JoSoTL

Journal of the Scholarship of Teaching and Learning

Volume 20, Number 1, April 2020 josotl.indiana.edu

T

JoSoTL

Journal of the Scholarship of Teaching and Learning

Volume 20	Number 1	April 2020

Journal of the Scholarship of Teaching and Learning, Vol. 20, No. 1, April 2020, pp.1-21.
doi: 10.14434/josotl.v20i1.24128

Impact of Team Formation Method on Student Performance, Attitudes, and Behaviors

Martiqua L. Post
U.S. Air Force Academy[1]
Martiqua.Post@usafa.edu

Anthony Barrett
U.S. Air Force Academy

Marlyse Williams
U.S. Air Force Academy

Lauren Scharff
U.S. Air Force Academy

Abstract: This project examined the effects of two team selection methods (self-selected and instructor-formed based on matched academic performance) on team and individual student performance and on self-reported attitudes and team behaviors in a freshman-level core-required introductory course. The data included mid and end-of-semester self-reports. Matched-performance groups had significantly higher grades on several performance measures, with a larger effect on the team grades than on the individual grades; however, overall the effect sizes were small. Although there were no group differences for most self-reported items, a key finding was that self-selected teams were significantly more likely to already have friends on their team, and a significant correlation showed that already having friends on a team was negatively correlated with many of the performance measures. In contrast, members of both types of teams reported equally high likelihood to make new friends, which was positively correlated with performance. Understanding the impact of different approaches to team formation may guide instructors and lead to more well-functioning teams, higher student learning, and greater student satisfaction.
Keywords: teams, team formation methods, team dynamics, student teams

Overview

Teams can make organizations more responsive, and they often become the primary unit of performance. In a military environment, the ability to work efficiently and effectively within a team can sometimes mean the difference between life and death. The importance of teamwork in the Air Force is reflected in one of the three broad categories of the United States Air Force Academy's (USAFA) institutional competencies, "Leading People and Teams" (USAFA, 2015). The pedagogical advantages of teamwork in an educational setting are also valuable. Team experiences can enhance student learning. For example, Oakley, Brent, Felder, and Elhajj (2004) found that "students taught in a manner that incorporates small-group learning achieve higher grades, [and] learn at a deeper level". Further, working in teams allows students to receive additional help and support from their peers (Feichtner & Davis, 1984) and share ideas holistically, increasing the performance of the entire class (Hernandez Nanclares, Rienties, & Van den Bossche, 2012).

[1] Disclaimer: The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the United States Air Force Academy, the Air Force, the Department of Defense or the U.S. Government.

While there seems to be broad agreement that teamwork has its advantages, these benefits are not automatic. Social factors such as communication, conflict resolution, and group time management are often required for high functioning teamwork, and in turn, these factors may be impacted by the particular combination of individuals on a team. Thus, how the teams are initially formed could impact the team's performance and overall experience. The objective of this investigation was to assess the impact of two different team formation methods (self-selected and instructor-formed based on grouping by student predicted performance) on team performance, individual student performance, and self-reported student attitudes, behaviors, and outcomes. The investigation was conducted within a freshmen-level introductory course required for all students as part of the core curriculum. This research focuses solely on the initial manner by which teams were formed, and not on any intentional efforts by instructors to develop the teams or teamwork skills after they were formed. The large sample size of this investigation (845 students across Fall and Spring semester offerings of the course), and the combination of academic performance as well as self-report items offer a meaningful contribution to the literature.

Background

To better understand why teams might provide benefits both academically and beyond, we need to understand what characteristics transform a group into a team, and what factors may impact team cohesiveness. The school of thought about teams summarized and extended by Katzenbach and Smith (1993) is that "a team is more than the sum of its parts" because teams produce discrete teamwork products that are better than individual products due to the joint contributions of their members. They suggest that a team is "a small number of people with complementary skills who are committed to a common purpose, set of performance goals, and approach for which they hold themselves mutually accountable" (Katzenback & Smith, 1993, 2). Therefore, teamwork is about common commitment. Without it, groups perform as individuals. With it, well-functioning teams can become a powerful unit of performance.

The positive relationship between team function and performance (Kozlowski and Ilgen, 2006; Katzenback & Smith, 1993), as well as the fact that many academic, government, and industry projects utilize teams, has led to considerable research to better understand the characteristics of well-functioning teams. Some identifiers of a well-functioning team include effective communication, constructive conflict, and shared commitment (Katzenback & Smith, 1993). Additional indicators include team members' positive attitudes, values, enjoyment, and high ratings of effectiveness of the team experience (Chapman, Meuter, Toy, & Wright, 2006). One factor that could influence the characteristics for team functionality is team composition, including team size, and myriad individual member characteristics: member academic experience, expertise and abilities, cultural background, gender, life experiences, and interpersonal skills (Connerley & Mael, 2001).

Given the known increased learning potential when students work in teams (e.g., Hernandez Nanclares et al., 2012; Oakley et al., 2004), some prior researchers (e.g., Chapman et al., 2006; Connerley & Mael, 2001) questioned to what extent the process of team formation may contribute to student satisfaction. Team formation methods in educational settings are divided in two main categories: self-selected and instructor-formed teams. Self-selected team formation allows students to choose their own teams. Instructor-formed teams can be formed in many ways, such as through randomization (Mahenthiran and Rouse 2000; Chapman et al., 2006) or by using known student characteristics such as prior academic achievement (the basis of the current study; Matta, Luce, and Ciavarro, 2010), gender, attitudes revealed by questionnaire results (McClough and Rogelberg, 2003), or student personality (Pociask, Gross, and Shih, 2017; Shen, Prior, White, and Karamanoglu, 2007).

Instructor-formed teams may lead to more diverse teams than self-selected teams, because students may have the opportunity to work with others with whom they would not typically work.

In the literature, there are strong advocates for each approach. For example, Bacon, Stewart, and Silver (1999) advocate "giving students a say in team assignments." In contrast, in their "profile for failure," Feichtner and Davis (1984) include "allowing students to form their own groups" on their list of procedures that are likely to lead to failure. A potential drawback to self-selected teams is that this process may lead to unselected team members, requiring that the instructor assign them to a team, and potentially reducing the positive dynamics of a self-selected team. Depending on the specific method the instructor uses to form the teams or add members, students could feel that the teams were not fairly formed based on perceptions of favoritism or prejudices.

While there are many discussions of findings in the literature, there is less empirical evidence to guide choice of team formation (McClough et al., 2003) and that evidence is mixed. Matta et al., (2010) found that the act of selecting their own teammates led to only a very small impact on students' level of satisfaction with the team. Likewise, Pociask et al., (2017) found student performance was similar regardless of team formation method, and therefore suggested that student self-selected teams can be a reasonable option for instructors to consider. Chapman et al. (2006) provide one of the few larger-scale empirical comparisons of self-selected and instructor-formed (using randomization) teams. Their end-of-semester, self-report questionnaire asked about team dynamics, participant attitudes about their teams, and team outcomes for semester-long project teams in upper-level business courses. The vast majority of questions for which there were significant differences favored the self-selected teams over the randomly-formed teams (e.g., better communication, more enthusiasm, more interest in teammates, better conflict resolution, higher confidence in teammates, greater perceived value, usefulness and effectiveness, higher likelihood of making new friends, greater enjoyment, greater pride in work). There were exceptions within their study, however, which showed randomly-formed teams benefited by being more task-oriented and more likely to get straight to work with less socializing.

Mahenthiran and Rouse (2000) compared the impact of a fully randomized team-formation method with a hybrid team-formation method that allowed students partial control over the process. The hybrid approach involved randomly combining self-selected pairs of friends. Student satisfaction was measured using a questionnaire. Mahenthiran and Rouse noted significantly higher performance on project grades by the hybrid teams, and this result was true regardless of their incoming grade point averages. They concluded that the best team formation method is to pair friends and then randomly combine pairs to form the team rather than using complete random assignment.

An important factor related to team formation is the diversity of the team members. Within academia, students often select teammates who are similar to themselves when given the opportunity (e.g., Pociask et al., 2017; Rienties, Alcott, & Jindal-Snape, 2014). Nonetheless, Scott and Pollock (2006) found that a diverse range of skills can be found in self-selected teams. Some of that discrepancy could be due to the two researchers using different characteristics to define diversity, since there will always be some diversity between members who are otherwise highly similar in many ways. If student academic ability is the characteristic of focus for diversity, research suggests that stronger students often seek out one another in self-selected teams, leaving the weaker ones to form their own teams (e.g., Oakley et al., 2004). Although such groupings reduce academic diversity in the teams, they can decrease the likelihood of weaker students "riding on the coattails of stronger students," allow teammates to be surrounded by others with similar levels of motivation, and prevent anyone within a team from feeling like they were falling behind (Bronson & Merryman, 2013; Carrell, Fullerton, & West, 2009). However, there are also benefits of having academic performance diversity enforced through instructor formation of teams. For example, in well-functioning academically diverse teams, weaker students can benefit from observing how stronger students set-up, approach, and solve problems and can even gain from being tutored by the stronger students. Stronger students benefit with increased understanding from having to explain the problem to another person.

The mix of support for the various ways to form teams suggests that additional research is needed to help guide the choice of team formation method. This study compares large sample sizes of self-selected teams with instructor-formed teams based on previous academic performance (heretofore referred to in this study as Matched-performance teams). We gathered self-reported attitudes, team dynamics, and outcomes via a questionnaire and several performance measures: two individual work averages, two team work averages, and the final course average. The questionnaires were administered at the mid-semester as well as at the end of the semester, which allowed us to investigate how team formation method might impact our measures over time. With these data, we were able to conduct group and time correlations to determine the effect of team formation on student performance and attitudes.

Research Questions and Hypotheses

Based on the literature reviewed, we formulated three research questions and hypotheses to give us some insight as to whether team formation method impacts student performance, attitudes, and behaviors. First, how do the two team-formation methods impact our performance measures? Because prior research (Chapman et al., 2006) indicates that self-selected teams are more prone to socializing and being less task-oriented, we predicted matched-performance teams would perform better academically than self-selected teams, and that group differences would be greater for team grades than individual grades.

Second, how do the two team-formation methods impact the self-reported attitudes, team dynamics, and outcomes? We predict a mix of results, depending on the question and the comparison group. In most cases, matched-performance and self-selected teams are likely to be academically similar (Oakley et al., 2004), and therefore might have similar reported attitudes and behaviors. Self-selected teams are more likely to have prior friends on their team (Pociask et al., 2017; Rienties, Alcott, & Jindal-Snape, 2014; Chapman et al., 2006), and therefore, may show more positive ratings of social factors. Randomly-selected teams are likely to be less similar than our two types of teams; thus we predict that our group differences will be smaller than those found when comparing self-selected and randomly-selected teams (e.g., Chapman et al., 2006).

In this study, the investigation of the time factor (mid-semester versus end-of-semester) was more exploratory, although we predicted there might be an increase in differences across time between the groups due to an enhancement of team dynamics as the teams' amount of time working together increased. We also chose to investigate how the individual self-reported questionnaire items would correlate with the performance measures. Specifically, this part of our investigation considers to what extent, if at all, the self-reported attitudes and team dynamic items predict the performance measures.

Method

Participants

A total of 845 first-year students enrolled in a core-required introductory engineering course participated in this study during the Fall 2015 and Spring 2016 semesters. The students represented 75% of the entire first-year class and were chosen because their instructors (N=18) used one of the two team-formation methods of this study. They represented a pseudo-random subset of their entire class because students do not choose their instructors or course sections at our institution. As is standard at our institution, student assignment to course sections was performed by the Registrar's Office, which used a random method with constraints (e.g., student intercollegiate athletes cannot enroll in late afternoon courses due to team practice schedules). Of the participating students, 617 were male (73%) and 228 were female (27%), an accurate reflection of the student population at the institution.

Design

The introductory engineering course used in this study was selected based on the large amount of teamwork, which comprised 40% of an individual's final grade in the course. Over the course of a semester, more than half of the contact hours during class time were spent working on team tasks, such as fabricating projects in a lab, field testing, working in a classroom environment on collaborative design, or presenting their team project results. In addition to the time working in groups during class, a considerable amount of out-of-class team work was expected as they completed the ten group-project deliverables required in the course.

Students in each section were organized in semester-long teams of four student members, a size that matched the tasks and scope of the course projects (Denton, 1996). There were two independent variables, team-formation method (Self-selected, Matched-performance) and time (Mid-semester, End-semester). Matched-performance teams used predicted performance to create teams with relatively similar prior semester GPA, when available, or Academic Composite scores[2] for first-semester students. Teams were formed by ranking students based on prior academic performance and then grouping the top four, then the next four, and so on.

To assess the impact of the two different team formation methods, this study used responses on a self-report questionnaire and five measures of student performance. The performance measures are summarized in Table 1 and include two team work averages, two individual averages, and the final course average.

Table 1. Academic Performance Measures

	Assignments & Exams included in Averages	Points
Mid-semester Team Average	4 team assignments	125 (14%)
Second-half Team Average	6 team assignments	230 (26%)
Mid-semester Individual Average	5 assignments & mid-term exam	235 (27%)
Second-half Individual Average	2 assignments & final exam	310 (33%)
Final Course Average	All of the above	900

[2]Academic composite scores are computed by USAFA's admissions office for each student prior to their arrival based on a number of factors, including high school coursework and activities and standardized test scores (ACT/SAT). The numerical value of the score ranges from zero to 4000, and is primarily used as an indicator of predicted academic performance, with higher Academic Composite scores predicting better student academic performance. For the participants in the study, the highest Academic Composite was 4000, the lowest score was 2411, and the average score was 3270 with a standard deviation of 332. These scores well represent the typical range found across our first-year students.

Materials

The self-report questionnaire contained 24 questions that assessed three categories: team attitudes (seven questions), team behaviors and dynamics (twelve questions), and team outcomes (five questions). Fifteen of the 24 questions were adapted from Chapman et al. (2006). Questions related to attitudes asked about overall attitude, fairness of team formation method, value of the teamwork, the scale to which the teamwork was enjoyable, and effectiveness of the team experience. Questions pertaining to team dynamics provided insight on enthusiasm, team communication, conflict resolution, work session dynamics, and methods teams used to share work responsibilities. Questions focused on team outcomes asked about friends on the team and quality of the work products. The full questions can be seen with data in Tables 5, 6, and 7.

Response options for all questions were multiple-choice with two-to-five Likert-scale response options each. For 20 of the 24 questions, the five response options represented a bipolar spectrum of agreement with the statement in the question or the student's assessment of their experience or perceived quality of work (e.g., strongly disagree / disagree / neutral / agree / strongly agree or terrible / bad / average / good / fantastic). The remaining four questions had between two and four valid choices (yes / no; unfair / doesn't matter to me / fair; never / once / twice / three or more times). Four questions were reverse coded to minimize influence of response bias.

Procedure

At the beginning of each term, instructors for the core engineering course identified the team formation method they would use to form teams within their section(s). Although they were free to select their method of choice, they were informed of the study and requested to consider using one of the methods to stimulate ample representation of these two methods, Self-selected or Matched-performance. In all cases, teams were formed within the first three lessons of the semester-long course. A small number of instructors used other methods (e.g., random assignment/alphabetical, assignment by dorm location, predicted performance to form teams with a diverse mix of Academic Composite scores). Teams formed by these other methods were not included in this study. Table 2 shows the total number of students, teams, sections and instructors corresponding to the two team-formation methods evaluated in the study. Two instructors used different methods for the Fall and Spring semesters, hence, the total number of instructors is 18 rather than 20.

Table 2. Team Formation Method Metrics

	Self-Selected	Matched-Performance	Total
Number of Sections	13 (42%)	18 (58%)	31
Number of Teams	87 (42%)	121 (58%)	208
Number of Students	355 (42%)	490 (58%)	845
Number of Instructors	7 (39%)	13 (61%)	18

Students remained on the same teams while completing all ten graded team assignments throughout the semester. A small percentage of participants (7%) were on teams that had some change to their team's composition during the semester. A change in team composition is defined as when a student moved to a new team or when a student was removed (change in section, disenrollment). These changes predominately occurred early in the semester, and any team transfers were initiated by the instructor to preserve 4-person team size (i.e., move one student to avoid having one 3-person team and one 5-person team). Neither the students nor the instructors were given any ancillary guidance on methods for developing higher functioning teams, and any team development encouraged by the instructor was informal and within what would normally be expected by educators at the institution. In a small number of situations, instructors intervened with dysfunctional teams to help them work better as a unit. No students transferred teams due to dysfunction.

The 24-item self-report questionnaire was administered to the students twice, once at mid-semester after they had completed four team assignments, and again at the end of the semester after they had completed nine (of the ten) team assignments. The questionnaires were administered in class using digital score sheets, and students were asked to provide identifying information to allow their responses to be linked for analyses. Once performance and questionnaire data were linked at the end of the semester, all identifying information was removed. Students were informed that the questionnaire was optional and that instructors could only have access to their responses after final grades were submitted.

Results and Discussion

At the completion of the two-semester study, all data were integrated for analysis, including attributing the following data to each individual participant: team identifier, section identifier, instructor identifier, team formation method, demographic information, all assignment scores and grade percentages, and questionnaire responses. After the data were integrated, all personally identifiable information was removed.

Prior to analyses being performed, participants who did not complete both mid-semester and end-of-semester questionnaires were removed from the data set. The missing questionnaire data were most often due to a class absence or the omission of a student identifier on the questionnaire response. It is worth noting that class attendance is mandatory at our institution; absences are typically due to illness, medical appointment or athletic team-related travel. Of the 845 participants in the study, 831 completed at least one of the two feedback forms (98%); however only 699 completed both feedback forms (83%). Using only those students with complete data sets, Self-selected teams accounted for 310 students (44%), while Matched-performance teams accounted for 389 students (56%).

Student Performance Analyses

Prior to conducting analyses of student performance measures, the Academic Composite score was investigated for use as a possible covariate. Academic Composite score significantly correlated with each of the five student performance measures at the $p < .01$ level for Mid-semester Team Average [$r(697)=.26$], Second-half Team Average [$r(697)=.27$], Mid-semester Individual Average [$r(697)=.49$], Second-half Individual Average [$r(697)=.56$], and Final Average [$r(699)=.57$]. Next, we investigated if there was a significant difference in Academic Composite between the two groups, which there was not, $t(1,697)= -0.28$, $p=.78$. Because of this, Academic Composite score was not used as a covariate in the subsequent analyses.

For each of the five performance measures, independent group t-tests and tests for effect sizes were completed. Table 3 presents mean and standard deviations for each group for each performance measure. For the team average measures, one group showed a significant difference, with the matched-performance teams having higher scores than the self-selected teams at the end of the semester;

however, effect sizes were small [Mid-semester Team Average: $t(1,697)=-1.33$, $p=.184$, partial eta squared =.003; Second-half Team Average: $t(1,697)=-4.00$, $p<.001$, partial eta squared =.022]. For the individual performance measures, Matched-performance teams had higher averages than those in the Self-selected teams, with that difference being significant at mid-semester, but only showing a trend for the second half of the semester [Mid-semester Individual Average: $t(697)=-2.16$, $p=.03$, partial eta squared =.007; Second-half Individual Average: $t(697)=-1.61$, $p=.11$, partial eta squared =.003]. The final course average showed a significant difference with the matched-performance groups having higher grades [$t(697)=-2.65$, $p=.008$, partial eta squared = .01]. Overall these findings support our first hypothesis that the manner in which teams are formed impacts student performance.

Table 3. Academic Performance Measures for the Two Methods of Team Formation

	Self-Selected		Matched-Performance	
	Mean	*(Std Dev)*	Mean	*(Std Dev)*
Academic Composite Score	3269	(*355*)	3276	(*326*)
Mid-semester Team Average (125pts)	88.2	(*6.7*)	88.9	(*8.1*)
Second-half Team Average (230pts) ***	86.7	(*5.1*)	88.3	(*5.3*)
Mid-semester Individual Average (235pts) *	87.7	(*6.9*)	88.8	(*6.4*)
Second-half Individual Average (310pts)	77.1	(*8.6*)	78.2	(*8.4*)
Final Course Average (900pts)**	83.9	(4.9)	85.0	(5.2)

* Significance of difference between groups: *$p \le .05$, ** $p<.01$, ***$p \le .001$.

Self-reported Attitude, Behavior, and Outcome Analyses

Our measures of self-reported team attitudes, team behaviors and dynamics, and team outcome were collected via the self-report questionnaire. Prior to analyses, all questionnaire responses were converted to a numeric rating based on a 5-point itemized scale using Table 4. Lower scores represented less positive responses.

Table 4. Feedback Response Numerical Conversion

Response Choices	Survey Response A	B	C	D	E
A, B, C, D, E	*Strongly Disagree* *1*	*Disagree* *2*	*Neutral* *3*	*Agree* *4*	*Strongly Agree* *5*
A, B, C, D, E (reversed scored)	*Strongly Agree* *1*	*Agree* *2*	*Neutral* *3*	*Disagree* *4*	*Strongly Disagree* *5*
A, B, C, D	*Never* *1*	*Once* *2.33*	*Twice* *3.67*	*3 or more times* *5*	*invalid*
A, B, C	*Unfair* *1*	*Doesn't Matter* *3*	*Fair* *5*	*invalid*	*invalid*
A or B	*No* *1*	*Yes* *5*	*invalid*	*invalid*	*invalid*

The objective of this portion of the research was to investigate whether team formation method affected team attitudes, behaviors, and outcomes. Tables 5, 6, and 7 show the results of a series of 2 Group (Self-selected and Matched-performance) X 2 Time (Mid-semester and End-semester) mixed ANOVAs for each question in team attitudes, dynamics, and outcomes respectively, as well as group means at each of the time periods for each question.

Table 5 presents the statistical findings for questions related to team attitudes. In most cases, the response averages were generally at the level of "agree" (ranging from 3.7 to 4.3). The slight exceptions were the end-semester ratings of satisfaction about the method of team formation (averages of 3.4 and 3.5). There were four significant main effects for Time, two significant main effects for Group, and one significant interaction. All had small effect sizes. Significant main effects for Time included a decrease in overall team attitude for both team formation methods, a decrease in satisfaction with the manner by which the team was formed, a decrease in reported enjoyment for working with the team, and a decrease in perceived effectiveness of working with the team. The two significant Group main effects indicated that Self-selected teams believed their team formation method was "more fair" than the Matched-performance teams, and they were less likely to agree that they wished the teams had been formed differently. These main effects for Group are consistent with the generally more positive team attitudes seen in the self-selected group from Chapman et al. (2006) and Mahenthiran et al. (2000), who also found improved student attitudes by giving students some control of the team selection process. This result supports our hypothesis that self-selected teams show more positive ratings of social factors. However, an interaction regarding overall satisfaction with how teams were formed showed a decrease for both groups with a greater drop in satisfaction for the Self-selected group. Thus, over time, Matched-performance teams may form bonds that offset the lack of choice in initial formation.

Table 5. Team Attitudes Measures

	Main Effects and Interaction						Response Means	
	Group Main Effect		Time Main Effect		Interaction		Self-Selected	Matched-Performance
Questions Relating to Team Attitudes	F η_P^2	p	F η_P^2	P	F η_P^2	p	*Mid-sem* *End-sem*	*Mid-sem* *End-sem*
My overall attitude toward my team was **positive**	1.99 <.01	0.16	**8.74** **0.01**	**<.01**	0.67 <.01	0.41	4.1 4.0	4.2 4.1
Overall, I believe my team formation method (instructor assigned, cadet choice) was **fair**	**9.08** **0.02**	**<.01**	0.38 <.01	0.54	0.59 <.01	0.44	4.3 4.3	4.1 4.0
Overall, I wish my team had *not* been **formed** in a different way[R]	**8.29** **0.01**	**<.01**	**13.5** **0.02**	**<.01**	**5.44** **0.01**	**0.02**	3.8 3.5	3.5 3.4
Working with my team was a **bad/good** experience	0.00 <.01	0.98	0.66 <.01	0.42	0.25 <.01	0.62	3.9 3.8	3.9 3.9
Working with my team was **valueless/valuable** for my learning	0.09 <.01	0.76	1.80 <.01	0.18	0.07 <.01	0.79	3.8 3.8	3.8 3.7
Working with my team was **not enjoyable/enjoyable**	0.06 <.01	0.80	**5.04** **0.01**	**0.03**	0.44 <.01	0.51	3.8 3.8	3.9 3.8
Working with my team was **ineffective/effective**	0.00 <.01	0.99	**7.37** **0.01**	**0.01**	3.39 <.01	0.07	3.8 3.8	3.9 3.7

[R] indicates question was reverse scored italic words added for analysis clarification
Bold values indicate a significant effect of at least p≤ .05.

For team behaviors and dynamics measures, shown in Table 6, the range of average scores was much wider (1.9 up to 4.2). Again, there were relatively few significant effects, and most of them had small effect sizes. Group main effects included the Self-selected group reporting higher levels of meeting outside of class and the Matched-performance group reporting more collaborative contributions. Significant main effects of Time included a decrease in reported worrying about grades on team projects, a decrease in being task oriented, a decrease in going "straight to work" in and out of class work sessions, an increase in meeting outside of class (medium effect size), and a decrease in reporting that they did not complete work for their team members. These Time main effects indicate that our entire population behaved similarly over time regardless of team formation method, and that team dynamics change in measurable ways across the semester. The one significant interaction indicated that the Self-selected group reported higher levels of good communication at mid-semester, but by the end of the semester, the Matched-performance group reported higher levels of good communication. This finding is in contrast to Chapman et al. (2006), who found that self-selected teams reported significantly higher levels of communication at the end of the semester.

Taken together, these results suggest student engagement in their groups show many shifts over time. Overall, these changes seem intuitive as the team members became acquainted over the course of the semester, leading to effects on teamwork measures (e.g., less worry, more socialization during meetings so less task-focused, more meetings outside of class, increase in completing work for teammates). The result that Self-selected teams met outside of class more often than the Matched-performance groups suggests that early familiarity with team members leads to an advantage with respect to making plans to work together outside of class time. Again, this supports our hypothesis that self-selected teams may show more positive ratings of social factors since they have friends on their teams. Matched-performance groups' greater levels of self-reported collaboration may be a result of greater similarity in academic ability, or perhaps greater similarity in their approach to academic tasks.

Table 6. Team Behaviors and Dynamics Measures

	Main Effects and Interaction						Response Means	
	Group Main Effect		Time Main Effect		Interaction		Self-Selected	Matched-Performance
Questions Relating to Team Behavior and Dynamics	F η_P^2	p	F η_P^2	p	F η_P^2	p	*Mid-sem* *End-sem*	*Mid-sem* *End-sem*
I was **enthusiastic** about working together with my team	0.49 <.01	0.40	0.24 <.01	0.62	0.00 <.01	0.99	3.9 3.9	3.9 3.9
I *was not* **worried** about my grade on team projects[R]	1.41 <.01	0.24	**15.9** **0.02**	**<.01**	2.02 <.01	0.16	3.3 3.1	3.3 3.2
My team had good **communication** (timely, respectful, informative)	0.13 <.01	0.72	0.14 <.01	0.71	**7.15** **0.01**	**0.01**	3.7 3.6	3.6 3.7
My team **resolved conflict** effectively	3.31 <.01	0.07	1.38 <.01	0.24	0.79 <.01	0.37	4.0 3.9	4.1 4.0
I asked other team members for **help** when needed	0.38 <.01	0.54	0.01 <.01	0.91	1.58 <.01	0.21	4.0 4.0	3.9 4.0
My team met outside of class	**17.3** **0.03**	**<.01**	**80.1** **0.11**	**<.01**	0.07 <.01	0.79	2.9 3.4	2.5 3.0
During in or out of class work sessions, my team was **task oriented**	3.27 <.01	0.07	**26.6** **0.04**	**<.01**	0.64 <.01	0.42	4.1 3.9	4.2 4.0
During in or out of class work sessions, my team went straight to work	3.26 <.01	0.07	**30.2** **0.04**	**<.01**	0.47 <.01	0.49	4.1 3.9	4.2 4.0
My team made **collaborative contributions**	**5.68** **0.01**	**0.02**	1.87 <.01	0.17	0.78 <.01	0.38	3.9 3.8	4.0 4.0
My team *did not* **divide-and-conquer** the tasks[R]	3.01 <.01	0.08	1.63 <.01	0.20	0.03 <.01	0.86	2.0 2.0	1.9 1.9
My team **reviewed our final turn-in product as a team** so we could make final revisions together	0.01 <.01	0.91	0.54 <.01	0.46	0.38 <.01	0.54	3.2 3.2	3.2 3.2
I *did not* complete work for other team members[R]	3.44 <.01	0.06	**7.70** **0.01**	**0.01**	0.03 <.01	0.86	2.9 2.7	3.0 2.9

[R] indicates question was reverse scored italic words added for analysis clarification
Bold values indicate a significant effect of at least p≤ .05.

Although there were no significant effects, the means for "my team did not divide-and-conquer the tasks" were very low (1.9 and 2.0) compared to all other responses (which ranged between 2.5 and 4.2), suggesting that dividing-and-conquering was a common strategy for both groups (and arguably also a non-ideal approach for team tasks). This finding could be an indication of confidence in team members' abilities or simply a strategy to cope with time constraints. More positively, some of the highest response means for both groups were for being task oriented, going straight to work, asking team members for help, and resolving conflict. All of these factors are indicators of healthy team behaviors.

Table 7 summarizes the statistical results of the team outcome measures. Again, there were few significant effects. We found no Group difference in making new friends, and overall our students reported high levels of making new friends. There was a significant Group difference for already having friends on the team; not surprisingly, the Self-selected group reported higher levels than the Matched-performance group. This was consistent with other research in the literature (Pociask et al., 2017; Rienties, Alcott, & Jindal-Snape, 2014; Chapman et al., 2006). We found no Group difference in the self-assessed level of the quality of the team work, but we did find a significant decrease over time. These results suggest that Matched-performance groups achieve similar team outcomes to Self-selected groups, even though they self-report starting out with fewer friends on their teams.

Table 7. Team Outcome Measures

	Main Effects and Interaction						Response Means	
	Group Main Effect		Time Main Effect		Interaction		Self-Selected	Matched-Performance
Questions Relating to Team Outcomes	F η_P^2	p	F η_P^2	p	F η_P^2	p	*Mid-sem* *End-sem*	*Mid-sem* *End-sem*
Some of my **friends** were on my team	**31.6** **0.05**	**<.01**	1.39 <.01	0.24	0.12 <.01	0.73	2.9 3.0	2.2 2.2
I made **new friends** on my team	0.24 <.01	0.63	0.74 <.01	0.39	1.91 <.01	0.17	4.5 4.5	4.5 4.6
My self-assessed quality of our **team work**	0.73 <.01	0.39	**3.90** **0.01**	**0.05**	0.30 <.01	0.59	3.9 3.8	3.9 3.9
My self-assessed quality of my **individual work** on the team	0.62 <.01	0.43	0.03 <.01	0.85	0.03 <.01	0.85	4.0 4.0	4.0 4.0
My self-assessed quality of my **individual work in this course**	0.06 <.01	0.81	1.98 <.01	0.16	0.01 <.01	0.91	4.1 4.0	4.1 4.0

Bold values indicate a significant effect of at least p≤ .05.

Table 8 shows our end-semester data alongside the matched question results from Chapman et al. (2006)[3]. This comparison allows us to evaluate how their randomized groups might compare to our matched-performance groups when comparing to self-selected groups (both studies). Chapman et al. found significant group differences for 11 of their 15 questions, with eight of them showing more positive ratings for their self-selected group compared to their randomized group. In contrast, we found no significant group differences for the 15 similar questions; however, we did find significant group effects for a few of our questions that did not match those from Chapman et al. Overall, the lack of group differences in the current study compared to those found by Chapman et al. support

[3]The results from Chapman et al. (2006) shown in Table 8 have been normalized to a 5-point scale for direct comparison with the current study. The study by Chapman et al. (2006) had 16 sections, with 583 end-semester survey respondents. The group size was 2-6 members with 4 as the mean. The composition of the study was 81% seniors, 14% juniors with 62% male and the remainder female.

our prediction that our group differences would be smaller than those found when comparing self-selected and randomly-selected teams.

In most cases, the overall mean scores for each question were similar across the two studies, except our students self-reported more positively (greater than a 0.3 point difference on the five-point scale) in terms of positive attitude, being enthusiastic to work together, worrying about grades on group projects, going straight to work, and completing the work of others. Overall, there was more than a 0.85 point difference with our students making new friends on the team. These more positive self-reported data could be due to the common shared experiences unique to that of a military academy.

Table 8. Self-Report Response Comparison with Chapman et al. (2006)

	Questions	Response Means (End-semester)			Chapman (2006) Response Means		
		Self-Selected	Matched-Performance	Overall	Self-Selected	Random Assignment	Overall
Team Attitude Measures	My overall attitude toward my team was **positive**	3.97	4.08	4.03	3.81*	3.67	3.73
	Working with my team was a **bad/good** experience	3.84	3.86	3.85	4.00	3.89	3.94
	Working with my team was **valueless/valuable** for my learning	3.77	3.74	3.75	3.79*	3.61	3.69
	Working with my team was **not enjoyable/enjoyable**	3.78	3.77	3.77	3.76	3.66	3.71
	Working with my team was **ineffective/effective**	3.79	3.72	3.76	3.89**	3.71	3.79
Team Dynamic Measures	I was **enthusiastic** about working together with my team	3.93	3.89	3.91	3.81***	3.44	3.59
	I **worried** about my grade on team projects	3.07	3.23	3.15	2.56	2.77	2.69
	My team had good **communication** (timely, respectful, informative)	3.58	3.72	3.65	4.13***	3.85	3.96
	My team **resolved conflict** effectively	3.92	4.05	3.98	4.19*	4.00	4.08
	I asked other team members for **help** when needed	3.97	3.99	3.98	4.09*	3.91	3.99
	During in or out of class work sessions, my team was **task oriented**	3.95	4.01	3.98	3.73	3.91*	3.84
	During in or out of class work sessions, my team went straight to work	3.91	3.98	3.94	3.33	3.68***	3.53
	I completed work for other team members	2.74	2.88	2.81	2.23	2.56**	2.42
Team Outcome Measures	I made **new friends** on my team	4.50	4.62	4.56	3.92***	3.53	3.69
	My self-assessed quality of our **team work**	3.80	3.87	3.83	4.51	4.48	4.49

For Chapman et al. (2006): Ratings were based on a 7-point itemized scale where 1 = unfavorable and 7 = favorable. These were normalized to a 5-point scale for direct comparison with the current study. Significance of difference between random and self-selected groups: *p ≤ .05, **p ≤ .01, ***p ≤ .001.

Overall, our data indicate a slight benefit on performance using Matched-performance to form teams, and a mixture with respect to impact on the self-reported measures. Because our design included both types of measures, and we were able to use identifiers to link the data, we are also able to investigate how these two types of measures correlate. Ultimately, our hope is to help instructors make choices that will enhance student learning. Thus, if we find that certain self-reported measures are more predictive of performance, then we should consider prioritizing team formation methods that would maximize those measures. Tables 9, 10, and 11 summarize the correlations for both the Self-selected and the Matched-performance groups between the two team and the two individual performance measures for each question in the three groupings of self-reported questions, respectively.

Several general trends are apparent, as are some interesting item relationships. For both groups, Attitudes measures (Table 9) and Behaviors and Dynamics measures (Table 10) show more and stronger significant correlations between the self-reported measures and team performance measures than between the self-reported measures and the individual performance measures. Overall, the correlations are more frequent and stronger for the Matched-performance group, which also showed more significant relationships between the self-reported measures and the individual performance measures. Finally, the correlations are stronger at the end of the semester than mid-semester. This latter observation suggests a growing influence of the team dynamics and influences.

Table 9. Team Attitudes and Performance Correlations

	Self-Selected					Matched-Performance				
Questions Relating to Team Attitudes	N Mid - semester N End - semester	Mid-sem Team Avg	2nd half Team Avg	Mid-sem Ind Avg	2nd half Ind Avg	N Mid - semester N End - semester	Mid-sem Team Avg	2nd half Team Avg	Mid-sem Ind Avg	2nd half Ind Avg
My overall attitude toward my team was **<u>positive</u>**	304 304	.18** .11	.08 .22***	.07 .12*	.07 .01	385 386	.16** .21***	.09 .20***	.07 .18**	.03 .12*
Overall, I believe my team formation method (instructor assigned, cadet choice) was **<u>fair</u>**	266 260	.16* .17*	.17** .22***	.10 .10	.01 .00	322 322	.16** .24***	.12* .15**	.08 .18***	.01 .08
Overall, I wish my team had *not* been **<u>formed</u>** in a different way[R]	306 304	.19** .21***	.13* .27***	.09 .09	.04 .03	385 384	.20*** .29***	.14** .23***	.14** .20***	.08 .15**
Working with my team was a **<u>bad/good</u>** experience	306 303	.23*** .14*	.11 .22***	.08 .12*	.03 -.02	385 385	.28*** .29***	.17** .32***	.17*** .23***	.17** .19***
Working with my team was **<u>valueless/valuable</u> for my learning**	306 303	.21*** .12*	.10 .16**	.10 .11*	.05 .02	386 384	.18*** .16**	.09 .16**	.07 .09	.06 .08
Working with my team was **<u>not enjoyable/enjoyable</u>**	304 304	.19*** .10	.06 .18**	.08 .15*	.04 .02	386 385	.15** .24***	.11* .23***	.08 .14**	.03 .14**
Working with my team was **<u>ineffective/effective</u>**	304 304	.24*** .20***	.15* .17**	.01 .12*	-.06 .00	386 384	.27*** .24***	.15** .25***	.12* .15**	.07 .12*

* p≤.0
** p≤.01
*** p≤.001

Table 10. Team Behavior and Dynamics and Performance Correlations

Questions Relating to Team Behaviors and Dynamics	Self-Selected					`Matched-Performance				
	N Mid - semester N End - semester	Mid-sem Team Avg	2nd half Team Avg	Mid-sem Ind Avg	2nd half Ind Avg	N Mid - semester N End - semester	Mid-sem Team Avg	2nd half Team Avg	Mid-sem Ind Avg	2nd half Ind Avg
I was **enthusiastic** about working together with my team	309 308	.16** .13*	.02 .25***	.07 .13*	-.01 -.01	387 388	.20*** .20***	.08 .21***	.14** .21***	.07 .12
I *did not* **worry** about my grade on team projects[R]	310 307	.31*** .25***	.11 .30***	-.04 .07	-.01 .03	388 387	.37*** .30***	.25*** .36***	.15** .13*	.05 .09
My team had good **communication** (timely, respectful, informative)	309 308	.23*** .06	.10 .23***	.07 .11*	-.06 -.08	387 386	.29*** .29***	.28*** .29***	.22** .15**	.15* .13*
My team **resolved conflict** effectively	309 309	.22*** .18**	.08 .15**	.02 .14*	-.01 .02	389 389	.08. .13*	.09 .14**	.06 .11*	.06 .10*
I asked other team members for **help** when needed	309 310	.10 .08	.07 .00	.02 .10	.04 .06	388 387	.07 .15**	-.04 .06	.01 .07	-.04 .00
My team met outside of class	296 286	-.11 -.11	.08 .09	.06 .06	-.08 -.07	367 367	-.06 -.17**	.10 -.10	.07 -.08	-.03 -.01
During in or out of class work sessions, my team was **task oriented**	310 310	.30** .07	.07 .10	.09 .16**	.07 .01	387 388	.16** .23***	.13* .21***	.14* .07	.11* .08
During in or out of class work sessions, my team went straight to work	310 310	.14* .05	.05 .05	.07 .14*	.05 -.01	388 387	.16** .16**	.08 .16**	.14** .07	.15** .09
My team made **collaborative contributions**	310 310	.18** .09	.01 .15*	.02 .15**	.01 -.02	389 389	.13* .16**	.10 .20***	.08 .14**	.02 .11*
My team *did not* **divide-and-conquer** the tasks[R]	310 310	-.18** -.07	-.16** -.19***	-.02 -.13*	-.03 -.02	389 389	-.16** -.19***	-.18** -.23***	-.07 -.14*	-.03 -.10
My team **reviewed our final turn-in product as a team** so we could make final revisions together	310 310	.23*** .01	.01 .09	-.02 .02	-.13* -.16**	389 389	.17** .09	.10* .08	.14** -.01	-.02 -.04
I *did not* complete work for other team members[R]	310 308	.06 .00	.08 .09	-.15* -.11	-.10 -.16**	389 387	.05 .10	.11* .15**	.02 .04	.05 .04

* p≤.0
** p≤.01
*** p≤.001

Two interesting item relationships are on division of tasks within the team and having or making new friends on the team. For the Behavior and Dynamics measures (Table 10), the self-reported measure "my team did not divide-and-conquer" had a negative correlation, suggesting that the more the team divided and conquered the task the better the performance, and not surprisingly, this was more connected with team performance measures. While we do not propose that dividing and conquering is the best team approach to completion of tasks, again this finding could be an indication of strong confidence in team members' abilities. It could also be a strategy for team members to cope with over-demanding time constraints.

For the Outcomes measures (Table 11), there are less obvious differences between the Self-selected and Matched-performance groups, or between the team and the individual performance measures. Rather, the correlations suggest that for both groups, forming new friends is more positively related to performance (team and individual) than already having friends on a team, which tended to negatively correlate with performance. Not surprisingly, self-assessment of work quality correlated strongly for both groups for both team and individual performance measures.

Overall, these correlation results suggest that many student perceptions about their teams capture aspects of team functioning that ultimately could impact team performance, and that in some cases, team perceptions seem related to factors that also impact performance in a broader manner. They reinforce the importance for instructors to establish healthy teams and provide structure and support to foster good team functioning.

Table 11. Team Outcome Measures and Performance Correlations

	Self-Selected					Matched-Performance				
Questions Relating to Team Outcome Measures	N Mid - semester N End - semester	Mid-sem Team Avg	2nd half Team Avg	Mid-sem Ind Avg	2nd half Ind Avg	N Mid - semester N End - semester	Mid-sem Team Avg	2nd half Team Avg	Mid-sem Ind Avg	2nd half Ind Avg
Some of my **friends** were on my team	293 288	-.13* -.05	-.05 .02	-.17** -.09	-.13* -.15*	371 372	.01 .01	-.03 -.06	-.03 -.05	-.05 -.10
I made **new friends** on my team	288 280	.20** .17**	.15* .15*	.13* .19**	.10 .13*	360 360	.13* .04	.11* .08	.15** .14**	.05 .08
My self-assessed quality of our **team work**	305 305	.29*** .11	.07 .23***	.05 .07	.10 -0.05	382 385	.29*** .29***	.22** .34***	.17** .21***	.10 .22***
My self-assessed quality of my **individual work** on the team	304 300	.25*** .20***	.17** .19***	.23*** .29**	.19*** .20***	383 383	.23*** .19***	.10* .19***	.25*** .17**	.20*** .12*
My self-assessed quality of my **individual work in this course**	292 295	.11 .17**	.17** .15*	.32*** .35***	.24*** .24***	365 376	.20*** .19***	.07 .13*	.27*** .27***	.24*** .20***

* p≤.0
** p≤.01
*** p≤.001

Conclusions

Our study compared groups formed by matching students based on similar academic potential with groups self-selected by students. Our inclusion of mid- and end-of-semester performance measures as well as self-reported attitudes and behaviors allowed an analysis of the developing dynamics of team formation and how they impact both individual as well as team assignment performance. The finding that many of the questionnaire items significantly correlated with performance measures, and that the correlations strengthened from mid to end of the semester, reinforces our conclusion that it is important to thoughtfully consider how teams are formed and to support the development of well-functioning teams.

As predicted, we found that the Matched-performance groups had significantly higher grades on several performance measures, with a larger effect on the team grades than the individual grades; however, overall the effect sizes were small, approximately 1% of the course grade. Although small, the fact that the impact on team grades increased across the semester suggests a growing positive influence when using Matched-performance teams. Students in the Matched-performance groups might take longer to bond, but once they do they become more effective with respect to the team activities.

The self-reported team attitudes, dynamics and outcomes responses supported our prediction that as a group our Matched-performance and Self-selected teams reported similar attitudes and behaviors. Furthermore, our group differences were smaller than those obtained when comparing self-selected and randomly-formed teams (e.g., Chapman, et al., 2006). We found significant group differences for some of questions we developed for this study, and these help us understand some of the differing influences of our two methods of team formation, especially when combined with the time data (mid versus end of semester) and when correlated with the performance measures. For example, both the sense of fairness (higher for Self-selected) and the likelihood to collaboratively contribute (higher for Matched-performance) showed significant group differences as well as significant correlations with several of the performance measures. Further, for both groups there were small but significant decreases from mid to end of semester in satisfaction with team formation method, general enjoyment, and task-orientation over time.

All of these time-related factors also showed significant correlations with the performance measures. Therefore, these factors seem like natural points for instructor attention when using teams in a course. For example, if an instructor chooses to use Matched-performance groups due to the potential benefits for performance and likelihood for collaboration, the instructor should explicitly address fairness issues.

One of the most salient correlation findings relates to the impact of already having friends on the team versus making new friends. Both groups reported similar levels of making new friends, and making new friends was strongly and positively related to better performance. However, already having friends on the team was our largest significant group difference (much more often true for Self-selected versus Matched-performance teams), and it was negatively correlated with performance. Based on this combination of factors, along with the performance results previously discussed, we recommend Matched-performance groups over Self-selected groups.

Although we make the recommendation for Matched-performance teams over Self-selected teams, we acknowledge that there are many other team formation techniques we did not study (e.g., the hybrid teams used by Mahenthiran & Rouse, 2000, randomly-formed teams), and that there are contextual factors that might also impact team performance and attitudes. For example, this study was conducted at the U.S. Air Force Academy (USAFA), a military institution, with all participants being first-year (freshman) students. Students in first-year courses are less likely to already know each other than students in upper-level courses within a major. Additionally, our institution has a rigorous course attendance policy, which makes it more likely that all members are present during team interactions in class. None of these factors are completely unique to our institution, but they may have influenced our results and should be kept in mind by others who may form teams in different contexts.

Future research can build on our work and that of other researchers to further investigate factors that might impact generalization, how team dynamics shift over time, and whether or not team dynamic interventions might influence team functioning. For example, instructors and other team leaders might incorporate intentional team-building activities or implement tracking of individual accountability on team efforts. Team performance and dynamics are complex, but given the key role teams play in academic, industry, military and other endeavors requiring cooperative productivity of individuals, it is important to continue research efforts that help us identify and enhance factors that contribute positively to team functioning and mitigate those factors that are detrimental to team functioning.

Acknowledgements

Many individuals who taught Introduction to Engineering at the United States Air Force Academy have contributed to this effort. The authors would like to extend a sincere thank you to the instructors and the students who made this paper possible.

References

Bacon, D., Stewart, K., & Silver, W. (1999). Lessons from the best and worst student team experiences: How a teacher can make the difference. *Journal of Management Education*, Vol. 23, No. 5, 467-488. http://journals.sagepub.com/doi/pdf/10.1177/105256299902300503

Bronson, P., & Merryman, A. (2013). Top Dog: The Science of Winning and Losing. New York.

Carrell, S., Fullerton, R., & West, J. (2009). Does Your Cohort Matter? Measuring Peer Effects in College Achievement. *Journal of Labor Economics*, Vol. 27, no. 3, 439-464.

Chapman, K., Meuter, M., Toy, D., & Wright, L. (2006). Can't We Pick our Own Groups? The Influence of Group Selection Method on Group Dynamics and Outcomes. *Journal of Management Education*, Volume 30, 557-569. http://journals.sagepub.com/doi/pdf/10.1177/1052562905284872

Connerley, M., & Mael, F. (2001) The Importance and Invasiveness of Student Team Selection Criteria. *Journal of Management Education*, Vol. 25, No. 5, 471-494. http://journals.sagepub.com/doi/abs/10.1177/105256290102500502

Denton, H.G. (1996) Developing design team working capacity: some planning factors emerging from a survey of engineering design courses. Loughborough University Institutional Repository, IDATER 1996 Conference, Loughborough University, 1-6.

Feichtner, S., & Davis, E. (1984). Why Some Groups Fail: A Survey of Students' Learning Experiences with Learning Groups. *Journal of Management Education*, Vol. 9, No. 4, 58-73. http://journals.sagepub.com/doi/abs/10.1177/105256298400900409

Hernandez Nanclares, N., Rienties, B., & Van den Bossche, P. (2012) A Longitudinal Analysis of Knowledge Spillovers in the Classroom. Chapter 11 of the book by Editors Van den Bossche, P., Gijselaers, W., and Milter, R., "Learning at the Crossroads of Theory and Practice: Research on Innovative Learning Practices," Volume 4, Springer, 157-175.

Katzenbach, J.R., & Smith, D.K. (1993). The Discipline of Teams. *Harvard Business Review*, reprint 2005. https://hbr.org/2005/07/the-discipline-of-teams

Kozlowski, S., & Ilgen, D. (2006). Enhancing the Effectiveness of Work Groups and Teams. *Psychological Science in the Public Interest*, Volume 7, Issue 3, 77-124. https://doi.org/10.1111/j.1529-1006.2006.00030.x

Mahenthiran, S., & Rouse, P. (2000). The impact of group selection on student performance and satisfaction. *The International Journal of Educational Management*, Vol. 14, Issue 6, 255 – 265. http://www.emeraldinsight.com/doi/abs/10.1108/09513540010348043

Matta, V., Luce, T., & Ciavarro, G. (2010). Exploring Impact of Self-selected Student Teams and Academic Potential on Satisfaction. 2010 ISECON Preceedings v27 n1304, Information Systems Educators Conferences, Nashville, TN. http://proc.edsig.org/2010/pdf/1304.pdf

McClough, A., & Rogelberg, S. (2003). Selection in Teams: An Exploration of the Teamwork Knowledge, Skills and Ability Test. *International Journal of Selection and Assessment*, Volume 11, Number 1, 56-66. http://onlinelibrary.wiley.com/doi/10.1111/1468-2389.00226/abstract

Oakley, B., Brent, R., Felder, R., & Elhajj, I. (2004). Turning Student Groups into Effective Teams. *Journal of Student Centered Learning*, Volume 2, No. 1, 9-34. http://www4.ncsu.edu/unity/lockers/users/f/felder/public/Papers/Oakley-paper(JSCL).pdf

Pociask, S., Gross, D., & Shih, M. (2017). Does Team Formation Impact Student Performance, Effort and Attitudes in a College Course Employing Collaborative Learning? *Journal of the Scholarship of Teaching and Learning*, Volume 12, No. 3, 19-33. https://josotl.indiana.edu/article/view/21364

Rienties, B., Alcott, P., & Jindal-Snape, D. (2014). To Let Students Self-Select of Not: That is the Questions for Teachers of Culturally Diverse Groups. *Journal of Studies in International Education*, Volume 18(1), 64-68. http://journals.sagepub.com/doi/abs/10.1177/1028315313513035

Shen, S., Prior, S., White, A., Karamanoglu, M. (2007). Using Personality Type Differences to Form Engineering Design Teams. *Engineering Education*, 2(2). http://www.tandfonline.com/doi/full/10.11120/ened.2007.02020054

Scott, E., & Pollock, M. (2006). Effectiveness of Self-selected Teams: A Systems Development Project Experience. *Issues in Informing Science and Information Technology*, Volume 3, 601-617. http://proceedings.informingscience.org/InSITE2006/IISITScot217.pdf

United States Air Force Academy Strategic Plan. (2015). https://www.usafa.edu/app/uploads/USAFA-Strategic-Plan.pdf

Journal of the Scholarship of Teaching and Learning, Vol. 20, No. 1, April 2020, pp.22-39.
doi: 10.14434/josotl.v20i1.24645

Evaluating Online Courses via Course-Related Competencies – A Mixed-Methods Quasi-Experiment Evaluation Study of an HIV Prevention Webcourses among College Students

Su-I Hou
University of Central Florida
su-i.hou@ucf.edu

Abstract: This convergent mixed method quasi-experiment study evaluates the effectiveness of an HIV Prevention Webcourses on reaching HIV related competencies among college students at a large public university in Florida. College students in health majors participated in the study, experiment group were students enrolled in the HIV Webcourses and comparison group were students who enrolled in non-HIV related courses from the same college. Six HIV competencies along with HIV knowledge were measured. Qualitative comments on own learning were also gathered from the experiment group of students. A total of 944 students participated (508 experimental and 436 comparison students). The reliability of the 6-item HIV competency scale showed satisfactory internal consistencies (Cronbach alpha = .914; CITCs ranged .670-.804). Regression analyses, controlling for their baseline scores, showed that students in the HIV Webcourses scored significantly higher than comparison group (all p<.001) on all the six competencies. Regression analyses also showed significant differences on objective HIV knowledge test, perceived HIV knowledge in general and about HIV testing specifically between groups (p<.001). Students commented they genuinely loved this course as it was extremely well organized, very useful, interesting and informative, and enjoyed the professor's passion of the topics. Students also commented how the personal stories and videos gave perspectives and provided life-changing lessons. Both the quantitative and qualitative data on student learnings convergently demonstrated the impact of this online course. Current study suggested effective design strategies and provided data to support the effectiveness of Webcourses on reaching course-related competencies among college students.

Keywords: competencies, HIV Prevention, Webcourses, mixed methods research, quasi-experiment

Introduction

The introduction of this article is structuring the following way. First, I brief the growing trend of online education. Second, I discuss challenges and some current best practices of online education. Third, I summarize current research on evaluating the effectiveness of online education. Fourth, I point out the importance and feasibility of measuring course-related competencies as student learning outcomes, using an HIV Prevention Webcourses as a case study example. Finally, I summarize gaps from existing research, and delineate the current study purpose on providing empirical data to demonstrate the effectiveness of an online HIV Prevention Webcourses on reaching course-related competencies among college students, as well as identifying effective online course design features based on this case study.

The Growing Trend of Online Education

Online courses are defined as those in which most or all the course content is delivered online, typically with no face-to-face meetings (Allen & Seaman, 2013). Data show that students in higher education taking at least one online class have significantly increased over years, from 1.6 million in 2002 to 6.7 million in 2011. This represents a compound annual growth rate of 17.3% (Allen & Seaman, 2013). While overall higher education enrollment has declined, many institutions have continued to add online learning programs. In fact, more than 60% of the higher education institutuions viewed that online education cruicial to their long-term strategic goals in the lst five years (Allen, Seaman, Poulin, & Strut, 2016).

Challenges and Strategies for Online Education

Despite the growing trend of online education, barriers and challenges exist both among students and instructors. Some of the common challenges or barriers regarding online education from students' perspective include, perception of lower quality education, isolation, and detachment of peer and instructor interactions. Studies show that students strongly expressed the importance of the presence of teacher (Richardson & Lowenthal, 2017; Tichavsky, Hunt, Driscoll, & Jicha, 2015). Challenges from instructors' perspectives include, time commitment to course development, potential of student cheating and collaboration, and technology frustrations. (Tichavsky et al., 2015).

Moore (1997) introduced the concept of transactional distance, the psychological and communications space separating between the learner and instructors which will need to be crossed especially in online learning (Moore, 1997). Thus, selecting appropriate communication medium taking learner and content characteristics into consideration when delivering the teaching and learning are critical (Moore, 1997). Some recommended ways to minimize transactional distance in online learning include creating a communicative learning culture such as online discussion groups, identifying preferred delivery format meet learners' needs, providing self-paced online resources to empowers learner take charge of own learning, and offering regular personalized feedback (Pappas, 2017).

Baran and Correia identified some best practices for successful online teaching for instructors. These include that instructors know the course content, know the students, have clear course design and structure, facilitate teacher-student relationships, guide student learning, evaluate online courses, and maintain teacher presence (Baran & Correia, 2014). The Community of Inquiry (COI), one of the most frequently used online learning frameworks, notes that a meaningful educational experience consists of teaching presence, social presence, and cognitive presence (Garrison & Anderson, 2000). Thus, online courses should thrive to create opportunities to enhance spontaneity and emergent design, coach students how to learn online, explore the use of diverse technologies for enhancing communication and social presence, and articulate and manage the expectations of the online community (Stodel, Thompson, & MacDonald, 2006).

Effectiveness of Online Education

Recent studies have focused on evaluating the effectiveness of online education through comparing student's grades achieved between online learning and traditional classroom learning (Shotwell & Apigian, 2015; Stack, 2015; Tichavsky et al., 2015). These studies showed mixed results on student grades. While a meta-analysis study reports that on average, student performance is higher for online sections than face-to-face learning (Means, Toyama, Murphy, Bakia, & Jones, 2009), another study show that exam scores are higher for traditional classes than online classes (Figlio, Rush, & Yin, 2013). In particular, Figlio and colleagues found that Hispanic students, male students, and lower-achieving students, showed the greatest score differences on the mode of course delivery. A recent study, controlling for self-selection effects on delivery mode and the proctoring of exams, shows that the academic performance of online students was in fact the same as the traditional students (Stack, 2015).

Course related Curriculum Competencies Measures

Competencies learned can demonstrate how well the course experience aligns with the scope of the academic curriculum (Ezeonwu, Berkowitz, & Vlasses 2013; Hou, & Pereira, 2017; Hou, 2009). Due to the need of curriculum alignment between competencies and workforce related skills, competencies (instead of grades) are increasingly used as an important indicator by academic accreditation bodies to assess the quality of academic programs. Course learning objectives, on the other hand, describe the knowledge and skills a student is expected to demonstrate upon completion of a specific course. When carefully designed, these objectives are intended to relate, in some discernable way, to the competencies of the overall program of study.

A major challenge to demonstrate the impact of a course has been to measure student outcomes. Behavior change or skill acquisition takes time and may not always be realistic or sensitive outcome indicators for evaluation immediately after a course. Behavioral scientists have come up with various indicators to predict behavioral outcomes. Among these, the Social Cognitive Theory (Bandura, 1986) is arguably the most significant and widely used and accepted theory that has been applied to various learning outcomes. According to the Social Cognitive Theory, how people behave can often be better predicted by the beliefs they hold about their capabilities, which are called *self-efficacy* beliefs. These self-perceptions or judgment of a person's capability to accomplish a certain level of performance help determine what individuals do with the knowledge and skills they have. Self-efficacy (SE) and academic competencies are indicators that have been successfully used to assess the impact of student learning after courses (Hou, & Pereira, 2017; Hou, 2009).

HIV Webcourses Case Study

College students are at the life stage of engaging in HIV-risk behaviors. Studies showed many college students have oral, vaginal and/or anal sex, low condom-use, and have multiple partners (Caico, 2014, Calloway, Long-White, & Corbin, 2014; Hou, 2009a; Hou, 2007). Besides protective sex via condom use, HIV testing is another important preventive strategy. Both CDC and The U.S. Preventive Services Task Force (USPSTF) recommend regular HIV testing among all adults and young people so everyone can be aware of own HIV status (Branson et al., 2006; CDC, 2017). For college students, some major barriers towards testing have been the low perceived risk of HIV infection, lack of testing-specific knowledge, fear of finding out if positive, and HIV/AIDS related stigma (CDC, 2015; Hoppel, 2012; Hou, 2009b; Hou, 2008; Hou & Luh, 2007; Hou, 2007; Hou, 2004;). A number of systematic reviews and meta-analyses studies have shown web-based HIV programs are effective to change HIV related knowledge and behavioral outcomes (Muessig, Nekkanti, Bauermeister, Bull, & Hightow-Weidman, 2015; Hosseini, 2013; Noar, et al., 2009; Noar, et al., 2010).

However, limited studies have evaluated HIV education programs using teaching and learning related outcomes. There is a lack of validated HIV related competency measurement for learning. One study examined HIV/AIDS-related competencies among nursing students in South Africa (Modeste & Adejumo, 2014). Using a qualitative approach and systematic research synthesis, this study identified core competencies related to HIV and AIDS for nursing graduate students, including HIV related knowledge, ethics, policy, interdisciplinary care, professional development, holistic safe practice, and health education. Yet, quantitative measurement on key HIV-related competency are not available and needs to be developed and tested.

Gap

Although existing literature show that using web-based program to deliver HIV prevention information can be effective and has the potential to reach larger audience especially among young people (Muessig, et al., 2015), there are several important research gaps. First, there is a lack of HIV related competency measurements to assess student curriculum competencies beyond individual knowledge or behavior change. Given colleges students are our next generation of change agents in the larger society and community, developing theses college-curriculum HIV related competency measurements are critical to ensure students can be a more effective change agent on HIV prevention and education efforts when they graduate. These could include, but not limited to comprehension on how HIV impact the global society, educate others about prevention, ethical-issues which are important for college students, in addition to behavior change outcomes. Second, effective web-based delivery education design and strategies needs to be identified for future program development and continue improvement. One major challenge of such web-based educational program was the difficulty to find the reliable and relevant health information on the web. In addition, it was hard to control the credibility and accuracy of the health information from the Web sites (Escoffery et al., 2005). Information delivered via academic Webcourse can be a credible way to provide organized and reliable information to educate college students. Third, online course like Webcourses has not yet been widely implemented or evaluated in terms of both the students learning outcomes and course design strategies. A gap exists because previous research has mostly provided a partial view by using either quantitative or qualitative approaches. There is a need for a more complete understanding through comparing and synthesizing both quantitative and qualitative data (Creswell & Clark, 2018). Mixed methods research studies can provide a more comprehensive picture to answer such questions, using data from both quantitative numbers and qualitative narratives to show evidence from a more holistic perspective (Creswell & Clark, 2018).

Purpose

The purpose of this mixed method research case study is to provide evidence to demonstrate the effectiveness of this HIV Prevention Webcourses on HIV related competencies and outcomes among college students at a large public university in Florida. In addition, this study aims to identifying effective design and learning strategies to deliver web-based educational information.

Quantitative measurements include a new 6 HIV Competency Scale, a previously validated HIV knowledge scale (Hou, 2008; Hou, 2004), as well as HIV testing and condom use intention and self-efficacy (Hou, 2009a; Hou, 2009b; Hou & Luh, 2007). Qualitative measures identify most helpful types of online course content, strategies and challenges in effective online learning, as well as learning impact among students. The integration of both quantitative and qualitative results can provide a value-added understanding on the impact and evidence-based learning outcomes to demonstrate the effectiveness of course, as well las identify effective design features of Webcourses to promote learning. Educators in higher education can see, using an HIV Prevention Webcourse as an example, how course-related competencies can be developed and validated to evaluate the effectiveness of student learning. In addition, online instructors can learn how some of the best practice of online education can be translated into concrete strategies from this HIV Webcourses case study, and apply or adapt to their own online courses. Finally, educators interested in scholarship of teaching and learning can see how mixed methods research design can be integrated with quasi-experiment study to evaluate online courses.

Methods

Design and Sample

A mixed-methods quasi-experiment study was conducted. College students in health-related majors participated in the study. Experiment group was students enrolled in the HIV Webcourses (n=508) and comparison group was students who enrolled in non-HIV related courses (n=436) from the same college during 2016-2017 academic years. Both groups of the students completed the same HIV prevention survey before and after their semester-long courses. In addition, students in the HIV Webcourses also completed course evaluation survey to provide comments and feedback regarding course design, content, and learning impact.

Design of the HIV Webcourse

The HIV Webcourses was a large semester-long fully online course (3-credits) offered to students in health service administration, health science (pre-clinical), or related programs. The average number of students enrolled is about 100-150 per course session. One key strength was the course instructor has extensive experience conducing HIV related preventive research and has deep passion and understanding of related issues. The instructor used the Understanding by Design framework (UbD framework) in the process of developing the HIV Webcourse (Mctighe and Wiggins, 2012). The course design was planned backward with the end or desired outcomes, course learning competencies, in mind. In addition, the design was focused on teaching and assessing for understanding, as well as continual improvement process (Mctighe and Wiggins, 2012). Based on the UbD framework and considering minimizing the transactional distance concept discussed earlier, the HIV Webcourses incorporates the following key unique design features to facilitate online learning:

(1) Weekly announcement to keep students on track of module progression. Weekly announcements were set up to be released on each Monday morning to help students get the weekly rhythm of the online course, as well as to facilitate the presence of course instructor (Richardson & Lowenthal, 2017; Tichavsky et al., 2015). The announcements outline the content activities, as well as noted the assignments or activities of the week to keep students on track.

(2) Easy-to-learn modules with up-to-date resourceful websites and multiple content sources to support different learning styles. Each module uses the same structure to facilitate students know what to expect and where to find information for clear course design and structure (Baran & Correia, 2014). Each module starts with an overview, module objectives, module content with required readings, required videos to supplement or complement required reading contents, required stories, then module assessment (assignment activities). The variety of multiple online sources aims to support different online learners' needs (Moore, 1997; Pappas, 2017). Due to the rapidly changing HIV/AIDS epidemics globally, all the module content and readings were based on the most current HIV/AIDS epidemics and statistics, via credible official webpage or reports from CDC, WHO, HIV.gov, etc.

(3) Instructor avatar throughout the online modules to provide guidance, online persona, and direction among module components. An instructor video was used to introduce the professional and personal background with an aim to close the online distance and increase accessibility to instructor (Moore, 1997; Pappas, 2017). In addition, various instructor avatars were created to increase online persona presence of the instructor, and to provide guidance and directions among module components. Such persona image icons also help break the text and paragraphs to facilitate student navigating the module contents.

(4) Real-life stories to humanize the disease & increase relevance. To keep students continuously engaged in the online module, personal stories from real people were provided in each module to help students

understand and emphasize the struggles faced among individuals living with HIV/AIDS (Mctighe and Wiggins, 2012). These stories were all hand-picked to deepen or complement the issues discussed in corresponding module content. For each personal story, a catchy heading was created to highlight the central theme of the story, with content edited to protect privacy and identify. Stories were meant to help engage the emotional dimension of the course content which is critical in addition to gaining factual knowledge. HIV related stigma and struggles can be conveyed more effectively in a relatable way through the power of real-life stories.

(5) Variety of assignment to keep students engaged and thrive. A total of four 25-item *online quizzes* were given throughout the semester to help hold students accountable and engaged in the course content (Mctighe and Wiggins, 2012). In addition to these automatically graded quizzes, three additional manually graded assignments were incorporated to encourage deeper thinking and reflections on student learning (Moore, 1997; Pappas, 2017). (1) A *current event article* was incorporated to help students further research on a course related current issues and share back with the class. (2) The *discussion post and respond assignment* was provided to encourage students discuss controversial topics, practice share perspectives with supporting evidence or citations, and exchange thoughts and comments with peer fellow students. The large online class was intentionally divided into 5-6 smaller groups for this discussion assignment. This allows each student being more comfortable with discussing issues and perspective in a smaller group of 20-25 student environment. This helps students reading and responding other students' posts and responses in a more manageable and less overwhelming online environment. (3) Each student was also given the choice to choose one of the 5 recommended books to read and reflect as their *book review project assignment*. All book choices are personal stories, ranging from the naked truth of HIV from a newly diagnosed young woman, how a celebrity with HIV promote combating stigma through advocating love is cure, personal journey of an AIDS doctors, to how varies characters cope with HIV in what looks like crazy in an ordinary day, etc.

(6) Prompt communication and feedbacks. Taking advantage of the online environment, students can communicate with the instructor via multiple channels (email, messages, chats, posts, etc.) whenever having questions or needing clarifications on content or assignments. Timely communications and feedback were provided throughout the course, along with grades and feedbacks for each assignment (Pappas, 2017; Mctighe and Wiggins, 2012; Moore, 1997).

Data Collection

This study used a mixed methods quasi-experiment study design. An anonymous online HIV Prevention survey via external link was conducted before and after the course among both experiment and comparison groups. Details of the survey measurements were described below. Students rated anonymously their agreement level with the HIV competency statements on a 5-point Likert-type scale, ranging from strongly agree (coded as 5) to strongly disagree (coded as 1), along with other survey items. To ensure anonymity of the participants, no personal information, such as name or student ID, were asked. Students were reminded that how they rated their confidence level on the course competency statements or other HIV-related knowledge items would NOT influence the grade they would receive. The survey took about 15 min. to complete, and extra bonus points were provided by course instructors to encourage survey participation among students in both experiment and comparison groups. A student self-generated ID was used to link students' before and after responses for comparison purposes. IRB approval was obtained at the PI's university before the study was conducted.

Measures

Quantitative measurements include a new 6-item HIV Competency Scale, a previously validated HIV knowledge scale (Hou, 2008; Hou, 2004), as well as HIV testing and condom use intention and self-efficacy (Hou, 2009a; Hou, 2009b). Qualitative measures identify most helpful types of online course content, strategies and challenges in effective online learning, as well as learning impact among students.

HIV Competency Scale (6-item). Item of the HIV Competency Scale were drafted based on the HIV Webcourses course competencies developed and approved by the PI's university curriculum committee. HIV Competency items ask students to rate, on a 5-point Likert scale, their perceived self-confidence levels for the following statements: discuss the global impact of the HIV/AIDS, its epidemics and prevention strategies, empathize the difficulty of maintaining treatment schedule, identify the needs of specific groups, recognize societal psychological and medical impacts among people living with HIV/AIDS, and describe related ethical and legal issues (see Table 1).

HIV knowledge. Hou's 14-item objective HIV knowledge test scale, which was validated among multiple samples of college students with satisfactory internal consistencies (alpha=.70), was adapted in the current study (Hou, 2008; Hou, 2004). Two subjective HIV knowledge items asked participants to rate on a five-point Likert scale to assess perceived levels of HIV knowledge, "How would you rate your knowledge about HIV/AIDS in general?" and "How would you rate your knowledge specifically related to HIV testing?" (Hou, 2008; Hou 2004).

HIV testing and intention. HIV testing intention measures intention, on a five-point Likert scale, towards obtaining an HIV test in the next 6 months. HIV testing behavior were yes/no items. At baseline asked whether students have ever had an HIV test. At post-survey, students were asked whether they had gotten HIV tested in the past 3 months (during the HIV Webcourses intervention period) (Hou, 2009b).

Condom use intention and self-efficacy. Condom use intention measures students' intention to use condom during next sexual activity. Self-efficacy asked participants' confidence of using condom or communicate condom use with partners (Hou, 2009a).

Qualitative items. Qualitative feedback on the Webcourses features, as well as reflections on own learning were gathered among students in the experiment group to gain deeper understanding. Responses from four main qualitative items were analyzed, including "most helpful types of HIV Webcourses content" "strategies in online learning" "challenges in own leaning" and "one thing I will remember in 5 years". These data were meant to complement the quantitative data collected to provide a more holistic picture of the students' learning, as well as to identify most effective course components of the HIV Webcourses.

Analyses

Data from the baseline survey were used to assess the reliabilities of the 6-item HIV course competency scale (HIV-Competency). Descriptive statistics, item-total correlation, and Cronbach's alpha coefficients were calculated to evaluate the internal consistencies. Data from before and after courses were used to compare HIV-Competency between experiment and comparison groups. In addition to HIV-Competency, other HIV-related outcomes including objective and subjective HIV knowledge (general and testing specific), testing intention and behavior, as well as condom use intention and self-efficacy were also compared before and after, and between the two groups of students. T-tests were used to compare continuous variables while chi-square tests were used to compare dichotomous variables between groups. Furthermore, for continuous variables, linear regression analyses were conducted to compare HIV related outcomes between experiment and comparison groups, while controlling for baseline levels. Logistic regression analyses were conducted to compare dichotomous variables between the two group, while controlling for baseline levels. For qualitative feedback and narratives, top three most frequently mentioned responses were highlighted, with student comments or reasons provided side-by-side via a joint-display table to illustrate quantitative and qualitative findings.

Results

Sample Characteristics

A total of 944 students participated (508 experimental and 436 comparison students), about 76% were females, and 94% indicated themselves as heterosexual. Over one-thirds (36%) were single but has boy or girl friends, another one-thirds (30%) were single and not in any relationship. There were about 6% indicated themselves have friends with benefits. About 20% were married or engaged, and the rest were others. The study sample involved diverse student racial/ethnicity, with only 42% were white, 22% self-identified as African Americans, 21% Hispanics, 9% Asians, and 7% multi-racial.

Overall study participates were active in sexual activities, with over 80% reported having had both oral and vaginal sex, and about 30% also reported ever had anal sex. Only about 40% have been tested for HIV. Only about one-third reported often or always use condoms using (vaginal) sexual behaviors. Almost half (48%) reported they have had sex with someone who has more than one sex partners. About 13% reported a sexually transmitted infection (STI) history.

HIV Competency

The reliability of the 6-item HIV competency scale (on 5-point Likert scales) showed satisfactory internal consistencies (Cronbach alpha = .914; CITC s ranged .670-.804). There were no significant differences between groups on any of the competencies, except that students enrolled in the HIV Webcourses scored a little higher on describing the epidemiology of HIV/AIDS competencies at baseline.

Table 1. Reliabilities of the 6-item HIV Course Competency Scale (HIV-Competency) (n=944).

Item Description	Mean (SD)	CITC	Alpha if deleted
(HIV-Competency 1). Discuss the global impact of the HIV/AIDS epidemic among developed and developing nations.	3.16 (1.20)	.738	.902
(HIV-Competency 2). Describe the epidemiology of HIV/AIDS, as well as prevention, clinical and treatment strategies.	3.20 (.124)	.772	.897
(HIV-Competency 3). Empathize with the difficulty of maintaining a treatment schedule and locate reliable sources for HIV/AIDS.	3.33 (1.20)	.670	.911
(HIV-Competency 4). Identify the needs of specific populations affected by HIV/AIDS.	3.18 (1.23)	.789	.893
(HIV-Competency 5). Identify societal, psychological, and medical impacts among people living with HIV/AIDS and their families.	3.25 (1.20)	.804	.893
(HIV-Competency 6). Describe social, ethical, and legal issues related to HIV/AIDS.	3.20 (1.24)	.771	.897
HIV-Competency Scale (6-item)	**Item mean = 3.22**	**Cronbach's Alpha = .914**	

- CITC = Corrected Item-Total Correlation

After controlling for baseline competency scores, regression analyses showed that students took the HIV Webcourses (experiment group) scored significantly higher than those who did not take the HIV Webcourses (comparison group) on all the six HIV Competencies (all p-value <.001). Overall students in the HIV Webcourses scored between 4.27 and 4.41, while comparison group of students scored between 3.34 and 3.49 per competency at post-test.

Table 2. HIV course competencies before and after the HIV Webcourses among experiment (n=436) and comparison (n=508) groups.

	HIV course competency		C- Mean (SD)	E- Mean (SD)	P-value
1	HIV-Competency 1.	before	3.13 (1.24)	3.21 (1.15)	.291
		after	3.36 (1.20)	4.27 (.79)	<.001**
2	HIV-Competency 2.	before	3.08 (1.29)	3.35 (1.15)	.001*
		after	3.37 (1.22)	4.36 (.78)	<.001**
3	HIV-Competency 3.	before	3.27 (1.23)	3.39 (1.15)	.111
		after	3.48 (1.21)	4.37 (.81)	<.001**
4	HIV-Competency 4.	before	3.15 (1.28)	3.23 (1.17)	.298
		after	3.34 (1.25)	4.30 (.82)	<.001**
5	HIV-Competency 5.	before	3.22 (1.26)	3.28 (1.14)	.419
		after	3.49 (1.17)	4.41 (.76)	<.001**
6	HIV-Competency 6.	before	3.19 (1.28)	3.21 (1.19)	.848
		after	3.44 (1.21)	4.33 (.80)	<.001**

- ** p<.001; *p<.05
- C = Comparison Group; E=Experiment Group

Objective and Subjective HIV Knowledge

At post-survey, objective HIV knowledge test scores were higher among students in the HIV Webcourses than students in the comparison group (11.03 vs. 9.47, p<.001). Similarly, data showed that students in the HIV Webcourses also perceived higher HIV knowledge than comparison group, with 3.82 on subjective HIV knowledge in general and 3.42 on testing specific knowledge among students in the experiment group, comparing with 3.06 and 2.89 respectively among students in the comparison group at post-survey (p<.001). After controlling for baseline differences on objective and subjective knowledge between groups, linear regression analyses result still showed significant differences on both objective and subjective HIV knowledge after the HIV Webcourses intervention (p<.001) (Table 3).

Table 3. Objective and subjective HIV knowledge, HIV testing, and condom use (CU) intention and self-efficacy, before and after the HIV Webcourses among experiment (n=436) and comparison (n=508) groups.

	HIV-related outcomes		C- Mean (SD)	E- Mean (SD)	P-value
1	Objective HIV knowledge test (14-item).	before	9.36 (2.96)	9.97 (2.67)	<.001**
		after	9.47 (3.65)	11.03 (2.91)	<.001**
2	Subjective knowledge (HIV/AIDS in general).	before	2.85 (.847)	2.99 (.768)	.007
		after	3.06 (.822)	3.82 (.666)	<.001**
3	Subjective knowledge (HIV testing specifically).	before	2.61 (.963)	2.61 (.925)	.974
		after	2.89 (.985)	3.42 (.824)	<.001**
4	HIV testing intention in the next 6 months.	before	2.56 (1.22)	2.67 (1.22)	.145
		after	2.52 (1.21)	2.73 (1.22)	.096
5	HIV testing (Yes%)	before	40.0%	36.8%	.347
		after	11.3%	13.6%	.629
6	CU intention (Yes%)	before	50.6%	58.7%	.015*
		after	49.4%	60.0%	.054
7	CU self-efficacy (Yes%)	before	66.8%	66.0%	.834
		after	56.3%	66.9%	.048*

- Condom use intention – intention to use condom next time during sexual activity.
- Condom use self-efficacy – confidence of using condom or communicate condom use with partners.
- HIV testing – (before) Ever had HIV test when taking the baseline survey. (after) Whether students had gotten HIV tested in the past 3 months (during the intervention period).
- ** p<.001; *p<.05
- C = Comparison Group; E=Experiment Group

HIV testing intention, condom use intention, and self-efficacy

Bivariate analyses showed no or borderline significant differences on HIV testing and condom use intention, or self-efficacy at baseline (Table 3). After controlling for baseline variables, regression analyses showed that students in the HIV Webcourses were 1.9 time more likely to have confidence on using condom or communicating condom use with partners (p=.009). Yet neither HIV testing nor condom use intention showed significant differences between groups (data now shown)

Qualitative comments and feedback

Four main qualitative questions were used to gather student feedback on the course components and reflections on own learning. The top three *most helpful types of HIV Webcourses content* identified by students were all story-based components including personal stories (33.3%), videos (18.0%), and the book project (12.8%). Students commented these contents gave the HIV survivor's perspective, and helped them better understand and emphasize personal real-life situations which individuals with HIV/AIDS face and going through. Such online course components help increase personal relevance, were easy to learn, and provided further insights and added emotion dimension to keep students interested. These real-life accounts help students relate to personal experience, and deepened their engagement in learning even in an online environment.

The top three *strategies students used in online course learning* were reading all course materials (28.2%), independent research on the course topics (9.9%), as well as taking notes as students reading through the online modules (9.9%). Students stated that reading all course contents was key to understand the topic, gain more related information, better perform in assignments, and clarify various concepts. In an online learning environment, reading to learn becomes even more critical; while taking notes can assist with fast recall.

Regarding *challenges in the learning process*, almost 40% of the students indicated none because the course setup was clear, logic, organized, interesting and informative. About 20% indicated time management was a challenge due to busy life schedule, procrastination, personal priorities, or not used to online learning. About 15% of the students indicated the extensive content was a challenge and time consuming.

Finally, top three *things students stated they will remember in five years* were all course information (28.5%), HIV stigma (14.1%), and the book project (12.8%). Students comments the course information was memorable, insightful, and can help close friends. HIV stigma was the most outstanding concept throughout the course, and the book project relates to personal experience and changed own perspectives (see Table 4).

Table 4. Qualitative comments from students' comments and feedback on the HIV Webcourses

	A. Top three most helpful types of the HIV Webcourses contents … because					
1.	Stories	33.3%	Personal relevance	Kept me interested	Real Life	New to me
2.	Videos	18.0%	Easy to learn	Added emotion	Visual	Extend the issue
3.	Book project	12.8%	Personal experience	Further insight	Stigma	
	B. My top three strategies in learning … because					
1.	Reading all course materials	28.2%	Help understand the topics	Helpful in assignments		
2.	Research the topics	9.9%	Help understand the topics	Understand people with HIV		
3.	Taking Notes	9.9%	Help with fast recall	Helpful in assignments		
	C. Top three challenges in my learning process … because					
1.	None	38.2%	Great			
2.	Time management	20.6%	Busy life schedule	Procrastination	Personal priorities	Not used to online courses
3.	Extensive content	14.7%	Time consuming	Personal priorities		
	D. Top three things I will remember in 5 years … because					
1.	All course information	28.3%	Memorable	Insightful	Informative	Can help close friends
2.	HIV stigma	14.1%	Most outstanding concept	How to stop stigma	Understanding people	Loving human
3.	Book project	12.8%	Personal experience	Changed my perceptions	Amazing	

Overall student course evaluation ratings were very high (4.5 on a 5-point Likert scale) on effectiveness of the course organization, explanation and communication, environment to conductive learning, provided useful feedback, helping students achieve course objectives, and overall instructor effectiveness. Students commented they genuinely loved this course as it was extremely well organized, information was very useful, interesting and informative, and enjoyed the passion of the professor. Students overwhelmingly voiced the need for all college students to be educated on HIV related issues. Students also commented how the personal stories and videos of those affected with HIV/AIDS opened their eyes and gave perspectives and provided life-changing lessons. In addition, students commented that the course had opened their eyes and really taught the impact of the disease and stigma people are facing, the professor truly cares for her students and provided wonderful and timely feedback, and everything was mapped out clearly from the beginning with perfect execution and instructions as if the instructor was instantaneous with her response to questions.

Discussion

Data from the current study show that the HIV Prevention Webcourses was effective in reaching both HIV-related competencies and HIV-related psychosocial outcomes related to behavioral change. Study showed that students took the HIV Webcourses (experiment group) scored significantly higher than those who did not take the course (comparison group) on all the six HIV Competencies (all p-value <.001) at the end of the semester. In addition, findings show students enrolled in the HIV Webcourses also scored significantly higher on both objective and subjective HIV knowledge, comparing with those who did not enrolled in the Webcourses (p<.001). Although intention towards HIV testing and condom use did not reveal significant differences, experiment group of students were 1.9 time more likely to report having confidence on using condoms or communicating condom use with partners (p=.009), compared with the comparison group.

The 6-item HIV competency scale also revealed satisfactory reliability with Cronbach alpha of .914, demonstrating good internal consistencies. The competencies measured in the current study share some of the key aspects of the competencies identified from a previous research using a qualitative approach among nursing students in South Africa (Modeste & Adejumo, 2014). These shared competencies including HIV/AIDS basic scientific knowledge, policy, and ethics related competencies. There are also some differences between key areas identified as the current competencies focus more from the broader public health and HIV prevention perspectives, such as HIV/AIDS global impact, epidemiology, social stigma, and special populations; instead of clinical care perspectives such as interdisciplinary care, professional development, or holistic safe practice related competencies emphasized in the previous study (Modeste & Adejumo, 2014).

Overall, the student evaluation for the HIV Webcourses was very high (4.5 on a 5-point Likert scale) in terms of the effectiveness of the course organization, the online learning environment and communication, as well as the overall instructor effectiveness. Specifically, students indicated the setup or design of the Webcourses was clear, logic, organized, interesting and informative, supporting the clear design structure best practice principle (Baran & Correia, 2014). The weekly announcement in particular help keep students on track and facilitate the presence of course instructor (Richardson & Lowenthal, 2017; Tichavsky et al., 2015), and the timely communication and feedback via discussion boards, assignment chats, and emails further demonstrated the instructor social presence (Stodel, Thompson, & MacDonald, 2006). About 40% indicated no challenges in their online learning for the course, while 20% indicated time management was a challenge due to own busy life schedule, procrastination tendency, or personal priorities.

Among the various design features, the top three favorite components identified were mostly involved with emotional engagement such as personal stories, videos, and the book project. Students indicated that videos and stories gave them more vivid feelings and understandings of the struggle from the HIV/AIDS patients' perspectives. These qualitative comments supported that the HIV Webcourses provided a good platform to address some of the key challenges identified from previous studies, including providing health information which is reliable, relevant, credible, engaging, and at appropriate depth (Escoffery et al., 2005). In addition, studies show that students can gain more information from the web, and were more motivated and willing to engage in learning. Current findings show that college students prefer such flexibility and independence during the learning process. In consistent with existing literature, the web-based learning model can also further enhance coordination, communication collaboration among students, and facilitate distance learning (Hosseini, 2013).

The current study is limited to its convenience sample. Study participants were invited from one college within a large public university, thus generalization of study findings may warrant attention. It would also be beneficial to compare the HIV course delivered as webcourse and non-webcourse, using the same base materials and resources. Unfortunately, the author's institution does not currently offer in-person HIV course sessions, and this could be an area for future research. Despite not being able to compare webcourse vs non-webcourse delivery, the current quasi-experiment research design with an equivalent online comparison group from the same college provided was a stronger design, as opposed to a commonly used single group before-after test. Furthermore, the clear and well-organized HIV Webcourses help ensure the quality and consistency of the online learning delivery. The convergent evidence from both quantitative statistics and qualitative narratives add values to the credibility of the effective learning outcomes observed.

Using Webcourses to deliver HIV Prevention knowledge can be effective and reach more college students (Muessig, et al., 2015; Hosseini, 2013; Noar, et al., 2009; Noar, et al., 2010). College students normally have busy schedule for classes and extra-curriculum activities, so their time is limited and less flexible. Webcourses can address their limited attention and time while allow flexibility and convenience in learning at own pace and location. Continuous attention is needed to ensure such Webcourses continue to provide updated information and address different student engagement issue. The reliable HIV Competency scale can be used in future study to assess the effectiveness of teaching and learning.

Conclusion

In short, findings show that the HIV Webcourses was well designed and effective in building students' HIV related competencies and HIV related outcomes, comparing with the comparison groups. Both the quantitative scores and qualitative comments on student learnings convergently demonstrated the impact and effectiveness of this online course, providing stronger and more robust evidence than either quantitative statics or qualitative narratives alone. Furthermore, the qualitative data highlighted some key design features for online learning, suggesting emotional involvement via personal stories or video are as critical as the factual information itself. Lessons learned have implications on translating best practices into concrete strategies for effective online education delivery, as well as incorporating mixed methods research design to evaluate online courses via course-related competencies.

Acknowledgement

The author appreciates students who participated in the study. Special thanks to UCF instructors Yara Asi, Vierne Placide, and Michelle Chandler, for their help in recruiting student participants. Appreciation also goes to Xian Cao (PAF-PhD student) for her help with literature reviews.

References

Allen, I. E., & Seaman, J. (2013). *Changing course: Ten years of tracking online education in the United States.* Babson Park, MA: Babson Survey Research. Group.Baran, E., & Correia, A.-P. (2014). A professional development framework for online teaching. *TechTrends, 58*(5), 95-101.

Allen, I. E., Seaman, J., Poulin, R., & Straut, T. T. (2016). Online report card: Tracking online education in the United States. Babson Park, MA: Babson Survey Research Group. Retrieved from Education Advisory Board Website: http://onlinelearningsurvey.com/reports/onlinereportcard.pdf

Baran, E., Correia, A.-P., & Thompson, A. (2013). Tracing successful online teaching in higher education: Voices of exemplary online teachers. *Teachers College Record, 115*(3), 1-41.

Branson, B. M., Handsfield, H. H., Lampe, M. A., Janssen, R. S., Taylor, A. W., Lyss, S. B., & Clark, J. E. (2006). Revised recommendations for HIV testing of adults, adolescents, and pregnant women in health-care settings. *Morbidity and Mortality Weekly Report: Recommendations and Reports, 55*(RR14), 1-17.

Caico, C. (2014). Sexually risky behavior in college-aged students. *Open Journal of Preventive Medicine, 4*(05), 354-364. Rehttp://doi.org/10.4236/ojpm.2014.45043

Calloway, D. S., Long-White, D. N., & Corbin, D. E. (2014). Reducing the risk of HIV/AIDS in African American college students: An exploratory investigation of the efficacy of a peer educator approach. *Health Promotion Practice, 15*(2), 181-188. https://doi.org/10.1177/1524839913504757

Centers for Disease Control and Prevention (CDC). (2015). HIV Basic Statistics. Retrived Octomber 28, 2017, from https://www.cdc.gov/actagainstaids/basics/statistics.html.

Centers for Disease Control and Prevention (CDC). (2017). CDC Vital Signs. Retrived Jan. 5, 2018, from https://www.cdc.gov/vitalsigns/

Creswell, J., & Clark V.L. (2018). Designing and conducting mixed methods research. (3rd ed.) Sage publications, Inc.

Escoffery, C., Miner, K. R., Adame, D. D., Butler, S., McCormick, L., & Mendell, E. (2005). Internet use for health information among college students. *Journal of American College Health, 53*(4), 183-188. https://doi.org/10.3200/JACH.53.4.183-188

Ezeonwu, M., Berkowitz, B., & Vlasses, F. R. (2013). Using an Academic-Community Partnership Model and Blended Learning to Advance Community Health Nursing Pedagogy. *Public Health Nursing*, 272-280. DOI:10.1111/phn.12060

Figlio, D., Rush, M., & Yin, L. (2013). Is it live or is it internet? Experimental estimates of the effects of online instruction on student learning. *Journal of Labor Economics, 31*(4), 763-784. https://doi.org/10.1086/669930

Garrison, D. R., & Anderson, T. (2000). Transforming and Enhancing University Teaching: Stronger and weaker technological influences. In T. Evans & D. Nation (Eds.) *Changing University Teaching: Reflections on creating educational technologies* (pp. 24-33). London: Kogan Page.

Hoppel, A. M. (2012). HIV: Still Epidemic after 30 yrs. *Clinician Reviews, 22*(6): cover, 11-14, 33-35.

Hosseini, H. (2013). The role of new technology in global health education. *International Journal of Business Anthropology, 4*(1), 108-120.

Hou, S. (2009a). HIV-related behaviors among black students attending Historically Black Colleges and Universities (HBCUs) versus white students attending a Traditionally White Institution (TWI). *AIDS Care, 21*(8), 1050-1057. https://doi.org/10.1080/09540120802626196

Hou, S. (2009b). Extending the use of the web-based HIV Testing Belief Inventory (wHITBI) to students attending Historically Black Colleges and Universities (HBCUs): An examination of reliability and validity. *AIDS Education & Prevention*, 21*(1), 80-90.* https://doi.org/10.1521/aeap.2009.21.1.80

Hou, S. (2009c). Measuring the infusion of service-learning on student outcomes – The Service-Learning Self-Efficacy Scale on Program Planning Competencies (SL-SEPP). *California Journal of Health Promotion, 7*(2), 123-128. https://doi.org/10.1177/1053825917699518

Hou, S. (2008). Using an online survey to assess knowledge of HIV in general and testing specifically among black and white college students. *Journal of National Medical Association, 100*(7), 826-832. https://doi.org/10.1016/S0027-9684(15)31377-8 Available at: http://www.nmanet.org/images/uploads/Publications/OC826.pdf

Hou, S. (2007). Alternative modes of measuring self-report on HIV-related behaviors among college students: Web-delivered mode versus paper-pencil mode. *American Journal of Health Education, 38*(1)*,* 9-15.

Hou, S. (2004). Objective and subjective knowledge and HIV testing among college students. *American Journal of Health Education, 35* (6), 328-335. https://doi.org/ 10.1080/19325037.2004.10604772

Hou, S. & Luh, W. (2007). The structure of a web-based HIV Testing Belief Inventory (wHITBI) for College Students: The evidence of construct validation. *Medical Informatics and the Internet in Medicine, 32*(2), 83-92. https://doi.org/10.1080/14639230601125134

Hou, S., Pereira, V. (2017). Measuring the infusion of service-learning (SL) on student outcomes – Validating an assessment suite on Program Development and Implementation Competencies (SL-SEPDI). *Journal of Experiential Education, 40*(2), 170-186. https://doi.org/10.1177/1053825917699518

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies. Project Report. Centre for Learning Technology. 1-66.

Modeste, R. R. M., & Adejumo, O. (2014). Identification of HIV and AIDS-related competencies for nurse graduates in South Africa: teaching and learning. *African Journal for Physical Health Education, Recreation and Dance, 20*(Supplement 1), 105-121.

Moore, M. (1997). Theory of transactional distance. In Keegan, D. (ed.). Theoretical Principles of Distance Education, Routledge, pp.22-38.

Muessig, K. E., Nekkanti, M., Bauermeister, J., Bull, S., & Hightow-Weidman, L. B. (2015). A systematic review of recent smartphone, Internet and Web 2.0 interventions to address the HIV continuum of care. *Current HIV/AIDS Reports, 12*(1), 173-190. https://doi.org/ 10.1007/s11904-014-0239-3

Noar, S. M., Black, H. G., & Pierce, L. B. (2009). Efficacy of computer technology-based HIV prevention interventions: a meta-analysis. *AIDS, 23*(1), 107-115. https://doi.org/ 10.1097/QAD.0b013e32831c5500

Noar, S. M., Pierce, L. B., & Black, H. G. (2010). Can computer-mediated interventions change theoretical mediators of safer sex? a meta-analysis. *Human Communication Research, 36*(3), 261-297. https://doi.org/10.1111/j.1468-2958.2010.01376.x

Pappas C. (2017). 8 tips to minimize transactional distance in eLearing. eLearning Industry. Available at: https://elearningindustry.com/tips-minimize-transactional-distance-elearning

Richardson, J. C., & Lowenthal, P. (2017). Instructor social preference: a neglected component of the community of inquiry. *Elearning & Software for Education, 2*. 531-536.

Shotwell, M., & Apigian, C. H. (2015). Student performance and success factors in learning business statistics in online vs. on-ground classes using a web-based assessment platform. *Journal of Statistics Education, 23*(1), 1-19. https://doi.org/ 10.1080/10691898.2015.11889727

Stack, S. (2015). Learning Outcomes in an online vs traditional course. *International Journal for the Scholarship of Teaching & Learning, 9*(1), 1-18. https://doi.org/10.20429/ijsotl.2015.090105

Stodel, E. J., Thompson, T. L., & MacDonald, C. J. (2006). Learners' perspectives on what is missing from online learning: Interpretations through the community of inquiry framework. *The International Review of Research in Open and Distributed Learning, 7*(3), 1-24. https://doi.org/10.19173/irrodl.v7i3.325

Tichavsky, L. P., Hunt, A. N., Driscoll, A., & Jicha, K. (2015). "It's just nice having a real teacher": student perceptions of online versus face-to-face Instruction. *International Journal for the Scholarship of Teaching and Learning, 9*(2), 1-8. https://doi.org/10.20429/ijsotl.2015.090202

Journal of the Scholarship of Teaching and Learning, Vol. 20, No. 1, April 2020, pp.40-64.
doi: 10.14434/josotl.v20i1.24688

Diverse Student Perceptions of Active Learning in a Large Enrollment STEM Course

Benikia Kressler
California State University, Fullerton
bkressler@fullerton.edu

Jochen Kressler
San Diego State University

Abstract: The concept of active learning as a superior mode of instruction has recently received great attention in the education research literature. It holds promise of steering students away from rote memorization towards higher order thinking (as defined by Bloom's taxonomy revised). However, few studies focus on student perceptions of higher order thinking activities and diverse student voices are all but absent in this regard. This study applies a combined approach of exploratory qualitative and supplementary quantitative analysis to address this gap. We examined perceptions of underrepresented and non-underrepresented students regarding their engagement in active learning to foster higher order thinking. The study was set within a large enrollment (198 students), undergraduate course in the area of science, technology, engineering and mathematics (STEM). The study sample comprised of 33 purposefully selected, ethnically and racially diverse students. Data sources included class attendance/participation, graded activity assignments, and a perception survey. Class attendance and graded assignments were used to triangulate responses on the perception surveys. The Generic Inductive Approach supported our qualitative analysis. Quantitative data were analyzed via two-way ANOVA, non-parametric Mann-Whitney Test (when assumptions did not hold) and simple linear regression. Findings include three themes that cut across groups; participants perceived their higher order thinking skills improved, that there were benefits and challenges to active learning and a fear of failing the course. Quantitative data from the active learning activities and attendance supported similar engagement and achievement in higher order thinking activities across race/ethnicity groups as differences failed to reach the a priori established significance threshold. This study extends the knowledge on active learning and demonstrates that it was possible to engage underrepresented and non-underrepresented students equally and effectively in higher order thinking actives in large enrollment courses and that students perceived this as beneficial.

Keywords: *higher order thinking, active learning, STEM, student perceptions, critical thinking*

The dreams, possibilities, and necessity of public education extends beyond K-12 systems. In higher education, educators must provide high quality and meaningful education to all students. This includes using research supported strategies that support the academic achievement of students typically underrepresented in higher education with diverse backgrounds. This study acknowledges this charge by examining student perceptions of their engagement in active learning activities that support higher order thinking (as defined by Bloom's taxonomy revised, (Anderson et al., 2001) within a large enrollment undergraduate course in the area of science, technology, engineering and mathematics (STEM).

Engaging in active learning strategies that support higher order thinking within higher education settings are considered best practices in educational literature (Bonwell & Eison, 1991; Casagrand & Semsar, 2017; Freeman et al., 2014; Sellami et al., 2017, White et al., 2016). Specifically, providing undergraduate students the opportunity to engage in complex course material using student centered, collaborative, multimodal techniques significantly increases student learning when compared to standard lecture (LoPresto & Slater, 2016). Examples of active learning include cooperative group activities, in-class worksheets, clickers, problem-based learning, and studio classrooms (Freeman et al., 2014).

Actively engaging students in large enrollment courses, specifically science, technology, engineering and mathematics (STEM) courses, is particularly important given the heavy cognitive load of the work, student failure rate and the tendency for rote memorization to study for tests (Biggs, 2011; Bligh, 1998; Gasiewski et al., 2012). Further, today's job market calls for 21st century skills that include ability to communicate, collaborate, think critically and be creative (Dede, 2010). It is imperative to arm students with 21st century skills to prepare them for entrance into one of the most competitive and diverse economic markets to date (Ali, 2017; Casner-Lotto & Barrington, 2006; Dede, 2010; Ghaffarzadegan et al., 2014; Maddux et al., 2014; Saavedra & Opfer, 2012). Consequently, this study examines student perceptions of their engagement in active learning within a large enrollment undergraduate course in a STEM course.

Literature Review

Many university instructors accept the notion that involving students in cognitive processing activities is more effective than instructor-led lectures (Dunkin, 1983; Eichler & Peeples, 2016; Mayer et al., 2009). The work of Freeman et al., (2014) (Freeman et al., 2014) provides strong scientific evidence in favor of active learning strategies over lecture-based instruction, at least at the undergraduate level in STEM disciplines. In a comprehensive meta-analysis of 158 randomized, controlled trials Freeman et al. 2014 (Freeman et al., 2014) reports that student performance (as assessed by examinations or concept inventories) improved by 0.5 SDs in active learning vs. lecture control. This effect is even larger for class sizes with 50 students or less, and the meta-analysis did not extend beyond common STEM fields (Freeman et al., 2014). In addition, sub-analysis of 67 studies found that the odds ratio for failing a course was almost twice as high (1.95) for lecture conditions. This Tier 1 (Harris et al., 2001) evidence clearly demonstrates the effectiveness of active learning for all students, including those with high risk of failure.

Critics posit that undergraduates in STEM are not receiving sufficient instruction (Bok, 2009; Stains, 2018). This critique has not been ignored. Active learning practices have gained traction within institutions of higher education. However, it has been a slow process. Studies show that 65% to 80% of university instructors continue to engage in teacher centered learning (i.e., lecturing) (Nelson & Crow, 2014).

There are many reasons that explain this discrepancy between the scientific evidence and incorporation of best practices into actual practice. There is a general lack of resources in support of pedagogical development with STEM disciplines (Baldwin, 2009; George & Bragg, 1996). Incentive structures usually do not reward instructors to study the pertinent literature regarding teaching and learning or to put extensive time and effort into improving the way they deliver courses to students (Baldwin, 2009; George & Bragg, 1996; National Research Council, 2003; Wieman, 2007). If teaching effectiveness is considered, universities mostly rely on student evaluations of teaching (SETs) which at best have no correlation to teaching effectiveness and at worst, can promote practices that are counterproductive (Braga et al., 2014; Carrell &West, 2010; Dunkin, 1990; Kornell & Hausman, 2016; Uttl et al., 2017). SETs may also reinforce many instructor's falsely held believes about their own teaching effectiveness and thereby obscure the need for a change in their teaching approach (Mazur, 2011). Indeed, fear of negative evaluations following substantial changes to teaching methodology causes many faculty to be reluctant to change (Eichler & Peeples, 2016; Ryan et al., 1980) because exposing students to a drastically different learning environment can result in decreased satisfaction (Eichler & Peeples, 2016; Gutwill-Wise, 2001). In addition, there is a lack of institutional commitment to developing, implementing and supporting research based teaching across a wide range of STEM courses (Wieman, 2007). Lastly, there are simple physical impediments. Many undergraduate STEM classes take place in large lecture halls constraining instructional practices through room architecture and seating arrangements (Baldwin, 2009).

This is of course extremely problematic because research show that lecture only approach is not effective for most students (Freeman et al., 2014; Wieman, 2007; Zoller, 1993), and is particularly ineffective for underrepresented students in large lecture classes with a reputation for high failure rates (Haak et al., 2011; Hrabowski, 2011; Mervis, 2010; Mulligan, 2000). Underrepresented and underserved students in college include Black, Hispanic, Native American and Pacific Islanders as well as first generation college students. These groups are less likely to complete four-year degree programs as White and Asian peers (US Department of Education, 2014). In addition, students in high failure courses are dissatisfied with the quality of STEM teaching, particularly when size limits student-teacher dialogue (Baldwin, 2009). Improving the persistence of underrepresented college students typically focus on advising, scholarships and tutoring with limited studies examining classroom based strategies (Winkelmes, Bernacki, Butler, Zochowski, Golanics & Weavil, 2016). Although limited, a study by Winkelmes (2013) found *transparency* regarding how college students learn, what they learn and why a course is structured a particular way "showed promise for improving underserved students' educational experiences" (p.1). Another study found that students perceived "technology-nested" instructional strategies as a strong component to content engagement and enjoyment (Lumpkin, Achen & Dodd, 2015 p. 12). These studies highlight the promise of explicit instruction or transparency and using technology in college courses for all students.

While active learning is more effective, problems can arise when student's perception of active leaning techniques become negative over time as the novelty wears off over the course of a semester (McDougall, 2013; Nguyen et al., 2016; Cooper, Downing, Brownell, 2018). The limited discussion of student perception in the literature reports that students perceive in-class activity as effective but did not specifically assess perceptions of higher order thinking (Nail, 2012; Nguyen et al., 2016). Certain types of active learning strategies such as using case studies and creative activities can be applied to pertinent situations to support higher order thinking (Bean, 2011). However, students in active learning environments resented the "intellectual effort" needed for successful completion of activities (Smith & Cardaciotto, 2012). Evidently, activities targeting higher order thinking require the most intellectual effort. This presents a stark contrast to the standard lecture format which encourages students to be mostly passive learners (National Research Council, 2003) who rely heavily on the lower levels of thinking such as memorization of facts or formulas (Baldwin, 2009; Brainard, 2007; National Research Council, 2003). Whether there are differences in the perception of active learning to support higher order thinking among underrepresented students is currently unknown.

Therefore, the purpose of this study was to examine both underrepresented and non-underrepresented student perceptions of collaborative, active learning activities that support higher order thinking in a large enrollment STEM course. The overarching research questions guiding this study are: "What are underrepresented and non-underrepresented students perceptions of active learning activities within a large lecture course?", "What are underrepresented and non-underrepresented student perceptions of their higher order thinking skills after engaging in the course?" and "Are there differences in these perceptions?".

Methods

This exploratory mixed methods study (Creswell et al., 2003) was drawn from a larger IRB approved research project. This study consisted of a sample of 33 students from a large (198 students), upper division (juniors and seniors) undergraduate, health science course at a major public university located on the west coast of the United States during Spring 2016. The sampled participants included 19 females and 14 males. Age ranged from 20-29 years. The sample was ethnically and racially diverse. Eleven Asian, Hispanic and White students were purposefully selected (Patton, 2015) as they represented the larger racial makeup of the course.[1] We focused on the three largest racial/ethnic groups in the course. Students self-identified their race/ethnicity and then were randomly selected within each group. Table 1 illustrates student identification in greater detail as well as other demographic information.

Table 1: Subject Demographics

Race/ Ethnicity	Details	Age	Gender	Mother's Highest Level of Education	1st Generation College Student	English 1st Language	Academic Level
	n/a	24	m	college graduate	yes	Yes	junior
	n/a	20	m	high school	no	Yes	junior
	n/a	22	f	college graduate	no	No	senior
	n/a	20	f	some college	no	Yes	senior
White / Caucasian	n/a	20	f	college graduate	no	yes	junior
	n/a	23	f	college graduate	no	yes	junior
	n/a	21	f	Unknown	yes	yes	senior
	n/a	21	f	college graduate	no	yes	junior
	n/a	21	m	college graduate	no	yes	junior
	n/a	23	f	some college	no	yes	senior
	n/a	22	f	some college	yes	yes	senior
	Mexican	20	m	college graduate	no	yes	junior
	n/a	23	f	some college	yes	yes	senior
	Mexican	23	m	some college	yes	yes	senior
	Portuguese	21	m	high school	yes	no	senior
	Chicano	21	m	less than high school	yes	no	senior
Hispanic /	n/a	22	f	high school	yes	yes	n/a

[1] The researchers recognize that these racial/ethnic categories are insufficient in capturing the diversity within these labels.

Latinx	n/a	24	m	college graduate	no	yes	junior
	Mexican	21	f	some college	no	yes	junior
	n/a	22	m	some college	no	yes	senior
	n/a	22	m	college graduate	no	no	senior
	n/a	20	f	high school	yes	yes	junior
	Filipino	21	f	college graduate	no	yes	senior
	n/a	21	f	high school	yes	yes	junior
	Thai	22	m	some college	yes	no	senior
	n/a	23	m	some college	no	yes	senior
	Japanese	22	m	some college	yes	no	junior
Asian	n/a	20	m	college graduate	no	yes	junior
	Chinese	20	f	less than high school	yes	no	junior
	Filipino	20	f	high school	no	yes	junior
	Filipino	21	f	college graduate	yes	yes	senior
	Filipino	23	f	some college	no	yes	n/a
	Filipino	29	m	college graduate	no	yes	junior

Course Description

The course for this study was Measurement and Evaluation in Kinesiology (ENS 305). This course is required for all emphases within the Kinesiology major, the largest major at the university (1,800 students, >5% of total undergrad enrollment) and part of the second largest field for undergraduate degrees conferred in the United States (National Center for Education Statistics, 2016). Other common names of this major include Exercise Science, Physical Education, Exercise and Sport Science, Health and Exercise Science, Exercise Science and Wellness, Exercise and Fitness, Kinesiological Sciences, and Exercise Physiology (Boone, 2000).

This course was chosen to be redesigned with a focus on technology and active collaborative learning to support student achievement. In previous semesters 20% of grades fell into the repeatable category. It has recently been the main bottleneck course leading to delays for students trying to advance through the major, impacting four year and six year graduation rates. The course builds on prerequisite statistics courses by using descriptive and inferential statistics to examine the quality of assessment tools, study designs and inferential analysis used by Exercise Science and Kinesiology professionals. The emphasis of the class is on the active use of higher order thinking skills to apply assessment principles for the determination of the quality of assessment tools and usefulness of the data for meaningful decision-making. The concepts of higher order thinking according to boom's taxonomy were explained to students on the first day of class (the only day with mandatory attendance), and students were reminded and made aware of what type of thinking they were expected to perform for the activities. The course met twice a week for one hour and 15 minutes each session.

Five graduate assistants (GAs) helped facilitate the active learning in class environment. In order to optimally guide deliberate practice during the active learning exercises and maximally engage student (particularly those who may be struggling) feedback, encouragement and supervision are crucial. To achieve this in a large class more than one person is necessary (Park, 2004). Activities that require higher order thinking and complex problem solving (i.e. "when the going gets tough") will have to be supported by accessible expert feedback (Harland, 2003). GAs were trained by the second author specifically to improve student engagement and higher order thinking. Training consistent of three meetings before the semester to familiarize GAs with Blooms taxonomy (revised), key aspects of active learning and its effectiveness compared to standard lecture formats as well as practice activities. During the semester, GAs and instructor met on Friday of each week to go over events and activities of the past week, the class material for next week, selected activities for the week and practiced every activity as intended for class before implementation. GAs acquired specific content expertise necessary to assist in the course, prepared active learning activities and provided analysis and evaluation of their effectiveness (i.e., reflection) (Aronson, 2011).

Data Sources and Data Analysis

Data sources included class attendance/participation as assessed by wireless audience response system (i>clicker2); 12 graded assignments of higher order thinking activities (worth 1 point each) completed in flexible groups; and a researcher developed, open ended perception survey (see Appendix 3). The students were randomly assigned to groups and stayed within that group for the duration a given topic was covered (i.e., concept of validity) which could span between 1-3 days and were reassigned to new groups for the next topic. Group assignments were based on student identifiers (i.e., last 4 digits of student ID) and seating charts projected on the screen. The survey was administered in class at the end of the semester. Response rate was 89% (177 out of 198 students). Class attendance and graded assignments were used to triangulate responses on the perception surveys (Patton, 2015).

The qualitative open-ended survey data was analyzed using an inductive approach (Charmaz, 2006). To analyze the survey data, we organized participant response using the qualitative data software NVivo (QSR, 2015). Using part of the question itself as a starting category (e.g., "What are your perceptions of active learning?" became the category "Student perception of Active Learning"), we organized open-ended student comments. Next, these open-ended comments were coded inductively, moving from more concrete 'open and focus' codes to more abstract 'theme' development (Charmaz, 2006). From this process, three themes emerged from the data: *improved higher order thinking skills*; *benefits and challenges of active learning in large enrollment courses;* and *fear of failing*.

Quantitative data (i.e., student points for activities and attendance) were analyzed to assess differences among race/ethnicities and sample groups via three univariate Analysis of Variance followed by post-hoc pairwise comparison with LSD adjustment and two sided independent sample *t*-tests, respectively. Assumption of normality was checked with Q-Q plots and the Shapiro-Wilk test. Assumption of homogeneity of error variances was assessed with Levene's test. If the assumptions did not hold logarithmic transformation was attempted, or non-parametric analysis of multiple groups (i.e., races/ethnicities) were assessed with the Kruskal-Wallis Test and between groups comparisons were assessed with the Mann-Whitney Test. Simple regression of attendance on class session was used to assess attendance over time and to identify differences among race/ ethnicities by comparing 95% confidence intervals of unstandardized coefficients. Adjustment of a-level for multiple comparison was done by the Bonferroni method were applicable. Level of significance was set a priori at $a = .05$ for all analysis.

Results

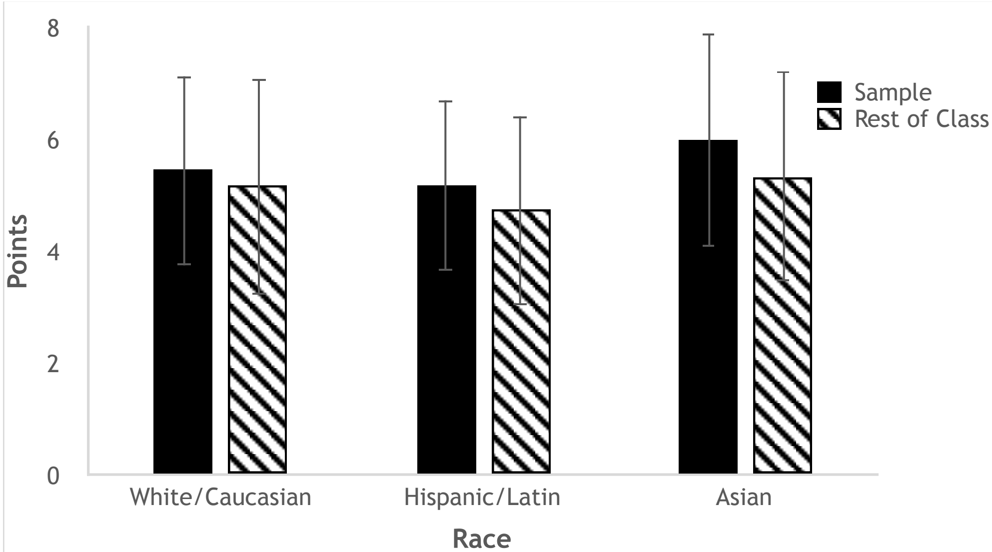
This exploratory mixed-methods study showed that across the three race/ethnicity groups, all students perceived that their higher order thinking improved, perceived both benefits and challenges of active learning in large enrollment courses and that throughout the process, they had a fear of failing (see appendix 1).

Improved higher order thinking skills

Although participants were juniors and seniors, many indicated that this course was the first time they were explicitly taught higher order thinking skills. Participants stated "No [they've never heard of higher order thinking]"; or "Never learned the details of higher order thinking." One student stated "I've never heard of higher order thinking. I'm used to standard lectures". These statements shed light on the lack of explicit attention to higher order thinking strategies and activities these students received in previous courses, regardless of ethnic background. Providing explicit instruction in higher order thinking gives students the tools they need to work in groups in meaningful ways. Because students were taught explicitly, most students regardless of ethnicity, perceived their higher order thinking skills had improved. For the question, "Do you believe your higher order thinking skills have improved? Please explain." One student stated: "I have applied these forms of higher order thinking skills in other classes this semester." Another participant said, "I do [feel that my higher order thinking improved]." "I did feel like I applied the knowledge I acquired towards assignments." Finally, a participant commented that working with others helped. The student stated, "[My higher order thinking skills improved], through partner activities and engaged learning." There were no discernable differences among the Race/ethnicity groups as all perceived they improved their higher order thinking skills in a similar manner and magnitude. This was supported by quantitative assessment of points obtained during the on-line class higher order thinking activities.

Quantitative analysis revealed no significant difference in points earned for activities across race/ethnicities across the sample ($F(2, 30) = 0.648$, $p = .530$, $\eta_p^2 = 0.04$) with absolute mean differences ranging from 0.3-0.8 points (Figure 1). There was also no significant difference for each race/ethnicity sample compared to the rest of the class i.e. Whites (*Medians* = 5, $U = 277.5$, $p = .663$, $r = 0.05$), Hispanic/Latino (*Medians* = 5, $U = 79.5$, $p = .667$, $r = 0.08$) and Asian (*Medians* = 6, $U = 151.5$, $p = .406$, $r = 0.13$).

Figure 1: Mean and SD of points gained in higher order thinking activities.

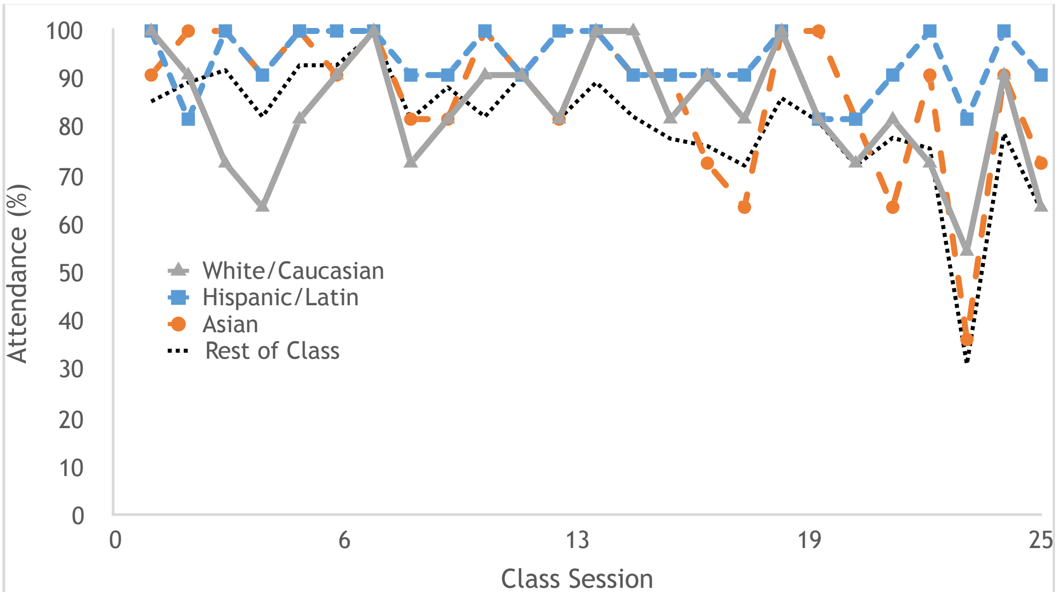


Benefits and Challenges in large enrollment course

Participants also describe benefits and challenges of engaging in higher order thinking skills through active learning activities within a large enrollment course. They stated that some of the challenges included working with unprepared students, not enough time and lack of feedback due to class size. One participant stated, "The size of the class makes getting deeper into concepts and materials difficult." This sentiment was expressed across several students and race/ethnicity in the study. One self-identified Asian male was particularly concerned about active learning in a large lecture setting. He stated, "I understand the purpose of this teaching method, but I believe that to have college students who have been taught another way their entire lives is difficult for us to shift gears. I learn better through the other way of teaching that we are used to. Yes, some students may "memorize" materials without fully understanding but others like myself, memorize and understand at the same time." This student clearly explains the struggle of trying to learn in a different course format. Further, despite the course's focus on Bloom's taxonomy, he believes that "understanding" course material is sufficient.

Participants also felt that at times the course was unorganized because there were so many students moving around getting into "active learning groups." Students were not familiar with moving around and engaging in group work in most of their courses, so it took some getting used to. One student commented: "At times, active learning activities in large lecture hall got a little chaotic and unorganized but I think it's worth the trouble." These sentiments about the benefits and challenges of active learning activities within a large enrollment course were shared across race/ethnicity groups, and this was again supported by quantitative analysis such as engagement with the course as measured by attendance. Attendance was generally very high (mean > 84% for all groups, Figure 2). Nevertheless, even at this high level, there was still a significant difference among groups (i.e., no ceiling effect, $\chi 2$ (2, N = 75), p = 0.15) with a mean rank of 28 for Whites, 47 for Hispanic/Latinos and 37 for Asians. There was no significant decline in attendance over the course of the semester for either Whites or Hispanic/Latinos group (b = -0.2 to -0.5, t(24) = -1.30 to -1.43, p = .165-.207) but attendance did decline for Asians (b = -1.0, t(24) = -2.80, p =.010).On course session 23, attendance dropped for all students and particularly so for Asians. However, there was no significant difference across trends for each group as indicated by overlapping confidence intervals (-1.20 to 0.22, -0.62 to 0.14 and -1.83 to -0.27 for Whites, Hispanic/Latinos, and Asians, respectively).

Figure 2: Mean attendance for each session across the semester.



Fear of Failure

The subject of points and how they negatively or positively affect grades came up quite often in these surveys. The active learning activities were completed for extra credit points. Students had a choice whether or not to complete these activities. In this way, students were encouraged to make their own decisions about their level of engagement. For these assignments, participants were concerned about being in "weak" groups, or groups they had to "carry." One participant stated, "I felt I had "carry" my classmates who did not show up or participate." Another student stated, "I didn't like working and depending on random people for some of the in-class assignments."

All participants in this study completed the activities but their reason for doing so mostly focused on the outcome, not the process. Some students stated that they appreciated the effort. "[The professor] is trying hard to change how students learn but that is hard to do when we have been "taught" to learn a certain way for 20+ years." Another student expressed frustration with never having a "right" answer. She stated, "I didn't like how some questions were answered by the professor with another question or with multiple questions. There was no clear answer."

Because of these issues, they were also afraid of failing the class/not passing the final. One student spoke to this, "I just hope I did well enough. Unfortunately, GPA is extremely important for grad school, so I just need an "A" for the class. I will not be evaluated on "what I learned in [this course]." The education system is built this way." All race/ethnicity groups shared this fear and expressed equal concern about their grade.

Discussion

This exploratory mixed-methods study examined underrepresented and non-underrepresented student perceptions of active learning in a large enrollment STEM course. It illustrates the complexity between integrating best practices in large enrollment STEM courses and the need for high grades so that all students have an opportunity to succeed in post-secondary education. The qualitative data indicated that the majority of participants across groups perceived active learning as beneficial but challenging in a large class, that their understanding and engagement in higher order thinking skills improved, and that they were very concerned about failing the course. The perceptions of improved higher order thinking skills were shared across all race/ethnicity groups, and the quantitative analysis supports this notion. All race/ethnicities showed similar achievement in higher order thinking activities, demonstrating wide applicability of this approach. This is of note because while lecture only approaches are ineffective for most students (Freeman et al., 2014; Wieman, 2007; Zoller, 1993), they are particularly ineffective for diverse students in large lecture classes with a reputation for high failure rates such as is often the case in many STEM courses (Haak et al., 2011; Hrabowski, 2011; Mervis, 2010). Thus, this approach provides an avenue for all students to engage in cognitively challenging concepts in critically meaningful ways. While it was beyond the scope of this study to objectively assess higher order thinking skills with a standardized, validated instrument, this outcome is nevertheless highly encouraging because students are capable of accurately reporting their own learning (Chesebro & McCroskey, 2000; Smith & Cardaciotto, 2012). In addition, their perceptions have the potential to influence learning outcomes (Lizzio et al., 2002; Smith & Cardaciotto, 2012). Indeed, students that perceive that their higher order thinking skills are improving are more likely to persist in engaging the skill than those that do not believe they are improving.

Other studies have also shown that, across all groups, purposeful and persistent practice of higher order thinking enhances these skills, the disposition to use them and their self-confidence (Miri et al., 2007; Nelson & Crow, 2014; Nguyen, 2016). In addition to similar achievement, engagement was also consistent across race/ethnicity with students from all assessed groups showing high levels of engagement (as assessed by attendance) with Hispanic/Latino students attending most regularly. As mentioned in the results, there was a noticeable but none significant drop in attendance during course session 23.This day was announced to be a specific "question and answer" session and did not involve activities for points which likely lowered attendance and participation rates. This is meaningful because it suggests that students may have been trained to attend class only if it affects their grade. This would indicate low internal motivation for attending class. If this data point is removed, then there is no more significant change in attendance over the course at the adjusted α-level. The qualitative data also supports this finding as students across race/ethnicity stated their frustration with the point system for the extra credit active learning activities. Students engaged in the extra credit activities to boost their grades because academic incentive structures value numerical indicators of improvement more than qualitative indicators. This is an important consideration given that problems can arise after initial implementation of active learning strategies when over the course of a week student perceptions become less favorable (McDougall, 2013). Based on the attendance data we did not observe such a tendency. This was in spite of the fact that students usually resent having to expend intellectual effort (Smith & Cardaciotto, 2012) without consistent, tangible reinforcements.

Implications and Recommendations for Practice

The current study extends the knowledge of active learning and examines the effects of best practice into a large, upper level undergraduate course with underrepresented students. A proper college education must go beyond mere knowledge acquisition and skill development in a specific discipline. This notion is not just philosophical but supported by stakeholders such as employers who expect students to gain proficiencies in non-material specific areas such as team work and higher order thinking (Hart Research Associates, 2015). In fact, a recent study of over 400 employers who hire a large percentages of college graduates indicates that more employers (25 vs. 15%) believe that long term career success depends on general knowledge and skills rather than knowledge and skills of a specific field or position (Hart Research Associates, 2015). This is contrary to the popular tendencies for colleges and university to emphasis 'job readiness,' and career focused/ specific training (Brand & Valent, 2013; Docherty, 2014; Fein, 2017). The highest valued learning outcomes by these employers include oral and written communication skills, teamwork, decision making, critical thinking and application of knowledge, analysis, evaluating information, innovation/ creativity numeracy and solving problems with people from different backgrounds (Hart Research Associates, 2015). One cannot help but notice that this list is almost entirely opposite to standard lecture based education in large classes where isolated, individual information memorization is the predominant style of learning and passivity abounds (Baldwin, 2009; Brainard, 2007; National Research Council, 2003). Despite the obvious antithesis of a standard, large, lecture course format to these desired learning outcomes STEM education is still dominated by this type of instruction (Stains, 2018). Also of note is that these desired learning outcomes are heavily weighted towards the higher rungs of Bloom's taxonomy of higher order thinking.

Recommendations regarding 'Improved higher order thinking skills'

Going forward, one can build on the positive perceptions expressed by students regarding engagement in active learning for higher order thinking. One should highlight the importance of these skills and how to achieve them. We provided students with detailed explanations on what these skills are using Bloom's taxonomy. We also highlighted how these skills are valued by key stakeholders such as employers. We emphasized higher order thinking as a targeted student learning outcome for the course as well as outlined which specifics skills where targeted with each activity.

Recommendations regarding 'Benefits and Challenges in large enrollment course'

The challenges expressed by students are difficult to overcome. Students were randomly assigned to groups as research has shown the benefits of heterogeneous groups, but most of the participants felt this was problematic. Problems such as students being unprepared or unwilling to contribute sufficiently to group activities were commonly expressed. Strategies to rectify this issue include assessments prior to engaging students in activities. Readiness assurance testing such as employed in team based learning formats offer one such option, and several approaches have been described (Antoun, J., Nasr, R., & Zgheib, N. K., 2015). Individual contributions to collaborative group activities are difficult to monitor in large classroom environments. Peer evaluation is likely the most feasible solution to this issue. Recently, a combined analytical mapping approach requiring minimal computational effort has been developed which allows for reliable individual grade assignment based on peer marks (Dijkstra et al., 2016; Spartar et al., 2015). We aim to investigate the incorporation of these strategies in future interactions of similar courses.

Further challenges such as the unfamiliarity of active learning may actually present a desirable learning opportunity. A certain level of uncomfortableness will likely be unavoidable when employing active learning for higher order thinking. Requiring student to step out of their comfort zone and engage in an unfamiliar activity is an important learning outcome in and of itself. Ideally, this can be done without introducing undue anxiety (Cooper, 2018). However, one should be cognizant and empathetic to student perceptions by being transparent in expectations and explicit in direction. Further, it is likely prudent to first engage in activities that more heavily rely on the lower levels of Bloom's to carefully scaffold the progression to higher levels in order to aid in the transition to the new format and way of thinking (Vygotsky, 1980).

Organizational issues perceived by students could be resolved with a simple training session at the beginning of the course. Group assignments were based on student identifiers and seating charts projected on the screen. Students were expected to easily transition to new groups at the beginning of each class. We did not anticipate that this would present a challenge to upper division students. In the future, a simple training sessions where students will be asked to find their groups quickly and efficiently several times with several different arrangements could likely alleviate much of this issue. We can envision a 'fun challenge' where students successfully try to improve the time it takes to arrange themselves into the appropriate groups based on varying the overhead display. After a few such practice runs, we would expect students will become highly efficient in the desired transition.

Recommendations regarding 'Fear of Failure'

The issue of fear of failing is a difficult one to resolve. It was not dependent on students fearing that they were unable to learn from the activities. Rather, it is to students worrying that they would not be rewarded with a high grade. An obvious solution would be to remove all points associated with the activities, but this may be counterproductive. Participation seemed to (in part) depend on the points attached to the activities. Therefore assigning fewer or no points would likely result in reduced participation. On the other hand, increasing the points allocated to the activities could further increase uncertainty and anxiety as this was already the case with just assigning extra credit points. A potential solution could be to allow more space for 'failure' while keeping incentives for participation. One such strategy is to only count a certain number of activities towards the course grade (e.g., the ten activities with the highest scores). As students do not know beforehand which activities they will get graded on and how high their score will be, this should reduce the stakes and still provide an incentive for participation. Concomitantly, being assigned to a 'weak' or poorly performing group would hopefully be perceived less impactful as activities from such days will likely not factor into the top ten scores. Lastly, perceived discomfort of not having clear cut, single solution answers is again an issue that is understandable but likely unavoidable. Clear cut, single solutions are nice and comforting but simply do not represent the real world, especially not for issues that require higher order thinking. While one should be aware of and empathize with students concerns, one should not shy away from exposing them to uncertainty or a multitude of possible solutions to any given problem if one cares to prepare students for more than taking simple examinations.

Conclusion

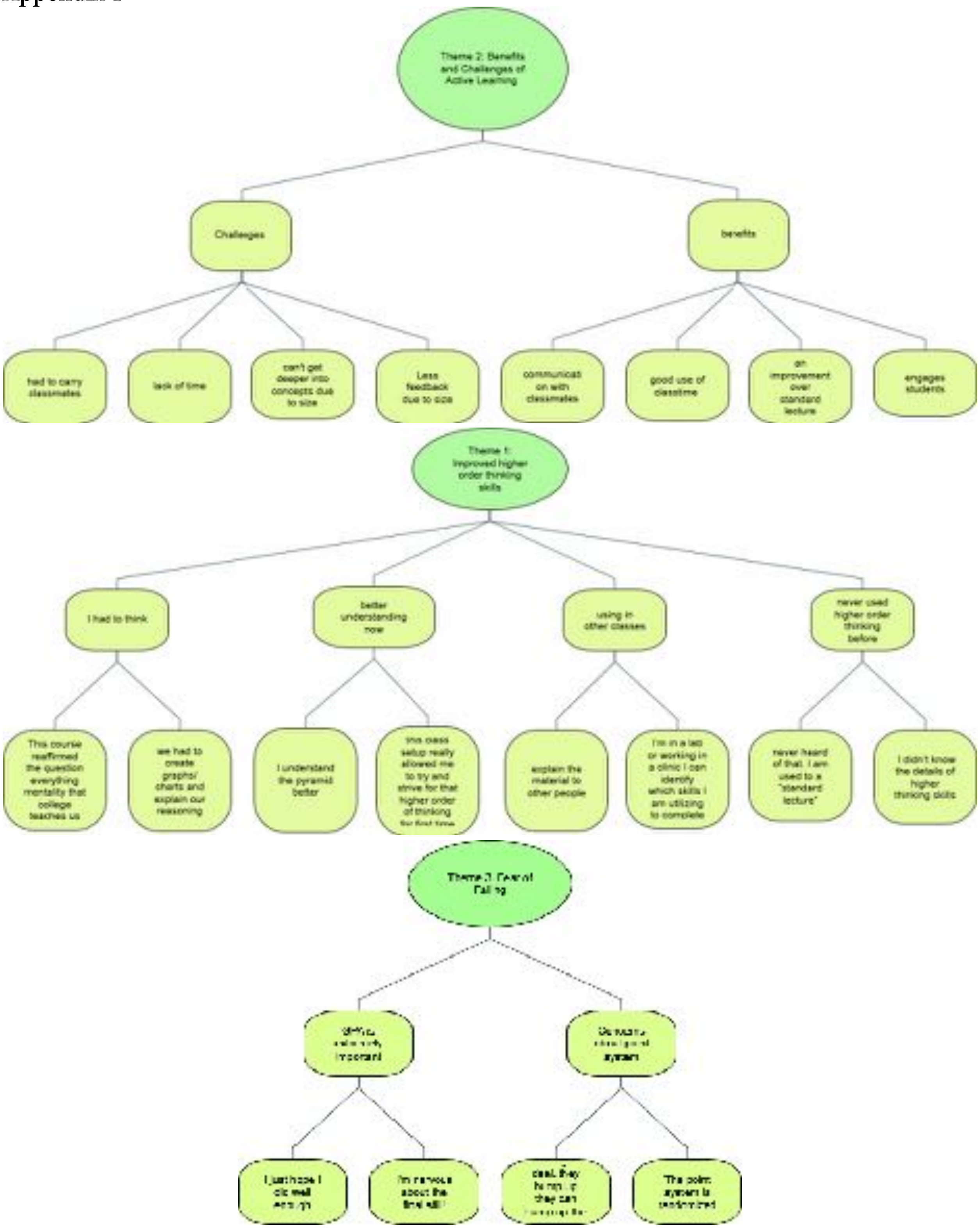
Active learning specifically targeted to higher order thinking was equally effective in engaging underrepresented as non-underrepresented students in this large enrollment STEM course (i.e., Measurement and Evaluation in Kinesiology). It also improved students' self-assessed higher order thinking skills. Clear definitions of higher order thinking and transparency regarding focus, importance, and impact of these skills were perceived as important elements for successful implementation by these students. Several challenges were perceived equally amongst students from all backgrounds regarding peer performance in group activities, organization difficulty due to class size, unfamiliarity of active learning approaches and grade pressures. Strategies addressing these challenges such as peer evaluation, organizational practice, progressive transition of activities and alternative grading strategies present promising targets for future investigations.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Appendix

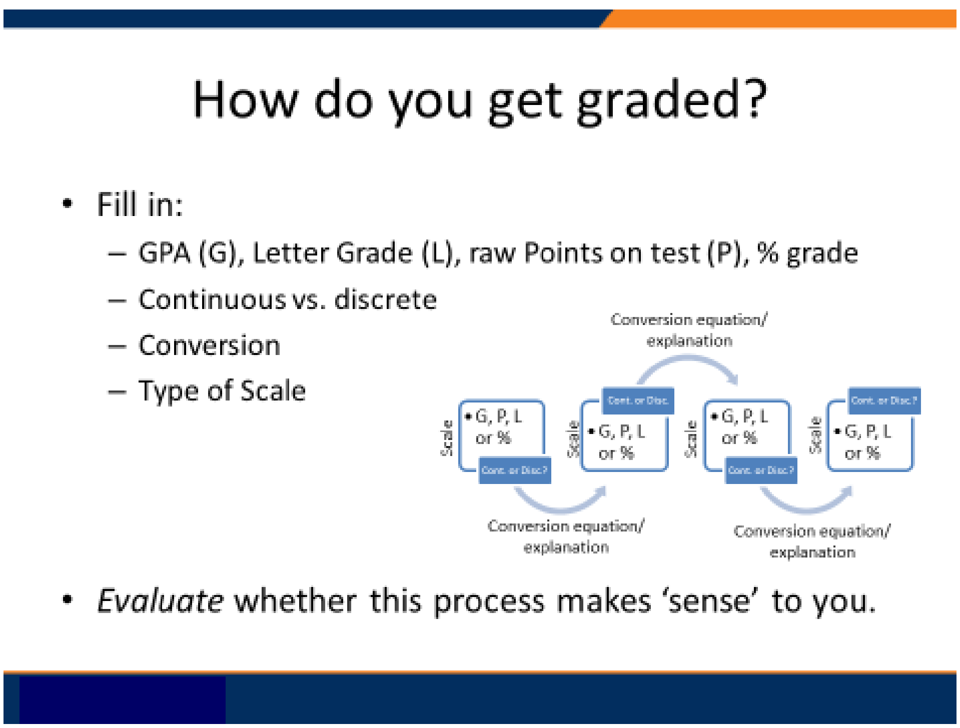
Appendix 1



Appendix 2

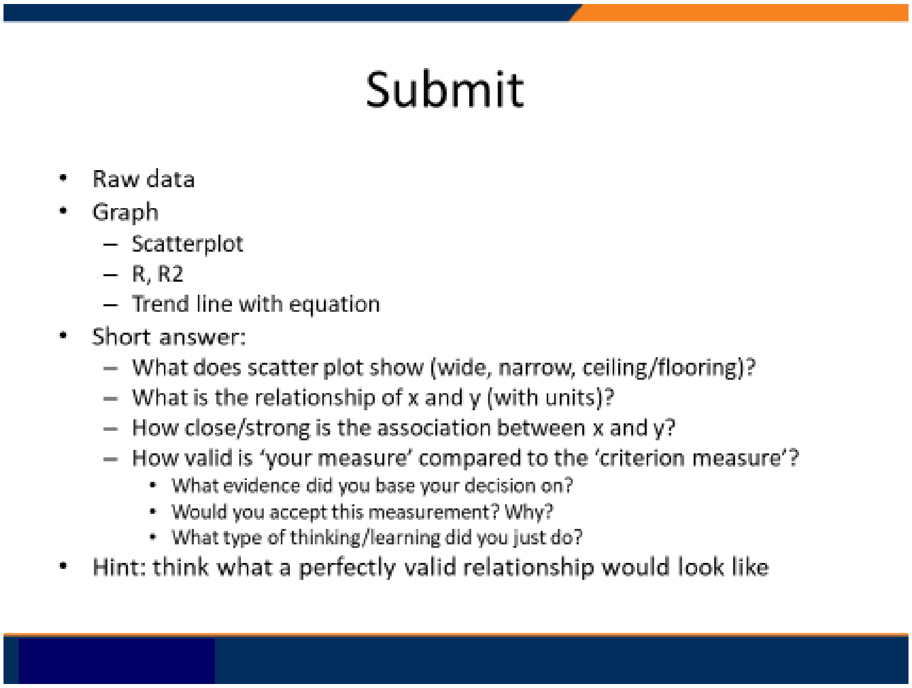
Activities were submitted for grading, and 1 point was awarded when activity specific criteria (outlined in examples below) were met. Each activity was worth 1 point. Total number of graded activities was 16.

Example of higher order thinking activity (early semester):
Students were tasked with applying terminology explained in the online lecture to a preset schematic on how grades are awarded at the University. Students were further tasked with evaluating whether this process makes sense to them in light of the concepts explained in the online lecture and discussed in class. The concepts of higher order thinking according to Bloom's taxonomy had been explained to students previously, and they were reminded and made aware of what type of thinking they were expected to perform for the activity (i.e., application, analysis, and evaluation). Students submitted their activity at the end of class. A point was awarded if the schematic was filled in correctly and if the written evaluation included discussion of issues regarding interconversion of continuous with discrete scores and ordinal with interval/ratio scales as well as a final position (irrespective of whether it was positive or negative).



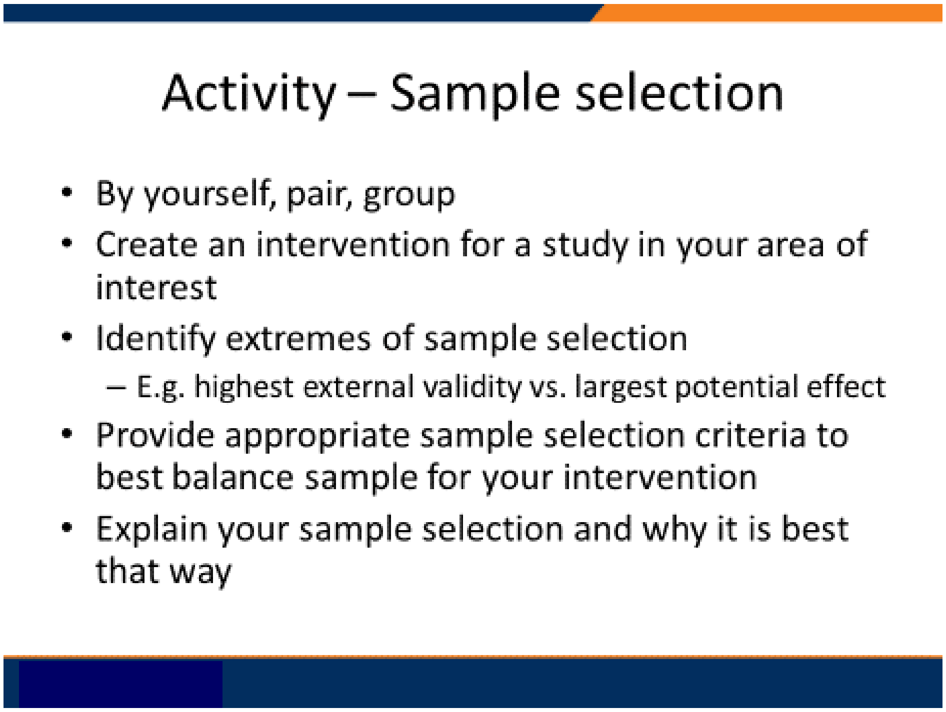
Example of higher order thinking activity (mid-semester):

Students were task to analyze data they had previously collected regarding different types of error applying simple, common analysis techniques such as correlation coefficients and scatter plots with trend lines. Students were further instructed to use these techniques evaluate the key measurement concepts of validity, culminating in the decision to either accept or reject their measurement. Students were also instructed to identify what type of higher order thinking they needed of accomplish this task. A point was awarded correct results and graphs were produced and appropriately discussed to support the final conclusion of whether or not the to accept the measurement.



Example of higher order thinking activity (late-semester):

As part of a multi-sectional exploration of the key components of study designs students were tasked to create an intervention, apply their knowledge of sample selection considerations, create sample selection criteria and evaluate their criteria. A point was awarded if the intervention was suitable for study, sample selection criteria were clearly outlined, and the explanation included appropriate discussions of the pros and cons of the criteria.


Activity – Sample selection
• By yourself, pair, group
• Create an intervention for a study in your area of interest
• Identify extremes of sample selection
– E.g. highest external validity vs. largest potential effect
• Provide appropriate sample selection criteria to best balance sample for your intervention
• Explain your sample selection and why it is best that way

Appendix 3

Informed Consent

Dear Student in ENS 305 Spring 2016,

This study is being conducted by Jochen Kressler, PhD, an assistant professor in the Exercise and Nutritional Science Department at SDSU and Benikia Kressler, PhD, assistant professor in the College of Education at CSU Fullerton. The aim of this study is to identify student perceptions of the course ENS 305 Measurement and Evaluation in Kinesiology for the spring 2016 semester. The course was redesigned for the spring 2016 semester to use I) a flipped format, II) a mastery learning progression approach and III) data based decision making to target instruction.

You are being asked to complete a survey because you are a student in the course. The results will be used as formative assessment to improve the course and potentially reported in an education journal.

The following survey includes questions that ask you to describe your perception of the course and instructor. The survey also includes questions about your demographic background. It will take about 30 minutes of your time to complete the survey.

Your participation in this study is voluntary. If you decide to participate, your responses will be confidential - that is, recorded without any identifying information that is linked to you. You will gain one point of extra credit for completing the survey. If you choose not to participate you will have the choice to engage in an alternative activity that will also include a one point extra credit opportunity.

If you have any questions regarding this survey, please contact Benikia Kressler at 657-278-3760.

Signing the consent form

I have read (or someone has read to me) this form and I am aware that I am being asked to participate in a research study. I have had the opportunity to ask questions and have had them answered to my satisfaction. I voluntarily agree to participate in this study.

I am not giving up any legal rights by signing this form. I will be given a copy of this form.

_______________________________ _______________________________
Printed name of participant **Signature of participant**

_______________________________ **AM/PM**
Date and time

DIRECTIONS: PLEASE ANSWER THE QUESTIONS BELOW HONESTLY AND TO YOUR BEST ABILITY

1) Demographics

a. Age in years: __________

b. Gender identity: ___________

c. Ethnicity/Race: ___________

d. Father's highest level of education (circle one):

less than high school   high school   some college   college graduate

e. Mother's highest level of education:

less than high school   high school   some college   college graduate

f. First generation college student: (circle one)?   YES   NO

g. Is English your first language (circle one)?   YES   NO

h. Current year in college (circle one):   Freshman   Sophomore   Junior   Senior

Major: __________

Emphasis: ___________

DIRECTIONS: PLEASE ANSWER THE QUESTIONS BELOW HONESTLY AND TO YOUR BEST ABILITY

2) Perception of Course

a. **What are your thoughts on active learning activities within the large lecture hall format? Please explain why.**

b. What are your perceptions of having the lectures online rather than in class? Please explain why.

c. Are there other things in this course that you enjoyed? Please explain why.

DIRECTIONS: PLEASE ANSWER THE QUESTIONS BELOW HONESTLY AND TO YOUR BEST ABILITY

d. Are there other things that you didn't like? Please explain why.

3) Higher Order Thinking Skills

a. Before ENS 305, did you know what higher order thinking skills were?

b. Do you believe your higher order thinking skills have improved? Please explain.

DIRECTIONS: PLEASE ANSWER THE QUESTIONS BELOW HONESTLY AND TO YOUR BEST ABILITY

5) Final Question

a. Do you think you got your "money's worth" in this course (and/or stakeholder's money's). Please explain why or why not.

THANK YOU FOR YOUR FEEDBACK!!

References

Ali, F. (2017). Global competitiveness report 2017-2018. Retrieved from http://www3.weforum.org/docs/GCR2017-2018/05FullReport/TheGlobalCompetitivenessReport2017%E2%80%932018.pdf

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., ... & Wittrock, M. C. (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives, abridged edition. *White Plains, NY: Longman*.

Antoun, J., Nasr, R., & Zgheib, N. K. (2015). Use of technology in the readiness assurance process of team based learning: paper, automated response system, or computer based testing. *Computers in Human Behavior*, *46*, 38-44.

Aronson, L. (2011). Twelve tips for teaching reflection at all levels of medical education. *Medical Teacher* 33(3): 200-205. DOI: 10.3109/0142159X.2010.507714

Baldwin, R. G. (2009). The climate for undergraduate teaching and learning in STEM fields. *New Directions for Teaching and Learning*, *117*, 9-17.

Bean, J. C. (2011). *Engaging ideas: The professor's guide to integrating writing, critical thinking, and active learning in the classroom.*: John Wiley & Sons.

Biggs, J. B. (2011). *Teaching for quality learning at university: What the student does.* McGraw-Hill Education (UK).

Bligh, D. A. (1998). *What's the use of lectures?* Exeter, England: Intellect books.

Bok, D. (2009). *Our underachieving colleges: A candid look at how much students learn and why they should be Learning More.* Princeton, New Jersey: Princeton University Press.

Bonwell, C. C. & Eison, J. A. (1991). *Active Learning: Creating Excitement in the Classroom. ASHE-ERIC* Higher Education Reports No. 1. Washington, D.C.

Braga, M., Paccagnella, M., & Pellizzari, M. (2014). Evaluating students' evaluations of professors. *Economics of Education Review 41*, 71-88.

Brainard, J. (2007). The tough road to better science teaching. *Chronicle of Higher Education 53*(48), A16.

Brand, B. & Valent, A. (2013). Improving college and career readiness for students with disabilities. *College and Career Readiness and Success Center.*

Carrell, S. E. & West, J. E. (2010). Does professor quality matter? Evidence from random assignment of students to professors. *Journal of Political Economy*, *118*(3), 409-432.

Casagrand, J. & Semsar, K. (2017). Redesigning a course to help students achieve higher-order cognitive thinking skills: from goals and mechanics to student outcomes. *Advances in Physiology Education, 41*(2), 194-202.

Casner-Lotto, J. & Barrington, L. (2006). *Are they really ready to work? Employers' perspectives on the basic knowledge and applied skills of new entrants to the 21st century US workforce.* Partnership for 21st Century Skills. 1 Massachusetts Avenue NW Suite 700, Washington, DC 20001.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. London; Thousand Oaks, Calif.: Sage Publications.

Chesebro J.L & McCroskey JC (2000). The relationship between students' reports of learning and their actual recall of lecture material: A validity test.

Cooper, K. M., Downing, V. R., & Brownell, S. E. (2018). The influence of active learning practices on student anxiety in large-enrollment college science classrooms. *International Journal of STEM Education*, *5*(1), 23.

Creswell, J. W., Plano, C., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. *Handbook of Mixed Methods in Social and Behavioral Research* 209: 240.

Dede, C. (2010). Comparing frameworks for 21st century skills. *21st Century Skills: Rethinking how Students Learn* 20: 51-76.
Dijkstra, J., Latijnhouwers, M., Norbart, A., & Tio, R. A. (2016). Assessing the "I" in group work assessment: State of the art and recommendations for practice. *Medical teacher*, *38*(7), 675-682.
Docherty, D. (2014). Universities must produce graduates who are ready for any workplace. *The Guardian.*
Dunkin, M. J. (1983). A review of research on lecturing. *Higher education research and development*, *2*(1), 63-78.
Dunkin, M. J. (1990). Willingness to obtain student evaluations as a criterion of academic staff performance. *Higher Education Research and Development*, *9*(1), 51-60.
Eichler, J. F. & Peeples, J. (2016). Flipped classroom modules for large enrollment general chemistry courses: a low barrier approach to increase active learning and improve student grades. *Chemistry Education Research and Practice, 17*(1): 197-208.
Fein, P. (2017). Better Marriage between College and Job Training. *Inside Higher Ed.*
Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., et al. (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.
Gasiewski, J. A., Eagan, M. K., Garcia, G. A., Hurtado, S., & Chang, M. J. (2012). From gatekeeping to engagement: A multicontextual, mixed method study of student academic engagement in introductory STEM courses. *Research in Higher Education, 53*(2): 229-261.
George, M. D. & Bragg, S. (1996). *Shaping the future: New expectations for undergraduate education in science, mathematics, engineering, and technology.* DIANE Publishing.
Ghaffarzadegan, N., Hawley, J., & Desai, A. (2014). Research workforce diversity: The case of balancing national versus international postdocs in US biomedical research. *Systems Research and Behavioral Science, 31*(2), 301-315.
Gutwill-Wise, J. P. (2001). The impact of active and context-based learning in introductory chemistry courses: An early evaluation of the modular approach. *J.Chem.Educ, 78*(5), 684.
Haak, D. C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science, 332*(6034), 1213-1216.
Harland, T. (2003). Vygotsky's zone of proximal development and problem-based learning: Linking a theoretical concept with practice through action research. *Teaching in Higher Education, 8*(2), 263-272.
Harris, R. P., Helfand, M., Woolf, S. H., Lohr, K. N., Mulrow, C. D., Teutsch, S. M., et al. (2001). Current methods of the US preventive services task force: a review of the process. *American Journal of Preventive Medicine, 20*(3), 21-35.
Hart Research Associates. (2015). *Falling Short? College Learning and Career Success.* Connecticut Avenue, NW Washington, DC. Retrieved from https://www.aacu.org/sites/default/files/files/LEAP/2015employerstudentsurvey.pdf
Hrabowski, F. A. (2011). Boosting minorities in science. *Science, 331*(6014), 125.
Kornell, N. & Hausman, H. (2016). Do the best teachers get the best ratings? *Frontiers in Psychology, 7*, 570.
Lizzio, A., Wilson, K., & Simons, R. (2002). University students' perceptions of the learning environment and academic outcomes: implications for theory and practice. *Studies in Higher Education, 27*(1), 27-52.
LoPresto, M. C. & Slater, T. F. (2016). A new comparison of active learning strategies to traditional lectures for teaching college astronomy. *Journal of Astronomy and Earth Sciences Education, 3*(1), 59.
Lou, Y., Abrami, P. C., & Spence, J. C. (2000). Effects of within-class grouping on student

achievement: An exploratory model. *The Journal of Educational Research*, *94*(2), 101-112.

Lumpkin, A., Achen, R. M., & Dodd, R. K. (2015). Using technology-nested instructional strategies to enhance student learning. *InSight: A Journal of Scholarly Teaching*, *10*, 114-125.

Maddux, W. W., Bivolaru, E., Hafenbrack, A. C., Tadmor, C. T., & Galinsky, A. D. (2014). Expanding opportunities by opening your mind: Multicultural engagement predicts job market success through longitudinal increases in integrative complexity. *Social Psychological and Personality Science, 5*(5), 608-615.

Mayer, R. E., Stull, A., DeLeeuw, K., Almeroth, K., Bimber, B., Chun, D., et al. (2009). Clickers in college classrooms: Fostering learning with questioning methods in large lecture classes. *Contemporary Educational Psychology,* 34(*1*), 51-57.

Mazur, E. (2011). *Memorization or understanding: Are we teaching the right thing?* Waterloo, ON, Canada, University of Waterloo.

McDougall, D. A. (N/A). *Doc's Flipped Classroom.* http://docsflippedclassroom.weebly.com/

Mervis, J. (2010). Undergraduate science. Better intro courses seen as key to reducing attrition of STEM majors. *Science, 330*(6002), 306.

Miri, B., David, B., & Uri, Z. (2007). Purposely teaching for the promotion of higher-order thinking skills: A case of critical thinking. *Research in Science Education, 37*(4), 353-369.

Mulligan, D., & Kirkpatrick, A. (2000). How much do they understand? Lectures, students and comprehension. *Higher Education Research & Development*, *19*(3), 311-335.

Nail, R. (2012). Education is undergoing a startling revolution — Let's support it! *Forbes*, Retrieved from https://www.forbes.com/sites/singularity/2012/06/27/education-is-undergoing-a-startling-revolution-lets-support-it/#75b0ece351ea

National Research Council. (2003). *Evaluating and Improving Undergraduate Teaching in Science, Technology, Engineering, and Mathematics.* National Academies Press.

Nelson, L. P. & Crow, M. L. (2014). Do active learning strategies improve students' critical thinking? *Higher Education Studies, 4*(2), 77-90.

Nguyen, B., Yu, X., Japutra, A., & Chen, C. S. (2016). Reverse teaching: Exploring student perceptions of "flip teaching". *Active Learning in Higher Education, 17*(1), 51-61.

Nguyen, T. T. B. (2016). *Critical thinking in a Vietnamese tertiary English as a foreign language context: Current practices and prospects*. (Doctoral dissertation).

Park, C. (2004). The graduate teaching assistant (GTA): Lessons from North American experience. *Teaching in Higher Education, 9*(3), 349-361.

Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice.* Thousand Oaks, CA: Sage Publications, Inc.

Ryan, J. J., Anderson, J. A., & Birchler, A. B. (1980). Student evaluation: The faculty responds. *Research in Higher Education*, *12*(4), 317-333.

Saavedra, A. R., & Opfer, V. D. (2012). Learning 21st-century skills requires 21st-century teaching. *Phi Delta Kappan, 94*(2), 8-13.

Sellami, N., Shaked, S., Laski, F. A., Eagan, K. M. & Sanders, E. R. (2017). Implementation of a learning assistant program improves student performance on higher-order assessments. *CBE Life Sciences Education, 16*(4), 10.1187/cbe.16-12-0341.

Smith, C. V. & Cardaciotto, L. (2012). Is active learning like broccoli? Student perceptions of active learning in large lecture classes. *Journal of the Scholarship of Teaching and Learning 11*(1), 53-61.

Spatar, C., Penna, N., Mills, H., Kutija, V., & Cooke, M. (2015). A robust approach for mapping group marks to individual marks using peer assessment. *Assessment & Evaluation in Higher Education*, *40*(3), 371-389.

Stains, M., Harshman, J., Barker, M. K., Chasteen, S. V., Cole, R., DeChenne-Peters, S. E., ... & Levis-Fitzgerald, M. (2018). Anatomy of STEM teaching in North American universities. *Science*, *359*(6383), 1468-1470.
US Department of Education. (2014). Graduation rates of first-time, full-time bachelor's degree seeking students at 4-year postsecondary institutions, by race/ethnicity, time to completion, sex, and control of institution: Selected cohort entry years, 1996 through 2006. Retrieved from https://nces.ed.gov/programs/digest/d13/tables/dt13_326.10.asp
Uttl, B., White, C. A., & Gonzalez, D. W. (2017). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation, 54*, 22-42.
Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Harvard university press.
Webb, N. M. (1980). An analysis of group interaction and mathematical errors in heterogeneous ability groups. *British Journal of Educational Psychology*, *50*(3), 266-276.
White, P. J., Larson, I., Styles, K., Yuriev, E., Evans, D. R., Rangachari, P. K., ... & Eise, N. (2016). Adopting an active learning approach to teaching in a research-intensive higher education context transformed staff teaching attitudes and behaviours. *Higher Education Research & Development*, *35*(3), 619-633.
Wieman, C. (2007). Why not try a scientific approach to science education? *Change: The Magazine of Higher Learning, 39*(5), 9-15.
Winkelmes, M. A. (2013). Transparency in teaching: Faculty share data and improve students' learning. *Liberal Education*, 99(2), n2.
Winkelmes, M., Bernacki, M., Butler, J., Zochowski, M., Golanics, J., & Weavil, K. (2016). A teaching intervention that increases underserved college student's success. *Peer Review, 18*(1/2) Retrieved from https://www.aacu.org/peerreview/2016/winter-spring/Winkelmes
Zoller, U. (1993). Are lecture and learning compatible? Maybe for LOCS: Unlikely for HOCS. *Journal of Chemical Education*, *70*(3), 195.

Journal of the Scholarship of Teaching and Learning, Vol. 20, No. 1, April 2020, pp.65-84.
doi: 10.14434/josotl.v20i1.24877

Field-based Learning in Higher Education: Exploring the Benefits and Possibilities

Heather N. Fedesco
Vanderbilt University
h.fedesco@vanderbilt.edu

Drew Cavin
Colorado College

Regina Henares
Purdue University

Abstract: Field-based learning in higher education is lacking both in practice at colleges and in research within the academic literature. This study aims to address these deficits by exploring the benefits and possibilities of executing field study in higher education across a variety of courses. We report the results of a qualitative research design that included the observation of five courses and an analysis of a field study database within the natural sciences, social sciences, and humanities. Approximately eight students per observed course were interviewed three times during their course to assess perceptions of the class, their peers and instructor, the field experiences, and their motivation throughout the course. In total, 130 individual interviews were conducted with 45 students and 721 field trips from 2015-2018 in the database were analyzed. Results revealed that field-based learning enhances the degree of relatedness students feel with their classmates and instructors, they have a greater degree of intrinsic motivation in the course, and these experiences facilitate learning in ways that may not be replicated in the traditional classroom. In addition, we created a typology of field-based learning, which includes eight different trips that could be employed in higher education courses. We also identified general strategies to improve the execution of these trips.

Keywords: field-based learning; experiential education; relatedness; motivation; self-determination theory

Introduction

Field trips are used as a common pedagogical tool in K-12 education; however, the use of this tool seems to dissipate by the time students reach college as many higher educational institutions fail to employ this type of experiential learning in their curricula (Behrendt & Franklin, 2014; DeGiacomo, 2002; DiConti, 2004; Kolb & Kolb, 2005). This trend is also mirrored in the field trip literature, where there is an abundance of research on field trips in elementary and secondary schools, but much less research in colleges and universities. The present study aims to address these deficits in the research and practice of field trips in higher education by exploring its benefits across multiple disciplines, identifying possibilities for a range of field experiences, and suggesting strategies for its successful execution.

Experiential Learning and Field-based Study

Experiential learning developed from the writings and philosophy of educational theorist, John Dewey. Dewey (1887) believed that "education must be conceived as a continuing reconstruction of experience; that the process and goal of education are one and the same thing" (p. 13). Guided by Dewey's philosophy, Kolb and Kolb (2005) describe experiential learning as a continuous holistic process that occurs as a result of experiencing the world first-hand and exploring it directly through the five senses. It can take several forms including internships, service-learning, cooperative education, undergraduate research experiences, study abroad, and of interest to the current paper, field trips (Moore, 2010).

Field trips can be defined as "any journey taken under the auspices of the school for educational purposes" (Sorrentino & Bell, 1970, p. 223). Much of the research on field trips agree that the intended educational outcomes of field trips focus on the following five areas: developing social and personal skills; developing observation and perception skills; adding relevance and meaning to learning; providing first-hand real-world experiences; and enhancing intrinsic motivation and interest in the subject (Behrendt & Franklin, 2014; Larsen, Walsh, Almond, & Myers, 2017; Tal & Morag, 2009).

When experiential learning is enhanced successfully through field trips, there are many beneficial outcomes, which are most often highlighted in the K-12 pedagogical literature. For example, in a study of sixth-grade students' perceptions and recall of an environmental education field trip, Nadelson and Jordan (2012) found that students were able to transfer their knowledge during this event, and a month after the field trip they were able to recall the lessons associated with novel and hands-on activities that occurred during the trip. A study by Lai (1999) found that high school students who went on geography field trips were able to relate the theories they were learning in class to reality, consider different perspectives, see the relevance of geography in their lives, and gain social experiences and an increased sense of autonomy. Hutson, Cooper, and Talbert (2011) also found that field trips can have an impact on at-risk youths' interests, pursuit of a certain academic subject, vocational choice, and future career.

Field-based Learning in Higher Education

The most recent reviews summarizing the benefits and best practices of conducting field trips focus on experiences in primary and second schools (Behrendt & Franklin, 2014; DeWitt & Storksdieck, 2008; Wilson, 2011). These reviews offer suggestions that can be applied to many field trips but their K-12 context and focus on issues that may be irrelevant for older students (e.g. behavior management and the use of chaperones) may send a subtle message that field trips are best suited to support primary and secondary education.

Yet educators who have incorporated field trips into their higher education courses have discovered that these experiences are just as beneficial to their students as they are for K-12 learners. For example, in a study of a nine-week intensive Introduction to Geology course where students went on field excursions almost every day (e.g., measuring water quality, identifying rocks), students gained a statistically significant improvement in geoscience concept knowledge (Elkins & Elkins, 2007). In a marketing class for tourism and hospitality students that included a hotel tour, results revealed that students had positive attitudes in regards to field trips citing that these experiences helped them understand the course material, helped them perform better on course assignments, and stimulated their interest in the subject matter among other benefits (Goh & Ritchie, 2011). In another example from a Construction Management course, students reported that their field trip to a construction site complemented the learning objectives and that the trip made the course material more relevant (Gunhan, 2015). Moreover, on end-of-semester course evaluations, students highlighted the site visit as a memorable and beneficial learning experience (Gunhan, 2015).

While this research is promising, the small amount of articles focusing on field study in higher education is still limited in several ways. Typically, this research consists of case studies focusing on one academic discipline (e.g., Elkins & Elkins, 2007; Healey & Jenkins, 2000; Marvell, 2008; Wright, 2000). Furthermore, the academic disciplines that are encompassed by this research have predominantly been within narrow specialties of the natural and social sciences, such as geology, geography, and sociology, with history being perhaps the sole discipline in the humanities and arts where field study research has been conducted (Sundermann, 2000). Due to the subject specificity of the current field trip research, the existing literature tends to focus on how field trips can enhance the learning of the subject matter of that specific discipline with little discussion on how other disciplines or higher education as a whole can benefit from field trips.

Limiting higher education field-based learning research to specific disciplines may mislead instructors from other fields to think that field-based experiences are reserved only for these particular courses. It may also make it harder for educators teaching in disciplines that do not typically include field study to predict the potential benefits of field trips for their students. However, across the current field trip literature in both K-12 and higher education institutions, it appears that one common positive outcome is the social development that field trips can provide students (e.g., Behrendt & Franklin, 2014; Rennie, 2007; Tal & Morag, 2006). In fact, Larsen, Walsh, Almond, and Myers (2017) found that personal and social development was the outcome that students valued the most in field trips, above the more academically-oriented outcomes such as providing first-hand experience and developing observation and perception skills. However, the implications of social development and how this may positively impact other outcomes for students are only briefly discussed in the study; the authors state that student motivation may be dependent on instructor and peer relationships but do not go into detail about how this may play out.

Focusing on the relational benefits that students might glean from field-based learning, and emphasizing how those benefits can enhance student motivation and related outcomes, may be a viable way for instructors to see past disciplinary boundaries in order to encourage those who do not typically include field trips in their courses to consider the possibilities. Exploring some of the more global benefits that students can experience across a variety of field trips in a variety of disciplines might further encourage instructors to take advantage of these rich learning activities.

Field-based Learning, Relatedness, Motivation, and Self-Determination Theory

One theory that elaborates on the relationship between instructor/peer relationships and motivation is self-determination theory (SDT). SDT posits that human motivation is best fostered when the basic psychological needs for autonomy, competence, and relatedness are satisfied (Ryan & Deci, 2017). Autonomy refers to having a choice in one's own individual behaviors and feeling that those behaviors stem from individual volition rather than from external pressure or control. Competence refers to perceiving one's own behaviors or actions as effective and efficient. Relatedness refers to feeling a sense of belonging, closeness, and support from others.

When students feel that their basic psychological needs of autonomy, competence and relatedness are met, then they are more intrinsically motivated and are more likely to perform behaviors in the course out of genuine interest (Deci & Ryan, 2000). Research consistently shows that across all levels of education, students who are more intrinsically motivated experience higher quality learning outcomes such as greater perceived transfer of knowledge, higher degrees of creativity, and greater performance in the class (Black & Deci, 2000; Grolnick, Ryan, & Deci, 1991; Jang, Reeve, Ryan, & Kim, 2009; Niemiec & Ryan, 2009; Richards, Levesque-Bristol, Zissimopoulos, Wang, & Yu, 2018; Williams & Deci, 1996).

Field trips can offer the space and setting for autonomy, competence, and certainly, relatedness to be fostered. Therefore, it may be that any field trip across the disciplines can foster perceived relatedness, and perhaps other basic psychological needs, which can positively impact student motivation, leading to a host of beneficial academic outcomes (e.g., academic performance; Ryan & Deci, 2017).

Visualizing Field-based Learning Possibilities

As encouraging as it may be to know that field trips might improve perceived relatedness and student motivation, instructors who are intimidated by the thought of incorporating field-based learning into their course may still struggle to identify what exactly they could do in their class, as well as how to do it, in order to reap these benefits. For those who turn to the literature, they may read about excursions to science museums, outcrops, and zoos, but for a professor teaching philosophy, this may not be useful. In addition, instructors teaching more theoretical versus applied courses may struggle to identify how they might incorporate field experiences into their courses. Orion's (1993) research supports this notion suggesting that teachers may avoid field trips because they simply are unfamiliar with conducting them. Thus, there is a need to continue exploring the possibilities that higher education instructors can consider in terms of the types of experiences they might build into their course, as well as some suggestions for ensuring these experiences are executed effectively.

The Present Study

The current study seeks to answer the call for more work in field study in higher education by taking a wider-ranging approach rather than focusing on discipline-specific courses. Doing so may continue to allow college and university instructors to realize that field study is not just a K-12 pedagogical tool. Using SDT as a theoretical lens, we aim to explore the general outcomes of field trips in higher education from a broader standpoint. By exploring these outcomes, we hope to provide educators who teach in disciplines that do not typically incorporate field study into their courses an idea of the range of positive experiences they could offer for their students through field study. Furthermore, we seek to create a typology of field study experiences for instructors who struggle to imagine the possibilities that could exist in their course. We also provide a list of best practices for anyone wishing to reap the benefits of a successfully led field trip experience in a higher educational setting. Thus, we aim to answer the following research questions:

> RQ1: What are the trans-disciplinary benefits of field study for higher education students?
> RQ2: What types of field trips can be incorporated into higher education courses across multiple disciplines?
> RQ3: What are the best practices and pitfalls when incorporating field trips in higher education courses?

Methodology

Overview

Multiple methods were employed at a small, private liberal arts college in the Southwest to address the three research questions. Given the unique structure of this particular institution, where students take one course at a time for at least three hours a day for eighteen days, instructors across all disciplines have more opportunities to incorporate field study into their courses, making it an ideal setting to conduct our research.

To answer RQ1, upon IRB approval, courses from multiple disciplines that included a variety of field study components were observed and students from each observed course were interviewed in order to identify the benefits of field study in higher education. To answer RQ2, a database maintained by the second author that tracks all the field trips that occur at the institution was analyzed to create a typology of field trips. Finally, to answer RQ3, data from classroom observations, student interviews, the field study database, and a review of the field study literature were used to generate a list of evidence-based practices for conducting field trips.

Classroom Observations

Procedures. The first author (HF) selected five courses taught during the 2016-2017 academic year that featured a field study component for inclusion in the study. These courses spanned multiple disciplines across the natural sciences, social sciences, and humanities. All instructors who were approached granted permission to have their course observed and their students interviewed. Instructors were paid $200 as an honorarium for their participation in the study. Instructors also received a summary of all aggregate data.

Approximately half the number of classes were observed, resulting in about nine classroom observations per course. Class observations occurred during field study and in-class experiences in order to get a better sense of the entire class experience. For more information on the courses that were observed, please see Table 1.

<table>
<tr><th>Course</th><th colspan="3">Description</th></tr>
<tr><td>Geology</td><td colspan="3">➢ Introductory course that fulfilled a lab requirement and thus was enrolled by students in various disciplines.
➢ Spent 8 days outside the classroom to practice geological field work, which included:
▪ A two-day, one-night trip in the mountains. The class spent the night in a cabin on college-owned property
▪ Two day trips exploring the local area
▪ A three-day, two-night camping trip in the mountains
▪ A day trip to a nature and science museum</td></tr>
<tr><td colspan="2">Class size = 22</td><td>Number of observations = 10</td><td>Hours spent observing = 78.5</td></tr>
<tr><td>Political Science</td><td colspan="3">➢ Introductory course that occurred right before the 2016 Presidential Election and thus was enrolled by students in various disciplines.
➢ Students volunteered by themselves or with classmates to help with each of the following political campaigns:
▪ Presidential election
▪ State election
▪ Local election</td></tr>
<tr><td colspan="2">Class size = 26</td><td>Number of observations = 9</td><td>Hours spent observing = 23</td></tr>
<tr><td>Environmental Studies</td><td colspan="3">➢ Introductory course that was required for environmental program majors and minors and thus was mostly enrolled by students interested in this discipline.
➢ Spent 2 days outside the classroom to talk with community members about sustainable development practices, which included:
▪ A visit to a net-zero energy house in town
▪ An on-campus food panel consisting of local young alumni who work to prevent food waste in town</td></tr>
<tr><td colspan="2">Class size = 27</td><td>Number of observations = 7</td><td>Hours spent observing = 17</td></tr>
</table>

Philosophy	➢ Introductory course that was cross-listed with feminist and gender studies and fulfilled the social inequality and writing requirements. The course was enrolled mostly by humanities or political science students. ➢ The class spent 5 days outside the classroom to immerse themselves in the material in the following location: ▪ An offsite campus in the mountains with limited access to internet or cell phone services		
Class size = 25		Number of observations = 7	Hours spent observing = 40
Comparative Literature	➢ Introductory course that took place entirely off campus at the Newberry Library in Chicago. Thus, it was mostly enrolled by students from various disciplines who were interested in the opportunity to take a course in a different city. ➢ In addition to taking place entirely off-campus, the class spent 4 days exploring the city they were studying in, which included: ▪ A class dinner at the library director's apartment overlooking Lake Michigan ▪ A visit to the Art Institute of Chicago ▪ A visit to the Chicago History Museum ▪ A class dinner at a restaurant eating local fare		
Class size = 12		Number of observations = 9	Hours spent observing = 24
Total Observation Hours = 182.5			

HF's role was that of "focused participant observer" such that she refrained from interacting in the learning space as much as possible (Tracy, 2013). Students were informed of the study at the beginning of the course and knew they were being observed. HF sat in the back of the class and took notes on a laptop or clipboard depending on the location of class. She did not engage in class discussions, answer questions posed by the instructor, or complete course assignments. In some instances, during field study experiences, HF acted as a "play participant" wherein she shadowed students, ate with them, and engaged in recreational activities and small talk during down time (Tracy, 2013).

Measures. An observation protocol was created by drawing from previously established protocols. This includes Reeve et al.'s (2004) observation sheet, which has been used in projects focusing on SDT, Richmond, Gorham, and McCroskey's (1987) list of teacher immediacy behaviors, and part of the Classroom Culture subset of the Reformed Teaching Observation Protocol (RTOP; Sawada et al., 2002). This observation protocol was pilot tested in four class periods across two different courses that took place during the summer 2016 term. Based on this testing, slight modifications were made to the final protocol (see Appendix 1 for the observation protocol). Observation sheets were uploaded to Dedoose (SocioCultural Research Consultants, 2017), a web-based application that allows researchers to organize and analyze qualitative data.

Student Interviews

Procedures. Semi-structured interviews were conducted with approximately eight students per observed class. Each student was interviewed three times throughout their course. The first interview lasted approximately 30-45 minutes and students were paid $25 for their participation. The second and third interview lasted approximately 25-30 minutes and students were paid $15 per interview for their participation. In total, 130 individual interviews were conducted with 45 students. All students completed all three interviews with the exception of three students from the geology class. Two of those students completed only one interview and one student completed two interviews.

Measures. The interview guide was developed by the first author, which was informed by existing literature and included questions that tracked student perceptions of the course, their peers and instructor, perceptions of the field experiences, and their motivation throughout the course (see Appendix 2 for interview guide). Questions were added each week based on what was observed during the class. HF was unable to access a sample of students to pilot test these questions before the start of the semester. However, four additional students from the geology class, which was the first class that was observed, were interviewed to test out the interview protocol. Some questions were dropped or refined following these initial interviews.

Interviews were audio-recorded, transcribed by a professional service that signed a non-disclosure agreement, and de-identified. Transcribed files were then uploaded to Dedoose. Descriptive statistics of participants can be found in Table 2.

Table 2. Descriptive Statistics of Student Interviewees

	Geology	Political Science	Environmental Studies	Philosophy	Comparative Literature	**Total**
Participants	13	9	8	8	7	**45**
Number of Interviews	34	27	24	24	21	**130**
Sex						
Male	6	5	5	5	2	**23**
Female	7	4	3	3	5	**22**
Year in School						
First Year	1	0	2	3	2	**8**
Sophomore	8	8	5	1	5	**27**
Junior	1	0	1	2	0	**4**
Senior	3	1	0	2	0	**6**
Interview Hours	19.47	13.65	11.12	13.03	10.94	**68.21**
Single-spaced Typed Pages	532	397	319	325	317	**1890**

Field Study Database

As part of his role as the Director of Field Study, the second author (DC) maintains a web-based application called Summit (Ideal-Logic, 2018), which tracks the field study experiences that instructors at the institution incorporate into their courses. Instructors who would like to include a field trip in their course must submit details of their trip into the Summit database. Records from 2015-2018 were included in the study, which contained 721 field trips from 30 different academic departments.

Data Analysis

Given the extensive data collection procedures that occurred during the fall 2016 semester, it was not possible to code transcripts as they came in, thus they were coded after the fall semester. A research assistant (GM), who was a recent graduate of the college, helped with the coding process and thus acted as a local expert to further explain the local context and customs and answered questions that arose from the data (Cornish, Gillespie, & Zittoun, 2013).

For the interview data, HF and GM first conducted primary-cycle coding (Charmaz, 2006; Glaser & Strauss, 1967; Tracy, 2013) using half the data randomly selected from each observed course. They met once or twice a week to discuss the data. An Excel file was created that contained codes and definitions that emerged as important to participants and/or were related to the research questions. The constant comparative method (Charmaz, 2006) was also used to compare data that fit with each code, and/or to modify or add code definitions to fit new data. Memos were taken throughout this process to capture initial analyses of participants' words. They then refined the codes and their definitions and both coders coded all of the data in Dedoose. GM coded the data used to develop the code book while HF coded the remaining data. Codes were modified or added as needed during this second round of coding.

Next, HF conducted second-cycle coding (Tracy, 2013) in order to synthesize and group the first-level codes into second-level codes by identifying ways the codes were related (or unrelated) to one another and/or to ideas from theory or the pedagogy literature. During this phase, data were reassembled such that codes were systematically grouped under a hierarchical category that made conceptual sense (Charmaz, 2006; Tracy, 2013). For the observation data, the same processes occurred, however only HF coded that data.

For the field study database records, DC conducted a thematic analysis by reading and rereading the descriptions of each field trip entered into the system by faculty members (Corbin & Strauss, 1993; Strauss, 1987). As a check on coding validity, the typology of field trips generated by the thematic analysis was tested against a set of data gathered by HF in a separate study in which students described various field trips they attended in their courses. All trips from this dataset were successfully categorized using the typology.

Results

Analyses revealed that students experience several benefits from engaging in field study opportunities. Three themes emerged regarding the benefits students receive from these activities: 1) they feel a deeper sense of connection with their instructor and peers following field study experiences, 2) field study has a positive influence on their motivation, and 3) field study facilitated their learning. In addition, we identified a typology of field study trips that can occur across multiple higher education courses, as well as suggestions for executing field study in higher education settings.

Field Study Benefits

Field study is associated with more classroom connection. Students in all classes attributed their sense of closeness with their peers and instructor to the field study opportunities they embarked on. When asked why these experiences might lead to deeper connections, students often stated that there were more opportunities for small talk and casual interactions, which allowed them to get to know each other better. Some examples include riding together in vans from outcrop to outcrop during the geology class, eating all meals together for a week during the philosophy class, or walking around town knocking on people's doors while campaigning during the political science class. As one sophomore from the political science class said, "That's another place where I get to know people more in the class, because, yeah, we're walking around canvasing together, and have conversations that aren't about just class. They're just about life in general." Many times, these experiences were in more intimate or casual settings. When compared to the rigidity of a classroom with desks and chairs, interacting in a place that is more comfortable encouraged students to loosen up a bit more and engage in more small talk. One senior philosophy student stated regarding his experience at the offsite campus:

> Just because of the feel that [the offsite campus] has and everything here is a little more laid back and detached from other routines, and patterns, and thought processes that get cemented on campus... It does feel more low stakes and low pressure, and more conversational and colloquial.

Students also felt like they actually shared a novel experience with their classmates in these field study settings. They formed a more prominent memory together, rather than just sitting in a classroom. A sophomore in the geology class said:

> It's amazing because throughout thirteen years of my life so far, we've been in this scheduled classroom setting. [Other students] and I were sitting on top of a rock today just looking at the view, like, "What!" The opportunity to do that, and go out into the field...That's just inspiring to me.

Field study positively affects motivation. Students also benefited from increased motivation because of field study experiences within their course. Some of those benefits followed directly from the increase in connection among students and instructors because of the bonding that occurred during their trips. Students had more opportunities to engage in small talk in these settings, which in a sense, warmed them up to engage in course-related conversations with their peers and instructor. As one sophomore in the political science class said:

> I think you kind of understand where people are coming from, from a better angle. When people that I now know a little bit better talk, I'm a little more interested in what they have to say because I know them a little better.

Similarly, a junior in the philosophy class added, "When you're more comfortable with people, I think you feel more okay sharing your ideas."

Including experiences outside the traditional classroom can also break up the monotony of a course. Varying class activities is an important pedagogical tool that can increase student energy and engagement so inserting these various activities kept the course fresher. For example, a first-year in the philosophy course said before the class headed to the offsite campus:

> Even though the material is engaging it's still kind of droning on in the same classroom for so long, for so much time. The fact that we're switching it up soon, made me really want to engage in the classroom…I mean today was our last real full class in the classroom.

In many cases these field study settings allowed students to have a deeper sense of autonomy, which improved their motivation. That is, students often had the choice to explore what they wanted to explore in these settings. For example, geology students could pick their own sample of rocks to identify and political science students could choose which campaign they wanted to volunteer for. One sophomore from the comparative literature class that took place at the Newberry Library in Chicago said:

> Being in the library, you just have a lot of freedom and it's very independent, you can go in to any of the reading rooms and you can kind of really get in the zone when you're looking at your sources and researching.

For experiences that include more field work, students were able to do the work of the expert. One sophomore in the geology class said:

> Just being able to use your hands and be a real geologist kind of, even if it's long. That's what it is being a geologist. That's what I think the best part is. You can actually do what people do with the knowledge you're getting.

They had a more authentic experience that allowed them to practice what they were learning and gave them a chance to discover things on their own, which enhanced their competence.

Field study facilitates learning. Students identified the ways in which field study experiences positively impacted their learning. For instance, these experiences were more memorable and thus were easier to recall later on. One student shared:

> I think with the rock identification part of the quiz, you had to first look at the rock and identify the mineral and then say how the rock got there and how it was formed. I don't think if I hadn't been out in the field and I was making connections to where other rocks are in relativity to the rock I was looking at, and if [Professor] hadn't shown me the mountains that were right by me, and he said, 'This rock came from those mountains.' If I hadn't of had that visual and that understanding, then I think it would have been more difficult on that part of the test to do well. (Sophomore, geology)

Often times the novelty of the setting created a more vivid image in their minds and they were better able to draw from the lessons learned in those settings. Conversely, in standard classrooms, it can be easier for lessons to blend together.

In some cases, field study experiences permit students to see the "bigger picture," which often allows them to make deeper connections within their course. In other words, students might have more "a-ha" moments outside the traditional classroom. As one sophomore in the geology class noted:

> I guess in class you're just basically learning out of a textbook for the most part or you're just listening to your professor lecture you, but then in the field you can actually see the evidence that you can't really see in class, which I think is really cool. You can see why that rock turned into a metamorphic rock. You can see the fossils within the rock structure, which…shows that it was formed under water. That's cool. You can see more of the large-scale processes, like I said, out in the field more so than in the classroom.

These experiences allow them to get a better sense of how what they are learning is actually relevant.

For example, a sophomore in the environmental studies class that visited a resident's net-zero energy house said, "It's cool to see that what we talk about in class can be applied in the real world. Whether or not it's realistic for everyone, I don't know, but for at least that guy it was cool." It also provides another frame of reference for students, which can make it easier to refer to when back in the classroom. Because all students had the same memorable experience, it can often lead to richer discussions in class as they dissect those moments together and connect it to other things they have learned.

In other cases, these experiences allow them to grow by removing them from their college "bubble"; students enjoyed gaining a sense of place from these experiences as they gain exposure to settings and communities they may otherwise not have known. One senior in the geology class put it best by saying:

> For me, a big takeaway that will be for this class, because I am never going to get down to science and the nitty-gritty of it just where I am at, but I am really interested in the sense of place that geology can cultivate and having more of an awareness of my surroundings and the state and I think that that is a big difference between looking at rocks and their little compartments and identifying them. Being in the field and doing the same thing, looking at the very small details and the minerals and the specifics of that, but it is also ... You are in this huge landscape and you are focusing on one small thing but you are also able to, just by being there, see and locate how it is part of this way bigger picture and this way bigger place. Just even getting to drive around and see more of the state, it makes me more excited to learn about this stuff than when I am in the lab and just like, "Ugh, what rock are you?!" It is more motivating and exciting to be like, "Okay, if I can understand this and comprehend this, it is going to tell me something about the place that I am in and can tell me more about the history and can give me a greater connection to this area.

Typology of Field Study in Higher Education

A typology of different field trips that have been used across multiple disciplines is presented below. It should be noted that while each field trip type varies in complexity, all of them require forethought and solid execution beyond what is described in the current paper; it is imperative that educators and instructors adapt what is outlined here to their own curriculum and students.

Collecting primary data/visiting primary sources. Being out in the field allows the students and instructors to be embedded *within* the material that they are learning about and to experience the messiness of data collection firsthand. The process of gathering data can sometimes lead to difficult and "unproductive" outcomes (e.g., if the data don't reveal anything significant). However, there is much learning to be had from these difficulties if instructors are open to the uncertainty.

Guided discovery of a site. In guided discovery, the professor brings students to a site that is familiar to the professor but new to the students and plans an activity that leads the students to uncover an intended outcome. In this environment, students must use tools or skills they learned beforehand to discover what is going on in the surrounding site. For example, geology students can visit a rock outcropping and use the skills they learned in class to identify the origin of those rocks. This technique is similar to collecting primary data/visiting primary sources described above in that it can be a more active way for students to practice research tools; however, in the case of guided discovery, the activity is conducted in a setting that has a relatively known and guaranteed outcome.

Backstage access. Backstage access is simply when the class has some sort of access to a site or place that the general public does not have, thus giving a special experience to the students (e.g., visiting a net-zero residence). This has the obvious benefits of gaining insights that are difficult to come by normally, but there is also a possibility for networking connections for students.

Show and tell. Show and tell involves the class and professor having access to a third-party expert or site where they might hear from the expert. For example, the class goes to a museum or a field site and they move around and hear from a curator or researcher on the subject at hand. This can be a useful technique when combined with the backstage access technique described above. However, instructors must be aware that it can also result in students sitting (or standing) for an uninspiring lecture from the third-party. Therefore, it is important for the instructor to do their research on or meet beforehand with the third-party expert and plan accordingly to ensure active engagement from their students.

Place-based immersion. Place-based immersion is when the class spends a significant amount of time in a place, investigating either a specific subject or an entire breadth of subjects tied to that place. In contrast to visiting one field site for specific analysis, place-based immersion encompasses being deeply involved with the place for an extended period of time and absorbing all the nuances of interconnectedness that exists in the area.

Community engagement. Community-engaged learning is a well-defined subgenre of field study (Driscoll, 2008). The professor and students work with a local partner over the span of the course or travel to a site to do time-bound projects or observations that students can then reflect on. Reciprocity and time commitment is important to this method to honor ethical concerns with community partners.

Retreats. Retreats are when the class gets away to a remote site for as little as a day or as much as a week to bond, to focus on the subject or a special project, and/or to write. There does not need to be a reason for the class visiting a certain site, although the retreat can be combined with the benefits of place-based immersion. The main objective is to garner the benefits of close proximity and focused time together and away.

Special events. Special events can be integrated into other methods, and involves the class traveling to a conference or special event (e.g., a speaker, a performance, etc.) that is pertinent to the course content or objectives.

General Strategies for Executing Field Trips

The following is by no means a comprehensive guide to best practices for executing field trips within higher education; however, evidence for these suggestions derives from data gathered via classroom observations, student interviews, and the field study database as well as a review of the field study literature. While design, learning outcomes, and facilitation all play an important role in the success of a trip, we would like to add five specific tips to those basic principles.

1. Beginning the trip with a full value contract (FVC) can be an incredibly powerful way to craft the learning culture in the class. This term, borrowed from outdoor education, is a contract written for the group, by the group, that ensures that each group member will be "fully valued" during the field experience (Curtis, 2008, June 22). The FVC can set expectations for appropriate and inappropriate behaviors during the trip, establish agreed-upon group norms, and provide a document that can be referred to should these group norms or behaviors be violated during the trip. Roberts (2016) suggests that "this not be a one-off event, but rather a living, breathing document that has a 'seat at the table' so to speak in every class" (p. 113). Having clear expectations builds safety and community, which are foundational to successful field outings (See Curtis, 2008, June 22 for a sample full value contract).

2. Logistics play an important role in the successfulness of a field trip. If a student is hot, cold, hungry, thirsty, looking into the sun, tired or lacking in any number of other basic human needs, they will not be able to absorb much of any knowledge on a field trip. Careful planning of the messy details around food is crucial for success. Being aware of and planning for students' food allergies and preferences, as well as ensuring that students have enough food to eat, all work to ensure a receptive environment. In addition, gauging weather forecasts and preparing students with proper clothing and gear is especially crucial for outdoor field trips.

3. Careful scheduling of the sequence of events can improve the outcomes of a field trip. In his primer on teaching in the outdoors, Roberts (2007, May 10) states, "Generally speaking, mornings are better for intellectual topics, afternoons are better for hands-on activities, and evenings are better for interpersonal discussions. Think AM-Brain, PM-Body, and Evening-Heart (para. 2)." This helpful framing can guide scheduling across many of the different pedagogical strategies listed previously. This also highlights the benefits of the often-overlooked evening time on reflection and group cohesion, a key part of any experiential strategy.

4. Balanced programming should also be taken into consideration when planning field trips. There is a consistent need to make good use of field trip time, and justifiably so. Field trips are resource-intensive so there is a tendency to feel the need to pack as much as possible into the trip. However, instructors run the risk of over-programming students to the point of saturation and an inability to take in any more information. The alternative risk is under-programming and boredom. Taking students into a setting where they are not busy enough has the effect of short-changing them on what could otherwise be a valuable and effective use of time. It can be challenging to balance these two opposites. Ideally, before the trip instructors who might have a tendency to over-program should identify activities that can be cut if students are saturated. Similarly, instructors who may have under-programmed their trip should have a set of additional activities, experiences, or assignments that they can implement if time allows. What is most important is that instructors continuously take the "pulse" of their class throughout the trip and make necessary adjustments to maximize the benefits of each outing.

5. Finally, instructors about to embark on a field trip should be prepared to engage in risk management. Mishandling an emergency incident, a minor illness, or mental health situation can be devastating to participants and to institutions. While a comprehensive look at risk management on field trips is beyond the scope of this paper, it is a critical piece. Field trip leaders should be familiar with resources available to them while they are away from campus. In addition to familiarity with emergency resources in the field location, leaders should be aware of and have contact information for campus security, administration, transportation, and mental health and sexual assault counselors before heading off on a field trip (Martin, Cashel, Wagstaff, & Breunig, 2006).

Discussion

This paper sought to identify global benefits of field study in higher education, the types of trips that can be included across a variety of higher education courses, and strategies for executing these trips. Data gathered from student interviews, classroom observations, and a field study database in courses from the natural sciences, social sciences, and humanities helped answer these questions.

Results revealed that field study had an overall positive impact on students' relationships with their peers and professor, their motivation, and their learning. SDT helps explain how field study enhances these positive outcomes. Field trips undoubtedly provide ideal environments for fostering the basic psychological needs of autonomy, competence, and especially relatedness. Across all classes, students stated that their field trips contributed to the enhanced closeness they felt with their classmates and professor. Through both the purposeful planned activities and the more mundane shared interactions, like eating meals together and sitting on a bus, students cited being able to make connections with one another and develop deeper senses of relatedness, which was largely due to the increased opportunities to engage in small talk with their peers and instructor.

A study by Murphy (2001), which explored the social interactions of backpackers, supports the idea that mundane shared experiences may aid in building deeper connections. Her study found that backpackers were more likely to build connections and have social interactions with other backpackers while eating or hanging out in hostel common rooms rather than doing planned activities. This suggests that the mere act of traveling and being in a new place with other people helps deepen relationships and may be the same effect that is helping to build relatedness between students during field trips. However, this is not to discount the effect that planned activities could also have in building relatedness. A study by Reissman, Aron, and Bergen (1993) found that in married couples, only spending time together was not enough to increase marital satisfaction, but it was the act of doing "exciting" and novel activities together that built the most marital satisfaction. This suggests that doing novel activities with other people, like engaging in planned activities on field trips, could enhance relatedness and relationship satisfaction. These studies support the notion that field trips, through shared novel activities as well as mundane interactions, can provide the ideal environment for relatedness to be fostered.

In an environment of fostered relatedness, positive academic outcomes, such as intrinsic motivation and learning, can be enhanced. Once students felt a deeper sense of connection and comfort with their peers, they were more willing to engage in the course. This supports the findings by Trenshaw, Revelo, Earl, & Herman (2016), which identified relatedness as the most salient of the three basic needs in promoting intrinsic motivation to learn in computer engineering students.

Field study experiences further promote increased motivation by satisfying the other basic psychological needs of autonomy and competence (Ryan & Deci, 2017). Students feel more autonomous when they are allowed to explore aspects of their surroundings that they deem most interesting (e.g., identifying the rocks they want to look at). Students might also feel more competent when they can engage in behaviors that practitioners of the discipline engage in. Beyond the satisfaction of basic needs, field trips also helped student engagement by including more varied instructional practices to keep the course fresh, which is an advised teaching practice, particularly within K-12 education (Ripp, 2016).

The higher levels of motivation that students gain because of these field trips may certainly increase their academic performance (Ryan & Deci, 2017), but other features of the experience seem to also impact learning. For example, the more vivid and novel settings help moments in class become more memorable, to which students can draw on those experiences later to help with their recall of information, a phenomenon supported by research on everyday memory conducted by a team of neuroscientists (Takeuchi et al., 2016). These experiences also allowed students to see the bigger picture and make deeper connections among the material they were learning in class, thus promoting the relevance of the course material, which is a crucial feature in enhancing student motivation and learning (Keller, 1987). Moreover, regardless of the material that students are learning about, in field trip experiences, they also gain a deeper sense of place and learn more about the community they are studying in.

Finally, in order to help instructors visualize the possibilities for field-based experiences in their courses, we identified a typology of field study experiences that can be employed in any college course. These include: a) collecting primary data/visiting primary sources; b) guided discovery of a site; c) back-stage access; d) show and tell; e) place-based immersion; f) community engagement; g) retreats; and h) special events. This typology hopefully provides faculty, especially those who teach in fields that do not typically include field study, with ideas they can incorporate into their courses so that students can reap the relational, motivational and learning benefits associated with these trips. They can also rely on some of the suggestions we provided to ensure their field study experiences go smoothly, such as including a full value contract, paying attention to logistics (e.g., making sure basic human needs are met), carefully scheduling events, being mindful of over- or under-programming, and engaging in risk management.

Limitations and Future Directions

There were a few limitations to this study that warrant discussion. First, the courses that were included in this study had varying numbers of days that were spent on field trips. As a result, some of the trips had a greater impact on students than others, especially when the quantity or duration of trips was higher, like in the geology class. Therefore, perhaps the experiences of someone who spent 8 days in the field cannot be compared to the experiences of someone who spent 2 days in the field. Despite this, however, we were still able to identify common themes shared by all students across classes, regardless of how many trips they went on.

Another limitation is that the extensive range of field trips were made possible because the institution where this research took place has the funds and unique course format that allows for more field study opportunities. Despite having a guide for how to execute certain field trips, instructors at other universities may still struggle with the logistics of incorporating these trips into the class due to funding or scheduling issues. The biggest barrier, then, for carrying out field trips may be less about knowing which types of trips they might include or how to execute these trips and more about a lack of resources and feasibility of including these trips into a course that exists in a more traditional format. Perhaps a first step would be to convince institutions of higher education of the value of these experiences in order to encourage the allocation of funds for such events (Goh & Ritchie, 2011). Students might also be advised when signing up for classes to arrange their schedule in such a way that will allow them to register for a class that includes more field-based learning. Similarly, instructors teaching three-hour long courses, which is typical of night classes, might take advantage of the extended meeting time to incorporate trips into their course.

Due to the qualitative nature of this research, generalizations cannot be made to the larger population. Therefore, future research could quantitatively test whether courses that include field study are associated with more benefits compared to more traditional courses. Additional research could explore the benefits of each type of field trip. Moreover, strategies for overcoming the barriers to implementing these trips within more traditional course structures (e.g., semester-based formats) in higher education could also be identified. One potential research project could include a cost-benefit analysis of the incorporation of field trips into a course. Given that field trips can be resource intensive, it would be useful to gather evidence as to whether the payoff in terms of positive student outcomes is worth the cost.

Acknowledgements

This work was supported by the Andrew W. Mellon Foundation's Grant to Strengthen Innovation in Engaged Teaching and Learning [grant number 31300614]. The authors would also like to thank

Georgina Mburu for her valuable contributions to the project.

Appendix

Appendix 1. Observation Protocol

1. Date
2. Observation #
3. Class
4. Instructor
5. Class Description: Where class took place; if outside of class describe what they were doing; number of students/mentors/instructors
6. Interpersonal Interactions
    a. *Instructors*: how they relate to students; instructors let students get to know them; expresses caring; patient; enjoys time with students; invests time/attention; knows students' names and interests; expresses affection; listens carefully; is physically close with students; is energetic/passionate; makes good eye contact
    b. *Students*: high proportion of student talk, especially between and among students; climate of respect; helped each other during class; students were friendly towards each other; students having fun together; knew each other's names; talked with a variety of peers (not just the same few); had positive interactions with instructors; weren't afraid to talk to instructors during or after class
7. Relevance Enhancing: Instructor promotes relevance/value/importance of topic/activity; activities are related to students' personal lives, future classes, careers, etc.
8. Motivation & Engagement
    a. *Instructors*: encourages active participation; encourages student conjectures; teacher acted as resource to support students; fosters student interaction/contributions rather than instructors lecturing; tries to get students excited about class; inspires students to try hard
    b. *Students*: exhibits signs they are enjoying the class; are interested/having fun; have focused attention; apply effort; talk/ask questions/discuss
9. Uncategorized comments
10. Other (not related to major hypotheses/RQs)

Appendix 2. Interview Guide

1. Why are you taking this course?
2. What do you think of the class so far? What are things you like about it? What are things you don't like about it? Would you change anything about the course so far? How does it compare to other classes you've taken so far?
3. Do you like your classmates? How close do you feel with them? Do you feel closer to people in this class compared to other similar classes you've taken? Why?
4. Have any activities encouraged you to get closer with your classmates? How?
5. Do you like your instructor? Are you getting to know him/her well? Do you think you are closer to him/her than other professors you've had? Why or why not?
6. Compare the days you spent in the classroom versus the days you spent in the field…
    a. Which experience do you like better? Why?

   b. Which environment do you feel more focused/engaged in? Why?
   c. Which environment do you feel like you're learning more in? Why?
7. In general, do you like going on field trips with your class? What about in this class? Do you like traveling with your classmates? Your instructors? What are the pros and cons of field trips?
8. If you could plan a class period, what would you do?
9. How relevant do you think the course material is? Do you think you will use the information you learn in this class in your career, other classes, or your personal life?
10. Are there elements of this course that help/hinder your motivation to learn/work hard in the class?

References

Behrendt, M., & Franklin, T. (2014). A review of research on school field trips and their value in education. *International Journal of Environmental and Science Education*, *9*, 235-245. http://doi.org/10.12973/ijese.2014.213a

Black, A. E., & Deci, E. L. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self-determination theory perspective. *Science Education, 84,* 740-756. https://doi.org/10.1002/1098-237x(200011)84:6<740::aid-sce4>3.0.co;2-3

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. London, England: Sage.

Corbin, J., & Strauss, A. (1993). *Methods of qualitative research*. California: Sage.

Cornish, F., Gillespie, A., & Zittoun, T. (2013). Collaborative analysis of qualitative data. In U. Flick (Ed.), *The SAGE handbook of qualitative data analysis* (pp. 79-93). London, England: SAGE Publications.

Curtis, R. (2008, June 22). Guide to the full value contract. Retrieved from https://www.outdoored.com/documents/guide-full-value-contract

DeGiacomo, J. A. (2002). Experiential learning in higher education. *The Forestry Chronicle*, *78*, 245-247. http://doi.org/10.5558/tfc78245-2

Dewey, J. (1897). *My pedagogic creed*. New York: E. L. Kellogg & Co.

DiConti, V. D. (2004). Experiential education in a knowledge-based economy: Is it time to reexamine the liberal arts? *The Journal of General Education*, *53*, 167-183. http://doi.org/10.1353/jge.2005.0003

Dillon, J., Morris, M., O'Donnell, L., Reid, A., Rickinson, M., & Scott, W. (2005). *Engaging and learning with the outdoors: The final report of the outdoor classroom in a rural context action research project.* Slough, England: National Foundation for Education Research.

Driscoll, A. (2008). Carnegie's community-engagement classification: Intentions and insights. *Change: The Magazine of Higher Learning, 40*, 38-41. https://doi.org/10.3200/chng.40.1.38-41

Elkins, J. T., & Elkins, N. M. (2007). Teaching geology in the field: Significant geoscience concept gains in entirely field-based introductory geology courses. *Journal of Geoscience Education*, *55*, 126-132. http://doi.org/10.5408/1089-9995-55.2.126

Fong, B. (2004). Looking forward: Liberal education in the 21st century. *Liberal Education*, *90*, 8–12.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory*. New York: Aldine de Gruyter.

Goh, E. & Ritchie, B. (2011) Using the Theory of Planned Behavior to understand student attitudes and constraints toward attending field trips. *Journal of Teaching in Travel & Tourism, 11*, 179-194. http://doi.org/10.1080/15313220.2011.575024

Gunhan, S. (2014). Collaborative learning experience in a construction project site trip. *Journal of Professional Issues in Engineering Education and Practice*, *141*, 1-5. http://doi.org/ 10.1061/(asce)ei.1943-5541.0000207

Grolnick, W. S., Ryan, R. M., & Deci, E. L. (1991). The inner resources for school performance: Motivational mediators of children's perceptions of their parents. *Journal of Educational Psychology, 83,* 508-517. https://doi.org/10.1037//0022-0663.83.4.508

Healey, M., & Jenkins, A. (2000). Kolb's experiential learning theory and its application in geography in higher education. *Journal of Geography*, *99*, 185-195. http://doi.org/10.1080/00221340008978967

Hutson, T., Cooper, S., & Talbert, T. (2011). Describing connections between science content and future careers: Implementing Texas curriculum for rural at-risk high school students using purposefully-designed field trips. *Rural Educator*, *33*, 37-47.

Ideal-Logic (2018). Summit [Computer Software]. Retrieved from https://apps.ideal-logic.com/website/
Jang, H., Reeve, J., Ryan, R. M., & Kim, A. (2009). Can self-determination theory explain what underlies the productive, satisfying learning experiences of collectivistically oriented Korean students? *Journal of Educational Psychology, 101,* 644-661. https://doi.org/10.1037/a0014241
Keller, J. M. (1987). Development and use of the ARCS model of instructional design. *Journal of Instructional Development, 10*, 2–10. http://doi.org/10.1007/bf02905780
Kolb, A. Y., & Kolb, D. A. (2005). Learning styles and learning spaces: Enhancing experiential learning in higher education. *Academy of Management Learning & Education*, *4*, 193-212. http://doi.org/10.5465/amle.2005.17268566
Lai, K. C. (1999). Freedom to learn: A study of the experiences of secondary school teachers and students in a geography field trip. *International Research in Geographical and Environmental Education*, *8*, 239-255. http://doi.org/10.1080/10382049908667614
Larsen, C., Walsh, C., Almond, N., & Myers, C. (2017). The "real value" of field trips in the early weeks of higher education: The student perspective. *Educational Studies*, *43*, 110-121. http://doi.org/10.1080/03055698.2016.1245604
Martin, B., Cashel, C., Wagstaff, M., & Breunig, M. (2006). *Outdoor leadership: Theory and practice.* Champaign, IL: Human Kinetics.
Marvell, A. (2008). Student-led presentations in situ: The challenges to presenting on the edge of a volcano. *Journal of Geography in Higher Education*, *32*, 321-335. http://doi.org/10.1080/03098260701514173
Moore, D. T. (2010). Forms and issues in experiential learning. *New Directions for Teaching and Learning*, *2010*, 3-13. http://doi.org/10.1002/tl.415
Murphy, L. (2001). Exploring social interactions of backpackers. *Annals of Tourism Research*, *28*, 50-67. http://doi.org/10.1016/s0160-7383(00)00003-7
Nabors, M. L., Edwards, L. C., & Murray, R. K. (2009). Making the case for field trips: What research tells us and what site coordinators have to say. *Education*, *129*, 661-667.
Nadelson, L. S., & Jordan, J. R. (2012). Student attitudes toward and recall of Outside Day: An environmental science field trip. *The Journal of Educational Research, 105*, 220-231. http://doi.org/10.1080/00220671.2011.576715
Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *Theory and Research in Education, 7,* 133-144. https://doi.org/10.1177/1477878509104318
Orion, N. (1993). A model for the development and implementation of field trips as an integral part of the science curriculum. *School Science and Mathematics, 93*, 325-331. https://doi.org/10.1111/j.1949-8594.1993.tb12254.x
Reeve, J., Jang, H., Carrell, D., Jeon, S., & Barch, J. (2004). Enhancing students' engagement by increasing teachers' autonomy support. *Motivation and Emotion, 28,* 147-169. http://doi.org/10.1023/b:moem.0000032312.95499.6f
Reissman, C., Aron, A., & Bergen, M. R. (1993). Shared activities and marital satisfaction: Causal direction and self-expansion versus boredom. *Journal of Social and Personal Relationships*, *10*, 243-254. http://doi.org/10.1177/026540759301000205
Rennie, L. J. (2007). Learning outside of school. In S.K. Abell and N.G. Lederman (eds.), *Handbook of Research on Science Education*. Mahwah, New Jersey: Erlbaum.
Richards, K., Levesque-Bristol, C., Zissimopoulos, A., Wang, C., & Yu, S. (2018). *Creation and validation of the perceived knowledge transfer scale (PKTS).* Manuscript submitted for publication.
Richmond, V. P., Gorham, J. S., & McCroskey, J. C. (1987). The relationship between selected immediacy behaviors and cognitive learning. *Annals of the International Communication*

Association, *10*, 574-590. http://doi.org/10.1080/23808985.1987.11678663
Ripp, P. (2016). *Passionate learners: How to engage and empower your students* (2nd ed.). New York: Routledge.
Roberts, J. W. (2016). *Experiential education in the college context: What it is, how it works, and why it matters*. New York: Routledge.
Roberts, J. W. (2007, May 10). *Teaching in the outdoors: A primer.* Retrieved from https://www.outdoored.com/articles/teaching-outdoors-primer
Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. New York: Guilford.
Sawada, D., Piburn, M. D., Judson, E., Turley, J., Falconer, K., Benford, R., & Bloom, I. (2002). Measuring reform practices in science and mathematics classrooms: The reformed teaching observation protocol. *School Science and Mathematics, 102,* 245-253. http://doi.org/10.1111/j.1949-8594.2002.tb17883.x
SocioCultural Research Consultants (2017). Dedoose (Version 7.6.6) [Computer Software]. Retrieved from http://www.dedoose.com
Sorrentino, A. V., & Bell, P. E. (1970). A comparison of attributed values with empirically determined values of secondary school science field trips. *Science Education*, *54*, 233-236. http://doi.org/10.1002/sce.3730540308
Strauss, A. L. (1987). *Qualitative analysis for social scientists.* Massachusetts: Cambridge University Press.
Sundermann, E. L. (2013). History lab for undergrads: A day at the museum. *The Social Studies*, *104*, 250-258. http://doi.org/10.1080/00377996.2012.735280
Takeuchi, T., Duszkiewicz, A. J., Sonneborn, A., Spooner, P. A., Yamasaki, M., Watanabe, M., ... & Morris, R. G. (2016). Locus coeruleus and dopaminergic consolidation of everyday memory. *Nature*, *537*, 357-362. http://doi.org/10.1038/nature19325
Tal, T., & Morag, O. (2009). Reflective practice as a means for preparing to teach outdoors in an ecological garden. *Journal of Science Teacher Education*, *20*, 245-262. http://doi.org/10.1007/s10972-009-9131-1
Tracy, S. (2013). *Qualitative research methods: Collecting evidence, crafting analysis, communicating impact.* Chichester, West Sussex, UK: Wiley-Blackwell.
Trenshaw, K. F., Revelo, R. A., Earl, K. A., & Herman, G. L. (2016). Using self-determination theory principles to promote engineering students' intrinsic motivation to learn. *International Journal of Engineering Education*, *32*, 1194-1207.
Williams, G. C., & Deci, E. L. (1996). Internalization of biopsychosocial values by medical students: A test of self-determination theory. *Journal of Personality and Social Psychology, 70,* 767-779. https://doi.org/10.1037//0022-3514.70.4.767
Wright, M. C. (2000). Getting more out of less: The benefits of short-term experiential learning in undergraduate sociology courses. *Teaching Sociology*, *28,* 116-126. http://doi.org/10.2307/1319259

Journal of the Scholarship of Teaching and Learning, Vol. 20, No. 1, April 2020, pp.85-104.
doi: 10.14434/josotl.v20i1.25423

The Impact of Undergraduate Research and Student Characteristics on Student Success Metrics at an Urban, Minority Serving, Commuter, Public Institution

Sheila I Baron
University of Iowa, Statistics Outreach Center
sheila-barron@uiowa.edu

Pamela Brown
New York City College of Technology, City University of New York
pbrown@citytech.cuny.edu

Tammie Cumming
Brooklyn College, City University of New York

Michelle Mengeling
University of Iowa, Statistics Outreach Center
michelle.mengeling@gmail.com

Abstract: Challenges to establishing and maintaining undergraduate research programs include how to demonstrate impact as evidence for future funding, establish eligibility criteria when resources are limited, and assess new components. To address these challenges, undergraduate researcher GPA, credit accumulation and time to graduation were evaluated longitudinally, at an urban, public, minority and Hispanic serving, commuter college. Students who participated in undergraduate research and matched peers were also compared. Evaluation revealed that all groups benefited from participation in undergraduate research, whether they had full- or part-time status, were STEM or non-STEM majors, or participated in single or multiple semesters of research. Addition of mandatory workshops after the fourth year of the seven years of students evaluated, correlated with longer participation in voluntary undergraduate research. Longer participation correlated with higher GPAs. Entering freshmen and transfer students, who began research with no College GPA, were more likely to have low GPAs during the semester of participation, suggesting that a successful semester at the college before eligibility may be an evidence-based criteria to implement.

Keywords: undergraduate research, underrepresented groups, part-time students, STEM, non-STEM, eligibility criteria

Introduction

A key factor to whether college students persist and thrive is the degree to which they participate in educationally effective activities that contribute to their learning, personal development and success (Kuh, 2001; Kuh, 2003; Lopatto, 2006). Undergraduate research is a recognized high impact pedagogical practice that enhances student development and results in increased retention and degree completion; it has been identified as particularly important to the academic success of under-represented groups. Undergraduate research experiences have been demonstrated to support STEM-related career aspirations and increase STEM graduation rates (Gregerman *et al*, 1998; Davis, 2009; Espinosa, 2011; Hu *et al*, 2008; Johnson, 2011; Schultz *et al*, 2011; Ishiyana and Hopkins, 2005; Seymour *et al*, 2004; Laursen *et al*, 2010).

This paper will explore who chose to participate in undergraduate research through the Emerging Scholars (ES) program at a Hispanic and minority serving, urban, public commuter college (New York City College of Technology or "City Tech,"), their academic outcomes, and evaluate results for program optimization and demonstration of its impact on student success. This work adds to the body of knowledge on the impact of undergraduate research experiences on part-time students, a topic of little research, despite the fact that 37% of undergraduates are enrolled part-time, including 61% of students at community colleges (College Board, 2014). Given that 84% of Hispanic students and 81% of black students, versus 72% of white students, enroll part-time at least one semester, a recent study has shown that if part-time students from underrepresented groups graduated at the same rate as their white counterparts, the achievement gap between black and white students would close by 13 points and for Hispanic students, it would close 7 points (EAB, 2019). These authors concluded that because underrepresented minority and first-generation students are more likely to attend part-time, student success initiatives that only target full-time students will not close the achievement gap as effectively as those that focus on part-time students.

Additional topics reported include the impact of the timing of research in a student's academic career, the duration of the undergraduate research experience and the introduction of professional development workshops for students. This work thus serves both as a model for assessment of the impact of undergraduate research experiences as well as provides guiding evidence demonstrating impact and for developing eligibility criteria when resources are limited.

ES Program Components

City Tech is an open access college, offering associate and bachelor's degrees. We have been sponsoring the Emerging Scholars (ES) undergraduate research program since fall 2006. The goals of the program include:

a. enhance the intellectual vitality of the college by providing students with the opportunity to apply what they learned in the classroom to discover new knowledge and solve problems,
b. promote student academic success through the opportunity to engage with faculty, and
c. provide faculty with an "extra pair of hands" to advance their scholarship, in lieu of graduate students.

The Emerging Scholars program is open to all students in good academic standing (minimum GPA 2.0) and provides full-time students (enrolled for 12 or more credits) with $500 stipends for working with faculty on their scholarly activities, approximately 50 hours each semester and part-time students (enrolled for 1-11 credits) with $250 stipends for working approximately 25 hours each semester. All students are required to complete CITI certification in Responsible Conduct in Research within the first month. While the program was initially envisioned as supporting promising students that faculty selected, a growing number of students now approach faculty to become involved.

To improve and help grow the undergraduate research program, the Honors Scholars Program and Undergraduate Research Committee, an interdisciplinary faculty committee, began offering mandatory workshops to promote student researcher professional growth, beginning in fall 2010. Workshop topics for first time student researchers include *Advancing Library Research Techniques, Writing Abstracts for Research Projects, Designing a Research Poster Presentation* and *Developing and Delivering Effective Research Presentations*. For returning researchers workshops include: *Advanced Writing Abstracts for Research Projects, Presentation Skills, How to Succeed in an Internship, ePortfolios for Academic and Career Advancement, NYC Fire Department C-14 Certificate of Fitness Preparation, How to Write a Personal Statement, Writing Effective Cover and Thank You Letters, Research Poster Design,* and *Getting your Poster into*

Academic Works (Academic Works is CUNY's open source platform for presenting publications and presentations). Students are required to attend 4 workshops each semester in order to receive the stipend. Workshops are offered by various groups around campus including the Professional Development Center, the Writing across the Curriculum program, faculty volunteers, librarians, etc. Introductory workshop goals include developing the most common research and communication skills needed in research, which reduces this responsibility for faculty mentors, so that they can focus on the research project and preparing students for external research opportunities. Other investigators have noted that faculty mentors have reported students' lack of academic writing experience as an obstacle to engaging students in undergraduate research (Myers, 2018). Advanced workshop goals are to further develop communication skills, and help students prepare for next steps – internships, employment and graduate school.

In the data used for this paper, 214 students (44.7%) participated in the Emerging Scholars progam prior to the addition of the workshop component and 265 students (55.3%) participated for the first time with the workshop component.

This work was part of a long-term plan to better understand and improve our undergraduate research program. Subsequent to this study, the Undergraduate Research Committee also created a mentoring handbook (Mentoring Handbook, 2018) to support faculty mentors, and a Mentor Brochure (Undergraduate Research Mentor Brochure, 2019), highlighting the research interests of faculty, to help students identify and connect with a mentor.

Methodology

Data for all students participating in the Emerging Scholars program from its introduction in the fall of 2006 through spring of 2013 were used in this study (n=479). Data include background variables (e.g., race/ethnicity, sex/gender, and age), pre-participation academic variables (e.g., major, cumulative GPA, cumulative credits earned), and post participation variables (e.g., enrollment, semester GPA, semester credits, and graduation). Post participation data on enrollment, GPA, credits earned, and graduation were available through spring of 2015.

For analyses involving semester GPA and semester credits, mixed model for repeated measures (MMRM) analysis was used to look for overall differences as well as differential change over time. For graduation and persistence, a combination of statistical methods was used. Rates at particular points in time (e.g., one year after participation) were compared using chi-square tests. Survival analysis, specifically discrete-time hazard modeling (Singer & Willett, 2003) was used to analyze longitudinal data. Implications for possible eligibility criteria were explored using chi-square tests and t-tests.

Each student participating in the ES program was matched with two nonparticipating peers based on similarity in terms of degree-level being pursued (associate or bachelor's), major, admission date, cumulative credits and cumulative GPA at the start of the first ES semester, as well as HS GPA (as measured by the College Admissions Average), age, gender and ethnicity. It was often not possible to to find students who matched on all of the criteria simultaneously. Top priority was given to degree-level being pursued and major. Then the closest two matches were chosen based on a weighted average of the other criterion variables. By design, the two groups were identical on enrollment at the college the semester of the ES participant's first research experience (the matching semester), degree being pursued (associate, bachelor's, or nondegree), and academic major. A comparison of the ES participant group and the matched controls found no statistically significant differences on sociodemographic or academic characteristics (Tables 1 and 2).

Table 1. ES Participant and Matched Control Sample, Sociodemographic Characteristics

	ES Participants (n=479)	Matched Sample (n=958)
Age		
Mean	23.75	23.39
SD	6.63	6.01
Gender		
Female	46.8%	44.8%
Male	53.2%	55.2%
Ethnicity		
Black, Non Hispanic	30.1%	32.6%
White, Non-Hispanic	26.5%	21.4%
Hispanic, Other	21.7%	22.3%
Asian or Pacific islander	21.3%	23.6%
Other	0.4%	0.1%

Table 2. ES Participant and Matched Control Sample, Academic Characteristics

	ES Participants (n=479)	Matched Sample (n=958)
Standing		
Freshman	30.7	32.1
Sophomore	44.3	42.6
Junior	11.5	11.9
Senior	13.6	13.4
Cumulative GPA		
Mean	3.33	3.34
SD	0.49	0.48
Credits Earned		
Mean	51.07	48.66
SD	35.39	33.73

A mixed model for repeated measures (MMRM) was used to assess differences in semester GPA over time for ES participants and their matched peers. Variables in the model included 'group', which identified ES participants (Group=1) and matched non-participant peers (Group=0); time measured in semesters with 0 being the semester prior to ES participation (for the participant group as well as the matched non-participant peers); baseline GPA, which is the cumulative GPA prior to first semester of ES participation; and status (full-time or part-time). In addition, the interactions between time and group, and time and status were included in the model. Of specific interest were: 1) the effect for group which would indicate a difference in semester GPA for participants and non-participant peers during the ES semester; and 2) the interaction between time and group which would indicate different growth trajectories for participants and non-participant peers after the first ES participation semester. The results indicated that none of the effects relating to status (full-time or part-time) were significant (p-value range = .70-.93), and the model was rerun without status. This finding suggests that benefits associated with participation where similar for both part-time and full-time participants.

Participant Characteristics

Table 3 compares the representation in the ES participant group to the overall enrollment at our College. A chi-square test indicated that Hispanic students were under-represented in the ES program whereas white students were over-represented ($p<.0001$).

Table 3. Comparison of ES Participant Race/Ethnicity with the Overall College Population

	ES Participants	Overall Enrollment+
Hispanic	21.7%	33.5%
Black, Non-Hispanic	30.1%	32.0%
Asian or Pacific Islander	21.3%	19.9%
White, Non-Hispanic	26.5%	11.4%
Other	0.4%	0.5%

+Spring 2013 Enrollment

Students may participate in the ES program at any time during their undergraduate studies. The most common year to participate for the first time was as a sophomore. Documentation for the ES progam indicates that preference will be given to full-time students (Undergraduate Research, 2019) and this is reflected in the breakdown of the ES participants with only 19% being part-time students, which is considerably lower than the average percent of students in the college population who were part-time (36%). The breakdown of students according to degree goal was quite similar to the general college population (38% of participants were pursuing a bachelor's degree compared to 36.5% of students ihe overall population). These results are summarized in Table 4.

Table 4. ES Participant Academic Characteristics as of 1st Semester of Participation (n=479)

Academic Characteristics	Number of Students	Percent
Standing		
Freshman	147	31
Sophomore	212	44
Junior	55	11
Senior	66	14
Degree Goal		
Associate	282	59
Bachelor's	183	38
Non-Degree	15	3
STEM Major		
No	273	57
Yes	206	43
Status		
Full-time	390	81
Part-time	90	19

The average cumulative GPA prior to ES particiption was 3.33 (n = 450, SD =.50). Not all 479 ES participants had a GPA prior to participation. Twenty-nine students (6%) participated in ES their first semester at our college and did not have a City Tech GPA involving coursework taken prior to participation in the research experience. Of these, the majority were first time freshmen (n=16; 55%). The rest were transfer students (n=9; 31%), non-degree students (n=3; 10%), and a readmission student (n=1, 3%).

Evaluation for Establishing Eligibility Criteria

Baseline cumulative GPA and baseline cumulative credits earned were examined as predictors of improvement in semester GPA and semester credits earned during and after ES participation. Students with low semester GPAs (Semester GPA < 2.0) or no semester GPA during ES participation (n=28) were more likely to come into the research semester with no cumulative City Tech GPA on record (14.3% of low/no semester GPA students had no cumulative GPA as compared to 5.5% of students with a semester GPA of 2.0 or higher, p=.04). Students with low/no semester GPA during participation who did have a cumulative GPA tended to have a lower cumulative GPA than other ES students prior to participation (2.86 for low/no semester GPA as compared to 3.35 for students with a semester GPA of 2.0 or higher, p<.0001). In addition, these students were younger on average than other students (21.1 years old for low/no semester GPA during participation, as compared to 23.9 years old for students who earned a semester GPA of 2.0 or higher, p=.03) during participation.

In constrast to the low/no semester GPA results, students who earned a lower number of semester credits ($\leq$ 6) during ES participation tended to be older than other ES participants (26.6 years old compared to 23.4 years old, respectively, p=.009), have similar cumulative GPAs (3.29 as compared to 3.33, respectively, p=.63), and have higher cumulative credits earned (62.8 credits compared to 49.7 credits, respectively, p=.03), and be part-time.

To summarize, subpar semester GPA during ES participation was associated with younger students and lower or no City Tech cumulative GPA prior to participation. Thus if a goal is to develop additional eligibility requirements to promote positive outcomes for ES participants, requiring at least one semester at the college prior to participation in addition to the current minimum GPA of 2.0 is a possibility. The students with few semester credits earned during participation have similar GPAs and tend to be older. Thus taking fewer credits may just be a sign of stage of life rather than a negative outcome.

Evaluation of the Impact of the Introduction of Student Professional Development Workshops

Semester GPA and semester credits earned were compared for students first involved in the ES program prior to the workshop introduction (semesters prior to Fall of 2010) and after the addition of the workshops (semesters starting in Fall of 2010). Semester GPAs were very similar for students prior to and after the addition of the workshop component (p=.72; see Figure 1). However, as can be seen in Figure 2, semester credits earned differed for the two groups (p=.009). Although the graph hints at an interaction where non-workshop students had a steeper fall off in credits passed after ES participation, the change was not large enough to be statistically significant (p=.37). Thus it appears there may be additional factors not captured by the current data resulting in students in the more recent years passing additional credits prior to, during, and after ES participation. This may be related to the workshop component in terms of the students who chose to participate under the altered ES program, but the data does not support an inference that students are differentially affected during and after participation.

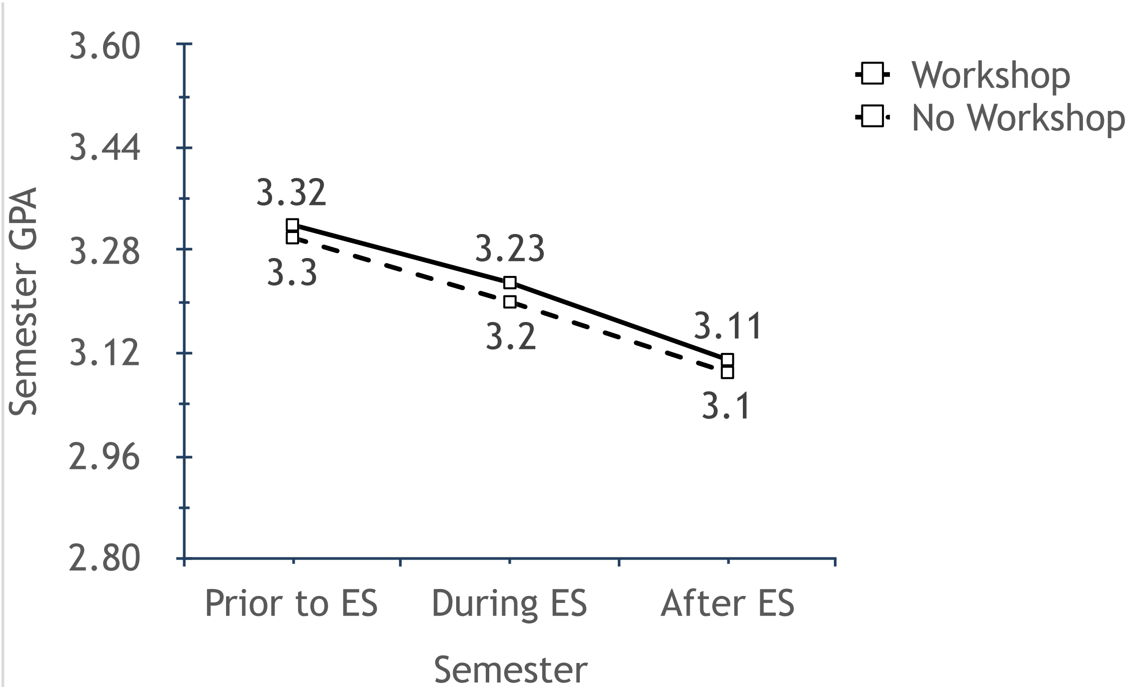


Figure 1. Semester GPA for workshop and non-workshop participants.

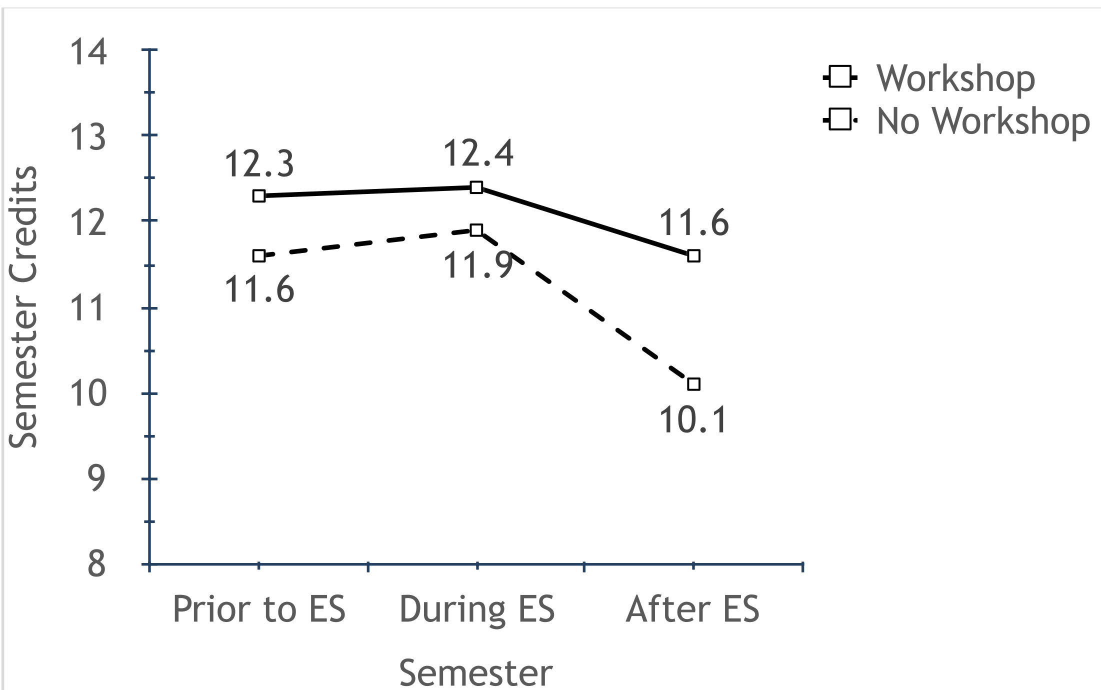


Figure 2. Semester credits earned for workshop and non-workshop participants

Another possible positive outcome of the workshop component may be increased likelihood of multiple semesters of participation in research. Whether or not a student participated in multiple semesters was compared for the non-workshop time period (fall 2006 – spring 2010) and the workshop time period (fall 2010 – fall 2012; spring 2013 was omitted from this analysis because it is the last semester of data on ES participation, and thus it is unknown whether or not these students participated in additional research semesters). More students participated in multiple semesters when their first experience involved the workshop than when it did not (32.8% and 23.8%, respectively; p=.045). Evaluation of other possible positive outcomes, such as developing a professional identify, were beyond the scope of this work

Evaluation of the Impact of the Length of Participation

As noted earlier, students are not limited to a single semester of partipation in the ES program and although most students participated only a single time (n=365), 24% (n=114) participated multiple semesters. Students who participated in the ES program once and students who participated multiple semesters were compared on semester GPA, semester credits earned, persistence, and graduation.

Semester GPA

The results indicated that none of the effects relating to status (full-time or part-time) were significant (p-value range = .75-.92), and the model was rerun without status. Neither multiple nor the multiple by time interactions were significant although multiple came close to statistical significance (p=.07 for multiple; p=.69 for the interaction). As can be seen in Figure 3 there was a relatively stable advantage on GPA for multiple semester participants. However, the difference did not exceed what could be explained by differences in prior GPA.

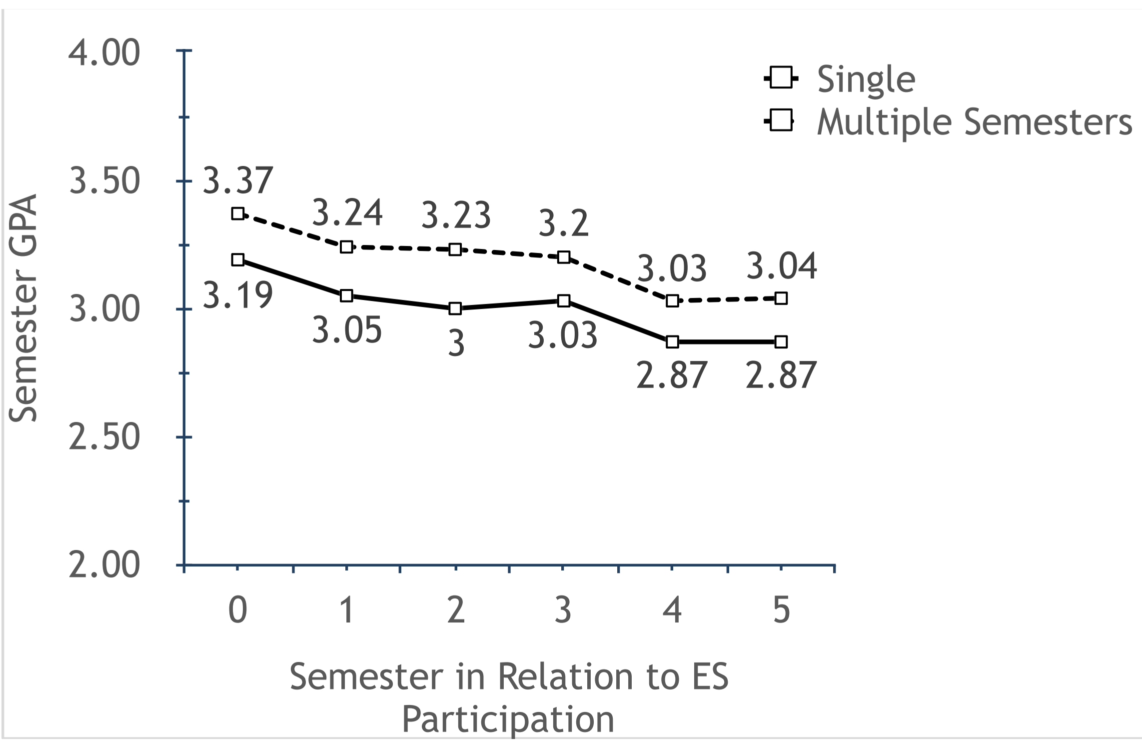


Figure 3. Semester GPA for Single- and Multiple Semester Research Participants

Semester Credits Earned

There were significant differences in the relationship between multiple semester participation and time depending on full-time or part-time status (p<.0001). For this reason the analysis was rerun separately for full- and part-time students.

For full-time students, there was a tendency for multiple semester participants to pass more credits per semester than single-semester ES participants (p=.0008). Although the difference between single semester participants and multiple-semester participants widens over time, the difference in the degree of decline for the two groups was not large enough to be statistically significant (p=.23). See Figure 4 for the trend in average semester credits earned for the two groups.

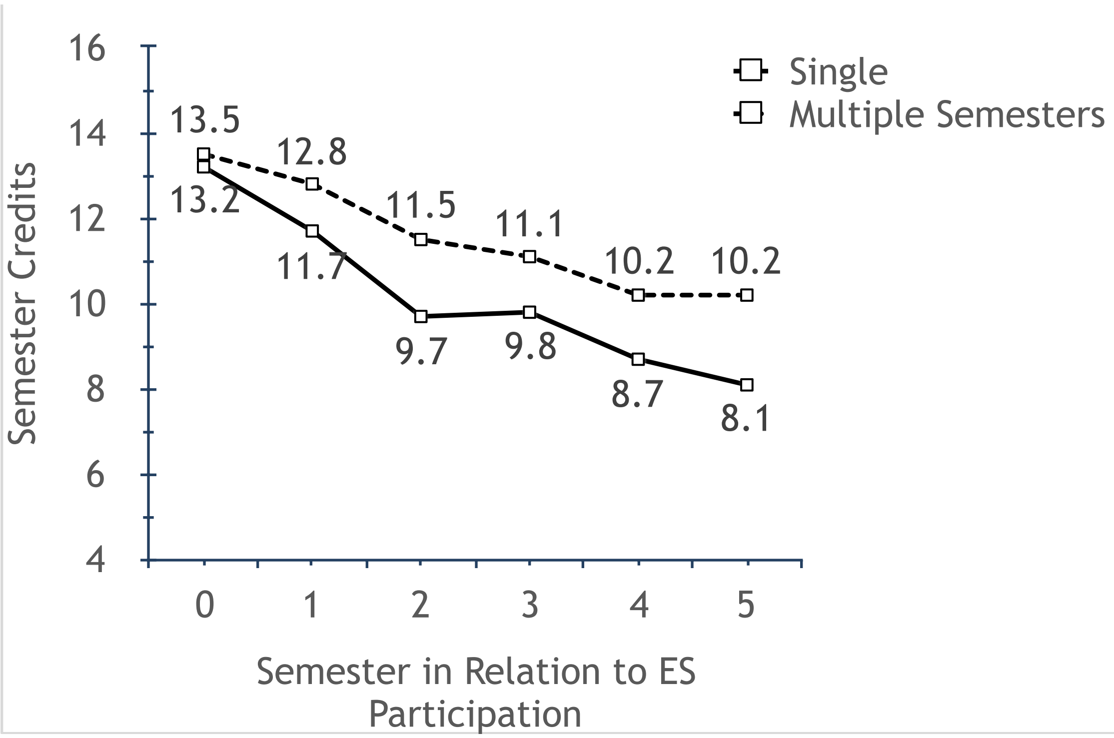


Figure 4. Semester Credits Earned for Full-Time Single- and Multiple-Semester Research Participants

For part-time students, multiple semester participants were not statistically different from single-semester participants in either average credits earned (p=.45) or in change over time in credits earned (p=.56). Thus for part-time students the determinants of semesters passed seem unrelated to whether or not the student participates in ES for a single semester or multiple semesters (see Figure 5). Thus, if resources are limited, limiting part-time students to a single semester of research, optimally promotes student success in terms of credits earned.

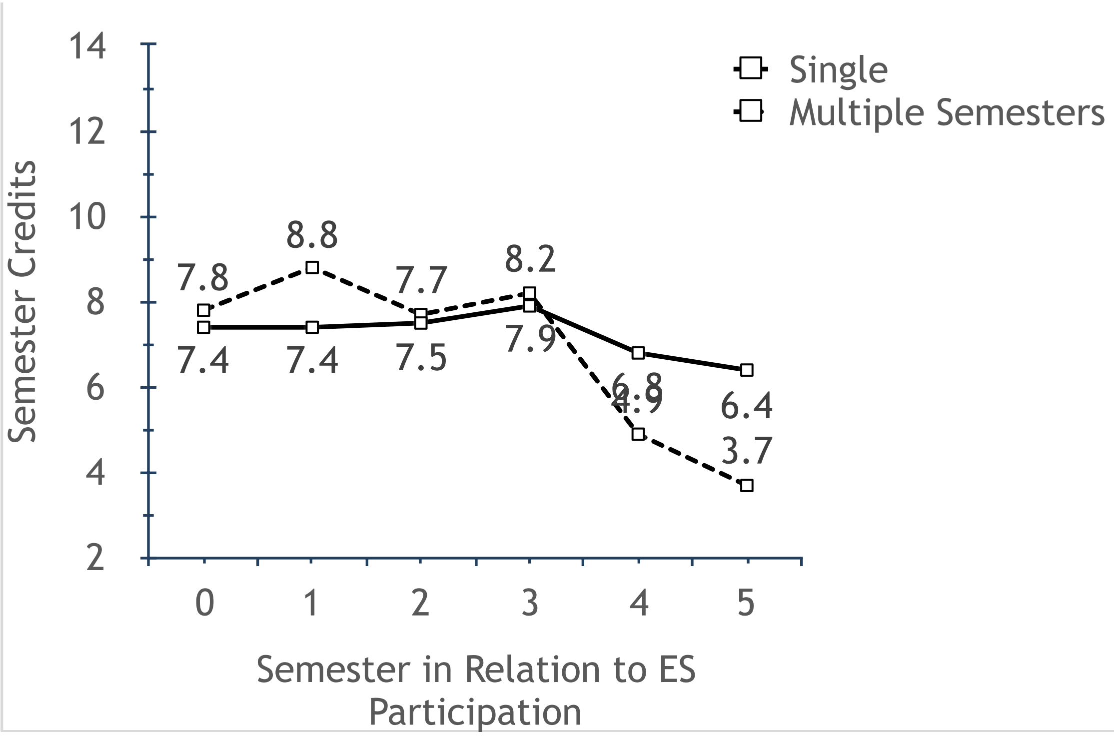


Figure 5. Semester Credits Earned for Part-Time Single- and Multiple-Semester Research Participants

Graduation

For this analysis only graduation involving the degree being pursued at the time of first participation in ES (or a higher degree) is counted as graduating. Also, spring 2013 participants were omitted from the analysis since whether or not they participated the following semester is unknown. Survival analysis was used to fit a discrete time hazard model to the data for single and multiple semester ES participants. Survival analysis indicated no significant difference between single semester participants and multiple semester participants in terms of time to graduate (p=.55). However, given that a student who is graduating soon has less opportunity for multiple semesters of participation, this analysis is difficult to interpret.

Overall, the graduation rate for the single semester participants (49.3%; n = 144) was significantly lower than the graduation rate for the multiple semester participants (60.5%; n = 69; p=.042) (see Figure 6). When these results were broken out by degree goal, the significance appeared to be driven by students pursuing associate degrees. The overall graduation rate for single semester ES participants pursuing an associate degree was 43.9% whereas the multiple semester students pursuing an associate degree had a graduation rate of 57.5%, p=.047). For bachelor's degree ES participants 60.6% of single semester particpants graduated whereas 65.9% of multiple semester participants graduated (p=.57).

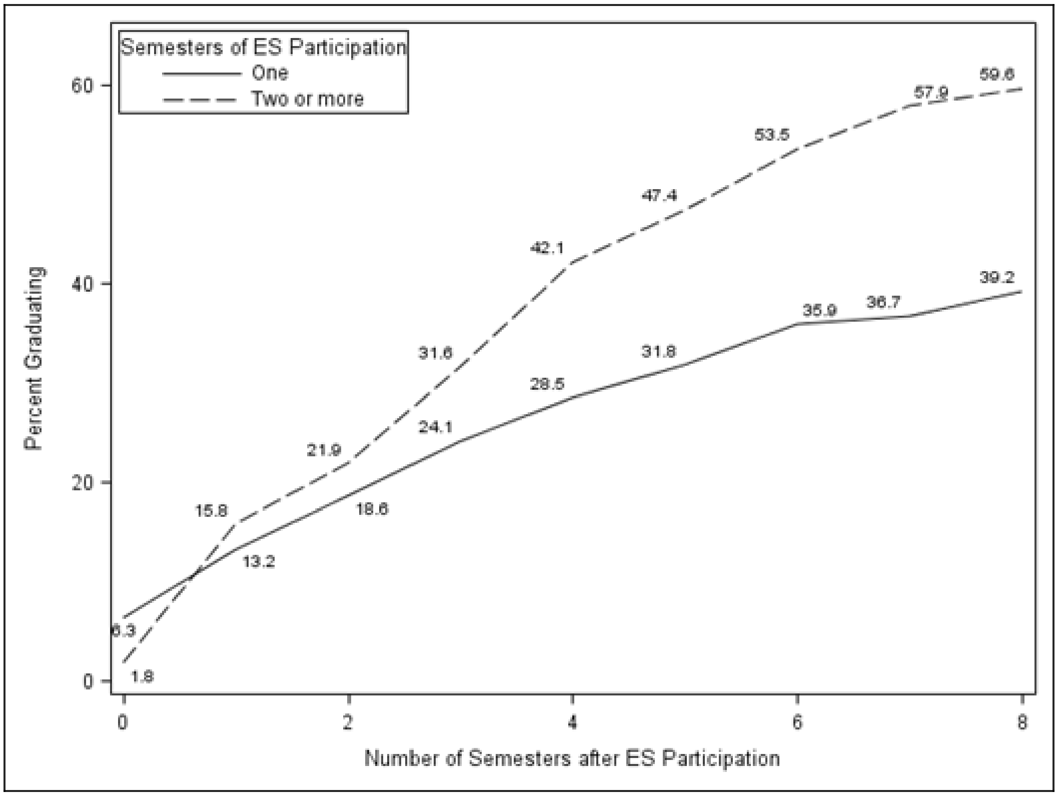


Figure 6. Graduation rate for single- and multiple-semester research participants.

Persistence

For students who do not graduate, it was of interest to look for difference in persistence between single- and multiple-semester participants. Multiple semester participants have higher persistence rates (see Table 5), but interpreting the meaning of the result is complicated. The only way for a student to participate in multiple semesters is if the students is enrolled at our college in multiple semesters, thus the persistence rate for students participating in multiple semesters of research is very high especially at 1 semester after the first ES participation. On the other hand, a single semester participant may not participate in research again because s/he does not wish to, or because they are not enrolled and thus don't have the opportunity. Whether persistence is responsible for multiple ES participation, or multiple ES participation is responsible for persistence, is unclear.

Table 5. Persistence of Single-semester and Multiple-semester ES Participants – Percent and Frequency

	Single Semester	Multiple Semesters
1 Semester*	69.4% (n=186)	93.7% (n=105)
1 Year*	52.9% (n=129)	79.2% (n=76)
2 Years**	46.7% (n=64)	59.1% (n=39)

*p<.0001
**p=.099

Comparison of Undergraduate Researchers and a Matched Comparison Group

Both group and the group by time interaction were significant ($p=.03$ and $p<.0001$, respectively). ES participants have higher GPAs during the ES semester (effect= .18 on the GPA scale), however over time the difference between the semester GPA for participants and non-participants narrows until at 4 semesters after participation the groups are not statistically different (see Figure 7). It is important to note the groups are matched on prior GPA and furthermore prior GPA was included in the model as a control variable. Therefore, participation is associated with higher GPAs in the semester of ES participation and in the several semesters after participation, and this effect is not explained by pre-existing GPA differences.

It is also important to note that the population represented in the analysis changes with time. All students in the ES participant group and the matched controls are present at time 0, however as time increases student are lost to the analysis through graduation or non-enrollment. Thus, the farther out on the time axis, the lower the sample sizes and the less the sample represents the original time 0 sample. Thus, the results out past 3 or 4 semesters should be interpreted with caution.

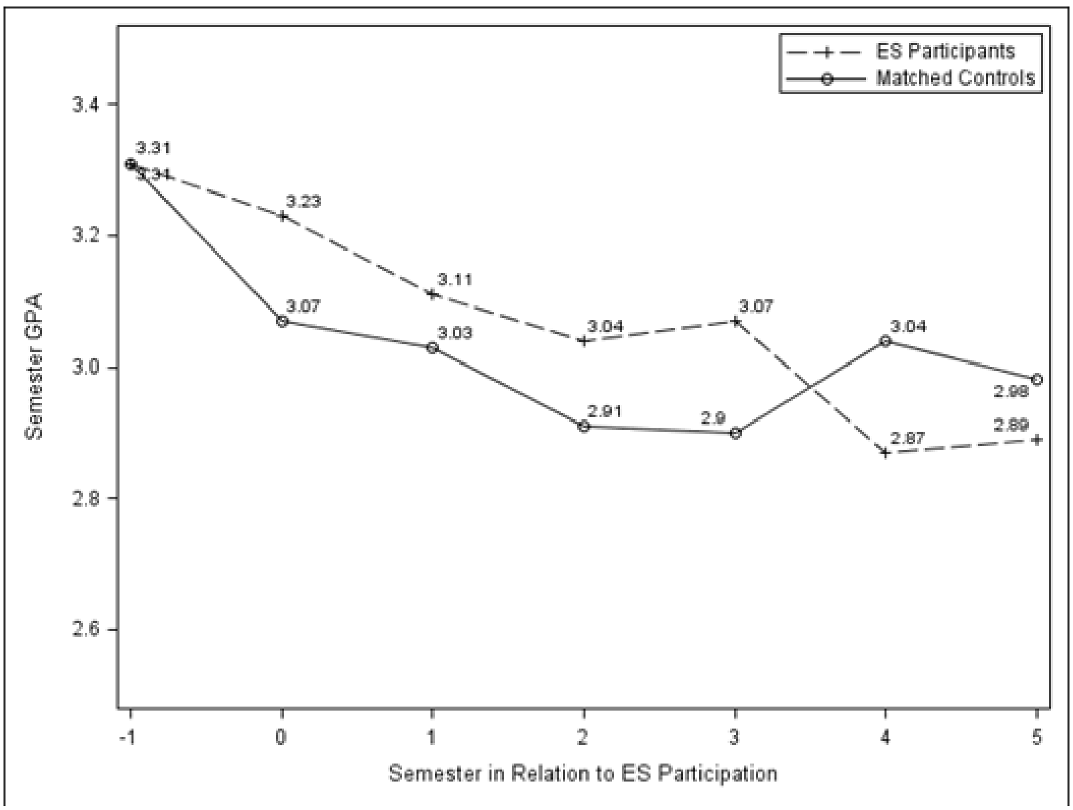


Figure 7. Semester GPA by group.

Credits Earned

The same analysis was conducted comparing semester credits earned for ES participants and their matched peers. In this analysis, whether or not a student was full-time had a significant effect on the relationship between time and group (Group x Time Interaction $p<.0001$). For this reason, the analysis was repeated separating the full-time and part-time students. For full-time students, group was significant (p=.05) but not the group by time interaction. On average, full-time ES participants earned more credits per semester than their matched non-ES peers (effect = 0.44 credits). Examining Figure 8, the difference between groups appears large in the ES semester and the semester after ES participation, however the fluctuation in the difference from one semester to another was not large enough to rule out it being due to random chance (p=.22). Once again, it is important to note the groups were matched on prior semester credits earned, and prior credits earned is included in the model as a control variable. Therefore, for full-time students, participation is associated with higher credits earned that is not due to pre-existing group differences. Again, as time increases the samples for both groups get smaller and we can have less confidence in the results being representative of the original groups.

For part-time students, group was not significant (p=.91) but there was a group by time interaction (p=.023). In the semester before, during, and after ES participation, the ES participants earned more credits than their matched peers (see Figure 9). However, at two, three and five semesters after ES participation there were not significant differences between the groups (p=.92, .998, and .40, respectively), and at four semesters after participation, the effect was reversed (p=.02). Once again, it is important to note the groups were matched on prior semester credits earned and prior credits earned is included in the model as a control variable. Therefore, for part-time students, participation is associated with higher credits earned in the semesters near ES participation that is not due to preexisting group differences in credit accumulation.

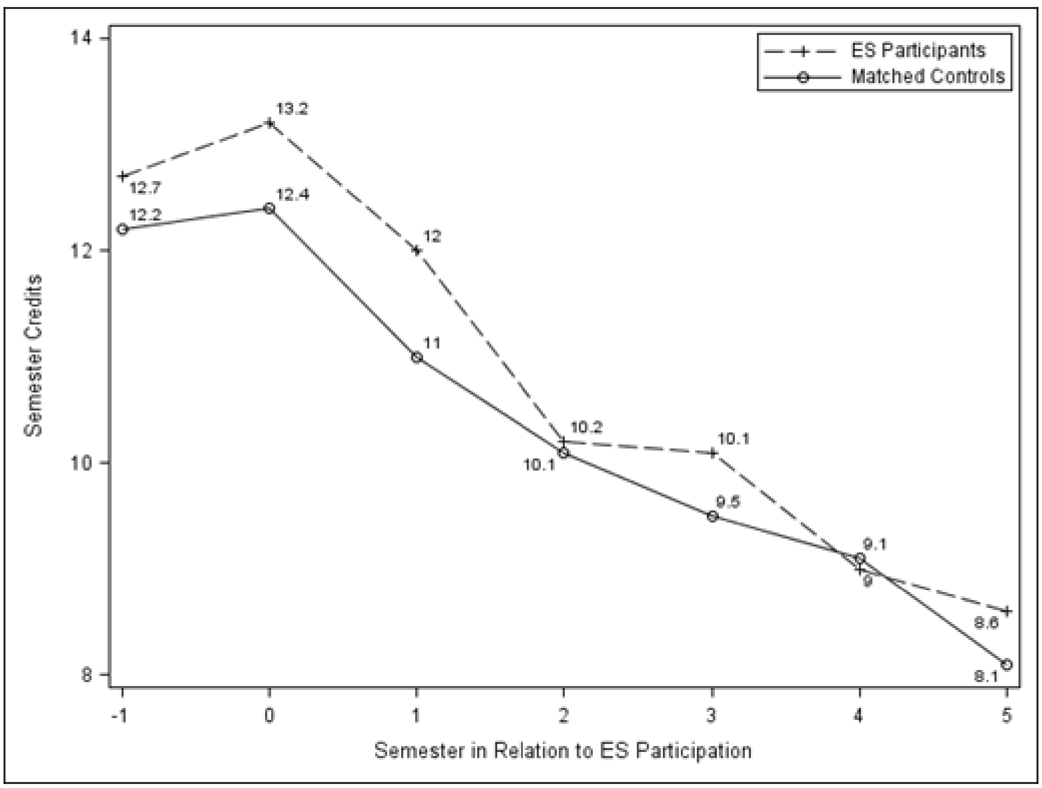


Figure 8 Semester credits passed by group for full-time students.

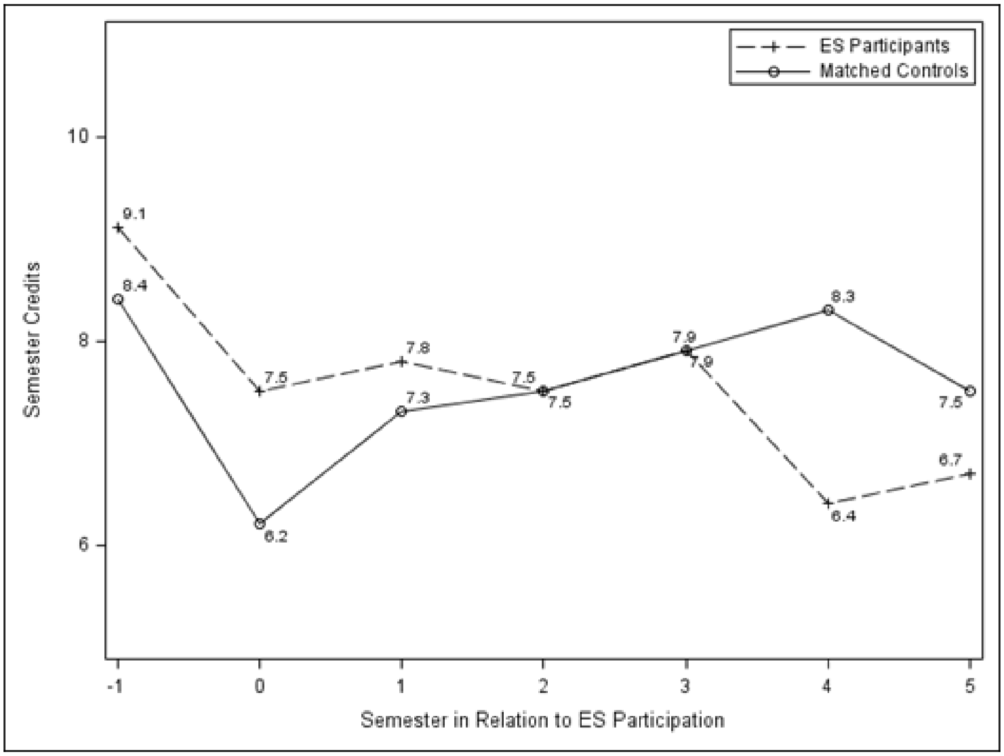


Figure 9. Semester credits passed by group for part-time students.

Graduation

Each semester after the first participation in the ES program, students are classified as 1) continuing; 2) graduated, or 3) not enrolled (and not graduated). Each consecutive semester the student continues to take classes is counted and used to measure persistence. The number of semesters until graduation are counted (including any semesters in which the student took no classes) to measure time until graduation.

Graduation is measured in two ways. The first method counts only graduation involving the degree being pursued at the time of first participation in ES (or a higher degree). Thus, a student pursuing a bachelor's degree who obtains an associate degree is not counted as graduated unless or until the student earns the bachelor's degree. However, a student pursuing an associate degree who earns a bachelor's degree is considered as graduating since a bachelor's degree is a higher-level degree than an associate degree. The second way of measuring graduation counts any degree as successful graduation. The results of both methods will be summarized.

Method 1 (Graduation with degree being pursued)

Overall, the graduation rate for the ES participants (50.9%; n = 244) was significantly higher than the graduation rate for the matched sample (44.8%; n = 429; p<.0001). When examined separately for associate degree students and bachelor's degree students, both groups were more likely to graduate if they participated in ES research (Odds-Ratio (OR) for associate degree students 1.24, p = .002; OR for bachelor's degree students 1.25, p=.01). See Figure 10.

Method 2 (Graduation with any degree)

Of the ES participants, 22 (4.6%) received an associate degree while working on a bachelor's degree. Of these students, 12 (55%) had not received a bachelor's degree as of spring 2015. Thus, the analysis counting all degrees as successful graduation has a slightly higher graduation rate for ES participants (53.4% (n=256) compared to 50.9%). Similar results were found for the matched control sample – 41 students (4.3%) received an associate degree while pursuing a bachelor's degree, and of these, 18 (43.9%) had not received a bachelor's degree as of spring 2015. The overall graduation rate for the matched sample went up to 46.6% (n=446) from 43.1% when all degrees were counted as successful outcomes. In summary, both graduation rates went up slightly when intermediate degrees are counted, but the gap between the groups remained (p=.016).

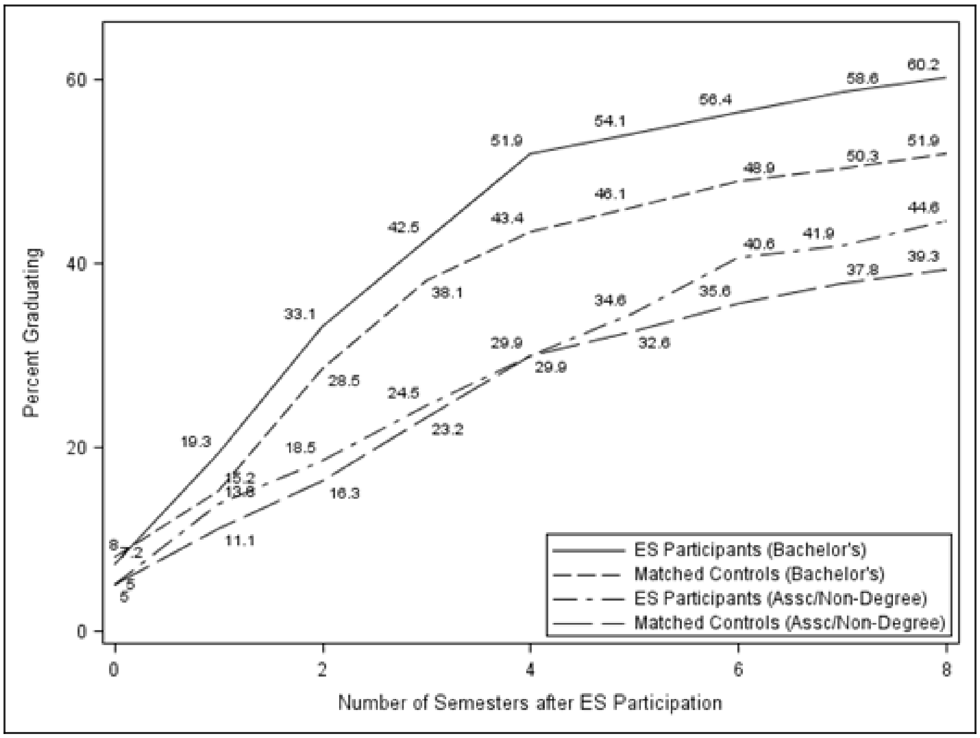


Figure 10. Graduation rate for associate and bachelor's students – ES participants and matched controls.

Persistence

When comparing persistence rates for the ES participants and the matched controls, differences were not statistically significant at any of the time periods examined (1 semester after ES semester (p=.80), 1 year after ES semester (p=.46), and 2 years after ES semester (p=.66)) (See Table 6 for details). Note, persistence rates are calculated without counting graduates as part of the numerator or denominator.

Table 6. Persistence of Emerging Scholars Participants and Matched Case Controls – Percent and Frequency

	ES Participants	Matched Controls
1 Semester	63.3% (n=285)	62.6% (n=563)
1 Year	51.4% (n=207)	49.1% (n=411)
2 Years	41.0% (n=98)	39.3% (n=178)

Is there evidence to guide eligibility criteria?

It appears that there may be benefits for students participating in undergraduate research based on our data demonstrating that ES students have higher semester GPA, earn more credits per semester, and graduate at higher rates than their matched non-participant peers. It was of interest to explore whether these potential benefits are larger for some groups of students than for others. This analysis differs from the earlier analysis due to the comparison with the matched sample. Thus, we are able to discern if the outcomes for ES participants are more positive than the outcomes for their matched peers for specific subsets of students like STEM majors and students earlier in their college education. The MMRM analysis conducted above was repeated with STEM classification added to the model. The interaction between STEM classification and group (ES participants or matched peers) was not significant, indicating that there was not a statistically reliable differential benefit for ES participation for STEM students compared to non-STEM students (p=.30 for semester GPA; p=.68 for semester credits).

Recall, that based on the earlier ES analysis, no City Tech cumulative GPA coming into the ES participation semester (ie freshmen and transfer students) was related to poorer outcomes than ES participants with a City Tech GPA. MMRM was used to test for differences in outcomes for these students (n=29) and their matched peers (n=58). In terms of semester GPA, ES participants had significantly higher GPA during the ES participation semester (p=.015) and in the two semesters following participation (p=.0003; p=.014). In subsequent semesters, the GPAs were not significantly different (p-value range = .07-.98).

The results for semester credits were similar. No prior-GPA ES participants earned significantly more credits during the ES participation semester (p<.0001) and in the two semesters following participation (p<.0001; p=.026) than their matched peers (see Table 7). In subsequent semesters, credits earned were not significantly different (p-value range = .22-.87) for the two groups. Graduation rates for the no prior GPA ES participants and their matched peers in this at-risk subgroup were low (16.7% for ES participants and 13.3% for matched peers). However, these statistics include the 25% of students from this group who were still enrolled in the spring of 2015 and thus we do not know their ultimate graduation status.

Table 7. Comparison of No Prior GPA Emerging Scholars Participants and their Matched Case Controls during 1st ES Participation Semester and in the Following Two Semesters

		Semester GPA				Credits Passed			
No Prior GPA	**n**	**ES Semester**	**+1**	**+2**	**Sig***	**ES Semester**	**+1**	**+2**	**Sig***
ES Participant	29	2.99	2.93	2.76	a	11.77	11.13	8.88	a
Matched Peer	58	2.45	2.03	2.09	b	7.63	6.62	6.32	b

* Groups with different letters are significantly different at $\alpha = .05$.

Persistence rates for the two groups were not significantly different although there was a trend for the no prior GPA matched peers to have higher persistence rates at 1 year and 2 years after the ES participation semester (see Table 8).

Table 8. Persistence No Prior GPA Emerging Scholars Participants and their Matched Case Controls - Percent and Frequency

Low/No Prior GPA	ES Participants	Matched Controls
1 Semester	62.1% (n=18)	65.0% (n=39)
1 Year	43.3% (n=13)	53.3% (n=32)
2 Years	42.9% (n=9)	58.5% (n=24)

Note: Graduates do not count in the numerator or denominator when calculating persistence rate.

Conclusions

This work presents a model for evaluation of the impact of an undergraduate research program that yields evidence of value and eligibility criteria and the impact of the introduction of a new component, professional development workshops. Evaluation of this data suggested important directions:

1. While African Americans participate at levels close to their representation in the student population (30.1% and 32.0%, respectively), bucking the trend that underrepresented minority students are less likely to participate than other groups (Finley and McNair, 2013), the underrepresentation of Hispanic students suggests that more intentional recruitment is needed.

2. Subpar semester GPA during undergraduate research participation (GPA< 2.00) was associated with younger students and lower prior cumulative GPA's or no City Tech cumulative GPA. However, when compared to their matched peers the students participating in research had significantly higher GPAs during the participation semester and in the two semesters following participation. The results for semester credits earned were similar. No prior-GPA research participants earned significantly more credits during the participation semester and in the two semesters following participation than their matched peers. Thus, if a goal is to conserve resources for those most likely to benefit, requiring at least one semester at the college prior to participation would be an evidence-based criteria. If resources are available even students with low/no GPAs are likely to benefit.

3. The addition of student professional development workshops correlated with multiple semester participation in the ES program, although it did not have a statistically significant impact on GPA or credit accumulation.

4. For full-time students, there was a tendency for multiple semester participants to earn more credits per semester than single-semester researchers. For part-time students, multiple semester participants were not statistically different from single-semester participants in either average credits earned in the research semester or over time. Thus, while there is no supporting evidence that full-time participation should be limited to one semester to conserve resources, no evidence of the benefits of multiple semester benefits for part-time students was found. Thus if resources are limited, limiting part-time students to one semester of participation may be a reasonable approach.

5. For full-time students, researchers earned more credits during and one semester after participation than their matched peers. For part-time students, ES participation was associated with higher credits earned in the semesters near ES participation than their matched peers, which was not due to preexisting group differences in credit accumulation. This points to the value of offering undergraduate research opportunities to part-time students.

6. The graduation rate for undergraduate research participants was significantly higher than that of the matched sample. This is evidence for sustaining an undergraduate research program.

7. The academic outcomes were similar for students majoring in STEM compared to non-STEM majors. This suggests that the benefits of ES participation are not isolated to a specific area of study but rather provide positive experiences for the larger undergraduate community, inclusive of a variety of major areas of study. This finding is in agreement with other studies (Healy and Jenkins, 2009; Ishiyama, 2002). Previous work has also shown that there are fewer opportunities for undergraduate research in the social sciences and humanities than in the natural sciences (Seymour *et al*, 2004).

Limitations

The primary limitation of this study is the observational nature of the data. Because students choose to be in the ES program and, in fact, put forth effort to get into the program, there are most likely substantial differences in the personal characteristics of ES participants and non-participants. A matched sample was drawn in an attempt to control as many of these personal characteristics as possible, but it is important to keep in mind that this is not a complete solution to the problem. There are likely unmeasured personal factors that are still influencing the results for the two groups and it is important not to over-interpret the results. This is true of any observational study, even though it is sometimes not acknowledged in research reports. Thus, the results of this study are suggestive of factors that contribute to student success, but they do not prove that it is the ES program responsible for higher success rates.

Acknowledgements

We gratefully acknowledge the support of the City University of New York Office of Academic Affairs, whose generous funding made this work possible.

References

College Board (2014). Trends in College Pricing 2013. https://trends.collegeboard.org/sites/default/files/college-pricing-2013-full-report.pdf (accessed August 2015).

Davis, D. I. (2009). The Academic Influence of Mentoring Upon African American Undergraduate Aspirants to the Professorate. *Urban Review 42*(2), 143-158.

EAB (2019), Why part-time student success may be the key to education equity. EAB Community College Blog. Retrieved September 16, 2019, from https://eab.com/insights/blogs/community-college/why-part-time-student-success-may-be-the-key-to-education-equity/

Espinosa, L.I. (2011). Pipelines and Pathways: Women of Color in Undergraduate STEM Majors and the College Experiences that Contribute to Persistence. *Harvard Educational Review 81*(2), 209-240.

Finley, A.; McNair, T. (2013). Assessing Underserved Students' Engagement in High-impact Practices**,** Association of American Colleges and Universities, Washington, DC.

Gregerman, S.R.; Lerner, J.S.; von Hippel, W.; Joinides, J.; Biren, A.N. (1998). Undergraduate Student-Faculty Research Partnerships Affect Student Retention. *Review of Higher Education 22*(1), 55-72.

Healey, M.; Jenkins, A. (2009). Developing Undergraduate Research and Inquiry. York: Higher Education Academy. Retrieved on September 16, 2019, from https://www.heacademy.ac.uk/knowledge-hub/developing-undergraduate-research-and-inquiry

Hu, S.; Kuh, G.D.; Li, S. (2008). The Effects of Engagement in Inquiry – Oriented Activities on Student Learning and Development. *Innovative Higher Education 33*(2), 71-81.

Ishiyama JT, Hopkins VM (2002). Assessing the impact of a graduate-school preparation program on first generation, low income college students at a public liberal arts university. Journal of College Student Retention 4: 393-405.

Kuh, G.D. (2001). Assessing what really matters to student learning: Inside the National Survey of Student Engagement. *Change, 33*(3), 10-17, 66.

Kuh, G. D. (2003). What we're learning about student engagement form NSSE. *Change, 35*(2), 24-32.

Laursen, SL; Hunter, A-B; Seymour, E.;Thiry, H.;Melton, G.(2010) Undergraduate Research in the Sciences: Engaging Students in Real Science, San Francisco: Josey-Bass.

Lopatto D (2006). Undergraduate research as a catalyst for liberal learning. *Peer Rev*, 8, 22-25

Johnson. R.C. (2011). Using Summer Research to Attract Pre-College Undergraduate Students to STEM Fields. *CUR Quarterly 31*(3), 7-15.

Mentoring Handbook (2018) New York City College of Technology Undergraduate Research Committee. Retrieved on September 16, 2019, from https://openlab.citytech.cuny.edu/undergraduateresearch/mentoring-handbook-download/

Myers, J., Sawyer, A., Dredger, K., Barnes, S., & Wilson, R. (2018). Making Time and Creating Space for Undergraduate Research. *Journal of the Scholarship of Teaching and Learning*, *18*(1), 136-149. Retrieved on September 16, 2019, from https://openlab.citytech.cuny.edu/undergraduateresearch/mentoring-handbook-download/https://doi.org/https://doi.org/10.14434/josotl.v18i1.22348

Schultz. P. W.; Hernandez, P.R.; Woodcock, A.; Estrada, M.; Chance, R.C.; Aguila, M.; Serpe, R.T. (2011). Patching the Pipeline: Reducing Educational Disparities in the Sciences through Minority Training Programs." *Educational Evaluation and Policy Analysis 33*(1), 95-114.

Seymour, E., Hunter, A. B., Laursen, S. L., & DeAntoni, T. (2004). Establishing the benefits of research experiences for undergraduates in sciences: First findings from a three-year study. Science Education, 88, 493-594. Retrieved on September 16, 2019, from https://openlab.citytech.cuny.edu/undergraduateresearch/mentoring-handbook-download/ http://dx.doi.org/10.1002/sce.10131

Undergraduate Research (2019). Retrieved on September 16, 2019, from http://www.citytech.cuny.edu/research/

Undergraduate Research Mentor Brochure (2019). Retrieved on September 16, 2019, from https://openlab.citytech.cuny.edu/undergraduateresearch/mentoring-handbook-download/https://openlab.citytech.cuny.edu/undergraduateresearch/find-a-mentor/

Journal of the Scholarship of Teaching and Learning, Vol. 20, No. 1, April 2020, pp.105-129.
doi: 10.14434/josotl.v20i1.24825

Exploring Organic Chemistry I Students' Responses to an Exam Wrappers Intervention

Evan Grandoit
Northwestern University

Rose Bergdoll
Brooklyn College

Edith Rosales
Brooklyn College

David Turbeville
Brooklyn College

Shoshana Mayer
Brooklyn College

Gail Horowitz
Brooklyn College

Abstract: Research has demonstrated that academically successful students are effective, self-regulated learners. Moreover, exam wrapper interventions have been shown to foster the development of self-regulated learning behaviors on the part of college students. In this naturalistic, qualitative, and exploratory study, an exam wrapper intervention was implemented in a key, gatekeeping STEM course at a diverse, public university. Student responses to a series of four exam wrappers were collected and analyzed. Results indicated that while many students were able to look critically at their study behaviors and course performance, these behaviors did not necessarily pay off, especially for weaker students. Notably, transfer and/or non-matriculated students were at greatest risk of withdrawal and failure. However, all students, both weak and strong, showed a lack of attention towards checking their answers and learning from their mistakes. Overall, the exam wrappers provided useful information regarding the self-regulated learning processes of these STEM students.

Keywords: Organic Chemistry, Self-Regulated Learning, STEM, Underrepresented Minorities

Introduction

The loss and attrition of qualified undergraduates from STEM majors is no longer an unfamiliar phenomenon. Unfortunately, it has become a well-known area of study, research, and exploration (Chang, Sharkness, Hurtado, & Newman, 2014; Hunter, 2016; Malcom & Feder, 2016; Seymour & Hewitt, 1997). Research indicates that among the most recent generation of college students, more than half of all students who enter college intending to major in STEM leave STEM or fail to graduate altogether (Eagan, Hurtado, Figueroa, & Hughes, 2014). For certain ethnic minority groups, attrition rates are even higher. Eagan and colleagues (2014) report that 76% of UREM (underrepresented minorities) who enter planning to major in STEM do not complete their degrees within six years.

The loss of women and UREM from STEM majors results in the underrepresentation of these populations in STEM fields (for example in chemical industry, biological research, engineering, etc.) (National Science Foundation, 2013), and also leads to their underrepresentation in vital health science fields such as Medicine and Dentistry (Association of American Medical Colleges, 2010; Smedley, Butler, & Bristow, 2004). Significantly, researchers have found that Organic Chemistry is a key "gatekeeping" course that deters students from moving forward in careers such as medicine (Barr, Matsui, Wanat, & Gonzalez, 2010).

Research indicates that problems with undergraduate science teaching, especially in introductory, so-called gatekeeping courses, are among the primary reasons why students switch out of STEM (Seymour & Hewitt, 1997). Problems cited include large lectures, where teaching is impersonal and primarily in a transmissive mode (Gasiewski, Eagan, Garcia, Hurtado, & Chang, 2012; Hunter, 2016), classroom climates that are competitive, chilly, and unwelcoming (Gasiewski et al., 2012; Rosser, 1997) and course content that is divorced from everyday life (Hunter, 2016).

Reform efforts such as Process Oriented Guided-Inquiry Learning (Moog & Farrell, 2006), Peer-Led Team Learning (Gosser, 2011), Team Based Learning (Michaelsen, Knight, & Fink, 2004), Flipped Classrooms (Talbert, 2017), and Peer Instruction (Watkins & Mazur, 2013) all attempt to create active, student centered learning environments where student cooperation is encouraged. These reforms have shown excellent promise in improving retention rates for all students, including women and UREM (Deslauriers, Schelew, & Wieman, 2011; Eberlein et al., 2008; Freeman et al., 2014; Hall, Curtin-Soydan, & Canelas, 2014; Lewis, 2011; Mooring, Mitchell, & Burrows, 2016). Unfortunately, their implementation has not been widespread for a number of reasons, including cost (e.g. for revamping classrooms, hiring additional instructors, professional development, release time, etc.), ideological resistance (to change in general and/or to student centered pedagogies in particular) (Burd et al, 2016; Chaskes & Anttonen, 2015; Kezar & Holcombe, 2016), deficit views of underrepresented groups (Castro, 2014), and/or simple inertia (Wieman, 2016).

Recently, researchers have begun to focus their attention on first-generation college students, another underrepresented population of students (Engle & Tinto, 2008; Jehangir, 2010; Stephens, Hamedani, & Destin, 2014; Stuber, 2012; Terenzini, Springer, Yaeger, Pascarella, & Nora, 1995; Uche, 2015). First-generation students are defined as either the first-generation in their family to earn a Bachelor's degree or as the first-generation in their family to enroll in college (Davis, 2010). National data show that first-generation college students are disproportionally from groups underrepresented in STEM, including women and UREMs, but are also disproportionally low-income students, non-traditional college age students and students with dependents (Engle & Tinto, 2008; Lohfink & Paulsen, 2005; Terenzini et al., 1995). A study by Saenz, Hurtado, Barreta, Wolf, & Yeung (2007) found that 87% of first-generation college students were students of color.

Disturbingly, data demonstrate that retention rates for first-generation college students in STEM are lower than those of continuing generation college students (Chen, 2013; Doerschuk et al., 2016) especially in key gatekeeping courses like General Biology and General Chemistry (Eddy & Hogan, 2014; Gregg-Jolly et al., 2016; Harackiewicz et al., 2014). Auspiciously however, interventions involving student centered pedagogies have shown excellent promise in combating these trends (Crimmins, 2017; Eddy & Hogan, 2014; Haak, HilleRisLambers, Pitre, & Freeman, 2011).

Background

Self-regulated learning (SRL) has been described as the regulatory processes that learners engage in while monitoring their own learning. These regulatory processes include: setting goals, selecting strategies and behaviors, monitoring one's performance, and adapting in response to feedback

(Zimmerman, 1990). Students who are more effective self-regulated learners are more successful in general (Plant, Ericsson, Hill, & Asberg, 2005; B. Zimmerman & Martinez-Pons, 1990), and in STEM in particular (Karabenick, 2003; Nandagopal & Ericsson, 2012; Szu et al., 2011; VanderStoep, Pintrich, & Fagerlin, 1996). Notably, research demonstrates that one of the key reasons why students struggle in key gatekeeping STEM courses is due to a lack of sophisticated study skills (Bunce et al., 2017; DiBenedetto & Bembenutty, 2013; Sebesta & Bray Speth, 2017).

Nonetheless, self-regulation is a learnable and trainable skill (Phillips, Clemmer, McCallum, & Zachariah, 2017; Weinstein & Acee, 2013). Discipline-specific training in metacognition and/or self-regulation has been shown to be effective in helping students improve in STEM subjects such as Chemistry (Zhao, Wardeska, McGuire, & Cook, 2014), Mathematics (Hudesman et al., 2014; Olszewski, 2016), and Biology (Azevedo & Cromley, 2004; Bernacki, Vosicka & Utz & 2017).

Exam wrappers (designed by Dr. Marsha Lovett of Carnegie Mellon University) are a specific pedagogical tool designed to encourage students to cultivate and improve their self-regulated learning skills (Lovett, 2013). Exam wrappers are typically written exercises that students complete (inside or outside of class) after receiving back a graded exam. These exercises encourage students to monitor their performance and to set goals and modify their behaviors because they invite students to reflect on their study behaviors and exam performance, to consider what strategies were or were not successful, and to state what behaviors they might want to initiate in preparing for the next exam (Lovett, 2013). Exam wrappers are sometimes also administered in a pre-test fashion to encourage students to reflect as they prepare for an upcoming exam. Students may be asked to assess their level of preparedness, to analyze the study strategies they have been employing and/or to consider making changes to their study strategies (Lovett, 2013).

Exam wrappers have been used in a variety of classroom settings (e.g. Achacoso, 2005; Soicher & Gurung, 2017; Thompson, 2012), but their use has been most frequently reported in STEM classrooms. Experimental studies have shown statistically significant improvements in student performance when exam wrapper assignments have been incorporated into Statistics and Engineering classrooms (Chen, Chaves, Ong, & Gunderson, 2017; Chew, Chen, Rieken, Turpin, & Sheppard, 2016). Similarly, studies in Biology and Physics classrooms have shown that when students are incentivized to correct their mistakes on exams these students show statistically significant improvements in their course performance (Brown, Mason, & Singh, 2016; Rozell, Johnson, Sexten, & Rhodes, 2017). Correlational or descriptive studies involving exam wrappers in Biology classrooms have shown weaker students have more limited SRL strategies and that these students are less able to implement changes in their SRL strategies (Stanton, Neider, Gallegos, & Clark, 2015).

In the naturalistic, qualitative study described below, exam wrappers were utilized in an exploratory fashion to investigate whether, with minimal intrusion onto the flow of a fairly traditional Organic Chemistry I lecture course, exam wrappers could encourage students to reflect upon their own learning and encourage them to make strategic and effective changes in study habits and behaviors. (We also conducted a separate, quasi-experimental intervention study utilizing exam wrappers in a different Organic Chemistry I course and hope to report on this study in the near future.)

Organic Chemistry I was chosen as the subject area for this study, in part, because of the crucial gatekeeping function that it plays for many STEM and pre-health science students. (It is also a course that has particularly low success rates at our institution.) Additionally, prior research at our institution (AUTHOR AND COLLEAGUES, 2013) had shown that students who were at risk for not passing Organic Chemistry I could overcome the odds and pass if they engaged in a particular self-regulated learning behavior: help-seeking -- specifically the making use of resources such as office hours and supplementary problem sessions. We hypothesized therefore, that if Organic

Chemistry students at our institution, at which there are large numbers of first-generation college students (2014 UNIVERSITY NAME STUDENT EXPERIENCE SURVEY), could be encouraged to be more self-reflective about their study habits, they might improve in their self-regulated learning abilities, modify their study behaviors and improve in their overall course performance. Thus, we chose to introduce exam wrappers into the Organic Chemistry I course because of their potential to impact the reflection and self-regulated learning behaviors of our students. We also chose exam wrappers as our pedagogical tool because they are a reform method that is fairly conventional, non-controversial, and easy to implement.

Method

Participants and Course Context

This IRB-approved study was conducted at a large, urban, public university located in the Northeastern United States. The undergraduate population at this institution is highly diverse. Approximately 40% of students are from ethnic groups underrepresented in STEM, about 50% are low income, about 30% are first-generation college students (neither parent has any college education) and approximately 40% speak English as a second or third language (2014 COLLEGE NAME STUDENT PROFILE; 2014 UNIVERSITY NAME STUDENT EXPERIENCE SURVEY).

This study was conducted in an Organic Chemistry I classroom in the spring of 2017. Including students who withdrew, the course enrolled a total of 176 students. Students met in a single classroom for lecture for a 75-minute class period twice a week (lectures were taught by AUTHOR). Additionally, students met in one of six smaller groups (of approximately 30 students each) for one 50-minute recitation period per week.

Student learning in the course was assessed via a combination of quizzes and exams. Five quizzes were administered over the course of the semester in recitation. Quizzes took approximately 20 minutes to complete and were standardized across all sections. Students also completed two 75-minute midterm exams and one 120-minute final exam (exams were administered in two large classrooms). Five percent of students' course grades were allotted towards completion of four assignments that were termed self-assessments. These self-assessments were actually two pre-exam and two post-exam exam wrappers. Students' course grades were determined using a mastery-based scheme, rather than a norm-referenced scheme (Popham, 1971).

Exam Wrappers

Exam wrapper 1 (see Appendix 1) was completed in recitation immediately before students took their first quiz (at approximately week three of the semester). Exam wrapper 2 (identical in content to exam wrapper 1) was administered similarly in that it was handed out and completed in recitation immediately before students took their second quiz (at approximately week four of the semester). One hundred sixty-six students (94%) completed exam wrapper 1 and one hundred sixty-three students (93%) completed exam wrapper 2.

Exam wrapper 3 (see Appendix 2) was completed online using the course management system available through the university. Exam wrapper 3 was only made available to students during week five of the semester (before the first midterm). Exam wrapper 4 (identical to exam wrapper 3) was similarly completed online and only available during week seven of the semester (in between the first and second midterm examinations). One hundred eleven students (63%) completed exam wrapper 3 and one hundred thirty-seven students (78%) completed exam wrapper 4. (Figure 1 shows the timing and sequence of the exam wrappers).

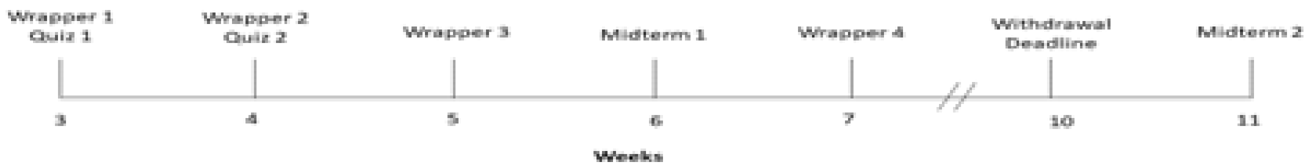


Figure 1. Timing of quizzes, midterm exams and exam wrappers

Exam wrappers 1 & 2 asked students to indicate how they felt about their competence in two areas: conceptual understanding and problem-solving ability.[1] These questions were intended to trigger self-reflection on the part of students (Zimmerman, 2002) and make them consider their degree of preparedness (Yuen-Reed & Reed, 2015). Students selected from among three choices (*strong*, *ok*, and *weak*) in two Likert-style questions. Additionally, students were asked to choose the extent to which they felt they could easily access help with the course when they needed it. This question was intended to lead students to consider whether or not they were seeking out and utilizing the resources available to them. Students selected from among three choices (*agree*, *unsure* and *disagree*) in a Likert-style question. Exam wrappers 1 and 2 were administered to students immediately before they took their first two quizzes (respectively) so that students who were underprepared but unaware of it would hopefully be hit with some dissonance if they put their expectations down on paper and then either struggled during the quiz or found out they were not in good shape when they received their quiz score.

Exam wrappers 3 and 4 asked students to report the extent to which they felt satisfied with their performance on their most recent assessment (quiz or midterm)[2]. Students selected from among five choices (*strongly agree, agree, neither agree nor disagree, disagree,* or *strongly disagree*) in a Likert-style question. Students were also asked to report how many hours per week they devoted to

[1] In the course syllabus, students were assigned readings (focused on conceptual understanding) and homework assignments (focused on problem solving) for each individual class period. The distinction between these two separate ideas or skills was also discussed in class. Thus, it was expected that students would understand and be able to differentiate between the terms conceptual understanding and problem solving ability.

[2] Typically within 24 hours of taking an exam, students' exam scores, along with a copy of the answer key, were posted online. However, typically they did not receive their graded exam papers back until their next recitation period. Quizzes took up to two weeks to be graded and were also returned to students in recitation. Quiz grades were also posted online. However, answer keys for quizzes were not provided.

studying and how many hours they spent problem-solving (with boxes provided for numerical responses.) Lastly, students were asked what they would do differently in the future to improve their course performance. (For this question, students wrote narrative responses of any length they wished and were able include as many study strategies as they wished.)

The first three questions asked in exam wrappers 3 and 4 were designed to encourage students to directly confront their feelings about how they were doing in the course, to determine how much time they were putting in overall, and to consider whether or not they were devoting sufficient time to problem-solving. The final question was intended to encourage students to self-reflect regarding their study behaviors, to consider what they ought to do differently, to set goals of changing their behaviors and to hopefully commit to those changes by asking them to put in writing what they intended to do differently from then on (Zimmerman, 2002).

In this study, the exam wrappers were utilized both as an intervention, intending to trigger students to reflect and possibly change their behavior, and also as a source of data, as a means for the researchers to examine and learn from the reflections and reported behaviors of students. Thus, with IRB permission, after the course was over the responses that students gave to the exam wrapper questions were collected, stripped of identifying information and utilized as data.

Course Syllabus with SRL Supplement

One significant modification was made to the course syllabus to reinforce the idea that students reflecting upon their own learning was an iterative process and that making decisions, being strategic, utilizing resources, and seeking help were all important aspects of being successful in the course. This modification took the form of an extra handout entitled: "How to Study for this Course". It contained a graphic display of a recommended iterative process for a student to follow and two subsections of guidance or advice entitled, "Ways to Assess Yourself" and "Help Seeking Guide". (See Appendix 3.) This handout was in addition to the usual guidance and information provided in the syllabus regarding how to be successful (e.g. spend 75% of your study time doing problem solving, come to class, get help right away if you get stuck, etc.) (See Appendix 4.) During the first day of lecture, AUTHOR reviewed the syllabus with students and explained this handout.

Website with SRL Emphasis

One factor relevant to this study is that AUTHOR maintains an extensive webpage (URL) devoted to providing students with additional resources for the Organic I course. The majority of these resources are supplementary problem sets. This is not unique as other Organic Chemistry instructors at her institution and elsewhere (Cortes, 2017; Reusch, 2017) also provide supplementary problem sets to students through webpages and/or course management software. However, the particular presentation of AUTHOR'S Organic I homepage gives a strong visual impression to students that problem solving is an extremely important aspect of the course. Furthermore, the homepage is organized such that students can quickly find either introductory or advanced practice problems (with answer keys) corresponding to each topic they learn about in lecture. Thus, students who utilize the website are encouraged to think about what topics they need to work on and are guided towards thinking about practicing in a strategic way where they ideally try to master simpler problems before they move to more challenging ones.

Another resource provided through the Organic I homepage is links to old exams (with answer keys) written by AUTHOR. The old exams are provided as additional practice problems for students. But they also illustrate what the format of exams will look like, what the level of difficulty of exams will be, and in what ways they will be expected to answer questions that require them to think across multiple chapters from their textbook while taking the exam.

One aspect of the old exams found on the website is that next to each exam question is listed the number of minutes a student should allocate toward completing it. This is a feature of AUTHOR'S exams that she includes to help students learn how to effectively pace themselves during exams, a skill that she finds is particularly challenging for many students at her institution.

Overall, there are a number of features in AUTHOR'S website that attempt to encourage students to strengthen and improve their self-regulation. However, these features have been in place on her website in a consistent fashion for a number of years. Therefore, for the purposes of this study, no modifications were made to the website.

Course Performance

Student scores on all assessments (quizzes, midterm examinations, and the final) were collected, stripped of identifying information, and utilized as data.

Prior Performance in Chemistry

Previously, AUTHOR and her colleagues (2013) had established that for students at her institution, the letter grades that they received in General Chemistry II (the immediate prerequisite course for Organic Chemistry I) were a good predictor of performance in Organic Chemistry I. Specifically if General Chemistry II was taken at our institution, 49% of the variability in performance in Organic Chemistry I was explained by performance in General Chemistry II. Therefore, for the purposes of the current study, we obtained transcripts from all students and collected General Chemistry II letter grades for students who completed General Chemistry II at our institution. (Letter grades for students who had taken General Chemistry II outside of our institution were not collected.) These letter grades were combined with the exam wrapper and course performance data described above and then all identifying information was removed.

Data Analysis

Exam wrappers 1 & 2. On exam wrappers 1 and 2, students were prompted to report the degree to which they felt competent (*strong, ok, weak*) in two areas: conceptual understanding and problem-solving ability. Students' self-perceptions were then compared to their actual quiz scores and then classified as either accurate, overestimates, or underestimates (see table 1).

A self-rating of *weak* on exam wrapper 1 was determined to be an underestimation if the student's quiz grade was above 45% and accurate if the student's grade was less than or equal to 45%. (An exam average of 45 was the approximate cutoff for passing the course.) A self-rating of *ok* was judged to be an overestimation if the student's grade was less than 46% and an underestimation if their grade was 80% or higher (and accurate if the grade was 46-79%). A self-rating of *strong* was gauged as an overestimation if the student's grade was less than 80% and deemed accurate if their grade was equal to or greater than 80%.

Because the subject matter of quiz 2 was more difficult than that of quiz 1 (and the student average on quiz 2 was 10 points lower than on quiz 1), more lenient criteria were utilized to define what was considered an overestimate, an underestimate or an accurate self-assessment for quiz 2. A self-rating of *weak* on exam wrapper 2 was gauged as an underestimation if the student's grade was above 35% and accurate if the student's grade was less than or equal to 35%. A self-rating of *ok* was categorized as an overestimation if the student's grade was less than 36% and an underestimation if their grade was 70% or higher (and accurate if the grade was 36-69%).

A self-rating of strong was gauged as an overestimation if the student's grade was less than 70% and deemed accurate if their grade was equal to or greater than 70%.

Table 1. Gauging students' accuracy in their assessments of their confidence in conceptual knowledge and problem-solving ability

		Overestimate	Accurate	Underestimate
Quiz 1				
	Weak	NA	≤ 45 %	> 45 %
	Ok	≤ 45 %	46-79 %	≥ 80%
	Strong	< 80%	≥ 80%	NA
Quiz 2				
	Weak	NA	≤ 35 %	> 35 %
	Ok	≤ 35 %	36-69 %	≥ 70%
	Strong	< 70%	≥ 70%	NA

Exam wrappers 3 & 4: Study time and problem-solving time. Questions on exam wrappers 3 and 4 prompted students to reflect on the number of hours they spent studying as well as the amount of time they spent practicing problems. The sample size, mean, standard deviation, and range of these reported hours were calculated. Outliers, participants whose reported hours were more than one standard deviation from the mean, were highlighted for possible further analysis. Differences between students' reported hours at the time of exam wrapper 3 and exam wrapper 4 were calculated.

Exam wrappers 3 & 4: Future plans. The final question of exam wrappers 3 and 4 asked students to describe what they would change or do differently in their future studying. Student responses were text-based and averaged approximately 30 words in length. Outliers, participants with word counts more than one standard deviation from the mean, were highlighted for possible further analysis.

Student written responses (to the question regarding what they planned to change in their study habits) were coded according to the following procedure. Four coders independently reviewed four different subsets of the student responses. (In total, one third of the student entries were reviewed.) In keeping with the explanatory nature of the study, the coders did not approach coding with preconceived or a-priori ideas of what the codes should be. Rather, we allowed themes to emerge from the data. Preliminary lists of themes and categories observed in the data were generated independently by each of the four coders. The coders then met and compared their preliminary lists. The preliminary lists were organized and collapsed into four codes, each of which contained a number of sub-codes. (See table 14 for a listing of the codes and sub-codes utilized in this study.) The four codes which emerged from the data were *Study Behaviors, Strategic Behaviors & Decisions, Help Seeking* and *What's Going on With Me.*

The code *Study Behaviors* referred to the many types of behaviors that students stated they planned to engage in, for example reading the textbook, working on practice problems, or reviewing their lecture notes. Each specific *Study Behavior* described by a student was given a separate sub-code. For example, the behaviors just described were assigned the sub-codes *Textbook, Problem Solve* and

Lecture Notes, respectively. Study behaviors were coded as *neutral, plus,* or *minus* depending if the student said he or she would *do it*, would spend *more time* doing it or would spend *less time* doing it, respectively.

The code *Strategic Behaviors & Decisions* was created to capture behaviors that students described which could best be characterized as being strategic about their studying. For example, choosing to do a little bit of problem solving every day, rather than saving it all for the weekend or choosing to read the textbook before lecture rather than after, were both coded under the code *Strategic Behaviors and Decisions* and under the sub-code *Timing Specific*, whereas choosing to work on advanced problems rather than simple problems was also coded under *Strategic Behaviors and Decisions,* but under the sub-code of *Which Problems*.

The code *Help Seeking* referred to the different ways in which students described how they would try to get help in the course. Sub-codes were created for the seeking of human help, *HHelp* – as in help from an instructor, TA, or tutor, electronic help, *EHelp* – as in help from an online resource like a tutorial or video, or help from a physical resource, *PHelp* - like a review book or molecular model set.

A code or category was created called *What's Going on With Me* to capture descriptions that students gave that did not fit under categories of study strategies or behaviors, but rather described emotional or psychological states. For example, a few of the sub-codes in this category were *Anxious, Unsure, Confident, Careless* and *Overwhelmed.*

A code book listing all the codes, sub-codes and their definitions was created. After this, the data entries were divided up into three equal portions. Each third of the data was coded independently by two coders (Six coders in total were utilized.) Afterwards, all six coders met as a group and went through each data entry one by one comparing the two sets of codes from the pair of coders against one another and against each data entry. Together, the group of six coders came to agreement on what the most complete and accurate codes should be for each data entry. Often the consensus or *agreed upon codes* matched the original codes assigned by the two coders, but occasionally errors or oversights were caught through this process. Therefore, it was determined that this method of meeting as a group of six and going through each data entry one by one was useful as it allowed for the most thorough, complete, and detailed analysis, without resulting in the coders reaching consensus prematurely. Occasionally as a result of this process, a few new sub-codes arose and had to be defined and created, and a few clarifications or refinements of existing codes had to be made. The code book was updated accordingly and modifications to the already coded data were made. After a complete listing of all the sub codes for each data entry was compiled, tallies were taken to determine the number of times each sub code was cited.

Exam 2 - 1. Student scores on exam 1 were subtracted from the scores on exam 2. Participants were then listed into four categories based on that difference: improved, worsened, no change, or NA (did not take exam 2).

Prior performance in chemistry. Students were coded as *at risk* for not succeeding in Organic Chemistry I if they had scored a grade of C plus or lower in General Chemistry II at our institution and as *not at risk* if they had scored a B minus or higher. Students who had not taken General Chemistry II at our institution (non-matriculated and transfer students) were coded as *unknown risk*.

Course performance. Students were grouped into categories based on their performance in the Organic course. For students entering the Organic course *not at risk*, satisfactory performance was defined as completing Organic with a grade of B minus or above.[3] For students entering the course *at risk* or with *unknown risk,* satisfactory performance was defined as completing the course with a C minus or above. Students who did not meet these criteria were categorized having unsatisfactory performance. (Students who withdrew from the course were categorized separately.)

Student categories. Nine categories of students were differentiated based on a) the level of risk of the students entering the course and b) their actual performance in the course (see table 12). These categories were analyzed and compared across areas of interest, such as confidence in conceptual understanding and problem-solving ability, accuracy of self-ratings compared to subsequent quiz performance, hours reported studying and practicing problems, and planned changes to study behaviors.

Overall Results

Exam Wrappers 1 & 2

Students were asked to describe their confidence in their conceptual understanding and problem-solving ability. In both exam wrappers 1 and 2 (see table 2), approximately 70% of students reported feeling *ok* about their understanding and ability. Additionally, the percentages of students who characterized their *conceptual understandings* as *weak* or *strong* changed only minimally from exam wrapper 1 to 2. However, with regards to confidence in *problem solving ability*, there was a notable increase from exam wrapper 1 to 2 in the percentage of students who felt they were *weak,* as well as a sizeable decrease in students who felt they were *strong*.

Table 2. Students' confidence in their conceptual understanding & problem-solving ability

	Wrapper 1				Wrapper 2			
	Concept		Problem		Concept		Problem	
	N	%	n	%	n	%	N	%
Weak	5	3	16	10	9	5	28	17
Ok	123	74	116	70	120	74	117	72
Strong	38	23	34	21	34	21	18	11

When comparing students' self-assessments to their actual quiz performances (see table 3), student accuracy was low (not on target), ranging from approximately thirty to forty percent. Student accuracy also decreased somewhat from exam wrapper 1 to 2. Furthermore, weak students (who scored 45 or below on quiz 1, or 35 or below on quiz 2) were highly likely to overestimate their abilities, while strong students (who scored 80 or above on quiz 1 or 70 or above on quiz 2) were highly likely to underestimate their abilities.

[3] We defined satisfactory as B- or better for *not at risk* students because we make the assumption that a C+ or worse will hinder future progress for these students (Hrabowski, 2016).

Table 3. Students' accuracy re their conceptual understanding & problem-solving ability

	Wrapper 1				Wrapper 2			
	Concept		Problem		Concept		Problem	
	n	%	n	%	n	%	n	%
Total Students who Are Accurate	67	40%	56	34%	51	31%	48	29%
Total Students who Overestimate	23	14%	25	15%	34	21%	24	15%
Total Students who Underestimate	76	46%	84	51%	78	48%	91	56%
Weak Students who Overestimate	12	100%	10	83%	26	93%	19	68%
Strong Students who Underestimate	74	73%	79	78%	72	71%	84	83%

The final question on exam wrappers 1 and 2 surveyed students' feelings about how easily they felt they could obtain help with the course material when needed (see table 4). Students were prompted to select a response of either *agree, unsure,* or *disagree* from a 3-point Likert-type scale. The most frequently selected response for both exam wrappers 3 (80%, n=166) and 4 (79%, n=163) was *agree*. Only 3% of participants chose *disagree* as their response to this question.

Table 4. Student responses to the statement that they are easily able to obtain help

	Wrapper 1		Wrapper 2	
	Help		Help	
	n	%	N	%
Agree	133	80	128	79
Unsure	28	17	30	18
Disagree	5	3	5	3

Exam Wrappers 3 & 4

In exam wrapper 3, student reports of satisfaction with their course performance (see table 5) spread in a bell-shaped distribution, with the majority of students reporting a neutral, mildly positive, or mildly negative attitude. (At this point in the semester, students had only received grades back on two quizzes, which had a combined average of 72 and counted only as 5-10% of their final course average.) However, by the time of exam wrapper 4, there was a large shift in student satisfaction with nearly 70% of students reporting dissatisfaction with their course performance. (Students filled out exam wrapper 4 shortly after receiving back their scores on exam 1 which had an average of 57% and counted as 20% of their final course average.)

Table 5. Student agreement that they are satisfied with their course performance

	Wrapper 3		Wrapper 4	
	n	%	n	%
Strongly Agree	16	14	2	1
Agree	28	25	20	15
Neither Agree nor Disagree	24	22	21	15
Disagree	27	24	40	29
Strongly Disagree	16	14.4	54	39

Table 6 shows student responses on exam wrappers 3 and 4, indicating number of hours spent studying and practicing problems. The changes in reported study and problem-solving time from wrapper 3 to 4 showed that on average, students only increased their study time by 0.2 hours and their problem-solving time by 0.5 hours (see table 7).

Table 6. Student reported study times and problem-solving times

	n	M	SD	Highest	Lowest
Wrapper 3					
Study Time (Hours)	111	8.2	3.5	25	1
Problem Solve Time (Hours)	111	5.7	3.5	25	1
Wrapper 4					
Study Time (Hours)	137	8.4	4.7	35	2
Problem Solve Time (Hours)	136	6.2	3.9	21	1

Table 7. Increase in reported study hours from exam wrapper 3 to 4?

	n	M	SD	Highest	Lowest
Study Time (Hours)	96	0.2	3.28	10	-11
Problem Solve Time (Hours)	96	0.5	3.13	14	-10

The length of participant responses to the question "What are you going to do differently from now on?" are reported in table 8. The average response was about 30 words in length. However, 21% and 23% of respondents (in wrappers 3 and 4, respectively) had word counts of less than 10 words.

Table 8. Length of student responses - what are you going to do differently?

	M	SD	Lowest	Highest
Wrapper 3				
Word Count	29	28	2	197
Wrapper 4				
Word Count	32	40	2	248

Because exam wrappers 3 and 4 asked students to report their overall study times and problem-solving times, students who reported an intention (on wrapper 3) to change their behavior by increasing their study time or problem-solving time were checked to see if they followed through on their intentions. Only 22% of students indicated that they would increase their overall study time and only about half of these students fulfilled their intention. Fifty percent of students indicated that they would increase their problem-solving time. Similarly, only about half of those students followed through on their intention (see table 9).

Table 9. Evidence (wrapper 4) of follow through of intended study plans (from wrapper 3)?

		n	%
Study Time			
	Did Follow Through	12	11
	Did Not Follow Through	11	11
	Not Applicable	82	78
Problem Solve Time			
	Did Follow Through	26	25
	Did Not Follow Through	26	25
	Not Applicable	49	47
	Unclear	3	4

Exam 2-Exam 1

Student performance on exams 1 and 2[4] was compared and differences were calculated (see table 10). Approximately half improved their scores and one-third worsened. An additional 15% did not take exam 2. (All students who did not take exam 2 also did not take the final exam.)

[4] Exam 2 covered more advanced and more challenging material than exam 1.

Table 10. Performance change from exam 1 to exam 2

	n	%
Improved	91	52
Worsened	57	32
No Change	2	1
NA (didn't take exam 2)	26	15

Prior Performance in Chemistry

Based on their prior performance in General Chemistry II, students were grouped into three categories. Categories indicated whether or not they were at risk of not succeeding in the Organic course. Each category contained approximately one-third of all students (see table 11).

Table 11. Students' at risk of not succeeding in Organic

	n	%
At Risk	54	31
Not At Risk	56	32
Unknown Risk	66	37

Course Performance

Overall, 63% of students performed satisfactorily in the Organic course, 29% performed unsatisfactorily, and 8% withdrew. Students of *unknown risk* were least likely to perform satisfactorily with only 55% of them successful, while 74% of *not at risk* and 71% of *at-risk* students were successful. (See table 12.)

Student Categories

Students were grouped into nine categories based on their risk when entering the course and their satisfactory performance in the course or lack thereof (see table 12). Seven of these nine categories were subjected to further analysis. Two categories were excluded (at risk & withdrawal, not at risk & withdrawal) because they each comprised only 1% of the student population.

	n	%

Table 12. Student categories*

	n	%
At Risk & Satisfactory	36	20
At Risk & Unsatisfactory	14	8
Not At Risk & Satisfactory	41	23
Not At Risk & Unsatisfactory	14	8
Unknown Risk & Satisfactory	31	18
Unknown Risk & Unsatisfactory	23	13
Unknown Risk & Withdrawal	11	6
At Risk & Withdrawal	2	1
Not At Risk & Withdrawal	2	1

* Two students, with grades of incomplete, were excluded and were not categorized.

Results for Student Categories

Exam Performance

Despite the fact that the second exam covered more challenging material than the first exam, the majority of *satisfactorily performing* students improved their performance from exam 1 to exam 2.

Estimation of Abilities

While results consistent with a Kruger-Dunning effect (1999) were observed for the overall population (see table 3), unsatisfactorily performing students, overall, did not tend to overestimate their abilities (see table 13). Satisfactorily performing students, however, were found to underestimate their abilities ($c^2(1) = 26.0867$, $p < 0.001$), while unknown risk students who withdrew were somewhat likely to overestimate their abilities ($c^2(1) = 16.264$, $p < 0.000055$, see table 13).

Reported Study Times

Satisfactorily performing students did not necessarily put in more study time or problem-solving time than unsatisfactorily performing ones. Furthermore, when unsatisfactorily performing students did increase their problem-solving time, this increase did not result in success. However, not at-risk students were somewhat more likely to indicate an intention (at wrapper 3) to increase the amount of time they were going to devote to problem solving. (See table 13.)

Exam Wrappers 3 & 4 and Unknown Risk Students

Unknown risk students differed from the other categories of students in a number of ways. Unknown risk, unsatisfactorily performing students was the only category to decrease their rate of completion of the exam wrappers from wrapper 3 to wrapper 4. At the time of wrapper 3, they were the most dissatisfied with their course performance as compared to the other categories. They reported the lowest average hours spent studying and doing practice problems. They also had the highest percentage of students who did not take either exam 2 nor the final exam.

The unknown risk, withdrawal students had the lowest completion rate of exam wrappers 3 and 4. Only three students (27%) completed wrappers 3 and 4. These students reported the greatest decrease in time spent studying and doing practice problems from wrapper 3 to 4, yet the greatest number of hours spent studying at the time of wrapper 3.

Planned Study Behaviors

Problem solving plans. At the time of wrapper 3, large numbers of all students in all categories indicated that they intended to devote more time to problem solving. However, by wrapper 4, almost none of the not at risk, satisfactorily performing students indicated that they needed to devote *additional* time to problem solving. Yet, over 70% of the not at risk, unsatisfactorily performing respondents reported that they *still* intended (and needed) to devote additional time to problem solving (see table 14).

Behaviors not reported. Plans to adopt behaviors such as joining a study group, attending office hours, reviewing lecture notes, checking one's answers against a key, or learning from one's mistakes were rarely (or never) reported by students of any category (see Appendix 1).

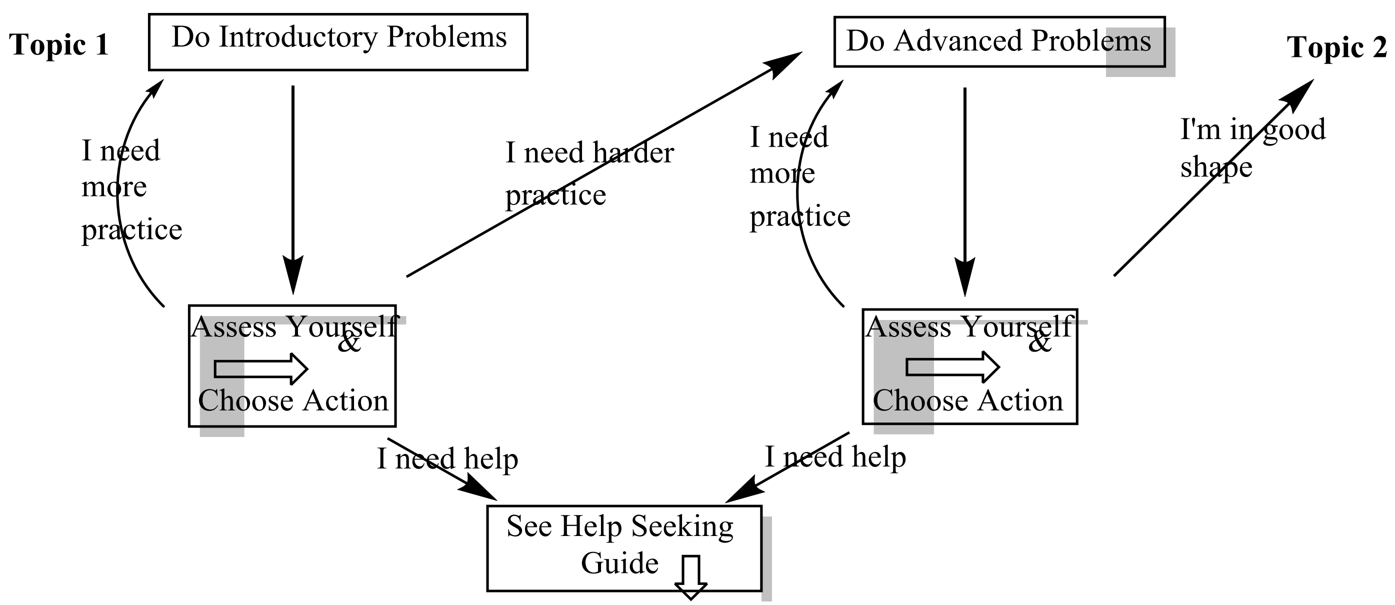
Appendix

Appendix 1: Comparison of Self-Assessments, Study Times, Follow Thru, Exam Performance Across the 7 Categories

	At Risk Satisfactory	At Risk Unsatis-factory	Not at Risk Satisfactory	Not at Risk Un-satisfactory	Unknown Risk Satisfactory	Unknown Risk Un-satisfactory	Unknown Risk Withdrawal
n in Category	36	14	41	14	31	23	11
% of Overall Population	20%	8%	23%	8%	18%	13%	6%
Self-Assessments of Conceptual & Problem Solving Abilities							
% Who were Accurate	64%	38%	31%	23%	41%	36%	42%
% Who Underestimate	7%	31%	66%	50%	51%	22%	11%
% Who Overestimate	29%	31%	2%	12%	8%	42%	47%
% Who Completed Wrapper 3	64%	57%	73%	64%	61%	57%	27%
% Who Completed Wrapper 4	86%	71%	95%	100%	81%	43%	27%
% Dissatisfied at Wrapper 3	30%	38%	23%	22%	53%	77%	33%
% Dissatisfied at Wrapper 4	71%	90%	63%	64%	52%	90%	100%
Mean Study Time (Hours/Wk)							
Reported in Wrapper 3	7.2	9.8	9.3	7.8	7.7	6.8	10.5
Reported in Wrapper 4	8.2	9.3	9.4	8.1	8.1	6.9	8.5
Wrapper 4 – 3	+1.0	-0.6	+0.1	+0.3	+0.4	+0.1	-2.0
Mean ProblemSolveTime (Hrs/Wk)							
Reported in Wrapper 3	5.0	7.8	6.1	5.9	4.9	4.7	8.7
Reported in Wrapper 4	5.6	8.0	6.3	6.3	6.1	5.3	5.5
Wrapper 4 – 3	+0.6	+1.0	+0.2	+0.4	+1.1	+0.6	-3.2
Average Word Count Wrapper 3	19	32	32	26	34	26	36

Average Word Count Wrapper 4	24	23	49	27	28	29	45
% of Category Who Planned to							
Increase Study Time	3%	29%	15%	7%	16%	13%	9%
Increase Problem Solve Time	28%	29%	41%	43%	29%	17%	9%
% of *Category* Who Followed Thru							
Increased Study Time	3%	14%	5%	7%	6%	4%	9%
Increased Problem Solve Time	8%	21%	19%	21%	23%	4%	9%
% Who Improve from Exam 1-2	67%	18%	90%	31%	68%	13%	0%
% Who Didn't Take Exam2 & Final	0%	14%	0%	7%	0%	30%	NA

Appendix 2: How to Study for Organic Chemistry



Ways to Assess Yourself:

- Check your work against an answer key.
- Try explaining your answers out loud to a friend or classmate.
- Try to convince a friend or classmate that your point of view is right (and theirs is wrong).
- Take the assigned question and modify it slightly but in a way that you think will matter and see if you can still answer it.

Help Seeking Guide:

1. Are you feeling lost about a specific topic or topics?
   a. A good source for simple explanations of Organic Chemistry concepts is Organic Chemistry as a Second Language by Klein.
   b. Sometimes a short, online video tutorial (no more than 10 minutes!) can help. See URLs listed above.
   c. Don't waste hours searching for or watching videos. If you don't find what you need quickly, ask a classmate or the course instructor for suggestions.
2. Are you getting some of the content, but missing bits and pieces, like parts of the HW you get right, parts you get wrong? Or you are not always sure why you get things right or wrong? In these kinds of situations (where you need quick, short explanations), it can be very helpful to
   a. Go to the tutors in the learning center.
   b. Ask questions of your recitation instructor, e.g. before or after class or during office hours.
   c. Ask questions of the lecture instructor, e.g. before or after class or during office hours.
   d. Ask questions of your lab instructor during quieter times in the lab (when there are waiting periods or when lab ends early).
3. Are you feeling completely lost in the course?
   This is a time to see someone like the course instructor (who is an expert at helping students succeed in Organic Chemistry) or another mentor/advisor that you know well and trust.

Appendix 3: How to Succeed In Organic Chemistry

How to Succeed In Organic Chemistry:

1. Set aside 10 hours per week of study time for this course.
2. Skim the textbook before class.
3. Attend class religiously and come on time to class.
4. Take notes in lecture.
5. Review your lecture notes as soon as you can after each class meeting.
6. Think of this class like a math class - problem solving is the most important thing.
   Spend the majority (at least 75%) of your study time doing problems, not reading!
   Do all the assigned homework problems (textbook and internet).
   Practice each topic until you have mastered it. Don't just stop when you have completed the assigned problems. Make sure you really understand what you are doing.
7. Study with a partner or in a group.
8. Don't be afraid to ask for help. Get help immediately if you get stuck.

References

Achacoso, M. V. (2004). Post-test analysis: A tool for developing students' metacognitive awareness and self-regulation. *New Directions for Teaching and Learning, 100*, 115-119.

Association of American Medical Colleges. (2010). *Diversity in the physician workforce: Facts and figures 2010*. Washington, D.C.: Association of American Medical Colleges.

Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology, 96*(3), 523-535. http://psycnet.apa.org/doi/10.1037/0022-0663.96.3.523

Barr, D. A., Matsui, J., Wanat, S. F., & Gonzalez, M. E. (2010). Chemistry courses as the turning point for premedical students. *Advances in Health Sciences Education, 15(1)*, 45-54. https://doi.org/10.1007/s10459-009-9165-3

Bell, P., & Volckmann, D. (2011). Knowledge surveys in General Chemistry: Confidence, overconfidence, and performance. *Journal of Chemical Education, 88*, 1469-1476. https://pubs.acs.org/doi/pdf/10.1021/ed100328c

Bernacki, M. L., Vosicka, L., & Utz, J. (2017). *Web-delivered training to improve learning and achievement for underrepresented and first-generation STEM learners*. Paper presented at the Annual Conference of the American Educational Research Association, San Antonio, Texas.

Bragg, D. D. (2012). Two-year college mathematics and student progression in STEM programs of study. In S. Olson & J. B. Labov (Eds.), *Community colleges in the evolving STEM education landscape* (pp. 80–101). Washington, DC: The National Academies Press.

Brown, B. R., Mason, A., & Singh, C. (2016). Improving performance in quantum mechanics with explicit incentives to correct mistakes. *Physical Review Physics Education Research, 12*(1), 1-20. https://doi.org/10.1103/PhysRevPhysEducRes.12.010121

Bunce, D. M., Komperda, R., Schroeder, M. J., Dillner, D. K., Lin, S., Teichert, M. A., & Hartman, J. R. (2017). Differential use of study approaches by students of different achievement levels. *Journal of Chemical Education, 94*(10), 1415-1424. https://pubs.acs.org/doi/pdf/10.1021/acs.jchemed.7b00202

Burd, G. D., Tomanek, D., Blowers, P., Bolger, M., Cox, J., Elfring, L., . . . Wallace, C. (2016). Developing faculty cultures for evidence-based teaching practices in STEM: A program report. In G. C. Weaver, W. D. Burgess, A. L. Childress, & L. Slakey (Eds.), *Transforming* *institutions: Undergraduate STEM education for the 21st century* (pp. 90-102). West Lafayette, Indiana: Purdue University Press.

Castro, E. L. (2014). "Underprepared" and "at-risk": Disrupting deficit discourses in undergraduate STEM recruitment and retention programming. *Journal of Student Affairs Research and Practice, 51*(4), 407-419. https://doi.org/10.1515/jsarp-2014-0041

Chang, M. J., Sharkness, J., Hurtado, S., & Newman, C. B. (2014). What matters in college for retaining aspiring scientists and engineers from underrepresented racial groups. *Journal of Research in Science Teaching, 51*(5), 555-580. https://doi.org/10.1002/tea.21146

Chaskes, J., & Anttonen, R. G. (2015). Advocating for first-year students. In M. L. Upcraft, J. N. Gardner, & B. O. Barefoot (Eds.). *The freshman year experience: Helping students survive and succeed in college* (pp. 191-203). San Francisco, California: Jossey Bass Publishers.

Chen, P., Chaves, O., Ong, D. C., & Gunderson, B. (2017). Strategic resource use for learning: A self-administered intervention that guides self-reflection on effective resource use enhances academic performance. *Psychological Science, 28*(6), 774-785. https://doi.org/10.1177/0956797617696456

Chen, X. (2013). STEM Attrition: College students' paths into and out of STEM fields (NCES 2014-001). Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Chew, K. J., Chen, H. L., Rieken, B., Turpin, A., & Sheppard, S. (2016). *Improving students' learning in statics skills: Using homework and exam wrappers to strengthen self-regulated learning.* Paper presented at the American Society for Engineering Education (ASEE) 123rd Annual Conference, New Orleans, LA.

Christianson, K., Mestre, J. P., & Luke, S. G. (2012). Practice makes (nearly) perfect: Solving 'students-and-professors'-type algebra word problems. *Applied Cognitive Psychology, 26*, 810-822. https://doi.org/10.1002/acp.2863

Cortes, S. (2017). Dr. Sergio Cortes' Organic Chemistry page. Retrieved from http://www.utdallas.edu/~scortes/ochem/

Crimmins, M. T. (2017). [Personal communication].

Davis, J. (2010). *The first-generation student experience: Implications for campus practice, and strategies for improving persistence and success.* Sterling, VA: Stylus Publishing, LLC.

Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrollment Physics class. *Science, 332(6031)*, 862-864. DOI: 10.1126/science.1201783

DiBenedetto, M. K., & Bembenutty, H. (2013). Within the pipeline: Self-regulated learning, self-efficacy, and socialization among college students in science courses. *Learning and Individual Differences, 23*, 218-224. https://doi.org/10.1016/j.lindif.2012.09.015

Doerschuk, P., Bahrim, C., Daniel, J., Kruger, J., Mann, J., & Martin, C. (2016). Closing the gaps and filling the STEM pipeline: A multidisciplinary approach. *Journal of Science Education and Technology, 25*(4), 682-695. https://doi.org/10.1007/s10956-016-9622-8

Dougherty, K. J. (1992). Community colleges and baccalaureate attainment. *Journal of Higher Education, 63*(2), 188-214. https://doi.org/10.1080/00221546.1992.11778349

Eagan, K., Hurtado, S., Figueroa, T., & Hughes, B. (2014). Examining STEM pathways among students who begin college at four-year institutions. Retrieved from http://sites.nationalacademies.org

Eberlein, T., Kampmeier, J., Minderhout, V., Moog, R. S., Platt, T., Varma-Nelson, P., & White, H. B. (2008). Pedagogies of engagement in science: A comparison of PBL, POGIL, and PLTL. *Biochemistry and Molecular Biology Education, 38*(4), 262-273. https://doi.org/10.1002/bmb.20204

Eddy, S. L., & Hogan, K. A. (2014). Getting under the hood: How and for whom does increasing course structure work? *CBE-Life Sciences Education, 13(3)*, 453-468. https://doi.org/10.1187/cbe.14-03-0050

Engle, J., & Tinto, V. (2008). *Moving beyond access: College success for low-income: First-generation students*. Retrieved from https://eric.ed.gov/?id=ED504448

Ertmer, P.A., & Newby, T.J. (1996). The expert learner: Strategic, self-regulated, and reflective. *Instructional Science, 24*, 1-24. https://doi.org/10.1007/BF00156001

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, 111*(23), 8410-8415. https://doi.org/10.1073/pnas.1319030111

Gasiewski, J. A., Eagan, M. K., Garcia, G. A., Hurtado, S., & Chang, M. J. (2012). From gatekeeping to engagement: A multicontextual, mixed method study of student academic engagement in introductory STEM courses. *Research in Higher Education, 53(2)*, 229-261. https://doi.org/10.1007/s11162-011-9247-y

Gosser, D. K. (2011). The PLTL boost: A critical review of research. *Progressions, 14*(1).

Gregg-Jolly, L., Swartz, J., Iverson, E., Stern, J., Brown, N., & Lopatto, D. (2016). Situating second-year success: Understanding second-year STEM experiences at a liberal arts college. *CBE-Life Sciences Education, 15*(3). https://doi.org/10.1187/cbe.16-01-0044

Haak, D. C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science, 332*(6304), 1213-1216. https://doi.org/10.1126/science.1204820

Hall, D. M., Curtin-Soydan, A. J., & Canelas, D. A. (2014). The science advancement through group engagement program: Leveling the playing field and increasing retention in science. *Journal of Chemical Education, 91*(1), 37-47. https://pubs.acs.org/doi/ipdf/10.1021/ed400075n

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Giffen, C. J., Blair, S. S., Rouse, D. I., & Hyde, J. S. (2014). Closing the social class achievement gap for first-generation students in undergraduate biology. *Journal of Educational Psychology, 106*(2), 375-389. http://psycnet.apa.org/doi/10.1037/a0034679

Hrabowski III, F. A. (2016). Transcript of keynote address by Freeman Hrabowski III at the 2014 Transforming Institutions Conference. In G. C. Weaver, W. D. Burgess, A. L. Childress, & L. Slakey (Eds.), *Transforming institutions: Undergraduate STEM education for the 21st century* (pp. 501-513). West Lafayette, Indiana: Purdue.

Hudesman, J., Crosby, S., Ziehmke, N., Everson, H., Isaac, S., Flugman, B., ... & Moylan, A. (2014). Using formative assessment and self-regulated learning to help developmental mathematics students achieve: A multi-campus program. *Journal on Excellence in College Teaching, 25*(2), 107-130.

Hunter, A.-B. (2016). *Factors contributing to undergraduate switching from STEM majors: Has anything changed in the last two decades?* Paper presented at the Annual Conference of the American Educational Research Association, Washington, D.C.

Jehangir, R. R. (2010). *Higher education and first-generation students: Cultivating community, voice, and place for the new majority*. New York, NY: Palgrave Macmillan.

Karabenick, S. A. (2003). Seeking help in large college classes: A person-centered approach. *Contemporary Educational Psychology, 28*(1), 37-58. https://doi.org/10.1016/S0361-476X(02)00012-7

Keeley, E.J. III & House, J.D. (1993). Transfer shock revisited: A longitudinal study of transfer academic performance. Paper presented at Association for Institutional Research Annual Forum, May 17, 1993, Chicago, Illinois.

Kezar, A., & Holcombe, E. (2016). Institutional transformation in STEM: Insights from change research and the Keck-PKAL project. In G. C. Weaver, W. D. Burgess, A. L. Childress, & L. Slakey (Eds.), *Transforming institutions: Undergraduate STEM education for the 21st century* (pp. 35-47). West Lafayette, Indiana: Purdue University Press.

Kobrak, M. (2006). *Surviving science at Brooklyn College: What every transfer student needs to know*. Brooklyn College. Brooklyn, New York.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121-1134. http://psycnet.apa.org/doi/10.1037/0022-3514.77.6.1121

Lewis, S. E. (2011). Retention and reform: An evaluation of peer-led team learning. *Journal of Chemical Education, 88(6)*, 703-707. https://pubs.acs.org/doi/ipdf/10.1021/ed100689m

Lindsey, B. A., & Nagel, M. L. (2015). Do students know what they know? Exploring the accuracy of students' self-assessments. *Physical Review Special Topics - Physics Education Research, 11*(2), 020103-020101. https://doi.org/10.1103/PhysRevSTPER.11.020103

Lohfink, M. M., & Paulsen, M. B. (2005). Comparing the determinants of persistence for first-generation and continuing-generation students. *Journal of College Student Development, 46*(4), 409-428. https://doi.org/10.1353/csd.2005.0040

Lopez, E., Nandagopal, K., Shavelson, R., Szu, E., & Penn, J. H. (2013). Self-regulated learning study strategies and academic performance in undergraduate organic chemistry: An investigation examining ethnically diverse students. *Journal of Research in Science Teaching, 50*(6), 660-676. https://doi.org/10.1002/tea.21095

Lovett, M. C. (2013). Making exams worth more than the grade. In M. Kaplan, N. Silver, D. LaVaque-Manty, & D. Meizlish (Eds.), *Using reflection and metacognition to improve student learning* (pp. 18-52). Sterling, VA: Stylus.

Mäkinen, J., Olkinuora, E., & Lonka, K. (2004). Students at risk: Students' general study orientations and abandoning/prolonging the course of studies. *Higher Education, 48(2)*, 173-188. https://doi.org/10.1023/B:HIGH.0000034312.79289.ab

Malcom, S., & Feder, M. (Eds.). (2016). *Barriers and opportunities for 2-year and 4-year STEM degrees: Systemic change to support students' diverse pathways*. Washington, D.C.: National Academies Press.

Mason, A., & Singh, C. (2010). Do advanced physics students learn from their mistakes without explicit intervention. *American Journal of Physics, 78*(760). doi: 10.1119/1.3318805

Michaelsen, L. K., Knight, A. B., & Fink, L. D. (Eds.). (2004). *Team-based learning: A transformative use of small groups*. Sterling, VA: Stylus Publishing.

Moog, R. S., & Farrell, J. J. (2006). *Chemistry: A Guided Inquiry* (3rd ed.). New York, NY: John Wiley and Sons.

Mooring, S. R., Mitchell, C. E., & Burrows, N. L. (2016). Evaluation of a flipped, large-enrollment organic chemistry course on student attitude and achievement. *Journal of Chemical Education, 93*(12), 1972-1983. https://pubs.acs.org/doi/ipdf/10.1021/acs.jchemed.6b00367

Nandagopal, K., & Ericsson, K. A. (2012). An expert performance approach to the study of individual differences in self-regulated learning activities in upper-level college students. *Learning and Individual Differences, 22(5)*, 597-609. https://doi.org/10.1016/j.lindif.2011.11.018

National Science Foundation. (2013). *Women, minorities, and person with disabilities in science and engineering: 2013*

Nuñez, A. M., & Yoshimi, J. (2017). A phenomenology of transfer: Students' experiences at a receiving institution. *Innovative Higher Education, 42(2)*, 173-187. https://doi: 10.1007//s10755-016-9374-7

Olszewski, P. (2016). Teaching millennials how to study under the 21st century sky. *Pyrex Journal of Educational Research and Reviews, 2*(1), 1-9.

Pazicni, S., & Bauer, C. F. (2014). Characterizing illusions of competence in introductory chemistry students. *Chemistry Education Research and Practice, 15*, 24-34. https://doi.org/10.1039/C3RP00106G

Phillips, J. A., Clemmer, K. W., McCallum, J. E. B., & Zachariah, T. M. (2017). A problem-solving framework to assist students and teachers in STEM courses. *Journal of College Science Teaching, 46*(4), 33-39.

Plant, E. A., Ericsson, K. A., Hill, L., & Asberg, K. (2005). Why study time does not predict grade points average across college students: Implications of deliberate practice for academic performance. *Contemporary Educational Psychology, 30(1)*, 96-116. https://doi.org/10.1016/j.cedpsych.2004.06.001

Popham, W. J. (Ed.) (1971). *Criterion-referenced measurement: An introduction*. Englewood Cliffs, NJ: Educational Technology Publications.

Reusch, W. (2017). Organic Chemistry practice problems at Michigan State University. Retrieved from https://www2.chemistry.msu.edu/faculty/reusch/virttxtjml/Questions/problems.htm

Rosser, S. V. (1997). *Re-Engineering female friendly science*. New York, NY: Teachers College Press.

Rozell, T. G., Johnson, J., Sexten, A., & Rhodes, A. E. (2017). Correcting missed exam questions as a learning tool in a physiology course. *Journal of College Science Teaching, 46*(5), 56-63.

Saenz, V. B., Hurtado, S., Barreta, D., Wolf, D. S., & Yeung, F. (2007). *First in my family: A profile of first-generation college students at four-year institutions since 1971*. Los Angeles: Reagents of the University of California.

Sebesta, A. J., & Bray Speth, E. B. (2017). How should I study for the exam? Self-regulated learning strategies and achievement in introductory biology. *CBE-Life Sciences Education, 16*(2). https://doi.org/10.1187/cbe.16-09-0269

Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences*: HarperCollins Publishers, Inc.

Smedley, B. D., Butler, A. S., & Bristow, L. R. (Eds.). (2004). *In the nation's compelling interest: Ensuring diversity in the health-care workforce*. Washington, D.C.: National Academies Press.

Soicher, R. N., & Gurung, R. A. R. (2017). Do exam wrappers increase metacognition and performance? A single course intervention. *Psychology Learning and Teaching, 16*(1), 64-73. https://doi.org/10.1177/1475725716661872

Stanton, J. D., Neider, X. N., Gallegos, I. J., & Clark, N. C. (2015). Differences in metacognitive regulation in introductory biology students: When prompts are not enough. *CBE - Life Sciences Education, 15*, 1-12. https://doi.org/10.1187/cbe.14-08-0135

Stephens, N. M., Hamedani, M. G., & Destin, M. (2014). Closing the social-class achievement gap: A difference-education intervention improves first-generation students' academic performance and all students' college transition. *Psychological Science, 25*(4), 943-953. https://doi.org/10.1177/0956797613518349

Stuber, J. M. (2012). *Inside the college gates: How class and culture matter in higher education*. New York, NY: Rowman & Littlefield Publishers, Inc.

Szu, E., Nandagopal, K., Shavelson, R., Lopez, E. J., Penn, J. H., Scharberg, M., & Hill, G. W. (2011). Understanding academic performance in Organic Chemistry. *Journal of Chemical Education, 88*(9), 1238-1242. https://pubs.acs.org/doi/ipdf/10.1021/ed900067m

Talbert, R. (2017). *Flipped learning*. Sterling, VA: Stylus.

Terenzini, P., Springer, L., Yaeger, P. M., Pascarella, E. T., & Nora, A. (1995). *First-generation college students: Characteristics, experiences, and cognitive development*. Paper presented at the Meeting of the Association for Institutional Research, Boston, MA.

Thompson, D. R. (2012). Promoting metacognitive skills in intermediate Spanish: Report of a classroom research project. *Foreign Language Annals, 45*(3), 447-462. https://doi.org/10.1111/j.1944-9720.2012.01199.x

Tipton, R.L., & Bender, P. (2006). From failure to success: Working with under-prepared transfer students. *Reference Services Review, 34*(3), 389-404. https:// doi.org/ 10.1108/00907320610685337

Townsend, B. K. (1995). Community college transfer students: A case study of survival. *The Review of Higher Education*, *18*(2), 175-193. https://doi: 10.1353/rhe.1995.0022

Townsend, B.K., & Wilson, K.B. (2006). The transfer mission: Tried and true, but troubled? *New Directions for Community Colleges, 2006*(136), 33-41. https://doi.org/10.1002/cc.257

Uche, A. R. (2015). *The retention of first-generation college students in STEM: An extension of Tinto's longitudinal model.* (Ph.D.), University of North Carolina, Charlotte.

VanderStoep, S. W., Pintrich, P. R., & Fagerlin, A. (1996). Disciplinary differences in self-regulated learning in college students. *Contemporary Educational Psychology, 21*(4), 345-362. https://doi.org/10.1006/ceps.1996.0026

West, C., & Sadoski, M. (2011). Do study strategies predict academic performance in medical school? *Medical Education, 45*(7), 696-703. https://doi:10.1111/j.1365-2923.2011.03929.x

Wieman, C. E. (2016). Foreword. In G. C. Weaver, W. D. Burgess, A. L. Childress, & L. Slakey (Eds.), *Transforming institutions: Undergraduate STEM education for the 21st century* (pp. ix-xiv). West Lafayette, Indiana: Purdue University Press.

Watkins, J., & Mazur, E. (2013). Retaining students in science, technology, engineering, and mathematics (STEM) majors. *Journal of College Science Teaching, 42*(5), 36-41. http://www.jstor.org/stable/43631580

Weinstein, C. E., & Acee, T. W. (2013). Helping college students become more strategic and self-regulated learners. In H. Bembenutty, T. J. Cleary, & A. Kitsantas (Eds.), *Applications of self-regulated learning across diverse disciplines* (pp. 197-236). Charlotte, NC: Information Age Publishing.

Yuen-Reed, G., & Reed, K. B. (2015). Engineering student self-assessment through confidence-based scoring. *Advances in Engineering Education*, *4*(4), 1-23.

Zhao, N., Wardeska, J. G., McGuire, S. Y., & Cook, E. (2014). Metacognition: An effective tool to promote success in college science learning. *Journal of College Science Teaching, 43*(4), 48-54. http://www.jstor.org/stable/43632012

Zimmerman, B.J., & Martinez-Pons, M. M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal, 23*(4), 614-628. https://doi.org/10.3102/00028312023004614

Zimmerman, B.J., Moylan, A., Hudesman, J., White, N., & Flugman, B. (2011). Enhancing self-reflection and mathematics achievement of at-risk urban technical college students. *Psychological Test and Assessment Modeling, 53*(1), 108-127

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*(1), 3-17. https://doi.org/10.1207/s15326985ep2501_2

Zimmerman, B.J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64-70. https://doi.org/10.1207/s15430421tip4102_2

Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology, 82*(1), 51-59. http://psycnet.apa.org/doi/10.1037/0022-0663.82.1.51

Journal of the Scholarship of Teaching and Learning, Vol. 20, No. 1, April 2020, pp.130-150.
doi: 10.14434/josotl.v20i1.25093

Meaningful Teaching Tool and/or 'Cool Factor'? Instructors' Perceptions of Using Film and Video within Teaching and Learning

Elizabeth Marquis
McMaster University

Cassia Wojcik
McMaster University

Effie Lin
McMaster University

Victoria McKinnon
McMaster University

Abstract: This study builds upon previous research that explores the pedagogical use of film and video by explicitly asking instructors about their attitudes towards and motivations for employing such texts in their teaching, as well as the challenges they face in the process. Data were gathered through an anonymous, online survey of instructors across disciplines at seven Ontario universities. Commonalities were found amongst participants in the purposes cited for using film and video as well as in the challenges that accompany use of this pedagogical tool. For example, instructors in four of our six Faculty groupings commonly noted drawing on film and video to engage student attention, and the two most frequently selected challenges in five of our six Faculty groupings were 'technical difficulties screening films' and 'problems finding appropriate materials'. We consider the implications of these findings for teaching and learning and suggest areas for future research.

Keywords: media pedagogy; film and video; instructor motivations; instructor attitudes; student engagement; technology in teaching

Introduction

The pedagogical potential of film has been increasingly recognized in recent years. On one hand, several scholars have argued that popular film and television exert a compelling instructional force within a media-saturated world, functioning as what Giroux (2004, 2008) has called a form of "public pedagogy." By constructing emotionally engaging and persuasive representations, such scholars suggest, film and television texts[1] contribute to shaping understandings of the world and constitute a site at which normative social discourses are (re)produced, negotiated, and sometimes contested (see, for example, Epstein, Mendick & Moreau, 2010; Garcia, 2015; Giroux, 2009; Marquis, 2018; Johnstone, Marquis, & Puri, 2018). At the same time, this educational capacity has also been recognized by educators who consciously deploy film and television as pedagogical tools within their classrooms and courses (Luccasen & Thomas, 2010; Sealey, 2008). As Myers and Abd-El-Khalick (2016) note, there has been long-standing attention to the ways in which film and television can be used as 'teaching aids' in science instruction, for example, although research on the effectiveness of

[1] Here, and throughout, we use the term 'text' as it is commonly deployed in film and cultural studies - to refer to a broad range of cultural/artistic/social artifacts, rather than to simply denote written works.

such tools is relatively scant. At the same time, recent studies suggest that the increasing availability and accessibility of media (Andrist, Chepp, Dean, & Miller, 2014; Holland, 2014) and growing student interest in learning environments that employ popular culture (Peacock et al., 2018) support the use of film and television in teaching within the contemporary moment. Further attention to how such media are incorporated within formal teaching contexts is thus merited and timely.

In this study, we therefore set out to examine how and why faculty across disciplines at seven postsecondary institutions draw on film and video within their teaching, and how they perceive the potential benefits and challenges of teaching with such media. Noting that much scholarship on instructor perspectives about teaching with film is relatively small scale and/or discipline-specific, we seek to expand on the evidence base in this area, offering important insights into faculty perspectives that may allow for improvements upon current teaching practices and shape further research.

Literature Review

The body of existing literature discussing the use of film and television within teaching suggests that educators have drawn on such materials in a range of contexts and for a variety of purposes. Scholars have described using both audio-visual materials and images to elicit and increase student empathy (Blasco & Moreto, 2012; Happel-Parkins & Esposito, 2015; Marcus & Stoddard, 2007), for instance, as well as to illustrate complicated and sometimes abstract concepts (Andrist et al., 2014; Calcagno, 2015; Pelton, 2013). Film and video have also been used in order to promote the development of professional skills (Ber & Alroy, 2002; Lumlertgul, Kijpaisalratana, Pityaratstian, & Wangsaturaka, 2009), and to support student critical thinking and deep approaches to learning (Bright, 2015; Olson, Autry, & Moe, 2016). Some research also describes using film to help develop students' media literacies, including in ways that encourage them to consider the dominant cultural discourses and representational biases encoded in popular texts (Holland, 2014; Huczynski and Buchanan, 2004; Sigler & Albandoz, 2014). Finally, several authors argue that incorporating film and other audio-visual material within teaching contexts can serve to increase student engagement or motivation (Algeo, 2007; Kabooha, 2016; Swimelar, 2013), supporting richer discussions by providing students with a familiar and relevant entry point (Madsen, 2014; Travis, 2016), making difficult or challenging ideas easier to approach (Bright, 2015; Calcagno, 2015; Madsen, 2014), and perhaps appealing to different learner groups or attracting new students to the content or the discipline (Brown, Smith, McAllister, & Joe, 2017; Luccasen & Thomas, 2010). Such scholarship indicates that instructors have used film and media to support a range of pedagogical goals.

Nevertheless, in spite of this pedagogical potential, existing research also begins to point to challenges and limitations of drawing on audio-visual media to help meet particular instructional objectives. A number of concerns and considerations have been raised in this regard, ranging from practical issues such as strains on class or instructor preparation time (McAllister, 2015; Sigler & Albandoz, 2014) and difficulties finding relevant materials (Huczynski & Buchanan, 2004; Kabooha, 2014), to the inadequacies and implications of film and television texts themselves. Some scholars note that films can reinforce problematic stereotypes or discourses if not properly critiqued, for example (Kuzma & Haney, 2001; Madsen, 2014), while others underline that films are a limited and partial source of knowledge that students might use inappropriately (Ansell, 2002; Madsen 2014; Marcus & Stoddard, 2009). Building on such concerns, some highlight that films or videos cannot simply be *shown*, but rather require clear instructor framing and objectives (Holland, 2014; Kabooha, 2016; Sigler & Albandoz, 2014), a task made more complicated by the fact that many faculty are not sufficiently trained in effective pedagogical use of audio-visual media (Peacock et al., 2018) and some draw on film and video in ways that have not always been carefully thought through (Hobbs, 2006). Perhaps most interestingly, a number of writers allude to concerns about the perceived frivolity of using film and television within classroom contexts, indicating possible faculty and student resistance to considering audio-visual media as meaningful sources of knowledge or significant pedagogical supports (e.g., Madsen, 2014; Peacock et al., 2018; Swimelar, 2013; Travis, 2016). Such challenges stand to impinge significantly on the uptake and efficacy of film and video as teaching tools, but they are often only discussed elliptically in existing research. There is a need to more directly explore instructors' perceptions of the limitations of film and media use and how these considerations affect their pedagogical practices and goals.

Moreover, as noted above, much of the existing scholarship in this area has been conducted within individual courses or disciplines, including geography (Algeo, 2007; Madsen, 2014), nursing (Kirkpatrick & Brown, 2014; McConville & Lane, 2006), economics (Leet & Houser, 2003; Sexton, 2006), languages (BavaHarji, Alavi & Letchumanan, 2014; Seferoglu, 2008), political science (Holland, 2014; Swimelar, 2013), medicine (Datta, 2009; Lumlertgul et al., 2009), and history (Marcus & Stoddard, 2009; Volk, 2015). While offering compelling insights about the ways in which audio-visual media can be used within particular disciplinary contexts, this single-discipline focus leaves open interesting questions about the commonality of the conclusions drawn and perspectives raised. One recent counterexample to this trend is a study by Peacock and colleagues (2018), which sought to explore faculty's attitudes towards and use of popular culture (including film) across disciplines at one American university. That research demonstrates clearly the value of examining faculty perceptions of film and video use within a range of subject areas, finding that participants across the disciplinary spectrum report drawing on popular culture relatively frequently in their teaching and express fairly high agreement with the argument that popular culture can be a meaningful support to the development of students' critical thinking. Nevertheless, significant differences were found between instructors in the Humanities and Social Sciences and those in the Natural Sciences and Math, with faculty in the former groupings indicating both more frequent employment of popular culture and more positive beliefs about its importance in the classroom. Such findings, which suggest points of contact and divergence in instructors' views of teaching with popular culture across different fields, underscore the potential value of further cross-disciplinary research in this area.

This exploratory study thus aims to contribute to the growing body of literature about the use of film within university teaching by examining how instructors across disciplines perceive and report employing film and video in their teaching activities. Since our primary goal is to examine the scope and transferability of faculty-reported experiences with and barriers to film and video pedagogy, we focused exclusively on instructors' perspectives in our survey rather than attending to student reactions or assessing learning outcomes. While such foci are surely important, more extensive knowledge of existing faculty practices and perspectives is likewise essential to understanding and enhancing the educational use of film and video, particularly given the central role instructors play in determining if and how film is incorporated into courses.

While Peacock et al. (2018) also attend to the experiences and understandings of faculty spanning multiple disciplines in higher education, we build on their work in several ways. Whereas their research focuses on instructors at one mid-sized public university in the United States, our study draws on responses from faculty across disciplines at seven Canadian universities, thereby bringing to bear a wider range of instructor perspectives and experiences. At the same time, we also narrow the focus, relative to Peacock et al. (2018); whereas they examined perceptions of popular culture broadly, we attend to the more specific question of film and video use, acknowledging that this still includes a wide variety of audio-visual media (from feature length films to brief YouTube videos). Moreover, while Peacock et al. (2018) focus primarily on the frequency with which instructors use popular culture in their teaching and on faculty attitudes towards its pedagogical significance, we supplement these considerations by exploring more directly instructors' motivations for drawing on film and video and their perceptions of the challenges and drawbacks of using these specific texts. We also draw on qualitative data from open-ended survey questions in order to offer further insight into instructors' perspectives, whereas Peacock et al. took a purely quantitative approach. Our work thus contributes significant information to the evidentiary base about faculty's perspectives on the role of such media within diverse university teaching contexts.

Methodology & Participants

In order to gather comparable information from a broad participant demographic, while also privileging participants' perspectives and opinions, we collected data via an anonymous online survey. This was in line with both our research aims and an underlying interpretivist methodology that understands realities as multiple and experiential (Merriam, 2009). In addition to basic demographic prompts, the survey contained both multiple choice and open-ended questions asking instructors to report how frequently they use film and/or video in their teaching, the purposes for which they use it, and the challenges they experience in this process. To get an inclusive picture of instructor perceptions of audio-visual media, we deliberately kept the focus somewhat broad, using 'film and/or video' in the question text, and including a question that explicitly asked respondents to select the types of media they use in their teaching (see Appendix 1 for the full survey instrument). We also asked respondents to rank on a Likert scale how useful they find film and video in their teaching, providing insight into their attitudes about the pedagogical potential of such texts. Before circulation to participants, the survey was sent to experienced pedagogical researchers for testing, and revisions were made based on their feedback.

The study then proceeded in two phases, both of which were cleared by the [university] Research Ethics Board. The first phase began by sending an email invitation to departmental administrators at the university with which we are associated, asking them to forward the invitation on to faculty and instructors within their respective departments. Given the limited yet provocative data generated by this first phase, we subsequently elected to expand our data pool to include additional universities within our province. We selected six universities, aiming to access a range of institutional types (e.g., medical-doctoral, comprehensive, primarily undergraduate) and geographical locations. We then gathered email addresses for instructors at these sites from public-facing institutional websites and emailed the invitation to complete the survey. A complete breakdown of survey participants across both phases of the study is provided in Table 1 below.

Once data collection was complete, responses were exported from the survey tool and basic descriptive statistics were computed for multiple choice and ranking questions. Using SPSS software, we also ran a Kruskal-Wallis H test to compare responses to ranking questions by discipline. The Kruskal-Wallis H test is a nonparametric alternative to the ANOVA that can be used to determine if there are statistically significant differences between two or more groups (here, respondents in different disciplinary groupings) in relation to a continuous or ordinal dependent variable (in this case, rankings of film's usefulness for teaching) when data are not normally distributed (as our ranking data were not). Finally, responses to open-ended questions were scrutinized to determine if and how they expanded on, corroborated, or qualified data gathered from the other question types. Typically, such responses offered further detail about the purposes for which instructors used film or helped to explain and justify participants' rankings of its utility.

Table 1: Participant demographics.

Institution	Number of Responses	Percentage of Total Responses
Algoma University	15	2.5
Lakehead University	42	7.1
McMaster University	64	10.8
Queen's University	98	16.5
University of Ottawa	160	27.1
University of Windsor	61	10.3
York University	137	23.2
Not Indicated	14	2.3
Discipline		
Business	36	6.1
Engineering	40	6.8
Health Sciences	50	8.4
Humanities	173	29.3
Interdisciplinary & Social Sciences	157	26.6
Sciences	125	21.2
Not Indicated	10	1.7
Years Teaching		
0-5	85	14.4
6-10	135	22.8
11-15	103	17.4
16-20	68	11.5
21-25	50	8.5
More than 25	138	23.4
Not Indicated	12	2

Findings

Attitudes

Our survey investigated instructors' perceptions of the potential value of using film and video as a teaching tool in post-secondary contexts. As in other studies (e.g., Peacock et al., 2018), the results were generally positive. Of the 588 instructors who responded, 479 (81%) identified that they currently use film and/or video in their teaching. Likewise, when asked about their sense of the pedagogical efficacy of film and video, respondents were extremely positive overall. Participants were asked to rate their agreement with the statement "film and/or video is useful in my teaching" on a 5-point Likert scale. The average rating of agreement with the statement was 4.26 with a median of 5, demonstrating broadly positive attitudes towards the usefulness of film and video as teaching tools. Many comments left by participants at the end of the survey give specificity to these positive perceptions. One instructor stated that "video has tremendous potential" as a teaching tool, for example, and a few participants stated that they expect to integrate more of it into their teaching in the coming years. Many comments praised the use of film and video, both as a tool in specific disciplines (e.g., "essential for teaching History-related courses"), and more widely (e.g., "it can be an extremely effective and powerful learning tool, it is an essential strategy to enhance education"). These findings make clear the perceived pedagogical value of audio-visual texts amongst respondents.

That being said, it should be noted that some participating instructors did not report positive attitudes toward film and video. Given the response bias inherent in a voluntary survey of this type, it is particularly important to look at these less positive responses in order to gain a clearer understanding of perceptions that might exist amongst a broader instructor population. Many respondents who indicated less positive agreement with the statement that film and video could be useful in their teaching justified their beliefs with comments in the next survey question. A line of thinking voiced by many participants is summed up by one comment in particular: "In order for students to benefit fully from videos in class, the professor needs to know the content of the video well and be able then to stimulate a discussion in class. It should not be a passive activity." Based on comments such as this, it is clear that some participants may still have misgivings about the potential for such materials to be used poorly or generate negative outcomes like student passivity.

Similarly, one respondent noted that: "In a 12 week term, it seems a bit like stealing your salary to rely on film in most courses," demonstrating that teaching with film might be seen as an abrogation of teaching duties rather than a meaningful pedagogical choice. Others offered similar responses, arguing that film and video are "counterproductive," "overused," and "relied upon," or take up "too much of limited student contact time". These comments were offered by instructors in humanities, business, health sciences, and engineering, demonstrating that similar beliefs about film and video are held across fields and are not necessarily correlated to subject matter, epistemology, or disciplinary approaches to teaching.

Still, it is useful to compare the data provided by respondents from different disciplinary groupings to determine if other distinctions exist. Within each area of study, results remained positive overall. The lowest rankings of film's utility came from Engineering, where instructors reported an average of 3.95/5 agreement with the statement "film and/or video is useful in my teaching". Respondents from other Faculties reported their agreement with the statement as indicated in Table 2:

Table 2: Mean and median discipline group rankings of agreement with the statement 'film and/or video can be effective tools for teaching and learning in my discipline.'

	Business	Health Sciences	Humanities	Interdisciplinary & Social Sciences	Engineering	Science
Mean	4.39	4.50	4.33	4.24	3.95	4.13
Median	5	5	5	5	4	4

A Kruskal-Wallis H test revealed a statistically significant difference in the mean rankings of usefulness between the different disciplinary groups (□ ^{2}(5)=17.142, p=0.004), with post hoc pairwise comparisons more specifically indicating a significant difference in the mean rankings of Engineering and Health Sciences instructors (adjusted significance, using the Bonferroni correction for multiple tests, p=0.036). This suggests that the mean ranking of film's utility among Engineering respondents was significantly lower than that of respondents from the Health Sciences. No other statistically significant differences in the mean rankings of film's usefulness were found between groups. Nevertheless, the general trend in our findings (Table 2) is consistent with disciplinary differences noted in the literature (Peacock et al., 2018), as respondents from Science fields tended to rank the utility of film somewhat lower than their counterparts in other subject areas. These findings thus offer some further corroboration for the argument that instructors in Engineering and Science may not be as convinced about the usefulness of film and video within their teaching, despite the fact that many faculty in these fields are still positive about film's pedagogical potential .

Purposes

It is important to understand not only if instructors are using film and video in teaching, but also for what purposes they are doing so. Our survey gave participants an option to choose from a list of purposes for which they use film and/or video in their courses, as well as space to expand on their selections by providing comments. The three most commonly selected purposes were: 'to engage student attention,' with 343 responses across two phases of the survey; 'to help students learn course concepts,' with 313 responses; and 'to provide variety in instructional methods,' with 305 responses (See Table 3). Documentaries, user-generated content from sites such as YouTube, and narrative feature films were the three most common types of film/video used to achieve these purposes, with narrative features also commonly used to evoke student emotion.

Generally speaking, respondents across disciplines reported similar reasons for using film and video in their teaching (Table 3). For example, the most commonly selected option for instructors in Health Sciences, Humanities, Interdisciplinary & Social Sciences, and Science was 'to engage student attention', and this was also the second most commonly selected option for respondents from Engineering. Other commonly selected options across disciplines included helping students learn course concepts, providing variety in instructional methods, engaging student attention, stimulating further discussion, and making abstract/theoretical ideas more concrete. These overlapping pedagogical rationales indicate that film and video are used across disciplines to similar ends.

Nevertheless, a few interesting distinctions between disciplines do emerge. Respondents in Engineering appeared far less likely to use film and video to stimulate further discussion than did those in other Faculty groupings, for instance, with only 30% of participants from Engineering selecting this response as compared to more than 60% of respondents from each of the other areas. Engineering instructors also tended to select fewer reasons for using film than did participants from other disciplines. Only three purposes (helping students learn course concepts, engaging student attention, and providing variety in instructional methods) were selected by more than half of Engineering participants, while five or more purposes were chosen by more than 50% of respondents from all other Faculties. The fact that Engineering instructors listed fewer reasons to use film and video may help to explain why these instructors also ranked its pedagogical potential the lowest out of all the Faculties. This could mean that Engineering faculty view film and video slightly less positively because they feel it offers the potential to fulfill fewer meaningful educational goals.

Table 3: Purposes for using film and/or video in teaching. Numbers in parentheses indicate the percentage of instructors that selected the item. Bus=Business; Eng=Engineering; HS=Health Sciences; Hum=Humanities; Inter & SocSci=Interdisciplinary & Social Sciences; Sci=Science; Not Ind=Not indicated.

Purpose	**Bus** (30)	**Eng** (23)	**HS** (41)	**Hum** (140)	**Inter & SocSci** (125)	**Sci** (91)	**Not Ind** (6)	**Total** (456)
To help students learn course concepts	26 (87%)	17 (74%)	26 (63%)	85 (61%)	87 (70%)	68 (75%)	4 (67%)	313 (69%)
	Example: "Flow visualisation films can show the transition from laminar to turbulent flow." (Engineering)							
To make abstract/theoretical ideas more concrete	21 (70%)	11 (48%)	22 (54%)	84 (60%)	89 (71%)	57 (63%)	5 (83%)	289 (63%)
	Example: "Used when disussing [sic] the concept of empathy versus sympathy." (Health Sciences)							
To engage student attention	20 (67%)	14 (61%)	31 (76%)	106 (76%)	92 (74%)	76 (84%)	4 (67%)	343 (75%)
	Example: "For instance: I'll start a class with something eye catching or engaging, or simply play music beforehand -- once the sound goes down, my class knows it's time to start." (Humanities)							
To evoke student emotion	6 (20%)	6 (26%)	17 (41%)	54 (39%)	63 (50%)	28 (31%)	4 (67%)	178 (39%)
	Example: "I use a short video on the impact of polio in teaching a section on diseases and vaccination." (Science)							
To provide variety in instructional methods	21 (70%)	14 (61%)	29 (71%)	97 (69%)	84 (67%)	57 (63%)	3 (50%)	305 (67%)
	Example: "I taught a course in a three hour block format, and would begin with lecture, often screen a short video in the middle, and then discuss the video, at least once every 3-4 weeks." (Social Sciences)							

To provide students with exposure to relevant procedures/experiences	11 (37%)	6 (26%)	16 (39%)	44 (31%)	33 (26%)	28 (31%)	2 (33%)	140 (31%)
	Example: "One of the lab sessions in one of my courses involves the pruning of ornamental trees. Prior to actually going outside and pruning real trees, I show several films on the hows and whys of pruning. I find these extremely useful for demonstrating why it is necessary to prune, and how to go about it properly." (Science)							
To demonstrate the application of course ideas in real world settings	20 (67%)	11 (48%)	25 (61%)	62 (44%)	88 (70%)	59 (65%)	4 (67%)	269 (59%)
	Example: "Feature film "Ingenous" [sic] shows how two entrepreneurs develop their product ideas, fail and eventually are succesful [sic]" (Engineering)							
To indicate the connections between course ideas and current events	19 (63%)	7 (30%)	14 (34%)	57 (41%)	75 (60%)	38 (42%)	5 (83%)	215 (47)
	Example: *"Awkward Black Girl to look at Black Lives Matter movement and debates over double-consciousness per W E B Du Bois." (Social Sciences)*							
To encourage analysis of how certain media types function in society	3 (10%)	0 (0%)	1 (2%)	53 (38%)	29 (23%)	5 (5%)	2 (33%)	93 (20%)
	Example: "comparing theatre and film as media, their intended or actual audiences, their cultural impact." (Humanities)							
To help students develop audio-visual literacies	0 (0%)	2 (9%)	2 (5%)	62 (44%)	27 (22%)	5 (5%)	2 (33%)	100 (22%)
	Example: "I encourage students to look at film as a document or source, and to assess it in the same way they should for fiction, Internet pages, scholarly texts, etc." (Humanities)							
To stimulate further discussion	23 (77%)	7 (30%)	28 (68%)	97 (69%)	83 (66%)	55 (60%)	4 (67%)	297 (65%)
	Example: "film "levels" the discussion quite often -- students will have read the course materials at different depths or unevenly, but showing film often helps students feel more confident." (Social Sciences)							
Other	4 (13%)	4 (17%)	9 (22%)	28 (20%)	20 (16%)	10 (11%)	0 (0%)	75 (16%)
	Example: "To let students actually see authors." (Business)							

Though engaging student attention and providing variety in instructional methods initially seem to be superficial reasons for using film, comments from some instructors make clear that their motivations for choosing to prioritize these goals are important. As one instructor commented, "different students learn best in different ways," suggesting that using film to vary instructional methods might in fact support the further goal of engaging and supporting students with different ability levels or preferred methods of learning. Similarly, comments about using film to engage student attention demonstrate how valuable this process can be to many instructors. While many people said they used videos to "break-up the format of the class" or because "films are more engaging than [their] talking," some intimated that student attention and engagement were precursors to or preconditions for meaningful learning. A respondent from the Health Sciences, for example, noted, "students find videos more engaging and are critical when watching them," positing a relationship between engagement and critical thinking. Likewise, an instructor from the Humanities wrote, "I would show music videos at the start of class that were connected to the ideas that we would be discussing that day. It helped to prime the students for the class." In this case, the instructor seems to view engaging student attention, promoting discussion, learning course concepts, and connecting the coursework to other ideas as potentially achieved simultaneously through the use of film texts.

Several other responses suggest that eliciting student attention might, to some extent, connect largely to promoting student satisfaction. For instance, one participant noted, "the "cool" factor is difficult to bring into class in any other way," while another wrote that students "tend to really like audio-visual aids." Bringing several of these ideas together, many felt that film and video were simply necessary, especially in an age of easy access to the internet, to keep students present and mitigate boredom. Exemplifying this idea, one wrote, "Students are easily bored and resort to cell phones and computers; use of varied teaching approaches helps to minimize this tendency."

Instructors also reported other reasons for using film and video beyond the options we provided in the survey. These included drawing on film and video "as texts to be analyzed," "to illustrate the difference between reading and performing a text," and "to evaluate competencies/abilities." One respondent also noted that film use "brings Indigenous perspectives into class, in own words [sic], so I don't speak for them." The wide range of purposes reported for using film shows how adaptable they may be to different teaching and learning contexts.

As indicated in Table 3, the purposes selected by the fewest instructors were: 'to encourage analysis of how certain media types function in society' (93 responses) and 'to help students develop audio-visual literacies' (100 responses). The fact that these two purposes were selected less commonly suggests that attention to film in many classrooms is largely focused on the film's content and how it relates to other course objectives, rather than on the form and social function of film itself. This appears to be slightly less the case in Humanities and Social Sciences, perhaps not surprisingly given that film, media, and cultural studies courses are housed in these areas. The above purposes were selected much more frequently by respondents from these two Faculties (though still by less than 50% of respondents in each case). These results again affirm some minimal disciplinary variation in the use of film for teaching and learning, while simultaneously underscoring that, across disciplinary groupings, film may largely be viewed as a supplement or a means to a different end, rather than a focus of attention or analysis in its own right.

Challenges

A final major focus of our survey was to understand the challenges participants ascribed to teaching effectively with film and/or video, given the comparative lack of attention to such issues in the existing literature. Participants who reported using film and/or video in their teaching were presented with a list of potential challenges from which to select, as well as the option of adding in further issues not included in the list. The results from this question are presented in Table 4.

Table 4: Challenges connected to using film and/or video in teaching. Numbers in parentheses indicate the percentage of instructors that selected the item. Short forms in heading labels as in Table 3 above.

Challenge	**Bus** (30)	**Eng** (23)	**HS** (41)	**Hum** (140)	**Inter & SocSci** (125)	**Sci** (91)	**Not Ind** (6)	**Total** (456)
Difficulty finding appropriate film/video materials	17 (57%)	17 (74%)	21 (51%)	60 (43%)	74 (59%)	49 (54%)	2 (33%)	240 (53%)
Student oversimplification/misunderstanding of concepts raised in films/videos	2 (7%)	3 (13%)	4 (10%)	44 (31%)	37 (30%)	20 (22%)	0 (0%)	110 (24%)
Student resistance to using film/video in educational contexts	0 (0%)	0 (0%)	2 (5%)	5 (4%)	5 (4%)	1 (1%)	1 (17%)	14 (3%)
Student inexperience with analysing films/videos	3 (10%)	1 (4%)	4 (10%)	55 (39%)	30 (24%)	7 (8%)	0 (0%)	100 (22%)
Your own inexperience with analysing films/videos	4 (13%)	1 (4%)	0 (0%)	12 (9%)	14 (11%)	7 (8%)	0 (0%)	38 (8%)
Technical difficulties attached to showing films/videos in the classroom	14 (47%)	6 (26%)	25 (61%)	74 (53%)	53 (42%)	38 (42%)	2 (33%)	212 (46%)
Inaccessibility of film/video for some students	2 (7%)	1(4%)	7 (17%)	34 (24%)	20 (16%)	11 (12%)	2 (33%)	77 (17%)
Copyright concerns	11 (37%)	7 (30%)	13 (32%)	59 (42%)	36 (29%)	27 (30%)	3 (50%)	156 (34%)
Student passivity during films/videos	7 (23%)	3 (13%)	6 (15%)	43 (31%)	50 (40%)	19 (21%)	2 (33%)	130 (29%)

Other	2 (7%)	3 (13%)	5 (12%)	10 (7%)	5 (4%)	3 (3%)	0 (0%)	61 (13%)
	Examples: "Student resistance to certain films." (Humanities); "Students who leave class when a video is about to be shown!" (Social Sciences)							

As Table 4 illustrates, the most commonly indicated challenge was 'difficulty finding appropriate film/video materials,' with 240 instructors selecting this option. While other challenges (e.g., 'technical difficulties attached to showing film/video in the classroom' and 'copyright concerns') were reported with some frequency, problems with finding relevant material was the only issue mentioned by more than 50% of respondents. At the same time, some potential challenges, such as 'student resistance to using film/video in educational contexts', and 'your own inexperience with analysing films/videos' were rarely selected, each being reported by fewer than 10% of the total pool of respondents. Together, these findings suggest that, across disciplines, survey participants not only view film use positively, but find it comparatively unproblematic to implement. Indeed, one instructor from the Social Sciences named this directly, stating, "I find no challenges." Surely, other participants were not this confident, but the fact remains that challenges seemed to resonate less with our participants than did potential benefits of using film and video.

A similar picture emerges when considering the patterns of response amongst various disciplinary groups. Instructors across subject areas often reported similar challenges with drawing on film and/or video in their teaching. Indeed, the same two issues (technical difficulties screening films and problems finding appropriate material) were the two most frequently selected items for participants in five of our six Faculty groupings. The one exception was respondents from Engineering, for whom problems finding film materials and copyright issues were most noted. In spite of this broad similarity, however, one area of potential disciplinary difference was seen in the number of challenges reported relatively frequently within different disciplinary groups. Instructors from the Humanities and Social Sciences reported a greater number of challenges with slightly higher frequency. For example, six of the challenge options were selected by more than 25% of Humanities instructors, while only three of the options were selected by more than a quarter of respondents from Business, Engineering, Health Sciences, and Science. This provides some preliminary evidence that perceived challenges may be more dispersed or divergent in some areas than others, and/or that instructors in Humanities and Social Sciences may be more aware of or concerned about a greater number of potential barriers to film use.

In written comments, participants highlighted a number of additional issues, ranging from "costs of accessing independent film and video" to "student distraction during video." Time was mentioned frequently, with participants highlighting both "time to find films" and "time for screening," as well as, in some cases, the time required to produce new video material for pedagogical purposes. Notably, many of these issues, much like those most commonly selected from the options provided, tend toward the logistical, indicating that the most prominent challenges for instructors in our study are practical issues rather than concerns about the potential capacity for film to meaningfully support their teaching goals.

Still, echoing the idea that film use might be seen as an 'unserious' pedagogical choice, a few instructors made comments which suggested that perceptions of the potential frivolity of teaching with film made it challenging to use effectively. One participant named "students interpreting a film screening as an 'off day'" as a challenge, for example, while another mentioned "resistance from colleagues." Expanding on this challenge, one respondent wrote:

> This may seem like a "strange" challenge but sometimes I feel like using film/videos is a cop-out, i.e. being used to get out of the work of "lecturing" as an educator. So is this an "ideological/pedagogical" challenge? It gets reinforced in subtle and informal ways between colleagues who, for example, may joke about a class being a "lighter day" because of showing a film, or as blatant as some colleagues suggesting others who use film/video regularly are being "lazy." I am pre-tenure and so I often think about what an ideal "balance" is between using film/video to enhance learning and to make a lecture come alive and more interesting [versus] being constructed as "lazy" and "relying" on film and video.

Such comments, while admittedly rare, gesture toward the ways in which broader student and faculty perceptions of film and video might impinge upon the extent to which they are taken up in the classroom.

Discussion

This study contributes to the existing literature about the use of film and video in university teaching and learning in several ways. Our survey is unique in its attention to instructors' motivations for and methods of using film and video in their teaching, as well as the challenges and obstacles they encounter in doing so. By drawing on a broader range of participants than is typical in research of this sort, including instructors from several institutions and disciplines, it also permits an initial view into the scope and transferability of issues reported in existing scholarship about the pedagogical potential of film and video. While the study is not without limitations, including the self-reported nature of the data, the potential for response bias, and the lack of direct attention to whether or not film is able to support student learning in the ways participants claim, it nevertheless generates a number of provocative considerations that might inform future research and practice.

Foremost amongst these considerations is the clear sense that respondents across disciplines generally appear to think highly of film's potential to support their pedagogical goals. While it is certainly likely that people who feel more positive about teaching with film would be inclined to respond to our survey, the fact remains that a relatively large number of participants, from a variety of disciplinary homes, value film highly as a teaching tool. In this respect, our study corroborates and extends both previous work that assesses instructors' perceptions of popular media within teaching across disciplines (Peacock et al., 2018), and the large body of work discussing film use within particular courses across fields (e.g., Algeo, 2002; McAllister, 2015; Kabooha, 2016; Holland, 2014). While, like Peacock and colleagues (2018), we found some minor variations between respondents from different disciplines, the general sentiment expressed in the data is one of widespread agreement about the vital role film and video can play in university teaching. In spite of the pervasive influence of disciplinary "teaching and learning regimes" (Trowler, 2008), which shape approaches to education in different subject areas, our data indicate that film can be adapted and deployed in broadly related ways in a number of subject areas and teaching contexts.

Perhaps especially interestingly in this regard, our examination of the stated purposes for which instructors use film and/or video in their teaching raises questions about the extent to which these texts are called on to serve significant pedagogical goals, such as the development of deep learning or the promotion of critical thinking (Bright, 2015; Olson, Autry, & Moe, 2016). This issue is particularly compelling given longstanding concerns about the potential for technology to drive pedagogy, rather than supporting it meaningfully (e.g., Ascough, 2002). The present data suggest that, in many cases, film and video usage are underpinned by broadly accepted pedagogical goals, such as helping students understand course content, making abstract concepts more concrete, and supporting critical discussion of ideas (see Andrist et al., 2014; Calcagno, 2015; and, Pelton, 2013, for similar claims). At the same time, the fact that the most commonly reported purposes for film and video use include engaging student attention and providing variety in instructional methods complicates this picture somewhat. Participants' comments about these purposes oscillate between underlining how they view factors like attention and variety as essential precursors to or components of learning, and suggesting, perhaps more problematically, that film and video can increase student satisfaction by making learning more fun. This ambiguity in the data, along with the commonality of responses connected to attention and engagement, indicate the potential value of further research exploring the relationships between media use and student engagement. Future studies might probe more deeply the reasons underpinning instructors' beliefs that interest and variety are important, and explore the ways in which film and video can be used to promote attention, interest, and engagement that connect to course goals rather than simply being entertaining.

Future work of this sort might also help to address the consideration, alluded to in our data as well as in existing literature (e.g., Swimelar, 2013; Travis, 2016), that film and video are perceived as unserious or frivolous and thus not indicative of 'actual' teaching. In spite of the overwhelmingly positive perception of film and video reported by instructors in our study, participants nevertheless noted concerns about students and faculty not taking such texts seriously, or criticizing their use as lazy and unprofessional. In fact, given that our survey likely did not access a large body of instructors who are relatively *less* positive about film and video use, such perceptions might well be widespread. Finding ways to develop and disseminate defensible, evidence-based approaches to using film and video in diverse disciplinary contexts (i.e., approaches that draw on film and video to meaningfully support course goals) might begin to counter such perceptions, while also supporting more effective deployment of these technologies within teaching.

Finally, participants' comments about other challenges they experience while using film and video likewise offer a range of insights into potential strategies for supporting instructors who seek to use film and video as part of their teaching repertoires. Most notably, the prevalence of logistical challenges in our data suggests attention ought to be paid to training, institutional strategies, or the development of resources that could help faculty navigate these practical problems. This is perhaps especially compelling for challenges that were raised commonly in the present study and have also been discussed in the existing literature, such as difficulty accessing appropriate materials (e.g., Huczynski & Buchanan, 2004; Kabooha, 2014), concerns about copyright (Leet & Houser, 2003; Sexton, 2006), and technical issues (Herman, 2006). If institutions wish to support the meaningful use of film and video within courses and programs, for instance, they might ensure that technological infrastructure is up-to-date and faculty are supported to use it well. Likewise, institutions, scholars, practitioners, and educational developers might contribute to assembling databases and bibliographies of relevant texts that can be used for particular pedagogical purposes (see Andrist et al., 2014, for one example of this type of resource development). Of course, the ways in which such texts can be used to actually support student learning in various disciplines must also be assessed more directly than it has to date.

This need for further evidence of student learning in situations where film and video are used is perhaps the most compelling task for future research in this area. While our study offers important insight into *how* and *why* instructors across disciplines draw on film and video in their teaching, and into potential barriers to even more extensive use, the effects of such choices on student learning need more immediate attention. The motivations, goals, and challenges reported here offer useful starting points for such work, making clear intended and perceived outcomes of film use that could be studied directly. Alongside such direct assessments of student learning, the preliminary research into instructor perceptions and motivations described here could also be meaningfully complemented by similar work exploring student perceptions of film/video use and/or by qualitative studies probing instructor and student motivations in more depth. Indeed, given the commonality of integrating film and video into pedagogy across disciplines, evidenced in this study and elsewhere, such research is both timely and pressing.

Appendix

Appendix 1. Survey Questions

1. With which Institution are you primarily affiliated:
   - ☐ Algoma University
   - ☐ Lakehead University
   - ☐ University of Ottawa
   - ☐ University of Windsor
   - ☐ Queen's University
   - ☐ York University

2. In which department(s)/programs do you teach: ________________________

3. How long have you been teaching at the University level:
   - ☐ Dropdown menu with options: Less than 1, 1, 2, … 25, More than 25

4. Do you use film and/or video in your teaching?
   - ☐ Yes
   - ☐ No

If 'yes', proceed to Questions 5a-11a. If 'no', proceed to Questions 5b-10b [The correct questions will display dynamically depending on the response to Q4]

5a. How frequently do you use film and/or video in your teaching?
   - ☐ Multiple times per class
   - ☐ Once per class
   - ☐ Once per week
   - ☐ Once per month
   - ☐ Once per semester
   - ☐ Less than once per semester
   - ☐

6a. In which types of courses/contexts do you use film and/or video? (Select all that apply)
   - ☐ Undergraduate courses of more than 50 students
   - ☐ Undergraduate courses of 50 students or fewer
   - ☐ Graduate courses
   - ☐ Undergraduate supervision
   - ☐ Graduate supervision
   - ☐ Resident training
   - ☐ Other (please specify): _________________

7a. For what purpose(s) do you use film and/or video in your teaching? Please select all that apply, and provide a brief example of how you use film and/or video in a way that achieves this purpose.
   - ☐ To help students learn course concepts (Example: _______________)
   - ☐ To make abstract/theoretical ideas more concrete (Example: _______________)
   - ☐ To engage student attention (Example: _______________)
   - ☐ To evoke student emotion (Example: _______________)

- To provide variety in instructional methods (Example: ______________)
- To provide students with exposure to relevant procedures/experiences (Example: ____)
- To demonstrate the application of course ideas in real world settings (Example: ______)
- To indicate the connections between course ideas and current events (Example: ______)
- To encourage analysis of how media texts function in society (Example: ___________)
- To help students develop audiovisual literacies (Example: ______________)
- To stimulate further discussion (Example: ______________)
- Other (please specify): _________________

8a. Which types of films and/or videos do you use for the purposes selected above? (Select all that apply) [Note: in the online survey instrument, the list of purposes participants select in 7a appears, with checkboxes representing the following options beside each purpose selected]

- Narrative feature films
- Documentaries
- Avant-garde/art films
- Instructional videos created specifically for teaching contexts
- TED talks or other filmed lectures/speeches
- User-generated content on YouTube or similar sites
- None of the types listed here: _________________

9a. Do you use any other types of film and/or video material for these purposes? If so, please list them next to the relevant purpose below:

- To help students learn course concepts: ______________
- To make abstract/theoretical ideas more concrete: ______________
- To engage student attention: ______________
- To evoke student emotion: ______________
- To provide variety in instructional methods: ______________
- To provide students with exposure to relevant procedures/experiences: ______________
- To demonstrate the application of course ideas in real world settings: ______________
- To indicate the connections between course ideas and current events: ______________
- To encourage analysis of how media texts function in society: ______________
- To help students develop audiovisual literacies: ______________
- To stimulate further discussion: ______________
- Other (please specify): _________________

10a. How do you typically incorporate film and/or video into your teaching? (Select all that apply)

- Brief clips/short videos integrated into a lecture
- Brief clips/short videos integrated into in-class group work
- In-class screening and discussion of longer films/videos (30 minutes +)
- Brief clips/short videos integrated into an online module
- Instructor-assigned films to be screened by students out of class
- Assignments that require students to find and view a relevant film/video
- Assignments that require students to analyse an instructor-assigned film/video
- Assignments that require students to produce a film/video text of their own
- Other (please specify): _________________

11a. What challenges, if any, have you experienced in using film/video in your teaching?

- Difficulty finding appropriate film/video materials
- Student oversimplification/misunderstanding of concepts raised in films/videos
- Student resistance to using film/video in educational contexts
- Student inexperience with analyzing films/videos
- Your own inexperience with analyzing films/videos
- Technical difficulties attached to showing films/videos in the classroom
- Inaccessibility of film/video for some students
- Copyright concerns
- Student passivity during films/videos
- Other (please specify): ________________

5b. Have you tried using film/video in your teaching in the past?

- Yes
- No

6b. Why don't you use film/video in your teaching currently?

- Difficulty finding appropriate film/video materials
- Student oversimplification/misunderstanding of concepts raised in films/videos
- Student resistance to using film/video in educational contexts
- Student inexperience with analyzing films/videos
- Your own inexperience with analyzing films/videos
- Technical difficulties attached to showing films/videos in the classroom
- Inaccessibility of film/video for some students
- Copyright concerns
- Film/video isn't relevant to your courses/teaching contexts
- Student passivity during films/videos
- Other (please specify): ________________

7b. In which types of courses/contexts, if any, do you think you might use film and/or video? (Select all that apply)

- Undergraduate courses of more than 50 students
- Undergraduate courses of 50 students or fewer
- Graduate courses
- Undergraduate supervision
- Graduate supervision
- Resident training
- None of the above
- Other (please specify): ________________

8b. For what purpose(s), if any, do you think film and/or video might be useful in your teaching? Please select all that apply, and provide a brief comment to explain.

- To help students learn course concepts: ______________
- To make abstract/theoretical ideas more concrete: ______________
- To engage student attention: ______________
- To evoke student emotion: ______________
- To provide variety in instructional methods: ______________
- To provide students with exposure to relevant procedures/experiences: ______________

- To demonstrate the application of course ideas in real world settings: _______________
- To indicate the connections between course ideas and current events: _______________
- To encourage analysis of how media texts function in society: _______________
- To help students develop audiovisual literacies: _______________
- To stimulate further discussion: _______________
- Film and/or video is not useful to my teaching: ____________
- Other (please specify): __________________

9b. Which types of films and/or videos do you think could be useful (within your teaching contexts) for the purposes described above? (Select all that apply) [Note: in the online survey instrument, the list of purposes participants select in 8b appears, with checkboxes representing the following options beside each purpose selected]

- Narrative feature films
- Documentaries
- Avant-garde/art films
- Instructional videos created specifically for teaching contexts
- TED talks or other filmed lectures/speeches
- User-generated content on YouTube or similar sites
- None of the types listed here: __________________

10b. Do you think any other types of film and/or video might be useful (within your teaching contexts) for these purposes? If so, please list them next to the relevant purpose below:

- To help students learn course concepts: _______________
- To make abstract/theoretical ideas more concrete: _______________
- To engage student attention: _______________
- To evoke student emotion: _______________
- To provide variety in instructional methods: _______________
- To provide students with exposure to relevant procedures/experiences: _______________
- To demonstrate the application of course ideas in real world settings: _______________
- To indicate the connections between course ideas and current events: _______________
- To encourage analysis of how media texts function in society: _______________
- To help students develop audiovisual literacies: _______________
- To stimulate further discussion: _______________
- Other (please specify): __________________

11b. If you were to use film and/or video, how might you incorporate it into your teaching? (Select all that apply)

- Brief clips/short videos integrated into a lecture
- Brief clips/short videos integrated into in-class group work
- In-class screening and discussion of longer films/videos (30 minutes +)
- Brief clips/short videos integrated into an online module
- Instructor-assigned films to be screened by students out of class
- Assignments that require students to find and view a relevant film/video
- Assignments that require students to analyse an instructor-assigned film/video
- Assignments that require students to produce a film/video text of their own
- Other (please specify): __________________

12. Please indicate your agreement with the statement below, using the following 5 point scale:
"Film and/or video can be effective tools for teaching and learning in my discipline"
1 Strongly disagree
2 Disagree
3 Neither agree nor disagree
4 Agree
5 Strongly agree

13. Please provide any other information you'd like to share about your perceptions or experiences of film and video as teaching and learning tools?

References

Algeo, K. (2007). Teaching cultural geography with *Bend it Like Beckham*. *Journal of Geography, 106*(3), 133-143. https://doi.org/10.1080/00221340701632781

Andrist, L., Chepp, V., Dean, P., & Miller, M. V. (2014). Toward a video pedagogy: A Teaching Typology with Learning Goals. *Teaching Sociology*, *42*(3), 196-206. https://doi.org/10.1177/0092055X14524962

Ansell, N. (2002). Using films in teaching about Africa. *Journal of Geography in Higher Educaiton, 26*(3), 355-368. https://doi.org/10.1080/0309826022000019927

Ascough, R.S. (2002). Designing for online distance education: Putting pedagogy before technology. *Teaching Theology & Religion, 5*(1), 17-29. https://doi.org/10.1111/1467-9647.00114

BavaHarji, M., Alavi, Z.K., & Letchumanan, K. (2014). Captioned Instructional Video: Effects on Content Comprehension, Vocabulary Acquisition and Language Proficiency. *English Language Teaching*. http://dx.doi.org/10.5539/elt.v7n5p1

Ber, R., & Alroy, G. (2002). Teaching professionalism with the aid of trigger films. *Medical Teacher*, *24*(5), 528–531. https://doi.org/10.1080/0142159021000012568

Blasco, P.G. & Moreto, G. (2012). Teaching empathy through movies: Reaching learners' affective domain in medical education. *Journal of Education and Learning, 1*(1), 22-34. http://dx.doi.org/10.5539/jel.v1n1p22

Bright, A. (2015). Why are we watching funny videos in our pedagogy course? Deconstructing humorous videos to foster social activism in educators. *European Journal of Humour Research 3*(4), 36-53. https://pdxscholar.library.pdx.edu/edu_fac/95/

Brown, S.P., Smith, J.W., McAllister, M., Joe L. (2017). Superhero physiology: the case for Captain America. *Advances in Physiology Education*, *41*, 16-24. https://doi.org/10.1152/advan.00106.2016

Calcagno, J. (2015). A current media approach to learning and teaching in psychology of women courses. *Psychology of Women Quarterly, 39*(2), 268-271. https://doi.org/10.1177/0361684314566406

Datta, V. (2009). Madness and the movies: An undergraduate module for medical students. *International Review of Psychiatry, 21*(2), 261-266. https://doi.org/10.1080/09540260902748001

Epstein, D., Mendick, H. & Moreau-M.P. (2010). Imagining the mathematician: Young people talking about popular representations of maths. *Discourse: Studies in the Cultural Politics of Education, 31*(1), 45-60. https://doi.org/10.1080/01596300903465419

García, J. (2015). Learning from bad teachers: The neoliberal agenda for education in popular media. *Critical Education, 6*(13). http://ices.library.ubc.ca/index.php/criticaled/article/view/184935

Giroux, H. (2004). Pedagogy, film, and the responsibility of intellectuals: A response. *Cinema Journal, 43*(2), 119-127. https://www.jstor.org/stable/1225923

Giroux, H. (2008). Hollywood film as public pedagogy: Education in the crossfire. *Afterimage, 35*(5), 7-13.
Giroux, H. (2009). The Iranian uprisings and the challenge of the new media: Rethinking the politics of representation. *Fast Capitalism, 5*(2). http://www.uta.edu/huma/agger/fastcapitalism/5_2/Giroux5_2
Greenberg, H.R. (2009). Caveat actor, caveat emptor: Some notes on the hazards of Tinseltown teaching. *International Review of Psychiatry, 21*(3), 241-244. https://doi.org/10.1080/09540260902746773
Happel-Parkins, A., Esposito, J. (2015). Using popular culture texts in the classroom to interrogate issues of gender transgression related bullying. *Educational Studies*, *51*(1), 3-16. https://doi.org/10.1080/00131946.2014.983638
Herman, J.W. (2006). Using film clips to enhance nursing education. *Nurse Educator, 31*(6), 264-269.
Hobbs, R. (2006). Non-optimal uses of video in the classroom. *Learning, Media, and Technology, 31*(1), 35-50. https://doi.org/10.1080/17439880500515457
Holland, J. (2014). Video use and the student learning experience in politics and international relations. *Politics, 34*(3), 263-274. https://doi.org/10.1111/1467-9256.12022
Huczynski, A. & Buchanan, D. (2004) Theory from fiction: A narrative process perspective on the pedagogical use of feature film. *Journal of Management Education*, *28*(6), 707-726. https://doi.org/10.1177/1052562903262163
Johnstone, K., Marquis, E., & Puri, V. (2018). Public pedagogy and representations of higher education in popular film: New ground for the scholarship of teaching and learning. *Teaching & Learning Inquiry, 6*(1). https://doi.org/10.20343/teachlearninqu.6.1.4
Kabooha, R.H. (2016). Using Movies in EFL Classrooms: A Study Conducted at the English Language Institute (ELI), King Abdul-Aziz University. *English Language Teaching*, 9(3), 248-257. https://files.eric.ed.gov/fulltext/EJ1095569.pdf
Kirkpatrick, M.K. & Brown, S. (2004). Narrative pedagogy: Teaching geriatric content with stories and the 'Make a Difference' Project (MADP). *Nursing Education Perspectives, 25*(4), 183-187.
Kuzma, L.M. & Haney, P.J. (2001). And… action! Using film to learn about foreign policy. *International Studies Perspectives, 2*(1), 33-50. https://doi.org/10.1111/1528-3577.00036
Leet, D. & Houser, S. (2003). Economics goes to Hollywood: Using classic films and documentaries to create an undergraduate economics course. *Journal of Economic Education, 34*(4), 326-332. https://doi.org/10.1080/00220480309595226
Luccasen, R.A., & Thomas, M.K. (2010). Simpsonomics: Teaching economics using episodes of *The Simpsons*. *The Journal of Economic Education*, *41*(2), 136–149. https://doi.org/10.1080/00220481003613847
Lumlertgul, N., Kijpaisalratana, N., Pityaratstian, N., & Wangsaturaka, D. (2009). Cinemeducation: A pilot student project using movies to help students learn medical professionalism. *Medical Teacher*, *31*(7), e327–e332. https://doi.org/10.1080/01421590802637941
Madsen, K.D. (2014). Blue Indians: teaching the political geography of imperialism with fictional film. *Journal of Geography, 113*(2), 47-57. https://doi.org/10.1080/00221341.2012.759994
Marcus, A.S. & Stoddard, J.D. (2007). Tinsel Town as Teacher: Hollywood Film in the High School Classrooms. *The History Teacher*, 40(3): 303-330. https://www.jstor.org/stable/30036826
Marcus, A.S., & Stoddard, J.D. (2009). The Inconvenient Truth about Teaching History with Documentary Film: Strategies for Presenting Multiple Perspectives and Teaching Controversial Issues. *The Social Studies*, *100*(6), 279–284. https://doi.org/10.1080/00377990903283957
Marquis, E. (2018). Beautiful minds and unruly bodies: Embodiment and academic identity in *Still Alice* and *The Theory of Everything*. *Discourse: Studies in the Cultural Politics of Education*, *39*(6), 829-840. https://doi.org/10.1080/01596306.2017.1301881

McAllister, M. (2015). Connecting narrative with mental health learning through discussion and analysis of selected contemporary films. *International Journal of Mental Health Nursing, 24*, 304-313. https://doi.org/10.1111/inm.12134
McConville, S.A. & Lane, A.M. (2006). Using on-line video clips to enhance self-efficacy toward dealing with difficult situations among nursing students. *Nurse Education Today, 26*, 200-208. https://doi.org/10.1016/j.nedt.2005.09.024
Myers, J.Y. & Abd-El-Khalick, F. (2016). 'A ton of faith in science!' Nature and role of assumptions in, and ideas about, science and epistemology generated upon watching a sci-fi film. *Journal of Research in Science Teaching, 53*(8), 1143-1171. https://doi.org/10.1002/tea.21324
Olson, J.S., Autry, L., & Moe, J. (2016). Moving beyond 'bookish knowledge': Using film-based assignments to promote deep learning. *Adult Learning, 27*(2), 60-67. https://doi.org/10.1177/1045159515595807
Peacock, J., Covino, R., Auchter, J., Boyd, J., Klug, H., Laing, C., & Irvin, L. (2018). University faculty perceptions and utilization of popular culture in the classroom. *Studies in Higher Education, 43*(4), 601-613. https://doi.org/10.1080/03075079.2016.1180673
Pelton, J. (2013). Seeing the theory is believing: Writing about film to reduce theory anxiety. *Teaching Sociology 41*(1), 106-120. https://doi.org/10.1177/0092055X12462142
Sealey, K.S. (2008). *Film, Politics, & Education: Cinematic Pedagogy Across the Disciplines*. Peter Lang.
Seferoglu, G. (2008). Using feature films in language classes. *Educational Studies, 34*(1), 1-9. https://doi.org/10.1080/03055690701785202
Sexton, R.L. (2006). Using short movie and television clips in the economics principles class. *Journal of Economic Education, 37*(4), 406-417. https://doi.org/10.3200/JECE.37.4.406-417
Sigler, T. & Albandoz, R.I. (2014). Beyond representation: Film as a pedagogical tool in urban geography. *Journal of Geography, 113*, 58-67. https://doi.org/10.1080/00221341.2013.764918
Swimelar, S. (2013). Visualizing international relations: Assessing student learning through film. *International Studies Perspectives, 14,* 14-38. https://doi.org/10.1111/j.1528-3585.2012.00467.x
Travis, M. (2016). Teaching professional ethics through popular culture. *The Law Teacher, 50*(2), 147-159. https://doi.org/10.1080/03069400.2015.1048500
Trowler, P. (2008). Beyond epistemological essentialism: Academic tribes in the twenty-first century. In C. Kreber (Ed.), *The university and its disciplines: Teaching and learning within and beyond disciplinary boundaries* (pp. 181-195). New York: Routledge.
Volk, S.S. (2015). How to navigate an 'upside-down' world: Using images in the history classroom. *New Directions for Teaching & Learning, 141*, 53-65. https://doi.org/10.1002/tl.20122

Journal of the Scholarship of Teaching and Learning, Vol. 20, No. 1, April 2020, pp.151-154.
doi: 10.14434/josotl.v20i1.25029

Class Assignments that Promote Openness to Diversity among Undergraduates at Predominantly White Universities

Jill K. Walls
Ball State University

Jennifer M. Weaver
Boise State University

Abstract: In this paper we provide two concrete examples of class assignments, developed by the authors, which challenge students to think more critically about themselves, their cultural values and beliefs, and the "invisible forces" that shape their perspectives. We include samples from students' work and our reflections to demonstrate the types of outcomes instructors might see if they utilized similar assignments in their classes. We emphasize the importance of diversity-focused assignment for programs operating within the broader context of a predominantly White university and provide suggestions for expanding our assignments to promote cultural competence among students planning to work directly with families and children.

Recent studies utilizing nationally representative survey data suggest that a significant proportion of freshmen begin their college careers with limited exposure to diverse populations and perspectives (Eagan, Stolzenberg, Bates, Aragon, Suchard, & Rios-Aguilar, 2015). It is not uncommon, for example, for first-year students to come from relatively homogenous communities, which can lead them to hold ethnocentric worldviews, minimize cultural differences, and lack an awareness of their own biases (Sandell & Tupy, 2015; Eagan et al., 2015). To prepare students for careers in a global society, many colleges and universities require coursework related to diversity as part of their undergraduate curriculum (Association of American Colleges & Universities, 2016). College provides students with an opportunity to explore diverse perspectives and, in some cases, exposure to information that challenges their preexisting belief systems about others and themselves. As a result, the college years can be transformative for many young adults; previously constructed identities and belief systems, based on one's cultural background and upbringing, can dramatically shift (Torres, 2009).

As teacher-scholars in the fields of family studies and child development, we are responsible for preparing our graduates to work directly with diverse families and professionals in their future careers. Transformative learning theory provides a useful framework for understanding the goals and potential utility of our class assignments because of its emphasis on broadening one's perspectives through critical reflection. It is theorized that transformative learning begins with students' lived experiences and develops from that reference point (Mezirow, 1997). The concept of *frame of reference* is defined as "the structures of assumptions through which we understand our experiences. They selectively shape and delimit expectations, perceptions, cognition, and feelings" (Mezirow, 1997, p. 5). From this perspective, we assume that our students' frame of reference when learning new content is rooted firmly in their family of origin experiences (Mezirow, 1991). The assignments described in this paper encourage students to think deeper about their frame of reference pertaining to family life practices and reflect on their assumptions and preferences. Both assignments were developed with a consideration for the broader contexts in which we both teach – predominantly White universities.

The Standpoint Paper – Developed by author 1

This assignment was developed for Family Relations, an undergraduate introductory family studies course that focuses on family dynamics and interactions across a variety of contexts. I assign this paper within the first three weeks of class because it aligns with the second textbook chapter titled "Social Status: Sex, Gender, Race, Ethnicity, and Social Class" (Seccombe, 2015). Prior to this assignment, we review the concept of social stratification and discuss how power and privilege shape experiences and opportunities within each category. Approximately 20 minutes of class time is devoted to reviewing the instructions, providing students with concrete examples, and responding to questions.

The term *standpoint* comes from Feminist Standpoint Theory, which traditionally has been used to understand and validate the perspectives of individuals in historically marginalized groups (Harding, 2004). For the purposes of this assignment, the term standpoint is defined as one's unique perspective of the world, and specifically of family life. Your standpoint guides how you think things *should* be and how you evaluate people and situations around you. The primary learning objective for this assignment is for students to gain a greater awareness of their beliefs about one aspect family life and an understanding of how those beliefs are shaped, in part, by their social identities. In addition, students are prompted to think beyond their personal experiences to speculate about how their beliefs might be different if they were members of other social groups. I have found that students from all backgrounds benefit from this assignment and the self-reflection that it requires.

The standpoint paper assignment is a two- to three-page paper about the effects of their social identities on their beliefs about one specific aspect of family life/relationships. First, students must identify their location in two social status categories (e.g., gender, race/ ethnicity, social class, sexual orientation, religious affiliation, or family structure). After identifying two social status categories that are salient to them, students explain how their position in those categories has shaped their beliefs about one topic related to family life. If students have a difficult time selecting a topic, I offer the following questions to get them thinking:

- When is the right time to start dating, get married, and have children?
- Should couples live together before they get married?
- What is "good" parenting?
- What is appropriate discipline?
- Is divorce an acceptable option for unhappily married couples?
- Under what conditions should mothers work outside the home?
- How should couples divide housework and childcare responsibilities?
- Should same-sex couples be allowed to marry/ raise children?
- What are the essentials of a healthy relationship?
- How should important decisions be made within families?

Common topics chosen by my students include beliefs about same-sex marriage, premarital sex, parenting practices/styles, and cohabitation. For example, one student wrote about how his sexual orientation (homosexual) and religious affiliation (Christian) shaped his beliefs about homosexuality and same-sex relationships.

Lastly, students are asked to speculate about how their beliefs might be different from a different standpoint. For example, a male student could discuss how his beliefs about maternal employment might be different if he was born a female. I have found that when students consider a standpoint other than their own, the unique influence of social identities and how they are tied to power and privilege become even more apparent. Papers are assessed with five criteria in mind: (1) identification of two social status categories, (2) description of one's beliefs about a family life topic,

(3) explanation of how social statuses shape beliefs, (4) speculation about a different standpoint, and (5) quality of writing.

Is a Bath Just a Bath? Activity – Developed by author 2

The baby bath time activity proceeds in a series of three phases and takes one 75-minute session to complete. This activity is introduced during the second week of an upper-division psychology course on Multicultural Children and Families. There are typically 30-40 students enrolled in the course. At this point in the semester, students have completed an introductory reading from Meredith Small's (1998) book, *Our Babies, Ourselves: How biology and culture shape the way we parent.* The primary learning objective of this activity is for students to recognize how culture shapes the most basic interactions we have with others, and to reflect upon how their own daily lives are influenced by embedded cultural values, mores and expectations.

On the day the activity takes place, students are placed in small discussion groups of approximately 4-6 students. In this first phase, students are given a discussion guide (Appendix A) and a baby "bath guide." The bath guide is taken from a popular American parenting website, "The Bump" (https://www.thebump.com/a/how-to-bathe-baby), and presents, in great detail, advice for parents on how to bathe their young child. Students spend time answering questions about the guide and the American values that might lead to the creation of such a guide.

Following the small group discussion, I show a short video clip from Margaret Mead's work that is available on YouTube (Bathing Babies in Three Cultures; https://www.youtube.com/watch?v=rmvqdDBSY0k). The video provides an opportunity to observe baby bath time in the United States, New Guinea and Bali. Although the footage is dated, it provides an interesting point from which to begin a large course discussion about the role of culture in shaping daily interactions and how cultural values are embedded in these practices. The American bath time depicted in the film is remarkably similar to current day practices. The bath practices in New Guinea and Bali, are, perhaps not surprisingly, quite different from American practices, which allows students to see and reflect upon how such a basic activity can be carried out in such diverse ways. To draw out these points, I then lead a large group discussion after the presentation of the film. In this classroom discussion, I have two primary goals: One is to allow students to come to an understanding of the role of culture in their own lives, and to draw their awareness toward how cultural biases can be difficult to detect. In discussing the guide as a larger group, I try to guide them in seeing parallels to their own experiences and expectations about bath time, which they are sometimes resistant to recognize. A second goals is to get students to recognize how cultural values and parenting goals guide behaviors that parents engage in with their children, and that this happens on a daily basis through simple interactions between parents and children.

Conclusion

It is our hope that through active participation in these assignments, students' meaning structures will be broadened such that they will be more willing to accept/respect deviations from their own beliefs/practices, with a deeper understanding that 1) their own perspectives have been shaped by their unique and personalized experiences, 2) not all people have had those same experiences, and 3) variations in cultural context will naturally lead to variations in human beliefs and behaviors. We acknowledge that one class assignment might not be the ultimate transformative "trigger," but perhaps one of several assignments that collectively broaden students' perspectives over the years.

Appendix

Appendix 1. Discussion Guide for the "Is a Bath Just a Bath" Activity
Is a bath just a bath?

Cultural Variations in Bathing Infants Instructions: Review the "Bathing Guide" that is from a popular parenting website in the US. Then, gather with 2-3 others and share your reactions to the guide. Nominate a note-taker in your group to answer the questions, and a spokesperson who will report back on what was discussed in your group.

1. Briefly share your initial reactions to the guide.
2. Analyze this Bathing Guide for evidence of cultural values. Discuss and answer the following questions;
3. Thinking of Meredith Small's Introduction, what aspects of this guide relate to basic biological necessity vs. what relates to "culturally driven directives"?
4. What does this guide say about "American" parenting?
5. How do you think people from, say, a non-Western society might react to this guide? Why?
6. What values do you think are being subscribed to in this guide? What type of parent-child interaction is being encouraged? Support your answer with examples.
7. Is this similar or different to how you remember being bathed as a child? Explain and similarities or differences. Is this how you would expect to (or do you) bathe your own child?
8. If you had to use 1-2 words to describe American bath time, what would they be?

References

Bensimon, E. M., Dowd, A. C., & Witham, K. (2016). Five principles for enacting equity by design. *Diversity and Democracy, 19* (1). Association of American Colleges & Universities. https://www.aacu.org/diversitydemocracy/2016/winter/bensimon

Eagan, K., Stolzenberg, E. B., Bates, A. K., Aragon, M. C., Suchard, M. R., & Rios-Aguilar, C. (2015). The American freshman: National norms fall 2015. Los Angeles: Higher Education Research Institute, UCLA.

Harding, S. G. (Ed.). (2004). *The feminist standpoint theory reader: Intellectual and political controversies*. Psychology Press.

Mezirow, J. (1991). *Transformative dimensions of adult learning*. Jossey-Bass, 350 Sansome Street, San Francisco, CA 94104-1310.

Mezirow, J. (1997). Transformative learning: Theory to practice. *New directions for adult and continuing education*, *1997*(74), 5-12. DOI: 10.1002/ace.7401

Sandell, E. J., & Tupy, S. J. (2015). Where cultural competency begins: Changes in undergraduates' intercultural competency. *International Journal of Teaching and Learning in Higher Education, 27*, 364-381.

Seccombe, K. (2015). *Exploring marriages and families (2nd Edition)*. Boston, MA: Pearson.

Small, M. F. (1998). *Our babies, ourselves: How biology and culture shape the way we parent*. New York, NY: Anchor Books.

Torres, V. 2009. The developmental dimensions of recognizing racist thoughts. *Journal of College Student Development, 50*(5), 504-520.

Journal of the Scholarship of Teaching and Learning, Vol. 20, No. 1, April 2020, pp.155-157.
doi: 10.14434/josotl.v20i1.24589

Breaking Through Student Bias with Creative Debate Assignments

Daniel M. Settlage
University of Arkansas-Fort Smith
dan.settlage@uafs.edu

Abstract: This article outlines a method used to successfully breakthrough student bias surrounding controversial issues in the classroom. The method uses a debate with randomized position assignments to encourage students to consider all sides of the topic. Student feedback is overwhelmingly positive and students appear to develop a newfound respect and deeper understanding of the assigned subject.

Keywords: debate, student bias, active learning, cooperative learning

Motivating students to become truly academically engaged with a topic while challenging their preconceived view of the word is a difficult task. Given the increasingly politicized and pundit-filled news cycle, inspiring a deep and critical analysis of both sides of a controversial issue is a challenging, yet vital, endeavor. Student opinions are often shaped more by their parents and peer groups than by a detailed factual analysis of the issues. The rigidity of their opinions frequently persists despite being confronted with additional, and often contrary, factual evidence. In a world full of 'fake news' it is more important than ever to endow our students with key critical thinking skills.

The job of an academic in the classroom is to break through this dogmatism and engage students in a thought process that removes barriers to learning and delves into all facets of a topic. In order to do this, I have experimented with the use of student debates in an upper division natural resource economics class. Debates have a long history of use in an academic setting. Debates as a teaching strategy dates back to ancient times, although modern high school and college students are rarely exposed to debates outside of participation on a debate team (Kennedy, 2007 and Kennedy, 2009). Research on debate as a pedagogical technique focuses on the role debate plays in fostering critical thinking skills in an active learning environment. In the healthcare field, Garrett and Hood (1996), Darby (2007), and Randolph (2007) all find debate as a teaching and learning technique is ranked highly by students. Kennedy (2007 and 2009) finds that debates foster active learning and students report an increase in critical thinking skills following their exposure to debates. Scott (2008) echoes these observations in a technological setting. Very little research has been done regarding debates in an economic setting, with Pernecky (1997) and Vo and Morris (2006) being the lone exceptions. Pernecky (1997) finds that research, written and oral communication, and critical thinking skills are all enhanced by incorporating debates into the economic classroom. Vo and Morris (2006) show that not only do debates engage students via active learning, but students find debate assignments helpful in the process of learning economics.

Natural resource economics is an ideal platform for debates as it is filled with a number of issues for which people often hold strong opinions. Drilling in the Arctic National Wildlife Refuge, delisting gray wolves from the Endangered Species Act, dam removal and habitat restoration on the Snake River in the Columbia River basin, off-highway vehicle use policy on public lands, and the silvicultural practice of clear-cutting on public lands are all controversial and multi-faceted topics that I have explored in this class via student debates.

The assignment I use is a modification of the popular *Public Forum Debate*, formalized in 2002 (Cossette, 2011). Groups of three or four students are assigned a current topic related to resource economics issues covered in class. Having students debate in teams rather than individuals engages the critical thinking and student learning benefits accorded to cooperative learning assignments (e.g. Slavin, 1996; Johnson and Johnson, 1999; Tsay and Brady, 2010; Yamarik 2007).

Students are instructed on the format, preparation, and grading rubric for the debate. Over the years, I have experimented with and fine-tuned the debate assignment. I initially assigned positions (either pro or con) by a coin flip on the day of the debate. The intent was to force each team to prepare a case for both sides of the debate, thus making them deeply consider both sides of the issue. To augment this debate preparation, groups were required to submit two graded position papers (both pro and con) with cited sources. These papers formed the foundation of the debate preparation and allowed an opportunity for the instructor to intervene if the group veered off track. The position papers formed a valuable component of the debate assignment, as various viewpoints and the validity of the group's sources could be discussed.

Although this was an effective setup, students balked at the last minute uncertainty regarding their debate positions, and the debates seemed to lack the 'passion' that truly engaged students often exhibit. In an effort to alleviate these issues, I switched to giving teams assigned positions determined by coin flip well before the debate. I continued to require the team pro and con position papers. In addition, I required a one-page bulleted debate plan, which included a summary of the team's main debate points, anticipated rebuttals, and closing debate points. These assignments help nudge students toward creating well-reasoned opening statements, rebuttals, and closing arguments. These changes lead to a more motivated and well-prepared debate experience.

Prior to the debates, class and library time was given for teams to meet and research their topic, both as a team and with the instructor. This allowed the instructor to interact with each team to make sure they were doing the necessary work and to answer any questions they may have. The team/instructor meetings are the perfect venue for the instructor to review the sources of information the team is currently using and to discuss ways of assessing the validity of those sources. Quite frequently, students gravitate towards weak or biased sources of information, and the one-on-one time is invaluable in teaching proper researching techniques.

The debate itself is worth a relatively modest 5% of the overall course grade. The format consists of an opening statement (two minutes for each team), rebuttal (three minutes for each team), group question and answer session where the class and instructor ask directed questions (three minutes total), and a closing statement (one minute for each team). There are 10 points awarded for each section plus an additional 10 points for full group participation during the debate. This totals 50 points for the assignment in a class with 1,000 possible points.

Not many students show up to a typical class eager to take a test or quiz. On debate day students invariably show up to class professionally dressed, meticulously prepared, excited and motivated. Demonstrated student interest, engagement, and effort expended on this assignment far exceed the 'value' the assignment carries in terms of its contribution to the overall grade in the class. In post-debate debriefs with students, most indicated they entered the assignment with a definite viewpoint regarding the issue they were assigned and fervently hoped they were assigned the side they believed in. After the debates, many students acknowledge they now have a new found understanding of all sides of the issue. By forcing students to examine issues in a critical light, they become more well-rounded and well-reasoned about their own beliefs. Even if their point of view regarding the issue did not change, they still garner both an appreciation of other viewpoints and a factual basis of support for their own position. Although students indicate they were initially very hesitant when first hearing about the debate assignment, they frequently cite the experience as one of the highlights of the course at the end of the semester. In the age of the five second politicized sound bite, it is quite an accomplishment to engage students in such deep level of thought, even if they were 'tricked' into doing it by a combination of a crafty assignment and their own competitive nature.

References

Cossette, N.A. (2011). *The Art of Debate.* (12th ed.). Lulu.com.

Darby, M. (2007). Debate: A Teaching-Learning Strategy for Developing Competence in Communication and Critical Thinking. *Journal of Dental Hygiene, 81(4)*, 1-12.

Garrett, M, Schoener, L, & Hood, L. (1996). Debate: A Teaching Strategy to Improve Verbal Communication and Critical-Thinking Skills. *Nurse Educator, 21(4)*, 37-40.

Johnson, D.W., & Johnson, R.T. (1999). Making cooperative learning work. *Theory Into Practice, 38(2)*, 67-73.

Kennedy, R. (2007). In-Class Debates: Fertile Ground for Active Learning and the Cultivation of Critical Thinking and Oral Communication Skills. *International Journal of Teaching and Learning in Higher Education, 19(2)*, 183-190.

Kennedy, R. (2009). The Power of In-Class Debates. *Active Learning in Higher Education 10(3)*, 225-236.

Pernecky, M. (1997). Debate for the Economics Class-and Others. *College Teaching, 45(4)*, 136-139.

Randolph, D.S. (2007). Student Perceptions of the Use of Debate as a Teaching Strategy in the Allied Health Professions. *Journal of Allied Health, 36(1),* 13-29.

Scott, S. (2008). Perceptions of Students' Learning Critical Thinking through Debate in a Technology Classroom: A Case Study. *Journal of Technological Studies, 34(1)*, 39-44.

Slavin, R.E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology, 21(1),* 43-69.

Tsay, M., & Brady, M. (2010). A case study of cooperative learning and communication pedagogy: Does Working in Teams Make a Difference? *Journal of the Scholarship of Teaching and Learning, 10(2)*, 78-89.

Vo, H.X. & Morris, R.L. (2006). Debate as a Tool in Teaching Economics: Rationale, Technique, and Some Evidence, *Journal of Education for Business, 81(6)*, 315-320.

Yamarik, S. (2007). Does Cooperative Learning Improve Student Learning Outcomes? *The Journal of Economic Education, 38(3)*, 259-277.

Journal of the Scholarship of Teaching and Learning, Vol. 20, No. 1, April 2020, pp.158-162.
doi: 10.14434/josotl.v20i1.24879

Wicked Problems through a New Lens: Combining Active Learning Strategies for Solutions-Oriented Teaching

Katherine E. Bishop-Williams
University of Guelph
University of Waterloo
kbishop@uoguelph.ca

Abstract: Wicked problems are large, complex problems involving multiple perspectives that present substantial future challenges. These challenges can be overwhelming for learners and pose difficulties in teaching for instructors. Herein a solutions-oriented teaching strategy that amalgamates proven active learning strategies is presented along with a step-by-step guide and materials list. Evidence of student learning is provided. This strategy provides students the opportunity to view complex, wicked problems from multiple perspectives and to visualize their role in future solutions.

Keywords: wicked problems, active learning, solutions-oriented

Teaching learners about "wicked problems" (Waltner-Toews, 2017) can position them to feel that the problems they face are insurmountable and impossible to solve. Wicked problems are complex issues involving many stakeholders and multiple perspectives, and often have incomplete problem definitions (Waltner-Toews, 2017). Climate change and health both present wicked problems, and together present one of the greatest challenges (Costello et al., 2009) and opportunities (Watts et al., 2015) of the twenty-first century.

As educators, it is critical to prepare learners to tackle these complex problems. Indeed, it may be the most important role of today's educators. However, limited research exists to inform educators of the best ways to teach about these problems, and to prepare students for these tasks. Teaching students about wicked problems (such as climate change) can have a negative impact on student mental health as existential concerns about humans and the environment can induce anxiety and depression (Maxwell & Blashki, 2016). Thus, to consider student needs and student mental health while simultaneously educating future leaders and global-citizens, educators should engage in solutions-oriented teaching of wicked problems. Solutions-oriented teaching will help learners consider ways to address these issues, and may empower them while also reducing some anxiety and depression.

This paper outlines an innovative approach to using well-established active learning strategies simultaneously to encourage students to be solutions-oriented in the classroom. Traditionally, think-pair-share activities involve an opportunity for learners to independently consider the answer to a question, then share ideas with a colleague, before the discussion is presented to the larger group (Silberman, 1996), theoretically creating an environment where learners are more comfortable and prepared to contribute in large group settings. Dotmocracy is a form of cumulative voting that allows students to vote for their favourite ideas by way of stickers on chart-paper (Diceman, 2010). Finally, personal reflections are short free-writing exercises that encourage students to consider their personal and professional connection to the material or lesson, completed either in-class or after class (Silberman, 1996).

The case study for this activity was a third-year Bachelor of Health Science and Bachelor of Public Health course on the Ecological Determinants of Health, designed following a recommendation from the *Working Group on the Ecological Determinants of Health in Canada* (Hancock, Spady, & Soskolne, 2016). The class consisted of 170 students. The activities (think-pair-share, dotmocracy, and personal reflections) were amalgamated into a large multi-phase activity (Figure 1). The first phase of the activity began with a broad question: *What can you do to protect the environment?* Students were given a few minutes to think about the question alone and were encouraged to write a list of ideas they had come up with. Next, students shared their ideas with a person sitting nearby. With this person, they were encouraged to broaden their scope to actions that other stakeholders could implement to protect the environment. After approximately 5 minutes working in pairs, students were given instructions for phase 3. Phase 3 of the activity involved students moving around the classroom to record their ideas on large flip-chart papers under headings for the various stakeholders involved in the issue. Students were encouraged to write, sketch, or depict their ideas in any way they felt comfortable. Students were given approximately 10 minutes to record their ideas and generate additional ideas to answer the question. The next phase of the activity involved providing students with stickers to mark the ideas that they shared, the ideas they liked the best, or the ideas that inspired them. At this point students were asked to return to their seats and a brief group discussion facilitated by the instructor focused on key observations and messages ensued. Finally, students were given approximately 15 minutes to reflect on the lesson and the activities. Students were encouraged to write freely, focused on ideas rather than on spelling or grammar. Prompting questions for the reflection included: *What were some of the most popular ideas? What inspired you? What, if anything, will you do differently because of this activity?*

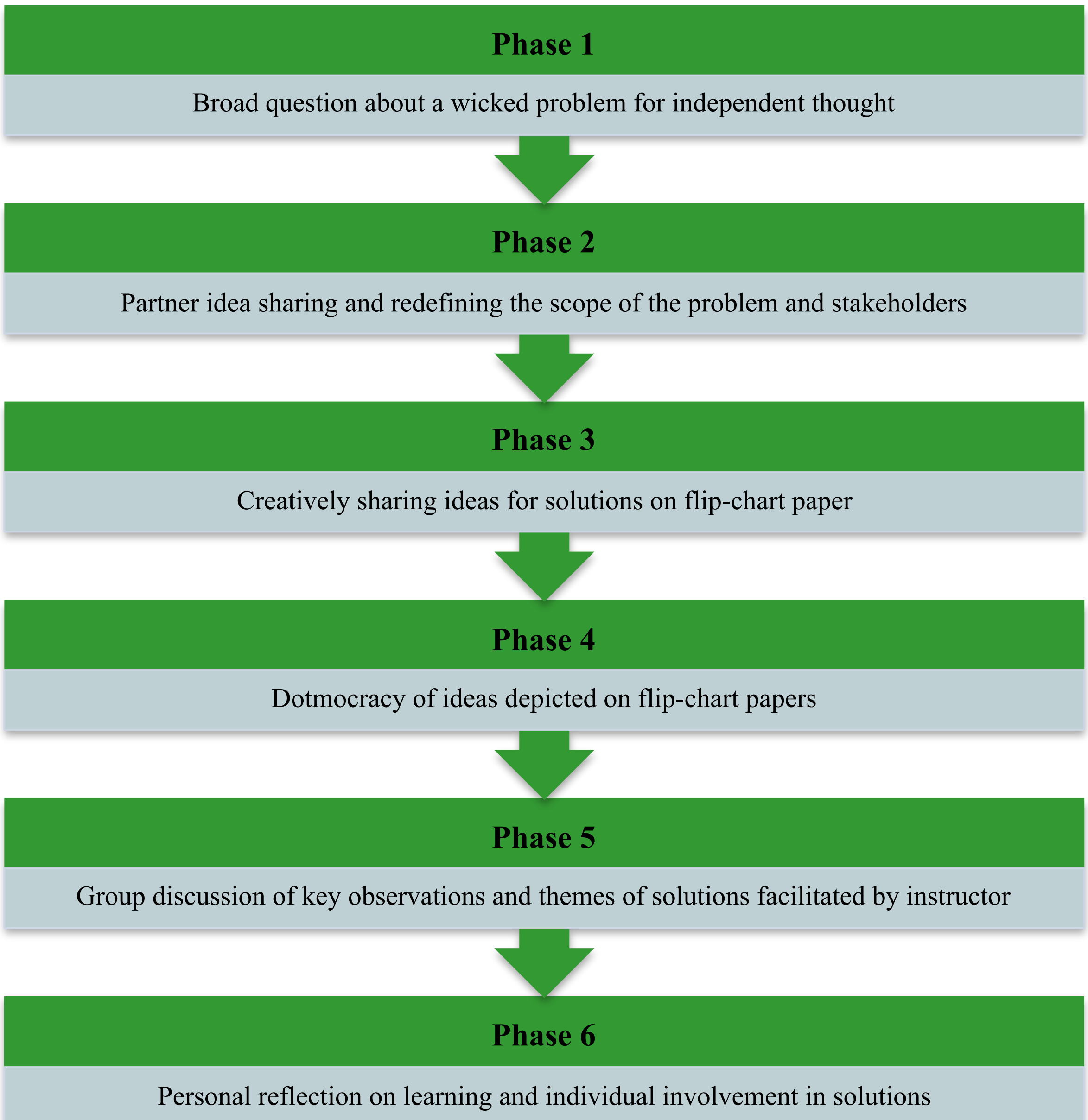


Figure 1. Step-by-step activity guide. Flow chart depicts the process for a multi-phase solutions-oriented teaching activity incorporating elements of think-pair-share, dotmocracy, and personal reflection.

A number of materials are needed for this activity. Educators will need flip-chart paper, tape, markers, and sticker dots. Students should come prepared with a pen and paper or a device that they can record their reflections on. Finally, the classroom must allow for students to move freely around the room. The case study class was held in a large lecture hall with row desks; however, sufficient aisle space made for a successful activity.

When reflecting on the lessons they had learned during the class, many students commented on the activity as well. The ideas expressed in the feedback are represented in a word-cloud (Figure 2). Key themes illustrated in the word-cloud include enjoyment, inspiration, engagement, interactive, and perspectives. Further, one student said:

> "It was inspiring to see my fellow peers contributing to this brainstorming/ idea-spreading process! It was a nice reminder that despite the dread (perhaps angst, even) I feel in regards to climate change, there are people willing to be more eco-friendly in their own lives."

Both the word-cloud and student comment suggest that the solutions-oriented nature of the activity and lesson helped to empower students in recognizing their potential and their power to contribute towards wicked problems like climate change.

Figure 2. Evidence of student learning. Word cloud illustrating common themes of the feedback on the multi-phase activity used in class.

The activity presented some challenges that could be addressed in the future. The class size was large (170 students), which sometimes limited access to the flip-chart papers. Potential solutions could include dividing the class into groups and having them move around the flip-chart papers as stations, reducing crowding at a single topic. One student suggested that free software such as "Google Docs" could also be considered as a solution to the crowding issue; however, this would limit student discussions as they move around the room. Weighing the pros and cons of digitizing this activity should be a part of the implementation process for instructors.

The problems that face tomorrow's graduates are large, complex, and sometimes, truly wicked. It is essential that students be prepared to tackle these problems. Further, it is important that students recognize there are solutions to these problems and that their colleagues and peers may be actively engaging in answering these questions with them. By providing solutions-oriented lessons and activities, educators can encourage, motivate, and inspire their students to be the global leaders our world needs.

Acknowledgments

Thank you to Kaitlin Roke for her constructive feedback on earlier drafts of this manuscript.

References

Costello, A., Abbas, M., Allen, A., Ball, S., Bell, S., Bellamy, R., . . . Kett, M. (2009). Managing the health effects of climate change: Lancet and University College London Institute for Global Health Commission. *The Lancet, 373*(9676), 1693-1733. DOI: https://doi.org/10.1016/S0140-6736(09)60935-1

Diceman, J. (2010). Dotmocracy handbook, version 2.2. In: Sur Internet:< http://www.dotmocracy.org.

Hancock, T., Spady, D. W., & Soskolne, C. L. (2016). *Global change and public health: addressing the ecological determinants of health*: Canadian Public Health Association.

Maxwell, J., & Blashki, G. (2016). Teaching about climate change in medical education: an opportunity. *Journal of public health research, 5*(1). DOI: https://10.4081/jphr.2016.673

Silberman, M. (1996). *Active Learning: 101 Strategies To Teach Any Subject*: ERIC.

Waltner-Toews, D. (2017). Zoonoses, One Health and complexity: wicked problems and constructive conflict. *Philosophical Transactions of the Royal Society B, 372*(1725), 20160171. DOI: https://10.1098/rstb.2016.0171

Watts, N., Adger, W. N., Agnolucci, P., Blackstock, J., Byass, P., Cai, W., . . . Cooper, A. (2015). Health and climate change: policy responses to protect public health. *The Lancet, 386*(10006), 1861-1914.

Contact Info for the Journal

JoSoTL Editorial Office

Indiana University
Bryan Hall, Room 203 B
107 S. Indiana Avenue
Bloomington, Indiana, 47405

josotl@iu.edu

ISSN: 1527-9316