainly studied by themselves or with students from the same classroom format when reading and preparing for class. Few students indicated they primarily worked with students in the other course section (e.g., students in the inverted section did not work with students in the traditional section and vice versa).

Beginning in the third study year, students in inverted classrooms completed weekly or daily quizzes based upon the material in the video for the class session. Many students indicated knowing that a quiz would be happening in class encouraged them to watch the videos and take notes or review topics in preparation for the quiz. The professor also indicated that students were more prepared for class with the inclusion of daily quizzes.

Student Learning and Metacognitive Gains

Do students in inverted classrooms show higher learning gains as compared to students in traditional classrooms? and *Do students in inverted classrooms demonstrate increased metacognitive gains as compared to students in traditional classrooms?*

Analyses were conducted using data combined across all three study years for each of the main assessments and projects for both the Engineering and Math courses. Using combined data, no significant differences were detected between students in the inverted courses and the traditional courses. The repeated measures ANOVA framework has been selected as an appropriate assessment for change, as the analysis tests whether mean differences from pre- to post- *differ* for treatment and control. As such, a statistically significant result would be interpreted as a difference in the mean change (i.e., growth), so this analytical framework should adequately assess the research questions. Rather than introducing a composite learning gain metric, we have reported results for each instrument separately for transparency's sake.

Engineering Results

Repeated Measures ANOVAs were calculated where appropriate as well as *t*-tests. Analyses were conducted on the TCI total score, the two main CTP subsections, and the combined weighted total for the TIP 1 and TIP 2. For the TCI analysis, scores were converted to standardized *z*-scores given that two different types of thermal concept assessments were used in the course of the study and using standardized scores allows for a more accurate comparison. As seen in Table 4, no significant differences were found. Analyses were conducted on the TCI total score ($F = .004$, $p > .05$); the CTP Identify & Formulate Problems ($F = .244$, $p > .05$); the CTP Apply Knowledge & Solve Problems $F = 2.36$, $p > .05$; and the TIP 1 and 2 weighted total score ($t = -1.56$, $p > .05$).

Table 4. Engineering TCI and CTP Repeated Measures ANOVA Analyses.

Measure	Traditional		Inverted		*df*	*F*	*p* value
	Pre Mean (*SD*)	Post Mean (*SD*)	Pre Mean (*SD*)	Post Mean (*SD*)			
TCI Total Score*	-.006 (.83)	.064 (.85)	.004 (1.12)	-.05 (1.10)	1, 110	.004	.949
CTP ID & Formulate Problems, Total	.564 (.88)	9.05 (1.58)	.500 (.68)	9.12 (1.03)	1, 121	.244	.622
CTP Apply Knowledge & Solve Problems, Total	.655 (.80)	8.21 (1.61)	.529 (.53)	8.49 (1.29)	1, 121	2.355	.128

*converted to *z*-scores

Math Results

Appropriate analyses were conducted on the combined Math data as well, including the course pretest and posttest, the homework composite score, the quiz composite score, and the final exam questions 4, 5, and 9 composite (measuring knowledge transfer). There was no significant difference between control and treatment students found for growth from pretest to posttest ($F = .021$, $p > .05$). No significant differences were detected for the other measures of student achievement either. These included the homework composite ($t = .041$, $p > .05$); the quiz composite ($t = -.892$, $p > .05$); and the transfer of knowledge items ($t = .216$, $p > .05$).

Table 5. Math Homework Composite, Quiz Composite, Transfer of Knowledge Analyses.

Measure	Traditional Mean (*SD*)	Inverted Mean (*SD*)	*t*	*df*	*p* value
Homework Composite	87.43 (9.79)	87.39 (10.00)	.041	456	.967
Quiz Composite	81.07 (13.42)	82.23 (14.52)	-.892	458	.373
Composite of Questions 4, 5, and 9	24.50 (5.75)	25.39 (5.70)	.216	458	.829

Analyses were conducted investigating potential differences in metacognitive gains, self-efficacy, and peer learning behavior gains. Again, using data from all three study years for Engineering and Math separately as well as together. No significant differences in metacognition were found for students in the Engineering course ($F = .032$, $p > .05$), students in the Math course ($F = .696$, $p > .05$), and combined across courses ($F = .242$, $p > .05$). Furthermore, no significant differences were found for peer learning or self-efficacy gains (see Table 7).

Table 6. Combined Metacognitive Gains.

Group	Traditional		Inverted		*df*	*F*	*p* value
	Pre Mean (*SD*)	Post Mean (*SD*)	Pre Mean (*SD*)	Post Mean (*SD*)			
Engineering	3.59 (.52)	3.64 (.58)	3.56 (.48)	3.64 (.47)	1, 117	.032	.859

Math	3.50 (.47)	3.56 (.52)	3.54 (.50)	3.57 (.54)	1, 360	.696	.405
Combined	3.63 (.44)	3.68 (.39)	3.79 (.39)	3.78 (.48)	1, 479	.242	.623

Table 7. Combined Peer Learning and Self-Efficacy Gains.

Group	Traditional		Inverted		*df*	*F*	*p* value
Engineering	Pre Mean (*SD*)	Post Mean (*SD*)	Pre Mean (*SD*)	Post Mean (*SD*)			
Peer Learning	3.74 (.56)	3.67 (.66)	3.74 (.54)	3.64 (.54)	1, 117	.120	.729
Self-Efficacy	3.58 (.79)	3.58 (.88)	3.81 (.67)	3.76 (.55)	1, 122	.151	.699
Math							
Peer Learning	3.56 (.61)	3.45 (.60)	3.64 (.49)	3.50 (.56)	1, 360	.365	.546
Self-Efficacy	3.60 (.77)	3.82 (.70)	3.72 (.80)	3.87 (.73)	1, 362	1.342	.248
Combined							
Peer Learning	3.60 (.59)	3.48 (.62)	3.66 (.51)	3.51 (.58)	1 ,479	.392	.532
Self-Efficacy	3.60 (.76)	3.76 (.73)	3.73 (.76)	3.84 (.70)	1, 486	2.156	.143

Gender Differences

Students' achievement and survey results for Engineering and Math students were also investigated in terms of potential differences between males and females. These differences were investigated in two ways. First, in terms of rate of growth from pretest to posttest compared between males and females. Second, in terms of differences in males' and females' responses on items that were not assessed using a pretest-posttest (e.g., certain questions only on the posttest student survey).

Data for all three years of the Engineering course administration (fall 2012, fall 2013, and fall 2014) were combined to investigate potential gender differences on assessments specific to the engineering course. Similarly, data for all three years of the Math course (spring 2013, spring 2014, and spring 2015) were combined to investigate potential gender differences on assessments specific to the math course. Student data for both Engineering and Math were combined across all three years for the analyses of the student survey composites.

Engineering

Males and females performed similarly on most measures used in the Engineering course, though one statistically significant difference was found such that females showed higher TIP 2 and TIP Total scores than males. Results remained similar when analyses were conducted separating students by course format as well. That is, when scores were analyzed solely for students in the inverted section and solely for students in the traditional section, there were no significant differences found between males and females.

The Thermal Inquiry Project (TIP) scores showed differences between males and females for both scores on the second TIP item as well as the total TIP score. Females scored higher than males for both of these (see Table 11).

Table 8. Gender Differences for Engineering Thermal Inquiry Project Scores.

Test and Administration Time	Gender	n	Mean (SD)	t-test	p-value
TIP 1	male	80	50.36 (2.55)	1.08	ns

	female	57	50.81 (2.13)		
TIP 2	male	80	49.86 (3.44)	**2.31**	**<.05**
	female	55	51.23 (2.60)		
Total TIP	male	80	100.22 (4.47)	**2.31**	**<.05**
	female	55	101.93 (3.79)		

There were no statistically significant differences males and females related to growth from pretest to posttest on the Thermal Concept Inventory (TCI) and Chemical and Thermal Process (CTP) inventory scores.

Math

Several statistically significant differences were found between males and females on Math 45 course assessments (see Table 9). Females did show higher average scores on homework compared to males, though this was only approaching significance (p = .08).

Table 9. Math 45 Only - Differences Found Between Males and Females.

Measure	***Difference Between Males/Females***	***Specific Difference***
Course Pretest/Posttest	Yes	Females showed greater gains from pretest to posttest
Homework	No	n/a
Quizzes	No	n/a
Knowledge Transfer (Questions 4, 5, 9)	Yes	Males showed significantly higher scores than females

As shown in Tables 10 and 11, females showed greater gains from pretest to posttest compared to males (F = 4.88, p <.05) and males showed significantly higher scores than females on knowledge transfer items (t = 2.46, p <.05).

Table 10. Gender Differences for Math Course Pretest-Posttest Assessment.

Assessment	***Males (n = 231)***		***Females (n = 215)***		***df***	***F***	***p value***
	Pre Mean (SD)	*Post Mean (SD)*	*Pre Mean (SD)*	*Post Mean (SD)*			
Course Pretest/Posttest	13.72 (9.30)	41.83 (6.21)	10.29 (9.15)	40.24 (6.10)	1, 444	**4.88**	**<.05**

Table 11. Gender Differences for Math Knowledge Transfer Items

Assessment	Gender	n	Mean (SD)	t-test	p-value
Knowledge Transfer Items	male	237	26.15 (5.69)	**2.46**	**<.05**
	female	217	24.85 (5.60)		

Engineering and Math

Analyses were conducted on the student survey composite scores for both Engineering and Math combined scores as well given that all students in both courses provided these scores each year. The one significant difference on the student survey was found in the area of self-efficacy. Females showed significantly higher growth in the area of self-efficacy from pretest to posttest compared to males (*F* = 10.73, *p* >.01). On average, males' score increased very slightly from pretest to posttest (pretest mean = 3.91, posttest mean = 3.96) while on average females' score increased by nearly a quarter point from pretest to posttest (pretest mean = 3.41, posttest mean = 3.64).

Table 12. Student Survey Construct Scores

Construct	***Males***		***Females***		***df***	***F***	***p value***
	Pre Mean (SD)	*Post Mean (SD)*	*Pre Mean (SD)*	*Post Mean (SD)*			
*Peer Learning**	3.60 (.59)	3.48 (.60)	3.65 (.52)	3.56 (.57)	1, 476	0.437	ns
*Self-Efficacy***	3.91 (.73)	3.96 (.70)	3.41 (.74)	3.64 (.69)	1, 483	**10.73**	**.001**
*Metacognition**	3.57 (.49)	3.59 (.53)	3.50 (.49)	3.58 (.51)	1, 476	0.122	ns

* - males, *n* = 252, females, *n* = 226
** - males, *n* = 255, females, *n* = 230

This difference was also found when males were compared to females in only the traditional sections but not when males were compared to females in the only inverted sections. That is, *females in the traditional section also showed higher growth in self-efficacy from pretest to posttest while males' scores decreased*. However, for students in the inverted sections, there was no significant difference in the rate of change from pretest to posttest for males compared to females – *both males' and females' scores increased slightly* (see Table 16).

Table 13. Gender Differences in Self-Efficacy Scores Separated by Course Format.

Self-Efficacy	***Males***		***Females***		***df***	***F***	***p value***
	Pre Mean (SD)	*Post Mean (SD)*	*Pre Mean (SD)*	*Post Mean (SD)*			
Traditional Students	3.87 (.71)	3.92 (.74)	3.31 (.74)	3.61 (.72)	1, 240	**10.88**	**<.01**
Inverted Students	3.95 (.74)	4.00 (.94)	3.52 (.74)	3.67 (.67)	1, 241	1.55	ns

Discussion

Throughout the three study years, students in the inverted sections showed mostly similar or equal results to students in the traditional sections with no significant differences detected between course formats for learning gains. The largest discrepancy was observed on one measure for the Engineering 82 students. The Engineering 82 professor changed the Thermal Concept Inventory instrument between the second year and the third year, resulting in students demonstrating significantly higher posttest scores compared to pretest scores. This was not observed with the Thermal Concept Inventory which was used in the first two years of the study. All other answers to evaluation questions remained similar or identical across study years. There were also very few differences detected between males and females for learning gains. However, females did show higher gains in self-efficacy compared to males from pretest to posttest.

While these findings still do not support original hypotheses of using the inverted model, there are possible explanations for these results. It is possible that the lack of differences stems from both particularities of the student population at the study site as well as an implementation model in which the two conditions are not distinct enough to influence significant differences. Students at Harvey Mudd College regularly work together outside of class; hence, peer learning may not be increased significantly over their baseline levels. Furthermore, HMC students are high achievers and strong independent learners, and may show high learning gains (on average; results are not uniform for students within a given condition) under a variety of teaching methods. Regarding the potential insufficient difference in conditions, this issue is especially acute for the Math course in which students in both sections had access to all course materials, including lecture videos designed specifically for inverted classroom students. However, given that the professor for the Engineering course implemented more distinctly different models during the third year, and there were still no significant differences between conditions, this may be attributable to another influence altogether – the influence of active learning. As described in Jensen, et al. (2015), the "flipped classroom" alone does not seem to result in higher learning gains or better student attitudes. Rather, implementing active learning, regardless of whether or not a class was traditional or inverted, resulted in significant learning gains. Jensen, et al., studied an active inverted classroom compared to an active traditional classroom and results showed equivalency between groups at the end of their study. Data collected from our study also seem to support the idea that students are impacted the most when an active-learning style of instruction is used, regardless of when they are introduced to new content (in the classroom or at home through video lecture). When content engages students, encourages them to explore concepts, and instructors explain new ideas or terminology to the students, they are using an active learning structure, which supports the students' learning. Freeman et al. (2014) also found support for this through their meta-analysis of 225 studies investigating student performance in STEM courses under traditional learning compared to active learning. Other research, for example studies conducted by Fagen, Crouch, and Mazur (2002) and Lasry, Mazur, and Watkins (2008) also highlight the benefit of implementing a classroom model which goes beyond just transfer of information or a traditional lecture format. Additionally, Bruff, Fisher, McEwen, and Smith (2013) state that successful classrooms which use online and in-class components engage in tight coupling between online and face-to-face components as well as a cohesive hybrid overall. This suggests that students benefit the most from using both home materials (e.g., lecture videos) as well as intentional in-class time.

References

Alexander, P. A., & Murphy, P. K. (1999). Nurturing the seeds of transfer: a domain-specific perspective. *International Journal of Educational Research, 31*, 561-576.

Baker, J. W., (2000). The "classroom flip": Using web course management tools to become the guide by the side. *Selected Papers from the 11th International Conference on College Teaching and Learning*.

Bishop, J. L., & Vergeler, M. A. (2013). *The inverted classroom: A survey of the research.* Paper presented at the 120th ASEE Annual Conference & Exposition (Paper ID #6219). Retrieved August 27, 2013, from http://www.asee.org/public/conferences/20/ papers/6219/view

Bruff, D.O., Fisher, D.H., McEwen, K.E., & Smith, B.E. (2013). Wrapping a MOOC: Student perceptions of an experiment in blended learning. *MERLOT Journal of Online Learning and Teaching, 9*(2), 187-199.

Carlisle, M. C., (2010). Using You Tube to enhance student class preparation in an introductory Java course. *SIGCSE '10 Proceedings of the 41st ACM Technical Symposium on Computer Science Education*, 470-474.

Day, J. A., & Foley, J. D. (2006). Evaluating a web lecture intervention in a human-computer interaction course. *IEEE Transactions of Education, 49*(4). 420 – 431.

Dollar, A., & Steif, P., (2009). A web-based statics course used in an inverted classroom. *Proceedings of the 2009 American Society for Engineering Education Annual Conference*.

Fagen, A.P, Crouch, C.H., Mazur, E. (2002) Peer instruction: Results from a range of classrooms. *The Physics Teacher, 40*, 206-209.

Foertsch, J., Moses, G., Strikwerda, J., & Litzkow, M., (2002). Reversing the lecture/homework paradigm using e-teach web-based streaming video software. *Journal of Engineering Education, 91*(3), 267-274.

Freeman, S., Eddy, S.L., McDonough, M., Smith, M.K., Okoroafor, N., Jordt, H., & Wenderoth, M.P. (2014) Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America, 111*(23), 8410-8415.

Jensen, J.L., Kummer, T.A., Godoy, P.D.d.M. (2015). Improvements from a flipped classroom may simply be the fruits of active learning. *CBE – Life Sciences Education, 14*, 1-12.

Lage, M., & Platt, G., (2000). The internet and the inverted classroom. *Journal of Economic Education, 31*(1), 11.

Lage, M., Platt, G., & Treglia, M., (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *Journal of Economic Education, 31*(1), 30-43.

Lape, N. K., R. Levy, D. Yong, K, Haushalter, R. Eddy, N. Hankel (2014). "Probing the Inverted Classroom: A Controlled Study of Teaching and Learning Outcomes in Undergraduate Engineering and Mathematics." *Proceedings of the 2014 American Society for Engineering Education Annual Conference*.

Lasry, N., Mazur, E., & Watkins, J. (2008). Peer instruction: From Harvard to the two-year college. *American Journal of Physics, 76*, 1066-1069.

Michaelsen, L.K., L. Bauman Knight, and L.D. Fink (2004). Team-Based Learning: A Transformative Use of Small Groups in College Teaching. Stylus Publishing.

Pugh, K. J., & Bergin, D. A., (2006). Motivational influences of transfer. *Educational Psychologist, 41*(3), 147-160.

Schraw, G., & Dennison, R. S., (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*, 460-475.