JoSoTL

Journal of the Scholarship of Teaching and Learning

Volume 13, Number 2, May 2013 *josotl.indiana.edu*

JOSOTL
JOURNAL OF THE SCHOLARSHIP OF TEACHING AND LEARNING

Journal of the Scholarship of Teaching and Learning, Vol. 13, No. 2, May 2013, pp. 1 – 22.

Standards-based grading in introductory university physics

Ian D. Beatty[1]

Abstract: Standards-based grading (SBG) is an approach to assessment and reporting in which scores are attached to the specific learning objectives of a course, rather than to assignments or tests. Each score represents a student's mastery of that learning objective, and may change over time in response to evidence that her level of understanding has changed. SBG is increasingly popular in K-12 education, but has been poorly documented and studied in a university context. I explored the practicality and effects of using SBG in a moderately large university class by incorporating it into the two successive courses of an introductory physics sequence. Although design flaws and logistical difficulties plagued these attempts, most students responded positively to the basic intent and elements of the approach. Our experiences revealed likely implementation errors and suggested some wise design choices. More interestingly, I found that SBG foregrounds and forces us to confront some fundamental tensions present but latent within most or all teaching.

Keywords: grading, standards-based grading, assessment, mastery learning

I. Introduction.

"Standards-based grading" (SBG), also called "standards-based assessment and reporting," is an alternative approach to assessing, tracking, reporting, and grading student learning in a course. It has garnered growing attention in recent years, as K-12 schools seek grading methods consonant with standards-based curriculum frameworks and assessment systems (Guskey, 2001), and as reformers seek to avoid the drawbacks of traditional grading (Wiggins, 1998).

Despite its growing popularity at the K-12 level, SBG seems largely neglected within higher education. A notable exception is Rundquist (2011), who reported a very positive experience implementing a pure SBG design SBG in a small upper-level physics course, with the novel feature that all assessment evidence had to include the student's voice. Thus, students were required to demonstrate their mastery of learning objectives via oral exams in class, one-on-one discussions with the instructor, or the submission of "pencast" videos in which they narrated a proof or a problem solution as they wrote it out. While intriguing, such an approach is clearly impractical in a large-enrollment course; Rundquist admitted that his grading load was heavy with only nine students.

During the spring and fall of 2012, I implemented SBG in the two semesters of an introductory calculus-based university physics sequence. In this article, I describe my SBG design and implementation, summarize student reactions to the approach, reflect upon the successes and difficulties we encountered, and draw some general lessons from the experience. My aim is to stimulate conversation about the benefits and drawbacks of SBG in higher education, and to assist instructors who might be inclined to try SBG.

[1] Department of Physics & Astronomy, University of North Carolina at Greensboro

II. Background and Motivation.

In a typical university course, each student is awarded a score, grade, or point total for assignments, tests, and other course components, and her overall course grade is determined from a sum or average of her component scores. In SBG, scores are not attached to specific assignments and tests. Instead, the instructor identifies a set of learning objectives or "standards" for the course and the student receives a mastery score for each standard. Every such score represents how well the student has mastered the standard, based on evidence from one or more assignments, test questions, in-person interactions, or other sources. Thus, as her "result" on an assignment or exam, a student would receive a set of scores for the standards it addressed rather than a single overall grade or point total (Marzano & Heflebower, 2011). At the end of a course, her scores for all standards can be combined to yield an overall course grade if one is required.

In an ideal SBG implementation, a student's score on any particular standard indicates how well he's mastered the standard *at that point in time*. This score may increase over time as he demonstrates increasing understanding and skill (an expected result of learning), and it may decrease if new evidence reveals previously-overlooked flaws in understanding. Thus, his set of standard scores provides a real-time snapshot of his skill and knowledge. In reality, each standard can be assessed only intermittently, but students still have some opportunity to persevere on standards until they reach an acceptable level of mastery.

Philosophically, SBG is predicated upon three principles. First, feedback from assessments should be linked to specific learning objectives in order to help students know and target what they need to learn (Marzano & Heflebower, 2011). Second, students should be permitted to remedy deficiencies in their learning when an assessment reveals them (Dueck, 2011). Third, standard scores should communicate students' degree of mastery of the course learning objectives, and not be confounded with other variables such as effort, good behavior, or rate of progress (Brookhart, 2011; O'Connor & Wormeli, 2011). The second and third principles, taken together, imply that poor performance on an early assignment or test should not forever weigh down a student's course grade, but should be completely overwritten by later evidence that the learning objectives have ultimately been met. Some SBG implementations step back from these ideals, for example by including an effort/behavior component in the calculation of the final course grade, by using a decaying average calculation for the various measurements on each standard so that later scores do not completely overwrite earlier scores, or by setting a deadline for the first attempt at a standard in order to discourage excessive procrastination.

SBG is intimately related to *mastery learning* (Block, 1971; Bloom, 1985; Guskey & Gates, 1986). In conventional learning, students study a particular topic or skill for a pre-determined time window, and then move on to subsequent topics regardless of their progress. Over the span of a course, all students attempt the same set of topics, and are differentiated by their average progress on each. In mastery learning, students persevere on each topic or skill until reaching a specified threshold of competence. Over the span of a course, all students achieve comparable levels of proficiency on those they study, and are differentiated by how many topics they master. Mastery learning designs may be sequential, with students focusing on one topic at a time; or overlapped, with students continuing to seek mastery on older topics while encountering new ones (e.g., Leonard, Hollot, & Gerace, 2008). Proponents of mastery learning claim that solidly mastering a core subset of a course's material is preferable to incompletely learning the entirety, and meta-analyses of mastery learning implementations often show positive impacts on student outlook variables (see discussion in Leonard et al., 2008).

To date, much of the argument in support of SBG is theoretical. Of the evidence that exists, much is anecdotal (e.g., Erickson, 2011). However, given the criticisms leveled against traditional grading practices—such as conflated variables and uninterpretable grades, inadequately timely or specific feedback, minimal incentive to learn from mistakes, and damage to intrinsic motivation (see discussions in Docan, 2006; Wiggins, 1998)—considering alternatives such as SBG seems worthwhile. Finding the principles underlying SBG persuasive, I sought to appraise SBG in introductory university physics. Rather than investigating specific research questions, I conducted a preliminary exploration of the terrain, guided by the general hypothesis that SBG can be practically implemented in a moderately large university physics course and provide an overall positive experience for the students and instructor.

III. Narrative: Course Design and Execution.

A. Physics 1: Diving in Deep.

At my university, Physics 291 and 292 constitute a two-semester sequence of *Introductory Physics with Calculus*, taught in the spring (291) and subsequent fall (292) semesters. The course serves physics, chemistry, computer science, mathematics, pre-engineering, and biochemistry majors. In recent years, typical initial enrollments have been 50–60 students for 291 and 30–40 for 292. I first taught the sequence in 2011, with conventional grading and a highly interactive pedagogical approach centered on "clicker"-mediated classroom discussion (Beatty, Gerace, Leonard, & Dufresne, 2006; Beatty & Gerace, 2009). I taught 291 again in 2012, aspiring to a full, pure SBG scheme but otherwise altering the curriculum and pedagogy as little as possible in order to observe how the switch to SBG altered the experience.

Initial Course Design. My 291 SBG implementation design had three components: a list of standards, a grading scheme, and an assessment plan. I retained the syllabus of topics and approximate schedule of coverage from the previous year, which organized sixteen chapters into five units. To develop a list of standards, I studied the textbook and identified specific skills or competencies that could be articulated, understood, and assessed more or less independently, and that seemed to form the essential core of the material. I sought a compromise between overly coarse-grained, general standards (which would be little better than topic headers) and overly fine-grained, specific standards (which could be logistically and administratively impractical). Following the general wisdom for SBG practice, I phrased each standard as an "I can…" statement in order to frame students' learning as the development of capacities rather then the retention of declarative knowledge. Before the course began I articulated 28 standards for the four chapters of the first unit, corresponding to an average pace of 3.5 standards per lecture meeting. Table 1 lists the standards for Chapter 2 as an example.

For the grading system, I chose a four-point scale to represent each student's mastery on each standard, as shown in Table 2. A student's score for a standard could change over time due to evidence from reassessment attempts, with a later score completely overwriting any earlier ones. At the conclusion of the semester, a student's latest scores for all standards would be averaged, combined with the lab score (see below), and the result mapped to a letter grade such that 4.00 yielded an A+, 3.75 an A, 3.50 an A–, and so on, with 1.00 or lower yielding an F. The syllabus asserted my right to adjust the thresholds, should the system prove too lenient or harsh in practice.

I scheduled five unit exams as the primary assessment of standard mastery. I intended brief in-class quizzes every few days on some of the more basic standards, hoping to reduce the

Table 1. Learning standards for Chapter 2, "Kinematics in One Dimension."

Standard
1. I can use the uniform motion model to analyze physical situations.
2. I can determine or reason about an object's instantaneous velocity at various instants during its motion, based on position vs. time information.
3. I can determine or reason about an object's position and displacement, based on velocity vs. time information.
4. I can use the constant acceleration model (including constant-acceleration kinematics formulae) to analyze physical situations.
5. I can explain and analyze cases of free-fall using the constant acceleration model.
6. I can use the inclined plane model (a special case of the constant acceleration model) to analyze situations involving a sloped surface.
7. I can determine or reason about an object's instantaneous acceleration.
8. I can produce, interpret, and interrelate graphs of position vs. time, velocity vs. time, and acceleration vs. time for various motion scenarios.
9. I can use basic calculus (derivatives and integrals) to interrelate expressions for position, velocity, and acceleration as a function of time.

number of standards covered by the exams. My plans for reassessment after unit exams were less clear. I intended to reassess a few of the most difficult earlier-unit standards on later-unit exams. I also hoped to let students reassess some standards via one-on-one oral quizzing outside of class, but due to the large enrollment I was reluctant to promise this. I reserved the final exam for last-chance reassessment. In the syllabus and in class, I stated that reassessment was not a guaranteed right, that not all standards would be available for reassessment, and that reassessment must be earned by demonstrating remedial work done (such as re-working an earlier exam for homework and articulating what learning had resulted).

Table 2. Rubric for assigning a mastery level score to each standard.

Mastery level	*Score*
Got it solidly!	4
Mostly got it. (Understand the idea well, but sometimes make small mistakes or get confused by subtleties.)	3
Making progress. (Definitely understand it somewhat, but still have misconceptions, gaps in knowledge, or make serious mistakes.)	2
Starting out. (Know a wee bit about this, but not enough to really use it for anything.)	1
Nothing yet. (Have no idea yet what this is.)	0
Not yet assessed.	—

The course included assigned textbook reading, corresponding sections of the accompanying *Student Workbook*, and homework problems drawn from the textbook and automatically evaluated by an online homework system. Although I stressed that completing these earnestly was "essential" to learning, they were not checked, did not contribute towards the course grade, and had due dates that meant nothing more than "keeping up with the course."

Using points and grades to coerce student behavior clashes with the SBG philosophy, so I relied on students appreciating the connection between doing the work and performing well on exams.

This course includes a one-credit laboratory section, with a single overall grade given to the combined lecture and lab. Thus, integrating the lab into the SBG system was essential. My undergraduate teaching assistants and I developed a list of 35 lab standards, divided into three categories: measurement and analysis, scientific communication, and experimentalism. Lab standards used the same four-point mastery scale as lecture standards, and at the end of the course would be averaged into an overall lab score that would then be combined with the overall average lecture score (weighted 1:3), with the result mapped to a letter grade as indicated above. Lab standards would be assessed through a variety of means—lab quizzes, lab notebooks, lab reports, and TA observation/interaction—with the details to be worked out as the semester unfolded.

Initial Course Execution. As the course began, students revealed no particularly strong reaction to the SBG aspect. My general teaching practice is to meta-communicate heavily about my pedagogical principles and tactics, so I spent considerable time explaining and justifying SBG, clicker question discussion, and group whiteboarding. Though few students appeared to grasp the specifics, all seemed amenable to the general idea. The presence of "reassessment" seemed reassuring to many, even though details were lacking. I suspect many interpreted it as "free second chances."

During the eight lecture meetings of Unit 1, I gave two quizzes, but scores were too low to omit the corresponding standards from the first unit exam. (I gave only one more quiz during the remainder of the term.) That exam assessed 22 of the 28 Unit 1 standards; one more was deferred to the second exam, and five were quietly dropped from the course. Even so, the exam was too long for the three hours provided.

Scores on the first exam were low. Of the 1254 scores (57 students times 22 standards), 22% were 4s and 14% were 3s. Eighteen students earned a 3 or 4 on at least half the standards, and eleven of these earned a 4 on at least half. Thus, most students needed reassessment on a significant fraction of the material, and a majority needed reassessment on almost everything. Had I given overall exam grades by averaging all of a student's standard scores, the course median would have been 1.95, corresponding to a low C–.

I generally teach demanding courses using unorthodox methods, and require an unusual level of engagement and responsibility from students; thus, the first exam is very often a wake-up call, with a lower average than subsequent exams. Additionally, this exam had been longer and harder than intended. As a result, I was not overly worried by these poor results, and I reassured the students that with some recalibration on all our parts, we had no need to panic.

However, the results of the second exam, three weeks later, were similarly dismal. Many students were becoming openly hostile or discouraged, and the class mood was souring badly. I realized that the course needed major intervention. After a frank discussion in class, I posted an extensive online questionnaire for students to complete anonymously. I motivated it by saying that I was seeking input in order to make significant improvements to the course.

The questionnaire contained 50 items: 23 short-answer responses, 21 scale ratings, and 6 numerical responses. 44 students started it and 34 completed it, although many who completed it did not respond to every item. Six of the items directly requested students' opinions about aspects of SBG, and a few others also provoked comments about the grading scheme.

Mid-Course Adjustment. In response to the survey responses, in-class discussions, and attendant one-on-one conversations with students, I made three major modifications to the course

design, which I announced during spring break (the approximate midpoint of the semester). The first was to include more worked problem examples and short lecture segments in class, in response to complaints that clicker-based discussions showed them what they didn't understand but "left them to learn the material on their own." The second was to change the course exam schedule by introducing two "reassessment exams" for Units 1 and 2, displacing the Unit 3 exam until later in the semester, abbreviating Units 4 and 5 due to schedule slippage, and merging the Unit 4 and 5 exams into one. The third modification was to designate fewer, broader standards for units 3–5.

The re-exams for Units 1 and 2 occurred on the first and second Wednesday evenings after spring break, respectively. Both consisted of a set of problems analogous to the original unit exam problems, each targeting one or a few related standards. Students were given in advance a list of the standards targeted by each problem, though not the specific problems. During the exam, they could choose which ones they wanted to try. Once they took and read a problem, they would be scored on it no matter how well or poorly they did, with those scores replacing the corresponding scores from their prior exam. In other words, either they believed they had improved their understanding on those standards and were willing to bet on themselves, or not. I dropped as logistically unmanageable the requirement that students rework and turn in the first exam in order to qualify for reassessment.

On both re-exams, almost all students chose to reassess on at least a few problems, and many tried most or all. Overall, scores increased significantly. For Unit 1 standards, the median of students' average scores was 2.91 (up from 1.95), on track for a final course grade between B– and B. 26 (up from 11) students earned a 4 or better on at least half of the 22 standards, and 39 (up from 18) earned a 3 or better on at least half.

As a result of these mid-course adjustments, the class mood improved noticeably. The primary drawbacks of the reassessment exams were the degree to which they distracted students from learning the Unit 3 material being taught concurrently, and the time they required of me to create and grade. Consequently, I told the class that giving reassessment exams after each future exam was impractical. Instead, I provided a realistic practice exam before each of the remaining two exams, since familiarity was likely one reason students improved on the re-exams. I also used the final exam period as a reassessment opportunity for the most-needed standards from Units 3–5, following the same choose-your-problems approach as for the Unit 1 and 2 re-exams.

The remainder of the course passed without major incident in a blur of practice exams, unit exams, and the reassessment final. By the end, the course had assessed 69 lecture and 35 lab standards. After submitting final grades, I created a second anonymous questionnaire soliciting students' reflections upon the mid-course adjustment and the course as a whole, repeating some of the questions from the mid-course questionnaire. I sent the invitation and questionnaire link by email after students had left campus, and received only 12 responses despite sending follow-up reminder messages.

B. Physics 2: Simplifying Greatly.

The following fall, I taught Physics 292, the second course in the sequence. Of the 46 students who had completed Physics 291, 23 continued into 292, joined by nine additional students for a starting roster of 32. My SBG design for the first course had proven to be barely manageable, and due to a double teaching load in the fall term I had to drastically simplify my SBG

implementation for the second course. In hindsight, this simplification resulted in a fatally thin, incomplete version of SBG.

Initial Course Design. I made six primary changes. First, I established fewer, broader learning standards: only 25 across the lecture and lab. Examples include "I can describe the particle-ray model of light, define and explain its elements, and cite empirical evidence to justify it"; "I can deploy the particle-ray model of light to explain phenomena, qualitatively and quantitatively analyze novel physical systems, predict their behaviors, and calculate the values of physical quantities"; and "I can interpret calculus and connect it to physical situations, for example by constructing integrals to mathematically represent infinite sums of infinitely small contributions in a physical context, or by interpreting the gradient of a scalar field."

Second, I replaced the large online homework sets with a much smaller number of challenging, extended homework problems and mini-projects—some from the textbook, and some of my own creation—to be turned in on paper. I provided written feedback on these, but no grades, and merely said that they would be "taken into account" when determining standard scores from exams.

Third, I included only two exams in the course: one midterm and one final.

Fourth, I allowed students to take each exam twice: once as a three-hour closed-book individual exam, and over the subsequent days as an untimed, open-book, collaborative take-home exam. (This is a strategy pioneered by my colleague William Gerace. We have found that it dramatically increases student learning from the exam process, and also helps fuse the class into a tightly-knit peer support cohort.) I assigned scores for each standard by looking at a student's in-class and take-home responses together and inferring what level of mastery they revealed. In practice, the in-class response usually determined the score, with a notably stronger or weaker take-home response nudging that score up or down.

Fifth, I provided no reassessment mechanism for standards assessed on the exams. Except for standards assessed in lab (through a mix of quizzes, hands-on practical challenges, lab notebooks, lab reports, and instructor observation), the exams provided the one and only score for each standard (with an occasional upward nudge in response to unusually convincing work on a homework problem or relevant lab report).

Sixth, I instituted a "brutally strict attendance policy." Three lecture absences or one lab absence put a student "on probation" (unofficially and only for purposes of this course), which means I'd give the student a dire warning and try to discuss the causes with them. Three more lecture absences or one more lab absence would, in principle, automatically cause the student to fail the course. In practice, I planned to offer exceptions to students who would promise to reform. My intent was to establish an expectation of 100% attendance and have some leverage to lean on students with an attendance problem, without doing violence to the spirit of SBG.

Course Execution. The course ran as designed, with only minor adjustments during the term. Homework completion was spotty, though it improved after the mid-term exam when I commented that several students had seen a slight improvement in one of their mastery scores based on their homework. I had great difficulty providing timely feedback on homework submissions. By the end of the course, the mood of the class seemed generally positive, with a highly interactive classroom dynamic, good discussion, and frequent jokes and laughter. Despite this, the course evaluations completed during the last class meeting contained many negative comments and lower ratings than I have previously received.

IV. Results.

This section summarizes students' reactions to the SBG aspect of the two courses (based primarily on their responses to the two questionnaires in the first course and their end-of-course evaluation forms), as well as my experiences as an instructor implementing SBG (based on notes I made during the courses and post-course reflections). Although this article focuses on SBG, the courses contained many other unusual elements, such as clicker- and whiteboard-based active engagement tasks during lecture, a strong focus on deep conceptual understanding and transfer, labs taught by undergraduate teaching assistants and focusing on open-ended experimental challenges rather than well-defined procedures, and take-home second chances on exams. Students are generally poor at discriminating between such things when they react and opine, so conflation is likely. Also, since questionnaire completion was voluntary and anonymous, some self-selection bias is possible. This is especially true for the post-course questionnaire, which had a response rate below 25%.

A. Student Opinions.

In general, many but not all students liked the SBG approach in Physics 291 despite the difficulties we encountered. Typical comments on the mid-course and post-course questionnaires and the course evaluation forms include:

> [I liked] The standards based grading
> I love the grading system.
> I love it, but sadly I know its hell on you when grading.
> the standards based grading is a little weird and trying to get use to it. may not like it because its different but i am trying to figure it out.
> [I dislike and would want to get rid of] Standard base grading

The questionnaires distinguished between three aspects of SBG: attaching scores to learning objectives rather than assignments, using grades only to represent achievement of learning objectives and not to coerce behavior, and permitting reassessment to change standard scores. For the first aspect—attaching scores to learning objectives rather than assignments—both the mid-course and post-course questionnaires asked:

> This semester, I'm trying a new grading system where I try to indicate and record how well you've mastered each "skill" or "piece" of the topic, rather than simply giving points and adding them up on each assignment. I know it's a bit more confusing and harder to answer the question "Overall, how well am I doing?" Aside from that, how do you like this approach?

Of the 32 responses on the mid-course version, 20 chose "I really like it," 6 chose "I like it a little," 2 chose "I'm indifferent; doesn't make much difference," 3 chose "I slightly dislike it," and 1 chose "I really dislike it!" On the post-course version, the corresponding response counts were 6, 3, 0, 1, and 0 (10 total). Representative comments are:

> I think that it makes it easier for a student to assess what he/she needs to work on.
> I know that for me, depending on the scores I receive for an exam, I will go back and decide whether or not I understand that standard and decide if I need to work on it more

> It kinda lets you know what you are golden on and what you suck at. It tells me what I need to look back at. If I bomb standard 2.1 but ace standards 2.2-2.4 then obviously there is something that I'm not getting from the earlier stuff. Using this approach I can go back and specifically target what I need to study much easier than if I were just handed a paper back with a 93% scrawled at the top in some undergrad/grad student's handwriting.

> I feel like it puts alot more pressure because there are many many standards to learn and each one could potentially mess up your score. If you try to reassess and don't do well then you get the worse score and there's not another chance to fix it so it puts alot of pressure instead of just getting an overall percentage of what you understand.

Following this, both questionnaires posed the open-ended prompt "Any other thoughts about scoring 'by standard' (by learning objective) rather than simply counting up points for each assignment?" 17 students offered comments on the mid-course version, and 7 did so on the post-course version. I aggregated those 24 aggregated responses, and coded them according to their essential point(s) with an emergent coding scheme. A few comments contained multiple distinct ideas and were multiply-coded. Of the positive comments, 8 expressed a non-specific liking (for scoring by standard); 6 appreciated the formative value of precise feedback; 2 appreciated the ability to focus their studying on learning, rather than jumping through hoops for points; and 2 thought the resulting assessment were more accurate indicators of understanding. Of the negative comments, 3 thought the approach was more stressful; 3 thought the standards and/or their scoring levels were unclear; 3 thought the scoring system was less precise than traditional percentage scores; 2 thought the effect was less motivating; 1 thought it was more demanding; and 1 disliked the possibility of having a standard score go down when reassessed.

For the second aspect of SBG—using grades only to represent achievement of learning objectives and not to coerce behavior—both questionnaires asked:

> As a matter of principle, I don't use points and grades to coerce you into doing things for your own good (like coming to class, completing the workbook, or doing the homework). I think grades should indicate how well you're learning the course objectives, and nothing else. I think it's insulting to treat students like they won't do any work that's not directly graded, and that they're not mature enough to recognize that they need to attend class, read the text, do the homework, etc. in order to get a decent grade on the eventual exams. Do you like this policy?

Of the 32 responses on the mid-course version, 18 chose "I very much like it," 7 chose "I kinda like it, I think," 2 chose "I'm neutral or ambivalent," 5 chose "I dislike it a bit," and none chose "I really don't like it." On the post-course version, the corresponding response counts were 4, 4, 1, 4, and 0 (10 total). Representative comments are:

> I study better figuring out the actual topics and focusing on those instead of completing homework just for the grade. the work book can help some times but i am a visual and auditory learner. I find videos online to drive the topics home. i think that's more beneficial than the work book for me.

> I have a hard time doing work that is not helpful to me learning regardless of whether or not it's graded. I'm here to learn, not to make straight A's (though obviously that would be nice) by doing pointless work that doesn't ultimately contribute to my education. I really appreciate that you trust us to make the decision of what is and isn't helpful for us and I really like your grading policy. I wish every class could be like that.
>
> I like being given several options / approaches and being able to find what works for me and not having to do something that isn't helping just for a grade, however there was obviously an adjustment period. At the beginning, attempting to go through all of the material to find what was going to help the most was tedious, but in the end worked out.

Despite generally liking this freedom, students reported mixed success at self-motivating themselves to complete ungraded work. The mid-course questionnaire asked, "How good are you at motivating yourself to do work well and on time without grades and deadlines to put pressure on you?" Of the 31 responses, 5 chose "no trouble, don't need grades and deadlines," 16 chose "a bit of difficulty, but overall do okay," 6 chose "struggle quite a bit, do significantly less work," 1 chose "won't do it if not due and graded. Period," and 3 chose "Something else…" In the free-text response for the "Something else…" option, one student referred to difficulty prioritizing physics over other courses with graded deadlines, and another claimed he or she "ALWAYS" does the work before the exam, but probably doesn't leave enough time to it adequately.

Several students wanted some kind of grade credit or reward for effort, separate from the benefits of learning the material and doing better on exams:

> I believe attendance policies like that, per me… they don't coerce me into doing well, I think you will find most students at this level truthfully want to do well. However, it's more about maintaining confidence, keeping your chin up, if all we have are tests, and everyone keeps bombing them (mostly) it will wreck moral. Additionally, no curve (not that I believe in them because I don't either) kinda does not help replace the void where typical "hoops" are used to pad grades.
>
> I am a self-motivated student and I don't need to be "bribed" with grades every step of the way. What I do need, is the feeling that my hard-work and dedication will pay off. I put more effort into this one class than into any of my other classes, and yet I am getting worse results than I ever have in my entire collegiate career. It would be nice to see some pay-off from that work, whether from homework, or quizzes or doing the work book. Perhaps those who complete assignments can have their work factored in when grading the exams, and those who don't do the hw/workbook will only be graded on their exams.
>
> I know you don't like bribing us with grades, but getting credit for getting all of our assignments done would be nice.

A related theme was that many students wanted some lower-stakes assessments in addition to the exams, both to provide earlier feedback about how well-prepared they were for the exams, and to reduce the stress of having "everything ride on exams."

> Essentially, the only time that our attempts "count" is during the exam, and that is very stressful.

Both the mid-course and post-course questionnaires included the open-ended prompt "Any other thoughts about using or not using grades and deadlines to 'make' you do the work of learning in a timely fashion?" Of the 23 relevant comments aggregated from the two rounds, 4 expressed a non-specific positive response to course's approach; 2 liked the freedom to choose what study tactics worked best for them; and 2 liked the flexibility to choose when to do physics work. On the other hand, 5 disliked the fact that physics studying took a back seat to other courses with homework deadlines and grades; 3 claimed they need deadlines to make them get work done; 2 missed the formative feedback that graded homework provides; 1 wanted extrinsic rewards for homework, 1 disliked the fact that the lack of any homework contribution towards grades made the exams higher-stakes, 1 wanted to see more direct payoff for doing homework, and 1 simply found the approach uncomfortably unfamiliar.

For the third aspect of SBG—permitting reassessment to change standard scores—students generally appreciated the "no history" aspect of SBG, where later evidence completely overwrote earlier evidence of mastery for each standard:

> It is the best method of evaluating student performance, in my mind, that I have ever encountered. Though the amount of reassessment we have is extreme now, I have always felt that it is wrong to penalize a student for doing poorly early in the semester and then succeeding very well by the end. The standards-based system takes the high-school "point system" mentality away from the learning experience so that you can focus more on getting the material down, and less on "how do I get a 90 or better"

> I like that we can reassess if we do poorly on an exam. A lot of the stress of grades comes from doing poorly on one exam and then having to hit a home-run on subsequent exams or the final.

> Also, the reassessing helps if you are having a bad day. Most teachers just drop one test if you have a bad day. But that keeps the student from learning from it.

On the other hand,

> I think the amount of reassessment is unnecessary. I think in the future, it would be fair to state that it was not a guaranteed right, but that should a student demonstrate significant improvement or effort, it may be granted as a privilege. And then try to include as much and as broad a selection of content on the final as possible, and use that as the "primary" opportunity to demonstrate that knowledge was, in fact, gained…

Despite my fears to the contrary, students claimed that the prospect of reassessment did not impact how well they prepared for the initial exams. The mid-course questionnaire asked:

> How has the possibility of "reassessing" standards later in the course, with later scores replacing earlier scores, impacted how you prepare for exams?

Of the 31 responses, 9 chose "It hasn't changed what I do at all," 18 chose "I'm a bit more relaxed and less stressed, but I still prepare and study just as much as I would if there weren't any reassessments," none chose either "I prepare and study a lot less than I would without reassessments, because I know these exams don't matter all that much" or "I don't bother preparing for the exams at all, since they don't really 'count'," and 2 chose "Something else…"

B. Student Difficulties.

In their questionnaire and course evaluation comments, students reported several common difficulties with the Physics 291 implementation of SBG. One difficulty was operationalizing standards:

> Also if the standards were explained more and feed back was given as to what could be worked on more to get a far better understanding would make this way of grading far better.
>
> The labs also had standards that were not clearly explained…

After the mid-course adjustment, I made a point of linking standards to specific textbook sections, and of listing representative problems for each standard.

> I liked how each standard had corresponding questions to it so we knew exactly which ones we needed to practice… I also liked [that] we either knew specifically or we generally knew which standards would be tested. That really helps in the preparation of exams.
>
> [GOOD changes to the second half of the course included] Having practice tests available with answers before taking the exams, as well as showing what each standard meant by giving page numbers from book as well as suggested practice problems.
>
> Instead of giving the practice exam key… maybe work one problem from practice exam per class period. Sometimes it is hard to know exactly what a standard wants until seeing a problem. Giving book problems for specific problems was good, but I feel I benefitted more when you wrote the questions.

Related complaints involved dislike of or difficulty interpreting the 1–4 mastery score scale.

> Another concern that I have is what a 4, 3, 2, or 1 for each standard is defined as. It was not clearly explained how exactly the numbers were given.
>
> If you want to grade by standard, then make it clear what each standard means. Answer each test with a 4, 3, 2, 1, and 0. Then compare the answers of our test to each test. Do we give a 4 pt answer or do we give a 2 pt answer? They use very similar type of grading in the essay writing portion of the SAT's.

> I think there should be more partials like 2.5 or 3.25 to really pinpoint our positions and if it's split into sub quarters determining our grade would be much easier because there would be a letter grade that corresponds to each quarter (including pluses and minuses).
>
> I do really enjoy the idea, but I have two main stipulations with it. The first is that this scale has the potential to be susceptible to more subjectivity than a traditional 0-100% grading scale. The area between a two and a three or a one and two can get rather gray sometimes…
>
> It seemed like there was kind of an overly-significant drop-off (4 to 3) for any small mistake, but in general it does give a pretty nice picture.

Several students requested more frequent feedback.

> I like being graded on assignments and having homework that counts for a grade. To me, I don't see this as carrot-dangling or demeaning or anything like that. I see grades as a running indicator of how well I'm doing in a class, and how well I am prepared for exams.
>
> I think there should be more deadlines and grades because if a student didn't do so well on the workbook then you can help them figure out what they did wrong and then they will know how to do it on the exam.
>
> At the same time, however, it is stressful that these standards, the only real "grades" we have in the course, are tried almost exclusively on the exams. It would be nice if we could knock out some of the standards with some sort of consistent homework or quiz schedule. Personally, I think that this would encourage and enable me to set a better pace for learning the material.

Often, students need more than feedback on specific standards or topics: They need an overall assessment of how well they are doing in the course and what sort of final grade they are headed towards. Some students are concerned that they are not ready for the course, and want an overall grade estimation in order to decide whether to drop the course. Others need a partial-term grade reported for athletic eligibility or fraternity/sorority reasons. Yet others simply want to know whether their current approach to the course is adequate or needs to be rethought, and cannot interpret an assembly of standard scores well enough to answer the question "Am I doing okay?"

C. Instructor Experience.

In the previous section, I summarized students' most significant reactions to my SBG implementations. In this section, I summarize major elements of my experience as instructor.

At the conclusion of the course, I followed my usual process of calculating prospective course grades according to the formula published in the syllabus; generating a list of students ordered by overall score; using my personal knowledge of several specific students' degree of understanding as points of reference to adjust the mapping from mastery score averages to letter

grades until most grades seemed appropriate; fine-tuning the cutoff thresholds between letter-grade bins to avoid dividing students based on meaninglessly small numerical differences; and, finally, adjusting the grades of any specific students with relevant extenuating circumstances.

During this process at the end of Physics 291, I found that although I did not have to re-tune the mapping drastically, the grading system made the A-level grades perhaps a bit too difficult to earn and a respectable C or C+ rather too easy to attain. Many of the students in the C regime were ones I knew to be relatively clueless. At the end of Physics 292, I again had to bump the top students up a bit and the bottom student down a touch to align with my sense of what grades ought to represent. In the future, I could alter the advertised rubric, or perhaps be stinger about awarding 1s and 2s and more generous with 4s while grading. However, my personal observation is that this particular cohort was light on A-level "star" students, which leaves me unsure of how well the grading scheme would work at that level. Similarly, most D-level students did not survive to Physics 292, so my data on how well my second implementation worked with low-end students is weak.

I discovered that I very much like the process of grading exams within SBG. In fact, that is perhaps my favorite aspect of the entire experiment. SBG allows me to scrutinize students' exam responses with one question in mind: "How well have they shown that they understand X?" The four-point mastery scale allows me to answer that question quickly. I don't need to ponder how many points I should take off for each error, or wonder how many points a sound solution to an incorrectly interpreted problem is worth. Given a suitable set of standards and exam problems, grading is relatively fast, easy, and communicative.

While grading, I did discover a need for more miscellaneous, crosscutting standards. For example, a standard for units and dimensions fluency would allow me to ding students for sloppy units and emphasize their importance, while still giving full credit for understanding the physics of a problem. Similarly, standards for using conventional notation, performing algebra reliably, and sanity-checking results could be helpful. Instead, my policy was to drop students from a 4 to a 3 for such errors, conflating imperfections of understanding with weaknesses of process.

From my experiences in prior courses, I have come to believe that the take-home open-book collaborative exam retake is perhaps the most powerful instructional innovation I have ever tried, and in neither semester did I find a satisfactory way to integrate it with SBG. In non-SBG courses, I simply average a student's in-class and take-home scores, which gives the take-home all the gravitas of a full exam. This is part of what makes the approach successful. (Since most students earn high scores on the take-home, and the class' take-home scores show little variance compared to their in-class scores, averaging the two rescales but does not significantly reorder student scores. Since I control the overall course grading scale, I can adjust for this effect. Thus, the practice is instructionally beneficial without being unfair.) During Physics 291, I indicated that a correct redo of the exam would be required for eligibility to reassess, and then backed off from that requirement for logistical reasons. During Physics 292, I found myself making difficult and unsatisfactory inferences about a student's understanding based on a weak in-class attempt and a strong take-home quite possibly copied from a peer's solution.

I discovered a steep learning curve to creating good SBG assessments. One challenge was "factoring" the standards: developing exam questions that isolated specific standards. While grading, simply giving half-credit to a partially correct problem is inadequate; I must identify what component knowledge or skill was lacking, and score the corresponding standard appropriately. I discovered that most "interesting" physics problems—meaning problems that are not simple isomorphs of the examples used in class or the textbook, and that demand some

robustness of knowledge rather than memorized procedures—require multiple skills and knowledge elements. Targeting more than one standard per problem (or section of a problem) does more than slow down grading: It also causes difficulty for reassessment, when one wants to let students reassess on specific standards.

I reacted to this in the second course by choosing very coarse, broad standards that encompassed almost any problem within a broad swath of content. Basically, my standards represented topic areas rather than specific competencies. This made exam creation and grading easier, but was unsatisfactory because it sacrificed much of the benefit SBG can offer by providing specific feedback to students and supporting targeted remediation and reassessment.

Within typical courses, few of us try to test every possible skill and piece of knowledge covered. Instead, we rely on a representative sampling to gauge student learning and determine grades. A second challenge I discovered with SBG was that a fine-grained standard system forces me to assess every bit of content and skill articulated in the standards list, rather than sampling a subset of learning objectives in order. This can lead to impossibly long or impractically frequent tests.

During both semesters, my solution (in addition to giving painfully long exams) was to quietly ignore some standards, neither assessing them nor including them in grade calculations. In the future, I would want to leave them in the official list of course standards, so students know that they are "responsible" for that material. I might explicitly state that students will be assessed on many but not all standards.

Wrestling with this did highlight a fundamental tension in instruction: We want to teach, and have students take seriously, more things than we can practically assesses. This problem is not unique to SBG, although the need to assess and reassess all standards exacerbates it.

While SBG considerations dominated my assessment and grading practices, I found that they had little effect on my actual teaching. I occasionally displayed the wording of a standard during class, in the context of clicker questions and whiteboard problems relevant to that standard, but otherwise I didn't use course standards to direct, organize, or frame class activity. In hindsight, that may have undermined the impact SBG could have had on student learning.

V. Reflection and Synthesis.

This section will distill general lessons about SBG from my experiences as summarized above. First, it will present three fundamental tensions that any instructor seeking to implement SBG will have to confront. Following that, it will list some suggestions that should help an instructor avoid or minimize at least some of the potential difficulties.

A. Three Fundamental Tensions.

Reassessment. I claim that reassessment is the heart of SBG: It is the most difficult aspect of SBG to implement sustainably and to get right, but is also the most crucial. Reassessment gives SBG the power to steer student learning, encourage perseverance, and improve learning outcomes. When I eliminated reassessment from my SBG design for Physics 292, I found that the result was little different than a traditionally-graded course.

On the other hand, the logistics of reassessment pose many challenges, as my Physics 291 course narrative demonstrated. These challenges include allocating sufficient class time, developing multiple assessments for each learning standard, managing the grading load, and

ensuring that students take initial assessments seriously. An additional difficulty is getting students to pay adequate attention to new material while concurrently preparing for reassessments of older material. Although the introduction of two reassessment exams after spring break may have rescued the first course and helped many students avoid disastrously low grades, it significantly distracted from Unit 3, slowed the pace of coverage, and ultimately caused me to abbreviate or cut several topics from later units. As a student commented on the Physics 291 post-course questionnaire:

> It wasn't really a change exactly, but the entire last half of the course was plagued with a problem of multi-tasking. I could barely get myself to care about the new material constantly being introduced, when I was faced with what could be my last assessment on earlier standards. I realize this wasn't entirely intentional, but it was my main difficulty.

I see this as more than a mere logistical difficulty. Rather, it reveals a fundamental tension between students' need to spend time and attention revisiting "past" material that they have not yet solidly mastered, and our need to drive course coverage forward at a rate sufficient to address what must be addressed in the time allotted.

Grain Size and the Dead Frog Problem. A second tension arises in the choice of learning standards for a course. This is a critical step, since the choice of standards will shape assessment requirements and (hopefully) direct students' learning efforts. The tension is one of "grain size": choosing many fine-grained, specific standards vs. choosing relatively few, coarse-grained ones encompassing much subsidiary knowledge and skill.

Fine-grained standards help students know exactly what they should be learning, can be linked very neatly to highly-targeted exemplar problems that help operationalize them, provide specific, diagnostic feedback to support remediation, and allow efficient reassessment of only what needs reassessing. On the other hand, fine-grained standards are inevitably numerous, which creates administrative headaches and can be difficult for students to digest and track. Creating assessment questions that target single standards without being trivial or repetitive is difficult. Also, a large number of explicit standards inhibits an instructor's ability to assess efficiently by sampling a subset of the covered knowledge. These are all difficulties I experienced during Physics 291.

Coarse-grained standards have the opposite benefits and drawbacks. They are easy to track, allow great flexibility in exam questions, and enable sampling for efficient assessment. However, they tend to represent topic areas rather than well-defined competencies, conflating multiple component skills and knowledge elements. As such, they are not easily operationalized, provide little specific guidance for students seeking to remedy deficiencies in their knowledge, and force reassessment of large swaths of content at a time. I encountered or hid from all of these difficulties during Physics 292.

Finer-grained standards would seem more in keeping with the spirit of SBG, suggesting that we aim for as many standards as we can practically manage. However, grain size is not just a practical issue. The more finely we factor a subject into distinct learning objectives to be separately assessed, the more strongly we encounter what I call the "dead frog problem." The name comes from a witty response a friend once gave when he made a joke and I, failing to understand it, asked him to explain. He said, "A joke is like a frog. You can dissect it, but it dies."

I maintain that physics, too—and almost any other academic subject—is like a frog. Physics is more than a collection of distinct knowledge bits and skills, and critical aspects of

"knowing" and "doing" physics do not reside in any specific bits. Rather, much of physics expertise resides in having a richly interwoven, multi-scale mental map of how the pieces of content interconnect. More resides in the capacity to identify which pieces of content are appropriate to any given situation and integrate them at need. I fear that SBG fails to be faithful to the interconnected nature of physics or amenable to assessment via rich, authentic, integrative problems and tasks (Wiggins, 1998).

One might imagine having a set of fine-grained learning standards for the "elements" of physics and augmenting them with additional standards articulating higher-level skills and perspective. I am not convinced that one could adequately capture these higher-level pieces in standards explicit enough to be reassessed fairly and to be grasped and operationalized by students. Additionally, such standards are likely to be broad and nebulous enough that factoring them into valid, repeatable assessment items seems daunting. While SBG may be a good approach to teaching lower-level knowledge and skills, a complete course may need to fuse it with another system for assessing and grading higher-level learning, one that honors the value of "putting it all together" while being consonant with the mechanisms and core principles of SBG.

Dissecting the physics content into focused, factored learning objectives can do more harm than just preventing us from assessing the holistic aspects of learning. It also plays into students' inclination to reduce a course down to an explicit list of relatively simple, clear things to learn how to do: a catalog of specific kinds of problem to solve in a standard way. I see a great danger that SBG could unwittingly and implicitly, but strongly, frame physics and learning undesirably.

As one facet of this, consider the question: Whose responsibility is it to make sense of the material as an organic whole by breaking it down, studying the bits, and then building it back up into a knowledge structure? The students', or the instructor's? Perhaps the work of unpacking this big thing called "physics" into separate pieces is something the students should be doing, to improve both their learning of the content and their ability to proactively learn future topics and subjects. One way to address this concern within a SBG approach might be to involve students in the identification of learning standards for the course. Clearly, such a change would require a drastic shift in class power structures and responsibilities, but the result might be beneficial in many ways (Rothstein & Santana, 2011).

The Attention Economy. A third tension unavoidable in any SBG implementation, at least in most current postsecondary environments, is between the SBG principle that grades only report mastery of learning objectives and the unavoidable fact that many students will not complete homework, attend lab meetings, or otherwise put in the requisite work without the carrot-and-stick of points and grades. Most college students perceive themselves as busy and stressed, constantly juggling many demands on their time, and they operate in an "attention economy" mediated by grades and deadlines. As one student put it on the Physics 291 mid-course questionnaire:

> The primary issue I have is that other classes don't have this mentality. So there are a large number of assignments due, that are required in order to get good grades. Which means that when I have big assignments due (which is, it seems, ALWAYS the case), it makes it very difficult to convince myself that my time would be better spent studying physics than working on some assignment that is due. I know that that is not the absolute truth, but I hope you understand how hard it is to put off something that's being graded, with the potential of not completing the assignment, or completing it at a sub-standard level, in order to study/work on

> something that isn't going to "directly" impact my grade. (Though I know that, in a sense, it will directly impact my physics grade, it is not an "immediate" impact, like the other assignments.)

Many other students expressed similar sentiments, appreciating the principle of "no coercion points" but finding difficulty with the reality. Overall, they seemed to be divided about whether they considered the SBG approach a net positive or negative. As another student candidly observed:

> I can honestly say I would get more done if it were graded, however I need to learn how to motivate myself and I believe this is helping.

Ultimately, this may pose the single biggest practical obstacle to successfully implementing SBG in an individual university course: realigning students' motivational structures in opposition to a conflicting environment.

B. Experience-Based Suggestions.

1. Develop assessments before, or with, standards. Committing to a set of standards before developing specific assessments for them invites anguish. Some standards sound great in the abstract, but are murder to develop practical assessments for (and even harder to develop reassessments for). Some groups of standards that sound distinct in the abstract may be difficult to separate on an exam. A set of standards that seems ambitious but not unreasonable can easily lead to a prohibitively long exam. Additionally, in the process of creating exam questions, I have invariably discovered requisite student knowledge or skills that I overlooked while designing standards. In the future, I shall discipline myself to develop at least one round of assessments (and preferably one or two rounds of reassessment as well) before finalizing the standards for a course.

2. Attend to topic weighting. Unless a grading system weights standard scores unequally for final grade calculation, topics will be weighted by the number of standards they contain. That may be undesirable, as one topic may naturally unpack into more or fewer learning objectives than another, equally-important topic. During both courses, this fact pushed me towards bifurcating or merging standards somewhat arbitrarily in order to get a comparable number per chapter or general topic. This is a tension between standards as tools for communicating and tracking desired learning outcomes, and standards as elements of a quantitative grading system. One solution is to assign each standard a weighting factor. Another is to group standards by larger topics, use the standard scores to determine an average score for each topic, and then combine those topic scores for an overall grade.

3. Include standards for crosscutting skills. Some skills cannot be factored out of assessment questions for other standards, but necessarily appear throughout exams. Examples include arithmetic, algebra, trigonometry, and other calculation skills; fidelity to units and dimensions; checking answers; adopting clear and conventional notation; communicating solutions clearly; and anything else an instructor might like to "ding" a student for erring on, but which is not the core content targeted by the question. Creating separate standards for these eases and clarifies grading and feedback, and also communicates to students that developing reliability in these skills is an important course objective. Any original or reassessment exam can then be used as a (re)assessment of such crosscutting standards, with a mastery score assigned holistically based on the entire exam.

4. Name standards usefully. If we want students to use our standards to organize and direct their thinking about the subject, we should give them conveniently short yet meaningful names. Neither my students nor I could reliably remember what "standard 6.02" referred to.

5. Keep (re)assessment efficient. The majority of current SBG activity seems to be in K-12 schools, and I've gleaned many good ideas from the weblogs of high school teachers (e.g., Noschese, 2012; O'Shea, 2012). However, the university teaching context differs from the K-12 context in some very profound structural ways, with implications for practical SBG implementations. Perhaps the most salient difference is that a university physics course is expected to cover more material than a high school course, in greater depth and with greater rigor, with far less in-class time. Consequently, much of the burden of learning is correspondingly shifted to the student outside of class. With class time at a premium, we cannot afford to dedicate much to assessment and reassessment; yet with a faster pace of coverage, we must assess material more rapidly. Thus, efficiency of assessment and especially of reassessment is critical. Many of the strategies that K-12 teachers employ, such as reassessing via one-on-one oral interviews or requiring students to qualify for reassessment by explaining what has been learned from homework or re-worked test questions, will be impractical for many of our courses. We may also have to eschew foundational, low-level skills in favor of more integrative competencies for our standards.

6. Invest in quizzing. My students' comments and my own reflections indicate that I should have persevered with my initial intention to give short, frequent, low-stakes quizzes in class. Such quizzes could have helped meet my students' needs for more frequent feedback and guidance, operationalization of standards, and familiarization with my exam question style. They might have allowed me to remove at least some standards from the overly-long unit exams. Additionally, they could very plausibly have helped motivate students to keep up with reading and homework, in the same way that graded homework would but without doing violence to the SBG philosophy. The precious class time devoted to the quizzes could be at least partially recouped by using the quiz questions as fodder for class discussion, in much the way that I use clicker questions and group whiteboarding tasks.

7. Schedule regular reassessment times. As an alternative to disrupting the course and distracting students with full-scale reassessment exams, one could designate one or two time windows a week for reassessment, and allow students to show up and reassess on a few standards at a time (perhaps with an appointment). By making any given standard available for reassessment only for one or a few consecutive days, an instructor could limit the opportunity for diffusion between students, especially if she has a few different questions ready for each standard.

8. Require students to qualify for reassessment. One way to motivate the students most in need of homework to complete it, without resorting to "coercion points," is to require relevant homework be completed to a high level of proficiency in order to qualify for reassessment on a standard. This could be implemented efficiently by means of an online homework system that allows students to redo problems until correct. Two additional benefits of doing this would be stressing the link between homework and assessment success, and reducing the number of "hail Mary" reassessment attempts by students who have not significantly improved since the original assessment.

9. Link standards to instruction. A carefully engineered set of standards can shape how students conceptualize the subject. Explicitly organizing our instruction around those standards reinforces that conceptualization and helps students understand the implications of each standard.

For example, if we present physics as a "toolkit" for thinking and analyzing, and we align standards with the various "tools," then we can structure class as an inventory of the tools and an exploration of how various tools apply to different problems. Teachers using both SBG and the *Modeling Instruction* pedagogy (Brewe, 2008; Hestenes, 1992) seem to do this particularly well, linking standards to the physics "models" (e.g., Noschese, 2012; O'Shea, 2012).

10. Build catch-up time into the syllabus. Learning rarely happens linearly, with each successive chunk of material learned adequately on first exposure and then never revisited. Rather than plan a traditional, linear course syllabus and then try to wedge SBG reassessment into and around it, an instructor could set aside two weeks or so—perhaps at the end of the term, perhaps distributed throughout it—for remediation and reassessment. Students who need additional help on specific topics could seek it, and reassessment could be carried out without competing with ongoing new instruction. Those students not needing significant reassessment could pursue advanced and/or elective topics semi-independently. (One could attach additional standards to those advanced topics, necessary to reach the very highest course grades.) Additionally, such a strategy could help us avoid the chronic trap of optimistically trying to cram just a little too much content into each course.

11. Don't overemphasize grades. As Docan (2006) observed, increasing students' awareness of course grades and grading can increase their stress. During Physics 291, my attempts to help students understand the novel grading system and allay their fears about it may have backfired by keeping them more conscious of the grading aspect of the course than they would otherwise have been.

VI. Summary and Discussion.

In this paper, I have described a teaching experiment in which I implemented standards-based grading differently in each course of a two-semester introductory calculus-based physics sequence. I've candidly narrated my design decisions, difficulties, and adjustments. I've summarized the students' reactions and my own experiences as an instructor. Based on that, I've identified three fundamental tensions that I claim must be negotiated by any instructor attempting to implement SBG at the university level, and I've made several suggestions for others who might wish to try SBG in their own courses. Overall, I have found the principles of SBG compelling and some of my experiences with it promising, but I have encountered serious practical challenges and pedagogical dilemmas.

I claim that many of the tensions we encounter when implementing SBG are not actually created by SBG itself. Rather, they are inherent within all our teaching, but are brought to a head by SBG because it forces us to be explicit about many things often left implicit: our learning objectives, how we prioritize remediation of older material vs. coverage of new, the purpose of homework, whose responsibility various aspects of learning are, and what grades communicate. For example, any instructional approach or assessment strategy can suffer from the dead frog problem, but SBG forces us to articulate our learning objectives precisely and link them to assessment, thus making our deconstruction of the subject clear. Similarly, we always hope our students will heed the feedback they get on assessments and remedy their knowledge of past topics, but SBG forces us to budget time and attention for that and to value it in the grading system.

Because of this, I have found that my explorations of SBG have clarified my thinking about teaching in general. I am confident that my future teaching will be better, whether or not it

includes "standards-based grading" per se. I am optimistic that consideration of SBG by reflective university instructors and education researchers will provoke valuable discussions about the more difficult and elusive aspects of the art of teaching.

Acknowledgements

While the professional literature is relatively silent on SBG in physics and in higher education, the weblogs and *Twitter* feeds of innovative instructors are not. I owe a great intellectual debt to SBG pioneers such as Andy Rundquist (2013), Joss Ives (2011), Frank Noschese (2012), Jason Buell (2013), and Kelly O'Shea (2012), who use these channels to chronicle their SBG-related thoughts and experiences openly and candidly. I am also grateful for the always-sage advice of my colleague William Gerace, and for the helpful suggestions of Ellie Sayre, Warren Christensen, and Sissi Li.

References

Beatty, I., Gerace, W. J., Leonard, W. J., & Dufresne, R. J. (2006). Designing effective questions for classroom response system teaching. *American Journal of Physics*, *74*(1), 31.

Beatty, I. D., & Gerace, W. J. (2009). Technology-enhanced formative assessment: A research-based pedagogy for teaching science with classroom response technology. *Journal of Science Education & Technology*, *18*(2), 146–162.

Block, J. H. (Ed.). (1971). *Mastery Learning: Theory and Practice*. Holt, Rinehart & Winston.

Bloom, B. S. (1985). *Developing Talent in Young People*. New York: Ballantine Books.

Brewe, E. (2008). Modeling theory applied: Modeling instruction in introductory physics. *American Journal of Physics*, *76*(12), 1155. http://dx.doi.org/10.1119/1.2983148

Brookhart, S. M. (2011). Starting the conversation about grading. *Educational Leadership*, *69*(3), 10–14. http://goo.gl/pKsL7

Buell, J. (2013). Standards-based grading implementation. *Always Formative* (weblog). http://alwaysformative.blogspot.com/p/standards-based-grading-implementation.html

Docan, T. N. (2006). Positive and negative incentives in the classroom: An analysis of grading systems and student motivation. *Journal of Scholarship of Teaching and Learning*, *6*(2), 21–40. http://josotl.indiana.edu/article/view/1668/1666

Dueck, M. (2011). How I broke my rule and learned to give retests. *Educational Leadership*, *69*(3), 72–75. http://goo.gl/GqupA

Erickson, J. A. (2011). How grading reform changed our school. *Educational Leadership*, *69*(3), 66–70. http://goo.gl/KufkC

Guskey, T. R. (2001). Helping standards make the grade. *Educational Leadership*, *59*(1), 20–27. http://goo.gl/nbOze

Guskey, T. R., & Gates, S. L. (1986). Synthesis of research on the effects of mastery learning in elementary and secondary classrooms. *Educational Leadership*, *43*(8), 73–80. http://goo.gl/lCYLR

Hestenes, D. (1992). Modeling games in the Newtonian world. *American Journal of Physics*, *60*(8), 732–748. http://dx.doi.org/10.1119/1.17080

Ives, J. (2011). Standards-Based Grading. *Science Learnification* (weblog). http://learnification.wordpress.com/category/course-components/standards-based-grading/

Leonard, W. J., Hollot, C. V., & Gerace, W. J. (2008). Mastering Circuit Analysis: An innovative approach to a foundational sequence. *38th ASEE/IEEE Frontiers in Education Conference*. F2H3–8. http://dx.doi.org/10.1109/FIE.2008.4720568

Marzano, R. J., & Heflebower, T. (2011). Grades that show what students know. *Educational Leadership*, *69*(3), 34–39. http://goo.gl/fXfMn

Noschese, F. (2012). Standards-Based Grading. *Action-Reaction* (weblog). http://fnoschese.wordpress.com/category/standards-based-grading-2/

O'Connor, K., & Wormeli, R. (2011). Reporting student learning. *Educational Leadership*, *69*(3), 40–44. http://goo.gl/Czm1m

O'Shea, K. (2012). SBG. *Physics! Blog!* (weblog). http://kellyoshea.wordpress.com/category/sbg/

Rothstein, D., & Santana, L. (2011). *Make Just One Change*. Cambridge MA: Harvard Education Press.

Rundquist, A. (2011). Standards-based grading with voice: Listening for students' understanding. (N. S. Rebello, P. V. Engelhardt, & C. Singh, Eds.) *AIP Conference Proceedings*, *1413*, 69–72. http://link.aip.org/link/doi/10.1063/1.3679996

Rundquist, A. (2013, January 6). SBG. *SuperFly Physics* (weblog). Retrieved February 11, 2013, from http://arundquist.wordpress.com/category/sbg/

Wiggins, G. P. (1998). *Educative Assessment*. San Francisco, CA: Jossey-Bass.

Journal of the Scholarship of Teaching and Learning, Vol. 13, No. 2, May 2013, pp. 23 – 30.

Interteaching: Discussion group size and course performance

Jacob C. Truelove, Bryan K. Saville, and Ryan Van Patten[1]

Abstract: Researchers have yet to examine whether discussion group size affects student performance in an interteaching-based course. In the current study, we addressed this question by manipulating discussion group size (smaller groups of 2 students vs. larger groups of 4 students) across 2 sections of an undergraduate psychology course. We found no significant differences between the sections on 6 unit exams, on a cumulative final exam, and in the total number of points earned across the semester.

Keywords: interteaching, group size, discussion, exam performance, behavior analysis

I. Introduction.

Interteaching is a relatively new method of classroom instruction that has its roots in behavior analysis (Boyce & Hineline, 2002). Like previous behavior-analytic teaching methods—which include precision teaching (Lindsley, 1964); programmed instruction (Skinner, 1968); direct instruction (Engelmann & Carnine, 1982); and, arguably the most well-known of these methods, Keller's (1968) personalized system of instruction (PSI)—interteaching attempts to improve student performance by identifying what behaviors students should emit to improve their course performance and then rearranging the reinforcement contingencies to produce those behaviors. Unlike previous behavior-analytic teaching methods, though, interteaching may be easier to implement in traditional classroom settings (Boyce & Hineline, 2002).

A typical interteaching session proceeds as follows (for more detail, see Boyce & Hineline, 2002; Saville, Lambert, & Robertson, 2011). Prior to class, students complete an instructor-created preparation (prep) guide that contains questions over a reading assignment. Each class typically begins with a lecture that lasts approximately one third of the class period and covers material from the previous class (see below). After the lecture, students divide into pairs and discuss the prep-guide questions they answered for class. During the discussions, the instructor moves around the classroom, answering questions and guiding discussion. After students finish their discussions, they complete a record sheet on which they note their partner's name, how well their discussion went (along with reasons why), and any questions they would like the instructor to review. The instructor then uses the information on the record sheets to prepare a lecture that begins the next class period and precedes discussion of the next prep guide.

Since Boyce and Hineline's (2002) introduction of interteaching, researchers have found that it typically produces better student-learning outcomes than lecture-based teaching methods (e.g., Saville, Zinn, & Elliott, 2005; Saville, Zinn, Neef, Van Norman, & Ferreri, 2006; Scoboria & Pascual-Leone, 2009; for a review, see Saville et al., 2011). Researchers have also conducted a small number of studies examining which components of interteaching contribute to its efficacy (Saville, Cox, O'Brien, & Vanderveldt, 2011; Saville & Zinn, 2009). To date, however, researchers have not studied the discussion component of interteaching.

[1] James Madison University, truelojc@dukes.jmu.edu

In their original description of interteaching, Boyce and Hineline (2002) suggested using pairs during the discussions to avoid social loafing (Latané, Williams, & Harkins, 1979). In contrast, Goto and Schneider (2010) reported that their students preferred working in larger groups of four students, which some researchers have suggested will provide superior outcomes in cooperative learning situations (e.g., Johnson & Johnson, 2009). Neither Boyce and Hineline (2002) nor Goto and Schneider (2010), however, reported any systematic performance data. Thus, the purpose of the present study was to examine discussion group size and student performance in an interteaching-based course. Specifically, we asked students to work in pair or in groups of four and then measured their performance on six unit exams and on a cumulative final exam; we also examined the total number of points students earned across the semester.

II. Method.

A. Participants.

Participants were 61 undergraduate students from James Madison University, a large, public university considered to be "more selective" by the Carnegie Foundation for the Advancement of Teaching (www.carnegiefoundation.org). The students in this study, the majority of whom were juniors (see Table 2), were enrolled in two sections of an undergraduate psychology of learning course taught by the second author. Section 1 contained 30 students (25 women, five men), and Section 2 contained 31 students (28 women, three men). Section 1 met on Tuesdays and Thursdays from 12:30-1:45 p.m., and Section 2 met on Tuesdays and Thursdays from 2:00-3:15 p.m.

B. Materials and Procedure.

The instructor assigned a prep guide for students to complete before each class. The prep guides usually covered 10 to 20 pages of textbook material and contained anywhere from eight to 12 items (each of which often contained multiple questions) that required students to define concepts, apply course material, and engage in higher-order thinking (see Appendix for a sample prep guide from the course). Once in class, students divided into groups (for more information, see below) and discussed their answers to the prep-guide questions. The instructor encouraged the students to choose different discussion partners each class period, but given the relatively small number of students in each section, it was not always possible for them to work with an entirely different set of partners each time. During the discussions, the instructor and a teaching assistant (TA) walked around the room and answered any questions that students had. After finishing their discussions, students completed a record sheet on which they listed their partner's name, how well their discussion went (along with reasons why it went well or poorly), and which material they wanted the instructor to clarify. Students who participated in the discussions and turned in a record sheet earned a small number of participation points that across the semester totaled 10% of their course grades (Boyce & Hineline, 2002). At the start of the next class period, the instructor lectured over material that the majority of students had listed on the record sheets. The class then got into groups and discussed the next prep guide.

There were six 45-point exams during the semester, each of which followed discussion of three or four prep guides. Each exam consisted of approximately 20 items, most of which were short-answer questions along with a few multiple-choice and fill-in-the-blank questions. The

questions were based on, but were not identical to, items from the prep guides and typically required students to solve problems, apply information, and show higher-level comprehension. For example, a sample prep-guide question was "Discuss the one-process and two-process theories of avoidance." whereas two related short-answer exam questions were "How would a one-process theory of avoidance explain a fearful person's tendency to avoid dogs?" and "How would two-process theory explain a person's fear of heights?" At the end of the semester, students took a 90-point cumulative final exam that covered all of the prep guides and contained short-answer, multiple-choice, and fill-in-the-blank questions.

To measure the impact of group size on student performance, we had Section 1 (Large Group) discuss the prep guides in groups of four students (cf. Goto & Schneider, 2010) and Section 2 (Small Group) discuss the prep guides in pairs (cf. Boyce & Hineline, 2002).[2] Because we could not randomly assign participants to the conditions, we took two steps to ensure that the groups were relatively equal prior to manipulating group size. First, at the beginning of the semester, we collected the following demographic data: gender, age, current year in school, cumulative GPA, number of psychology classes taken so far, number of credits taken during the semester, and employment status. Second, prior to the first exam, we had both sections complete their discussions in pairs (which is the way interteaching was originally described by Boyce & Hineline, 2002). These measures provided a baseline against which we could compare the sections after our manipulation.

C. Interobserver Agreement.

For each exam, one TA graded all 61 exams, while a second TA graded a subset of 15 exams. To determine interobserver reliability (IOR), we divided the number of questions on which the TAs gave the same number of points by the total number of questions on the exam and multiplied by 100. The average IOR across the six exams was 87% (range = 83-92%). When the TAs disagreed on a question, they discussed the item and came to agreement on the final score.

III. Results and Discussion.

We first examined students' demographic information. One student in the Large-Group section (Section 1) only provided her gender on the demographic questionnaire. In the Small-Group section (Section 2), one student provided no information other than gender, another student did not report the number of credits she was taking, and a third student did not report her age and GPA. Our demographic analyses are thus based on the remaining data that participants provided. In sum, we found no significant differences between the sections on any of the demographic measures (all ps > .30, Table 1).

We next examined students' performances on the first exam (which, along with the demographic information, served as a baseline). One student in the Large-Group section did not take Exam 1. Thus the following analysis is based on the scores of 29 students in the Large-Group section and 31 students in the Small-Group section. We found no significant difference between sections on Exam 1, $t(58) = 0.50$, $p = .62$ (Large Group mean = 81%, Small Group mean = 80%). Together with the demographic information, this finding suggests that the sections were relatively similar prior to our manipulation.

[2] In Section 1 (Large Groups), depending on attendance, we sometimes had to let one or more groups have five students. Similarly, in Section 2 (Small Groups), we sometimes had to let one group have three students.

Table 1. Demographic information for the Large-Group (Section 1) and Small-Group (Section 2) sections.

	Section 1 (Large Group)	Section 2 (Small Group)
Gender		
Male	5	3
Female	25	28
Age (in years)	*M* = 20.90 (SD=0.94)	*M* = 21.03 (SD=0.73)
Year in School		
Junior	27	26
Senior	3	4
GPA (out of 4.00)	*M* = 3.37 (SD=0.41)	*M* = 3.40 (*SD* = 0.33)
Psychology Courses Taken	*M* = 10.86 (*SD* = 4.00)	*M* = 9.79 (*SD* = 3.83)
Semester Credits	*M* = 14.66 (*SD* = 2.50)	*M* = 14.69 (*SD* = 1.92)
Employed		
Yes	14	14
No	15	16

For Exams 2 through 6 (which students took after we manipulated group size), we conducted a 2 (group size) x 5 (exam) mixed ANOVA. We found a main effect for exam, $F(4, 236) = 2.61$, $p = .04$, $\eta_p^2 = .04$ (which, for the purposes of this study, is of little importance), but no main effect for group size, $F(1, 59) = 0.01$, $p = .92$. The average score across all of the unit exams was approximately 83% for both sections. In addition, we found no interaction between group size and exam, $F(4, 236) = 0.66$, $p = .62$ (see Figure 1 for all exam scores). Finally, we did not find a significant difference between sections on the cumulative final exam, $t(59) = 1.22$, $p = .23$ (Large Group mean = 84%, Small Group mean = 86%) or in the total number of points (unit exams plus cumulative final exam) earned across the semester, $t(59) = .527$, $p = .60$ (Large Group mean = 296 of 360 possible points, Small Group mean = 301 of 360 possible points).

In sum, there were no significant differences between the Large-Group section (groups of four students) and the Small-Group section (pairs) on any of the unit exams, on the cumulative final exam, and in the cumulative number of exam points earned across the semester. In their original recommendations on how to implement interteaching, Boyce and Hineline (2002) suggested using pair discussions to minimize social loafing. Goto and Schneider (2010), however, reported that students in their interteaching-based course preferred larger groups of four students. Neither Boyce and Hineline nor Goto and Schneider provided any systematic data, however, to show whether discussion group size affected performance. The present results suggest that when teaching an interteaching-based course, using smaller groups of two (or three) students or larger groups of four (or five) students may not result in differential course performance, at least as measured by exam performance and the related measure of cumulative exam points.

There are at least two possible reasons why we did not find significant differences in the present study. First, as Williams, Harkins, and Latané (1981) demonstrated, identifiability of an individual's contribution can help deter social loafing in larger groups. With interteaching, one's contribution to the discussion is often recorded by other group members on their record sheets and may also be apparent to the instructor as he or she roams the classroom answering questions. These publicly viewable events may function to increase participation (i.e., eliminate social loafing), even in larger groups of four or five students. If social loafing was reduced in the

Large-Group condition, then it may not be surprising that exam performance between the two conditions was similar in the present study. Given, however, that identifiability often becomes more difficult with increased group size (see Guerin, 1994), there may be a point where performance begins to deteriorate. Future research may thus wish to examine what happens to performance in an interteaching-based course when group size increases beyond four or five members. Specifically, researchers could replicate the present study but include a greater number of students (e.g., six to eight or more) in the Large-Group condition. This would provide more information on whether there is a point at which a "large" group becomes too large.

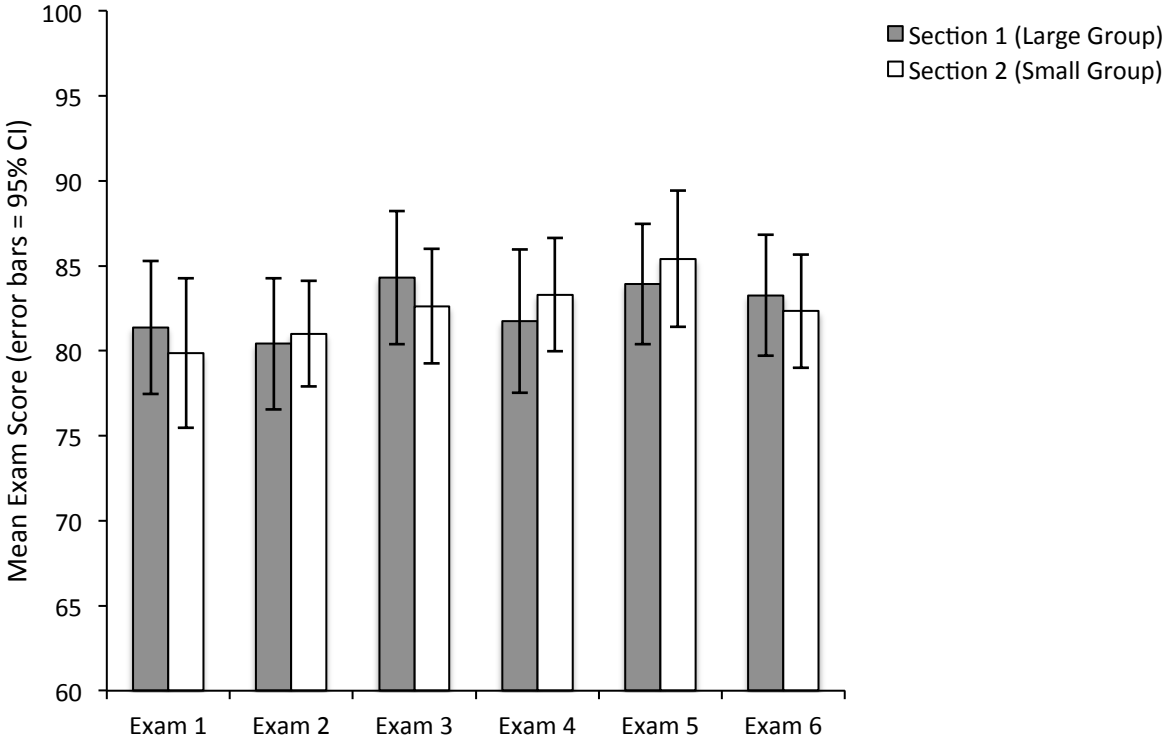


Figure 1. Shows the mean exam scores and 95% confidence intervals for students in the Large-Group (Section 1) and Small-Group (Section 2) sections. Students in both sections worked in pairs prior to Exam 1.

Second, it is also possible that the discussions simply are not an important component of interteaching. If this is the case, one would not expect differences in group size to produce differences in course performance. Unfortunately, determining the contribution of group discussions to interteaching is not possible from the present study because it lacked a true "no-discussion" control condition. Future researchers could examine this possibility more directly by exposing two groups to interteaching but eliminating the discussions in one group. Specifically, rather than discussing the prep guides with another student, students in the no-discussion group could use class time to re-study the prep guides they completed before class.

There are other factors that future researchers might wish to examine more closely as well. As noted earlier, the institution at which we collected our data is considered by the Carnegie Foundation to be "more selective" with its admissions criteria (meaning that our students typically have strong academic backgrounds). Our samples also consisted largely of

women, who tend to perform better in college than men (e.g., Mau & Lynn, 2001). Although these variables most likely did not *differentially* affect our results (as shown by our demographic analyses), it is possible that having these types of students in our study may have produced a small ceiling effect, which clouded our ability to identify significant differences (see also Saville & Zinn, 2009). In short, it would be interesting to see if our results might change when the samples studied are more diverse in nature.

Ultimately, if future research determines that the discussions *are* an important component of interteaching, instructors might then wish to consider student preferences when deciding whether to use smaller or larger groups (Wolf, 1978). If students prefer one group size to the other, allowing them to determine how many partners they have may increase their enjoyment of interteaching, which may improve course performance further.

Acknowledgements

The authors would like to thank Tracy Zinn and Kevin Apple for their excellent comments on an earlier version of this paper.

Appendix. Shows a Sample Preparation Guide from the Course.

Preparation Guide #10

Based on: Ch. 7, pp. 267-280

1. What is a schedule of reinforcement? Discuss the difference between continuous and intermittent reinforcement schedules. Identify some behaviors that are reinforced continuously and some that are reinforced intermittently. Which of these is more representative of the types of schedule that operate in our daily lives?

2. What is the relation between the response requirement and the postreinforcement pause in fixed-ratio (FR) schedules? Imagine you are a business owner who is trying to get your employees to be more productive. How might you incorporate a FR schedule to do this? Would you use a small FR schedule or a big FR schedule? Explain your answer.

3. In what way are VR and FR schedules similar? Different? Give some real-life examples of behavior maintained by VR schedules. How might you use these schedules to modify your own behavior?

4. How are the patterns of behavior produced by FI and FR schedules different? What are some behaviors that are maintained by FI schedules?

5. If you owned a casino and wanted visitors to gamble a lot, would you program your slot machines to pay off according to a FI, VI, FR, or VR schedule? Be sure to discuss what pattern of behavior (i.e., pulling the "arm" of the slot machine) each schedule would produce.

6. What are noncontingent schedules, and how do they differ from contingent schedules? How do these schedules possibly account for superstitious behaviors? Also, discuss how noncontingent schedules are likely involved in the development of "learned laziness."

7. In recent years, there has been a push to increase children's self-esteem (called the "self-esteem movement") by making sure that, for example, every child gets a trophy or award or even good grades, regardless of how well they actually perform. The belief is that receiving these "rewards" will make children feel good about themselves, which will then result in improved performance. Unfortunately, studies are showing that the "self-esteem movement" is having negative effects on children's performance. Based on what you know about noncontingent schedules of reinforcement, explain why this is not surprising.

8. If most of our daily behaviors are reinforced under complex schedules (e.g., conjunctive schedules, adjusting schedules, chained schedules), why do you think psychologists have spent so much time studying simple schedules?

References

Boyce, T. E., & Hineline, P. N. (2002). Interteaching: A strategy for enhancing the user-friendliness of behavioral arrangements in the college classroom. *The Behavior Analyst, 25,* 215-226.

Engelmann, S., & Carnine, D. W. (1982). *Theory of Instruction: Principles and Application.* New York: Irvington.

Goto, K., & Schneider, J. (2010). Learning through teaching: Challenges and opportunities in facilitating student learning in food science and nutrition by using the interteaching approach. *Journal of Food Science Education, 9,* 31-35. doi:10.1111/j.1541-4329.2009.00087.x

Guerin, B. (1994). *Analyzing Social Behavior: Behavior Analysis and the Social Sciences.* Reno, NV: Context Press.

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher, 38,* 365-379. doi:10.3102/0013189X09339057

Keller, F. S. (1968). Good-bye teacher… *Journal of Applied Behavior Analysis, 1,* 79-89. doi:10.1901/jaba.1968.1-79

Latané, B., Williams K. D., & Harkins S. G. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology, 37,* 822-832. doi:10.1037/0022-3514.37.6.822

Lindsley, O. R. (1964). Direct measurement and prosthesis of retarded behavior. *Journal of Education, 147,* 62-81.

Mau, W-C., & Lynn, R. (2001). Gender differences on the Scholastic Aptitude Test, the American College Test and college grades. *Educational Psychology, 21,* 133-146.

Saville, B. K., Cox, T., O'Brien, S., & Vanderveldt, A. (2011). Interteaching: The impact of lectures on student performance. *Journal of Applied Behavior Analysis, 44,* 937-941. doi:10.1901/jaba.2011.44-937

Saville, B. K., Lambert, T., & Robertson, S. (2011). Interteaching: Bringing behavioral education into the 21st century. *The Psychological Record, 61,* 153-166.

Saville, B. K., & Zinn, T. E. (2009). Interteaching: The effects of quality points on exam scores. *Journal of Applied Behavior Analysis, 42,* 369-374. doi:10.1901/jaba.2009.42-369

Saville, B. K., Zinn, T. E., & Elliott, M. P. (2005). Interteaching versus traditional methods of instruction: A preliminary analysis. *Teaching of Psychology, 32,* 161-163. doi:10.1207/s15328023top3203_6

Saville, B. K., Zinn, T. E., Neef, N. A., Van Norman, R., & Ferreri, S. J. (2006). A comparison of interteaching and lecture in the college classroom. *Journal of Applied Behavior Analysis, 39,* 49-61. doi:10.1901/jaba.2006.42-05

Scoboria, A., & Pascual-Leone, A. (2009). An 'interteaching' informed approach to instructing large undergraduate classes. *Journal of the Scholarship of Teaching and Learning, 9,* 29-37.

Skinner, B. F. (1968). *The technology of teaching.* New York: Appleton-Century-Crofts.

Williams, K. D., Harkins, S. G., & Latané, B. (1981). Identifiability as a deterrent to social loafing: Two cheering experiments. *Journal of Personality and Social Psychology, 40,* 303-311. doi:10.1037/0022-3514.40.2.303

Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11,* 203-214. doi:10.1901/jaba.1978.11-203

Journal of the Scholarship of Teaching and Learning, Vol. 13, No. 2, May 2013, pp. 31 – 48.

Experiential learning: Exploring its long-term impact on socially responsible behavior

Jay Caulfield[1] and Treesa Woods

Abstract: Today's students are exposed to world events that require considerable cross cultural understanding and recognition that education is far more than learning facts about specific disciplines and diverse groups while sitting in a classroom. For the past several decades, research in education has repeatedly demonstrated that adults learn effectively through experience. However, does experiential learning, when designed specifically to heighten awareness of a significant social problem, evoke socially responsible behavior specific to that problem in the long run? Employing a qualitative longitudinal research design involving 25 graduate students as participants, this study explored that question. Findings indicated that 94.7% of participants who reported a high impact learning experience when participating in experiential learning while enrolled in a graduate class also reported engaging in socially responsible behavior because of that learning experience. In some instances, the socially responsible behavior continued for as long as three years after the class had ended.

Keywords: experiential education, experiential learning, high impact learning, qualitative longitudinal research, social responsibility

Today's students are exposed to world events that require considerable cross cultural understanding and recognition that education is far more than learning facts about specific disciplines and diverse groups when sitting in a classroom. College curricula are becoming more interdisciplinary in practice as evidenced by the interweaving of topics such as civic engagement, ethics, global studies, leadership and social responsibility within course work across curricula, with the intended goal of influencing students to become more socially responsible global citizens. Increasingly more learning is taking place within communities versus solely within classrooms. In fact in the UK, a national council has been formed for learning outside of the classroom (http://www.lotc.org.uk/) and the international "Association for Experiential Education" (AEE) (http://www.aee.org/membership/) reports over 1,200 members from 31 countries across the globe.

The purpose of this study was to explore the long-term impact of learning about social problems in environments outside of the classroom. In the educational literature, learning occurring outside of the classroom frequently meets the definition of *experiential learning*.

I. Background for Experiential Learning and Applied Theoretical Framework.

Historically, psychologists who are behaviorists generally believe that learning occurs through conditioning and conditioning occurs when individuals interact with their environment (Hergenhahn & Olson, 2000). "For behaviorists, observable interactive behavior demonstrates

[1] Marquette University, jay.caulfield@marquette.edu

learning" (Caulfield, 2011, p. 33). While the field of psychology looked to behaviorism to explain learning, around the same time, the philosopher John Dewey became a major contributor to the idea of *learning by experience*. He wrote extensively of its importance to democratic society in works such as *Democracy and Education* (1916) and *Experience and Education* (1938), where he proposed:

> . . . since democracy stands in principle for free interchange, for social continuity, it must develop a theory of knowledge which sees in knowledge the method by which one experience is made available in giving direction and meaning to another. (1916, p. 248)

Providing practical guidelines for such educational delivery in experience, he argued for "a sound philosophy of experience" (Dewey, 1938, p. 91) and defined the roles of educators as facilitators connecting learning to students' experiences, helping shape student understanding through "cooperative enterprise, not dictation" (Dewey, 1938, p. 72).

Although Dewey laid the preliminary framework for experiential learning, he did not refer to it as such. Hoover and Whitehead (1975) came forward with the following definition of experiential learning published in the "Association for Business Simulation and Experiential Learning" following its second annual conference proceedings: "Experiential learning exists when a personally responsible participant cognitively, affectively, and behaviorally processes knowledge, skills, and/or attitudes in a learning situation characterized by a high level of active involvement" (p. 25).

It is the well-known work of Kolb (1984), however, that really brought experiential learning to the forefront in the educational literature. Kolb postulates, "Learning is the process whereby knowledge is created through the transformation of experience" (p. 38). Kolb theorized that learning takes place when individuals engage in a novel experience, reflect upon it, conceptualize it and then test its authenticity by applying it to similar future experiences. Kolb's theory of learning is particularly interesting, as it integrates learning styles with the cyclical stages of how he proposed experiential learning occurs.

Boulding (2000), renowned for her extensive academic work in the study and promotion of peace, has noted the power of experiential education and service learning, in particular in promoting peaceable communities (p. 232). She expresses her appreciation of experiential education for its ability to connect students to "real-life situations," and expand students' "personal development and capacity for intellectual analysis" (p. 231). Such activities are part of open learning systems, which Boulding identifies as rooted in relationships with others and with the planet.

Eyler and Giles (1999) also noted that "service learning is a form of experiential education where learning occurs through a cycle of action and reflection as students work with others through a process of applying what they are learning to community problems" (p. 14). Eyler and Giles (2002) believe that service learning is an especially effective teaching strategy for students to learn about complex issues within larger environments, and to discover how complex systems are integrated and reliant on each other. Students gain experience and learning transfers because "students apply concepts repeatedly in real or realistic settings" (p. 148). Service learning mediated courses have reliably reported high levels of student engagement (Turner, 2002). Research suggests that experiential learning can help students develop positive attitudes toward life, encourage acceptance of responsibility, promote community involvement, develop power of thought and help them understand their strengths and weaknesses in a real

world context, thereby inspiring personal growth through the development of global competencies (Eyler & Giles, 1999).

The more recent work of Fenwick (2000, 2001, 2003) further develops the concept of experiential learning in terms of five perspectives, specifically constructivist; psychoanalytic; situative; critical cultural and enactivist ecological. Fenwick explains that a major conception of experiential education assumes that a learner will reflect on concrete experience to build new understandings. Within Fenwick's constructivist perspective, the learner reflects on lived experiences; these experiences are generalized to form mental structures. It is these structures of knowledge that are stored in memory as concepts that may be represented, expressed, and transferred to new situations, affecting how a learner perceives and interprets information. Fenwick's situative perspective identifies learning as being rooted in the situation in which learners find themselves, including specific interactions that occur within social settings. The psychoanalytic perspective stresses the necessity for learners to work through inner conflicts enabling them to engage fully in the learning process. The critical cultural perspective emphasizes the imbalance of power and inequity within social structures, which may have a direct impact on social transformations of experiential learning. Finally, the enactivist ecological perspective proposes that learning may best be understood through co-emergence of knowledge with other participants. Throughout this study, the researchers identify instances where the interview data is reflective of Fenwick's perspectives.

Today experiential education is popular on college campuses, often taking place outside of the classroom. The AEE (2012, ¶ 2) offers the following definition for experiential education:

> Experiential education is a philosophy that informs many methodologies in which educators purposively engage with learners in direct experience and focused reflection in order to increase knowledge, develop skills, clarify values, and develop people's capacity to contribute to their communities. http://www.aee.org/about/whatIsEE

Drawing from this definition, in the context of this study, socially responsible behavior refers to behavior that enhances social well-being within communities.

Leaders in both the "American Association of State Colleges and Universities" (AASCU) and the "Association of American Colleges and Universities" (AAC&U) have advocated for experiential education as a means of fostering civic and global engagement. In a May, 2008 concept paper for AASCU, George Mehaffy, Vice President for Academic Leadership and Change, and Harry Boyte, the Co-Director of the Center for Democracy and Citizenship at the University of Minnesota, challenged institutions of higher education to remain connected to the communities in which their students are themselves engaged (p. 3). Boyte and Mehaffy (2008) suggest that higher education should be promoting "citizen learning" that connects students to places and gives them opportunities to develop their "skills and learning habits" to live fully and integrally in community (Boyte & Mehaffy, 2008, p. 5).

College curricula apply experiential learning to assist students in achieving learning outcomes. Research in education for the past several decades (Jarvis, Holford, & Griffin, 1999; Kolb, 1984; Fenwick, 2000, 2001, 2003) has repeatedly demonstrated that adults learn effectively through experience. *However, does experiential learning, when designed specifically to heighten awareness of a significant social problem, evoke socially responsible behavior specific to that problem in the long run?* The purpose of this qualitative longitudinal study is to explore that question, thereby hypothesizing that experiential learning may contribute to

students' awareness of social issues, potentially evoking socially responsible behaviors that remain active well after a formal learning experience has ended.

II. Methodology.

A. Design.

The research question drove the design choice for the study. Farrall (2006) writes:

> QLR [qualitative longitudinal research] embodies a range of mainly in-depth interview-based studies which involve returning to interviewees to measure and explore changes which occur over time and the processes associated with these changes. The approach is particularly useful if one is studying a process which has a notion of a 'career' of some sort or *which involves a developmental process* [emphasis added]. (p. 2)

Molloy and Woodfield (2002) explain, "Longitudinal qualitative approaches have been used extensively in the fields of sociological research, ethnography and social history to explore individuals' changing life experiences and life course patterns" (p.5). They further state, "The chief goal of longitudinal methods is to explore change (relating to the unit of enquiry, at an individual or case study level) over time" (p. 10).

Creswell (2007) suggests that qualitative research may be accomplished by talking directly with participants, which supports the use of semi-structured interviews as a means of data collection for this study. Thus, employing a QLR design to study the research question posed seemed most appropriate for obtaining an understanding of what the interviewees experienced when engaged in experiential learning and whether, at a future point in time, these experiences influenced them to become social advocates for change specific to the social problem they had studied. The research was unfunded.

B. Sampling Method.

The researchers employed non-probability, purposive sampling to select 25 voluntary participants who were either graduate students or alumni previously enrolled in one of three applied social science elective classes offered in two graduate degree programs; in one instance, a participant was enrolled in two of the three classes. The median age of students in one of the degree programs was 26, and 19 of the 25 participants were from that program. The median age of students in the other degree program was 30. There were 18 female participants. Racial origin was as follows: eighteen (18) Caucasians, four African Americans, two Hispanics and one Asian. The programs were offered at a highly ranked research and teaching university located in the Midwest. All three classes were approximately 14 weeks in duration, designed to include significant experiential learning. Participation in the study was voluntary by written consent and all who consented were included in the research.

C. Data Collection.

In the first class of 20 students, 13 (65%) chose to participate in the study. As indicated by the course syllabus, the purpose of the experiential learning component was to increase awareness of the issues surrounding urban poverty and the strategies to sustain existence for those living in

urban poverty, utilizing community social support services that were available. To maintain confidentiality of both students and professor, the name of the professor and the course has not been disclosed. All students enrolled in the class were given mock identities of graduate students who were living independently and earning minimum wage. The objective of their mock identity was to avoid becoming homeless; they were encouraged to explore and utilize community social support organizations available to low income individuals. Students did weekly journal entries regarding their experiences and their perceptions of those experiences. Any classroom activities were predominantly discussion based and focused on student perceptions of urban poverty as lived in their mock identities.

At the conclusion of the class, the students did a final presentation that was in the form of a filmed documentary whereby they explained what they had learned while living their mock identities throughout the semester and whether they had met their objective of avoiding homelessness. During their presentations, students identified their perceptions of the lived experiences and, in some cases, identified specific socially responsible actions they intended to take because of the impact of the learning experience. The documentary served as the first point of data collection in the study. Approximately three years following the conclusion of the class, the 13 participants were interviewed, which was the second and final point of data collection for the participants who had enrolled in this class.

The second and third classes had the same course numbers and titles and were taught by the same professor, but were offered approximately 18 months apart. In the second class of 12 students, four (33%) chose to participate in the study. In the third class of 18 students, nine (50%) chose to participate. The topic studied in these classes was sustainability. Sustainability is defined as an "international movement for meeting the needs of current generations while preserving the capacity of future generations to meet their needs" (Sustainable World Coalition, 2010). The course description stated that the "interdisciplinary course explores the values, knowledge base, and organizational models of sustainability, an international movement for meeting the needs of current generations while preserving the capacity of future generations to meet their needs." To maintain confidentiality of both students and professor, the name of the professor and course has not been disclosed. For these two classes, students visited multiple community organizations that were actively engaged in sustainability projects. They spoke extensively about sustainability with leaders from within these organizations. Any classroom activities occurring were predominantly discussion based and focused on their perceptions of the onsite visits. At the conclusion of both of these classes, students were asked to write a *bold proposal* that identified what they intended to do personally in the future to promote sustainable environments. The bold proposals served as the first point of data collection for the study. Participants from the second class were interviewed approximately two years following the termination of the class while students from the third class were interviewed approximately nine months following the termination of the class; the interview was the second and final point of data collection for participants from these two classes.

D. Format of Semi-Structured Interviews.

According to Rubin and Rubin (2005) open-ended questions encourage the interviewee to describe specific events and experiences. Rubin and Rubin (2005) further explain that probing is a technique used to keep a discussion going while providing clarification. This allows the interviewee to keep talking about the matter at hand, complete ideas, fill in blanks, or request

additional information on what was said. The semi- structured open-ended questions that follow were prepared by the researchers prior to interviewing. For the one student who had enrolled in both classes, questions were asked about the urban poverty class and the sustainability class. The researchers interviewed participants from each class over a period of no more than four weeks, recording and transcribing the interviews.

E. Interview Questions.

Introductory question for all participants: To the best of your ability, please list the classes that you were enrolled in as a graduate student. Please identify which three classes most impacted your life in some way.

Questions for those participants who *did* identify the urban poverty or sustainability class as being one of the three most impactful classes (n = 19):

1. Explain how your knowledge of social issues [urban poverty or sustainability] has been altered because of completing this class.
2. Please explain how this class impacted your life.
3. Here are the program learning outcomes and the student learning objectives for the class. Please explain how the learning experiences in the class related to or did not relate to the program outcomes and the student learning objectives.
4. Has your behavior changed because of taking this class? If so, how has it changed? If not, why not? Please cite specific examples.
5. You identified two other classes that were impactful to you. Please explain what learning activities made each of these classes impactful. Has your behavior changed because of taking these classes? If so, how has it changed? If not, why not? Please cite specific examples.

Questions for those participants who *did not* identify the urban poverty or sustainability class as being one of the three most impactful classes (for all three classes, n = 7):

1. You did not select the class [urban poverty or sustainability] as one of the classes that most impacted your life. As best as you can, please explain why you did not choose this class.
2. Please explain what would have made this class [urban poverty or sustainability] more impactful to you.
3. You identified three other classes that were impactful to you. Please explain what learning activities made each of these classes impactful to you. Has your behavior changed because of taking these classes? If so, how has it changed? If not, why not? Please cite specific examples.

In summary, the overarching question of interest was whether experiential learning, when designed specifically to heighten awareness of a significant social problem, evokes socially responsible behavior specific to that problem over the long run. The research design was QLR. The initial data collection point for the first class studying urban poverty originated from *filmed documentary* of participants' experiences at the conclusion of the class while the initial data collection for the two classes studying sustainability originated from the participants' documented *bold proposal* at the conclusion of the class. Semi-structured interviews of participants occurring at nine (9), 24 or 36 months following the termination of the class they had enrolled in were the second data point. At point of interview, 60% of the participants had already graduated.

III. Findings.

In response to the introductory interview statement, "Please identify which three courses most impacted your life in some way," participants identified 22 different classes as their top three choices. Of those 22 classes, nine were mentioned three or more times. For a rank order summary of the nine classes most frequently identified by participants as those that "most impacted their lives in some way," please refer to Table 1.

Table 1. Top Nine Classes Ranked as Most Impactful.

Topic Description	Top Choice	Second Choice	Third Choice	Total Choice
1. Sustainability (n = 4 , 9)	5 (1)*	3 (1)*	(2)*	8 (4)*
2. Leadership & Ethics	5	2	2	9
3. Urban Poverty (n = 13)	5	1	1	7
4. Ethics in Public Sector	2	4	1	7
5. Social Entrepreneurship	1	1	2	4
6. Social Justice		1	3	4
7. Police Leadership		3		3
8. Public Policy	2		2	3
9. Research Methods		1	2	3

*The number in parentheses indicates the number of participants from the class where interviews occurred 24 months following termination of the class (n = 4). Remaining interviews occurred at nine (9) months following termination of the class (n = 8).

A. Sustainability Classes.

As noted in Table 1, the sustainability class was the top choice for most participants and had the total highest number in the "Total Choice" column for participants, with only one participant from both of the classes in sustainability that did *not* select the class as one of the top three most impactful classes. That one participant who was interviewed nine months following the termination of the sustainability class stated the following: "… just because I did not choose this class [sustainability] it does not mean it did not impact my life, because I see sustainability in everything that I do in my daily life." He further explained that the top three classes that he did choose were ones that taught him how to make a bigger impact on *others' lives* (by profession he was a counselor) whereas in the sustainability class, he felt that the impact from the class pertained mostly to his own life.

Knowledge of social issue, impact and actions. In all cases, participants in the sustainability classes indicated during their interview that they had a heightened sense of awareness regarding sustainability and that this newfound awareness impacted their lives in one or more ways. In general, the participants reported making a conscious effort to actively think about sustainability when carrying out daily tasks of living, such as when buying groceries for their family, watering the lawn and plants (conserving water), recycling, and riding their bike more versus driving a car. One participant contacted the local "Office of Sustainability" to learn more about sustainable work environments. Another participant stated:

> While I was taking the class I was really self-conscious about how much water I was using and simple practical things; some of that wore off a little especially as I became less immediately connected to a lot of the literature telling me how much

> water I was wasting so once the shame factor lessened I didn't become as self-conscious about it, but I think it increased my desire to learn a lot more [about sustainability].

Another said:

> I remember first feeling incredibly guilty about personal things I was doing like using something as small as a paper towel. I would feel guilty like I was not living the way we all needed to be living and moving in a certain direction. Then I kind of readjusted and realized that all these problems are so institutionally driven in such a large scale that anything I do, I can't really beat myself up about, because we got to make changes on a huge scalable level like on an institutional level, a federal level, on a larger scale.

These two quotes make evident Fenwick's (2000, 2001, 2003) psychoanalytic perspective, which stresses the necessity for learners to work through inner conflicts before they are able to engage fully in the learning process.

The participants' bold proposals written at the conclusion of the class identified personal actions to take. Although all participants indicated engaging in actions to promote sustainable environments, those actions did not necessarily match the ones identified in their bold proposals. Based on the nature of the bold proposal and the time of interview, participants may not have had time to act upon what they had identified in their bold proposal. Table 2 contains data from bold proposals and interviews of six participants, mainly direct quotes regarding knowledge of the social issue, impact of acquired knowledge, actions identified in bold proposals and reported actions engaged in at time of interview.

Note that several of the direct quotes in Table 2 may be applied to the perspectives in Fenwick's (2000, 2001, 2003) experiential framework. As an example of the constructivist perspective, consider the following comment from Participant 6: "I now see sustainability as a social issue locally (community development) and at a global level, especially as related to sustainable cultures." The quote is indicative of how the student's reflection on a learning experience expanded her mental structure of sustainability, which in turn impacted the socially responsible actions she identified taking. Another example of the constructivist perspective is illustrated in the following quote taken from Participant 5: "I would say this class basically blew up my understanding of social issues. Not that I was unaware, for I have always been aware of social problems. But, it gave more urgency." Again, this participant's mental structure of sustainability had changed because of participating in the class. By time of interview, he had organized a successful community clean-up day, demonstrating the urgency to act on his expanded view of sustainability.

Learning objectives and program outcomes. In all cases, interviewees reported during interview that they saw alignment with what they had learned in the class, the learning objectives as stated in the syllabi for the two classes and at least a few of the degree program learning outcomes. In most instances, participants reported alignment with all five program outcomes.

B. Urban Poverty Class.

As noted in Table 1, the urban poverty class was the third "Top Choice*"* for most participants and was third (along with public service ethics) in the "Total Choice" column for number of participants. For those seven (7) students that did choose the urban poverty class, during interview, six (6) reported socially responsible actions taken because of enrolling in the class.

Table 3 contains data from the documentary filmed at the conclusion of the class and data from interviews of five participants enrolled in the class. As in the case of Table 2, the entries are mainly direct quotes regarding knowledge of the social issue, impact of acquired knowledge, actions or impressions identified in the filmed documentary and reported actions engaged in at time of interview approximately 36 months following the termination of the class.

Table 2. Sustainability: Knowledge of social issue, impact and actions.

Participant 1 (24)*	
Knowledge of Social Issue	"I'm ashamed to say that until registering for the class, I've not given much thought about sustaining our natural capital. I didn't really have a foundation for sustainability . . . it [the class] opened my eyes to the things that were going on in our own community that I pay no attention to."
Impact of Acquired Knowledge	"But to go and sit in front of these people, listen to them [leaders in sustainability] talk about their passion, listen to them talk about why they do what they do, have them take us through their experience, I think that was the most beneficial thing that he [professor] could have done for someone like me and I would recommend the class to people because of that."
Actions Identified in Bold Proposals	"Create a program for student athletes focused on personal awareness, reflection and gratitude [method of creating heightened awareness of sustainability] as the foundation."
Actions Taken by Time of Interview	Talked with three individuals in department about building a "life skills" program for student athletes and incorporating content regarding sustainable futures into the program. Two of those individuals have begun to build the program.
Participant 2 (24)*	
Knowledge of Social Issue	"I think a lot of what we covered in the class I had a general interest in going into the course, but I can't say that I had the necessary knowledge" [referring to sustainability].
Impact of Acquired Knowledge	" . . . we heard from someone talking about local foods and eating local . . . I am more mindful of where the milk's coming from when I purchase . . .we went to a fair trade market . . . just understanding where the products I use are coming from . . .those are the areas where I have probably been impacted the most."
Actions Identified in Bold Proposals	"Begin with an element of sustainability in our next advancement event [worked in university advancement] to begin the process and build momentum."
Action Taken by Time of Interview	"I assisted the College of Business with a sustainability panel discussion followed by round table discussions around the topic and so, having the knowledge from this class was actually incredibly beneficial to be a part of that. And then, through my work here whenever we have roundtable programs we have been able to infuse that sustainability topic in a number of ways so, either bringing it in as one of our roundtable topics that our attendees can choose from or just continuing to help raise that awareness and have that be a part of the different programs we offer. ". . . being mindful of the food that I am eating and where it's grown, how it's produced, just being mindful of my environment and I think just being aware of where I choose to spend my money reflects how I feel, and knowing that that's a way I can communicate what I feel strongly about I guess, where I choose to buy products or organizations that I support can have an impact."
Participant 3 (9)*	
Knowledge of Social Issue	"Prior to taking the class, I was not familiar with the sustainability concept. The course offered me the ability to understand in more detail how our society is set-up as a consumer based society."
Impact of Acquired Knowledge	"The professor took us on many field trips which allowed me to view the world from a way or perspective I never really viewed the world before."
Actions Identified in	"I propose a national service program that would require all 18 year old Americans to serve their country for two years in a local, state, national or international "Peace Corps" or military

Bold Proposals	unit."
Action Taken by Time of Interview	"I'm more aware of what I do each day. Even with buying things . . . we were having issues with our dryer. Instead of buying a new one, which would have been easier, I took it [dryer] apart and it was just a $20 fuse."
Participant 4 (9)*	
Knowledge of Social Issue	"I had some knowledge prior to this because I use to teach Environmental Education in California, but I still did change; the class changed my thinking with more collaboration on community building."
Impact of Acquired Knowledge	"Where I worked [previously to becoming a student], we were in the middle of nowhere and people came to us. So I was not part of a community . . . this class gave me a lot of resources again and thoughts on how to build a community, and being the solution, being more locally focused moving towards a solution as it relates to sustainability."
Actions Identified in Bold Proposals	"Partner with environmental organizations that can go to schools for environmental education and promotion of a week at an environmental camp."
Action Taken by Time of Interview	"trying to find new partnerships with other [sustainability] programs . . . the urban ecology center is making a new place right next to the stadium, and I have gotten their e-mails so now I'm trying to start a collaboration with them, because we could actually bike there and we have biked along the adjoining trail before. . . I sent out e-mails to the community center to start thinking about ways to collaborate with them because they are so close to us now, and we can do community service to help with the building of it and with the programs later . . . with just having the class I think more about community building and using resources that are right here to help form partnerships."
Participant 5 (9)*	
Knowledge of Social Issue	"I would say this class basically blew up my understanding of social issues. Not that I was unaware, for I have always been aware of social problems. But, it gave more urgency."
Impact of Acquired Knowledge	" . . getting more in-depth with the issues and how immediate some of them are. So basically this class gave me a feeling that "yes" I can help with this (sustainability) and be one of the players who tries to fix the issues. . . every other issue is resting on top of the ultimate social issue, which is just take care of our momma, which is mother earth, because none of this is even possible--there are no other social issues if we can't sustain life on this planet . . . every other social issue that I found important [went] back to this foundation of sustainability, with sustainability being the ultimate battle that we will be faced with in the next few decades."
Actions Identified in Bold Proposals	"Volunteers compete in teams for a chance to win meaningful prizes and gift cards by being the team who completes the most improvement challenges and clears the most pounds of trash and recyclables from the streets of our community."
Action Taken by Time of Interview	"The [bold proposal] event was a great success! I am still in a celebratory mood and it's definitely sinking in how my 'bold proposal' turned out to be such a successful real-life venture. Here's some of the outcomes from the day... Garbage and recyclables = 2,972 lbs. Illegally Dumped Materials = 403 lbs. Total = 3375 lbs. Total Cigarette Butts = 3,682 Total attendance = 28 teams of 4 or 5 (approximately 126 community members) Approximately 150-180 people during the celebration part of the afternoon"
Participant 6 (9)*	
Knowledge of Social Issue	"I now see sustainability as a social issue locally (community development) and at a global level, especially as related to sustainable cultures."
Impact of Acquired Knowledge	"I learned about the urban ecology center and did an internship there. I now volunteer there on a regular basis."
Actions Identified in	"Include sustainability courses as part of the mandatory curriculum in schools from k-12."

Bold Proposals	
Action Taken by Time of Interview	"I make a conscious decision to ride my bike more than previously, I make conscious decisions about water consumption, practice recycling, think about what I use—like every time I grab for the paper toweling, I think about whether I need to use it, and groceries—where to buy them and what to buy and how what I buy affects the world. For example, apples grown in the State of Washington may be shipped out of the country to become waxed and then are returned for sale."

*Number in parentheses indicates the number of months following the termination of the class that the interview occurred.

Table 3. Urban Poverty: Knowledge of social issue, impact and actions.

Participant 1	
Knowledge of Social Issue	"I learned so much about the rich history of the city. I learned to see parts of the city and its inhabitants in an entirely different way."
Impact of Acquired Knowledge	"The class changed my life in that I have a better understanding of different social classes and different ethnicities."
Actions/Impressions Identified in Documentary	"I learned to respect the dignity of those individuals" [referring to the homeless population] … "I began to better understand my role as a servant of the city" [participant was a police lieutenant].
Actions Taken by Time of Interview	"I teach my officers to respect all classes of people and all neighborhoods … I tend to look at crimes committed and more carefully consider the need for punishment to fit the crime."
Participant 2	
Knowledge of Social Issue	"It [the class experience] kind of made me realize how quick a couple of life events can really turn your life upside-down … I have a better understanding of the struggling single mother and how she has to deal with things."
Impact of Acquired Knowledge	"It made me a lot more conscious with my own personal budget—trying to make sure I stayed on top of things."
Actions/Impressions Identified in Documentary	"When I get those bags from the Boy Scouts to put food in I used to throw them away or use them for garbage. Now I put food in them for other people."
Action Taken by Time of Interview	"I feel more knowledgeable when people make a comment about something like a single mom who is on welfare, and I am able to say that it is really hard to make it out there on a minimum wage. I feel more empathy for people in those situations. "
Participant 3	
Knowledge of Social Issue	"In my undergrad I did take a lot of sociology classes so they did focus a lot on social issues and I think for this class it was just very nice to see how the different migration patterns impact a city and with that it also brings a lot of social issues based on the lower working class or the middle class."
Impact of Acquired Knowledge	"I just think that overall talking about those issues and then seeing like the homeless and putting everything together was very significant for me."
Actions/Impressions Identified in Documentary	"…I can volunteer my time to help people that go through these situations" [poverty, homelessness].
Action Taken by Time of Interview	"Last spring we actually did a volunteer program and we did a soup kitchen … it's just really interesting how this class really applied to the real world and how many people are not aware of some of the bigger social issues in the city and it's sad that we have to come up as an office and go there as opposed to us taking a personal interest and learning about this type of issue."
Participant 4	
Knowledge of Social Issue	We learned about the diverse ways that urban cultures and subcultures settled in urban areas … I definitely look at the health care programs that the state would offer infants and young children in a much different light now. This is something that I might not have thought of before taking this class, because I did not have kids. This is something that I would remember from this class. All these different social programs that you had to look up

	for yourself, to see if you could apply for or be considered for. This was a different way of thinking for me, a different way of making me look at life."
Impact of Acquired Knowledge	"I thought this was one of my best classes ever in college life, which took me by surprise, because it was not anything directly related to law enforcement. I think that this class made me a tremendously better supervisor at work."
Actions/Impressions Identified in Documentary	Speaks of learning to budget money very carefully. "At $7.00 an hour after taxes you are basically working an hour to pay for a gallon of gasoline and that gallon of gasoline isn't going to take you far."
Action Taken by Time of Interview	"I am a lot more sensitive to programs that the city would offer. For instance the church I go to … we will form groups and go to the meal program as volunteers, and it is a very humbling experience. I know that if you don't know where your next meal is coming from, it is a tough life. Especially when you start seeing kids coming through, and when the night is over and you see kids walking into the darkness with mom and dad, this can really affect you as a person. I understand that even more so after taking this class; this class has made me much more aware of that. The times I spent volunteering at church, this kind of stuff really does affect a person."
Participant 5	
Knowledge of Social Issue	"This class helped me to understand the problem of homelessness. Again, those services [referring to social support services] are needed for the homeless. That experiment we did like interviewing for the shelter opened my eyes. These experiences are not fun, and they do affect people. Not having a stable home can change a person's life, and having resources can help people. This class really allowed one to think about homelessness and how people become homeless … It also brought together other cultures as well. So you got to see what's similar and what's not similar**.**"
Impact of Acquired Knowledge	"What made this class very impactful was when we did a social experiment. Each student was assigned a mock identity, and the goal was to prevent ourselves from becoming homeless. I actually interviewed with the head person of the shelter who asked me many questions like how did I get into this situation I was in. So basically she asked me the same questions that they would ask someone who was really homeless, which I thought it was a good experience, because it brings about an awareness of how people can actually become homeless."
Actions/Impressions Identified in Documentary	"Just doing this experiment is something I think everybody should do to make them more aware of homelessness."
Action Taken by Time of Interview	"I have a German co-worker who talks to her parents in German for that is their native culture. I also have a friend who is of Latina decent and I have been trying to learn that language, and a lot of what me and this friend do, our culture is very similar … So, basically I mix and mingle with my co-workers a bit with the Spanish and German decent." Participant explains that she realizes that getting to know other cultures better promotes an understanding of those cultures.

Several quotes of participants reflected a perspective highlighted in Fenwick's (2000, 2001, 2003) experiential framework. For example, "I came into this experiment pretty naïve, thinking that homeless people wanted it to be that way. I learned to respect the dignity of those individuals" [referring to the homeless population] … "I began to better understand my role as a servant of the city," is reflective of the critical cultural perspective, which emphasizes the imbalance of power and inequity within social structures that may have a direct impact on social transformations of experiential learning. The enactivist ecological perspective, which proposes that learning may best be understood through co-emergence of knowledge with other participants, is reflected in the following quote from Participant 5: "What made this class very impactful was when we did a social experiment. Each student was assigned a mock identity, and the goal was to prevent ourselves from becoming homeless … and everyone had to report on their situation and how they resolved things."

Six (6) of the 13 participants enrolled in the urban poverty class did *not* choose the class as one of the three most impactful classes. The reasons stated for *not* identifying the class as most impactful were as follows:

Reason 1: "I was not from the area [city where university was located] and the class focused on that one city, specifically."

Reason 2: "It was too limiting, focusing only on one city. The purpose of the class was not clearly identified."

Reason 3: "The class was not realistic. There were too many restrictions that the professor made us conform to that I don't think would have happened in real life."

Reason 4: "I didn't select the class because the topic was too related to the work that I already do."

Reason 5: "More than anything, it provided a good history of the city. I did not see its relevance to public service."

Reason 6: "The book was excellent, but I felt pressured to participate in the role play."

It is interesting to note that Reasons 3 and 6 are in themselves reflective of Fenwick's (2000, 2001 and 2003) critical cultural perspective in that the identified reasons indicate a perceived imbalance of power between instructor and students and, in this case, that perception likely had a negative impact on experiential learning.

Learning objectives and program outcomes. In all but one case, interviewees reported during interview that they saw alignment with what they had learned in the class, the learning objectives as stated in the syllabus for the class and at minimum a few of the degree program learning outcomes. In most instances, participants reported alignment with all program outcomes.

Aggregate findings. Table 4 reports the findings per class and the aggregate findings. As noted in reviewing Table 4 aggregate data, for this limited sample in the context of these three classes, experiential learning is reported by those participants identifying the experiential learning class as having high impact to have evoked socially responsible behavior specific to the social problem they studied 94.7% of the time. Again, interviews occurred at nine (9), 24 or 36 months following termination of the class, indicating that the reported socially responsible actions continued to occur for a relatively long period following termination of the class.

Table 4. Aggregate Findings.

Class	No. of months interviewed after class	No. of participants interviewed	No. who identified class as impactful	Of those impacted, number reporting action	Of those impacted, percent reporting action
1. Sustainability	9	9	8	8	88.9
2. Sustainability	24	4	4	4	100.0
3. Urban Poverty	36	13	7	6	85.7
Totals		26*	19	18	94.7

* One participant enrolled in one sustainability class and the urban poverty class.

IV. Discussion and Implications.

Study findings imply that experiential learning may lead to socially responsible behavior that continues to occur well after the formal learning experience ends. Secondly, several quotations seemingly validated the perspectives that Fenwick (2000, 2001, 2003) identifies in her experiential learning framework. In the case of the urban poverty class, based on participants' stated reasons for *not* selecting the class, one could argue that the participants may have been distracted by specific pedagogical design features of the class versus having a lack of concern regarding the social problem being studied. Two quotes from the six participants who did *not* choose the class for one of the reasons previously identified implied this. "I now realize how difficult it is for some single parents, so I volunteer my time to the food pantry and I donate food too. I take time to learn about resources that my church has to offer" [for those in need]. Another participant stated, "The fact of standing in line and waiting for appointments-- and I don't know why social programs have appointments in the middle of the day and you have to wait for two hours. You feel like a cow in line waiting for the slaughter." These quotations indicate the impact that studying the social problem had on these two participants, and in one case, the participant identifies socially responsible behavior as related to urban poverty.

From a pedagogical standpoint, the reasons that participants gave for not selecting the class demonstrates the importance of sharing these types of class design criticisms in course evaluations so that teachers become aware of them and may decide whether design changes might improve the overall learning experience. Teachers may motivate students to provide this type of constructive feedback by explaining that it will be used to improve future classes (Caulfield, 2007). It is interesting to note that in two cases, participants stated that the class focused too much on one specific city, indicating that these participants likely did not make the connection that much of the content related to urban poverty, homelessness and the social experiment, would likely apply to other urban cities in the area. Thus, emphasizing that potential applicability to other cities may have been helpful for participants not residing in the area where the social experiment took place.

One final consideration when designing a class with significant experiential learning is that it may be more time intensive and more costly. It may also place students at a higher risk in certain types of environments outside of the classroom with no certainty that experiential learning is more effective over time than other forms of classroom learning.

A. Two Remaining Higher Ranked Classes.

It is interesting to note that the class titled, "Leadership & Ethics," which was the top choice for the second highest number of participants and the second highest number in total choice for participants, did have an experiential learning component; but it was not as significant a portion of the class as that of the urban poverty and sustainability classes. With the permission of the professor teaching the leadership and ethics class, the class evaluations were reviewed. In the comment section of the class many students mentioned that the leadership panel discussion that the students organized and facilitated with community leaders they had selected (the experiential component of the class), was the most valuable learning experience occurring during the class. As one student put it, "I felt that the most instructional portion of the class was the leadership panel … I found their ideas and suggestions much more applicable than many of the texts for my own leadership capacity." This quote again highlights what Fenwick (2000, 2001, 2003)

identifies as the enactivist ecological perspective, which proposes that learning may be best understood through co-emergence of knowledge with other participants.

The "Ethics in Public Service" class also had an experiential component where students had to apply ethical concepts to a current work experience in which they found themselves facing an ethical dilemma. As one student participant who had chosen the ethics class as his top choice indicated, "Ethics in public service was very impactful because I had a job in public service. I had the ability to take what I learned in the classroom and directly apply it to my job. Also, the instructor allowed us to give real world examples in the class." The preceding quote is a good example of the situative perspective that Fenwick (2000, 2001, 2003) identifies in her framework in that this perspective identifies learning as being rooted in the situation in which learners find themselves.

B. Additional Implications.

The remaining five classes, although discussion based, did not have a major experiential learning component that the researchers were aware of as indicated by reviewing the syllabi for the courses. However, it should be noted that many of the participants stated that the reason they identified a class as impactful was due to its immediate applicability to their respective work settings. Another words, experiential learning was occurring on the job at the same time as concepts were being learned in the classroom. This observation implies that full time students who are unemployed may benefit from internships and practicum experiences designed into the curriculum to facilitate experiential learning that complements other types of learning activities.

Finally, it appears that the point in time when a course is taken makes a difference in the degree of impact. The following three quotes provide evidence of this. "I think it [the class] was for me at that time of my life and career really important." Another participant states, "It kind of gave me training. I needed to work out a budget and find a home for myself and do food … at that time I was still living at home with my parents … but less than a year later I moved out of my parents' house and I now have to do my own budget." The final participant states, "I think part of it [the impact of the class] was about the time of my life. I was going through the death of someone very close to me … " Thus, it seems that significant life events make a difference in the depth of reflection occurring at the time or at some point following the class, which may also increase the impact of experiential learning.

C. Limitations to Research.

Clearly the specific context of the classes within this study limits transferability to other contexts. Secondly, the external environment in which the participants were engaged with over the course of 14 weeks was ever changing. Thus, shared perspectives may have changed during the time of the study, decreasing the dependability of the perspectives shared. Another limitation was the researchers' inability to determine whether the difference between selecting the urban poverty class and the sustainability classes as most impactful was due to the length of time from the termination of the class to time of interview, or if the differences were the results of other factors identified by participants who had enrolled in the urban poverty class and had *not* selected it as a high impact class. Lastly, lack of comparison of findings to other similar studies further limits transferability.

D. Future Research.

To substantiate further the work of Eyler and Giles (2002), future research in experiential learning may explore how integrated classroom education and authentic community learning experiences may influence students to seek a better understanding of the world in which they live by helping them to make sense of unpredictable and complex interactions. To add credibility to the AASCU and the AAC&U stated beliefs that experiential education is a means of fostering civic and global engagement, future research could explore how experiential learning may help students accept responsibility, promote community involvement and develop community collaborations that will aid in students' learning experiences. Lastly, future studies could continue to explore the relationship between experiential learning and other forms of learning and their overall effectiveness.

E. Conclusion.

Based on the generally accepted premise that learning occurs through experience, experiential learning has been and continues to be a vital way of learning for adults. As highlighted by Montrose (2008):

> The most important reason for integrating experiential learning into existing programs is the benefit that students derive from it. The opportunity to take an individual 'peak experience' and guide students to transcend the mere memory of a situation, to coach them to analyze, theorize, and to better understand themselves and others is an opportunity that many of us welcome and appreciate. (Conclusion, ¶ 3)

The purpose of this qualitative longitudinal study was to explore whether experiential learning about significant social problems outside of the classroom evoked socially responsible behavior specific to the social problems being studied. Findings from the study involving 25 graduate students as participants imply that experiential learning does lead to socially responsible behavior that continues to occur as much as three years after the formal learning experience has ended. Furthermore, interview data from the study seemingly validate the five perspectives identified in Fenwick's (2000, 2001, 2003) experiential framework, indicating its usefulness in studying experiential learning.

It is the researchers' hope that these study findings will encourage further study of experiential learning in college curricula and community organizations. As indicated by past research that complements findings from this study, experiential learning exposes learners to authentic social experiences that are more likely to engage them in future socially responsible behaviors, transforming them into agents of positive social change within their communities.

References

Association of American Colleges and Universities. (2002). *Greater expectations: A new vision for learning as a nation goes to college*. Washington, D.C. Retrieved from http://greaterexpectations.org/

Association of Experiential Education (2012). Retrieved from http://www.aee.org/about/whatIsEE

Boulding, E. (2000). *Cultures of Peace: The Hidden Side of History*. Syracuse, NY: Syracuse University Press.

Boyte, H. C., & Mehaffy, G. (2008). *The Civic Agency Initiative*. Washington, DC: American Association of State Colleges and Universities.

Caulfield, J. (2007). What motivates students to provide feedback to teachers about teaching and learning? An expectancy theory perspective. *International Journal for the Scholarship of Teaching and Learning,* 1(1), 1 – 19. Retrieved from http://academics.georgiasouthern.edu/ijsotl/v1n1/caulfield/IJ_Caulfield.pdf

Caulfield, J. (2011). *How to Design and Teach a Hybrid Course: Achieving Student-centered Learning through Blended Classroom, Online and Experiential activities*, Sterling, VA: Stylus.

Creswell J. (2007). *Qualitative Inquiry & Research Design: Choosing among Five Approaches*. Thousand Oaks, CA: Sage Publications.

Denzin, N.K., & Lincoln Y.S. (1994). *The Handbook of Qualitative Research*. Thousand Oaks, CA: Sage Publications.

Dewey, J. (1938). *Experience and Education*. New York: Touchstone.

Dewey, J. (1916). *Democracy and Education*. London: The Macmillan Company.

Eyler, J., & Giles, D. E. (1999). *Where's the Learning in Service-learning?* San Francisco: Jossey-Bass.

Eyler, J., & Giles, D. E. (2002). Beyond surveys: Using the problem solving interview to assess the impact of service-learning on understanding and critical thinking. In A. Furco and S. Billig (eds.), *Service-Learning: The Essence of the Pedagogy* (pp. 147-159). Greenwich, CT: Information Age Publishing.

Farrall, S. (2006) What is qualitative longitudinal research? *Papers in Social Research Methods, Qualitative Series no. 11.* London School of Economics and Political Science Methodology Institute: London, U.K. Retrieved from http://www2.lse.ac.uk/methodology/pdf/QualPapers/Stephen-Farrall-Qual%20Longitudinal%20Res.pdf

Fenwick, T. J. (2000). Expanding conceptions of experiential learning: A review of the five contemporary perspectives on cognition. *Adult Education Quarterly, 50*(4), 243.

Fenwick, T.J. (2001). *Experiential Learning: A Theoretical Critique from Five Perspectives.* Columbus, OH: Centre on Education and Training for Employment.

Fenwick, T. J. (2003). *Learning through Experience: Troubling Orthodoxies and Intersecting Questions.* Malabar, FL: Krieger Publishing Company.

Hergenhahn, B. R., & Olson, M. (2000). *An Introduction to Theories of Learning* (6th ed.) Upper Saddle River, NJ: Prentice Hall.

Hoover, J. D., and Whitehead, C. (1975), An experiential-cognitive methodology in the first course in management: Some preliminary results. In R. H. Buskirk (ed.), *Simulation Games and Experiential Learning in Action*, Richard H. Buskirk (ed.), pp. 25-30. Austin, TX: Bureau of Business Research, University of Texas at Austin.

Jarvis, P., Holford, J., & Griffin, C. (1999). *The Theory and Practice of Learning.* Sterling, VA: Stylus.

Kolb, D. (1984). *Experiential Learning: Experience as the Source of Learning and Development.* Englewood Cliffs, NJ: Prentice Hall.

Molloy, D., & Woodfield, K. (2002). *Longitudinal Qualitative Research Approaches in Evaluation Studies.* Department for Work and Pensions: London, U.K. Retrieved from http://research.dwp.gov.uk/asd/asd5/WP7.pdf

Montrose, L. (2008) International study and experiential learning: The academic context. *Frontiers: The Interdisciplinary Journal of Study Abroad*. Retrieved from: http://www.frontiersjournal.com/issues/vol8/vol8-08_montrose.htm

Rubin, H. J., & Rubin, I. S. (2005). *Qualitative Interviewing* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Sustainable World Coalition. (2010). *Sustainable World Sourcebook: Critical Issues, Viable Solutions, Resources for Action*. Berkeley, CA: Sustainable World Coalition.

Turner, R.M. (2002). A pragmatic approach to educating: Combining problem-based learning with service learning. *Abstracts*. ERIC number ED471964.

Journal of the Scholarship of Teaching and Learning, Vol. 13, No. 2, May 2013, pp. 49 – 71.

Changing attitudes and facilitating understanding in the undergraduate statistics classroom: A collaborative learning approach

Erin Curran[1], Kerri Carlson[2], and Dayius Turvold Celotta[3]

Abstract: Collaborative and problem-based learning strategies are theorized to be effective methods for strengthening undergraduate science, technology, engineering, and mathematics education. Peer-Led Team Learning (PLTL) is a collaborative learning technique that engages students in problem solving and discussion under the guidance of a trained peer facilitator. This comparative study investigates the impact of a PLTL-based learning community program on both content mastery and dispositions of undergraduate students taking an introductory course in applied statistics. Results suggest that students participating in the learning community program acquired significantly greater content mastery in statistics when compared to non-participating peers. Moreover, the learning community experience may provide students with a buffer against developing the negative attitudes and perceptions that often pervade the undergraduate applied statistics classroom.

Keywords: statistics education, collaborative learning, peer-led team learning, student attitudes, learning outcomes

Statistics educators face significant challenges with engaging undergraduate students in applied statistics courses. These students may have insufficient mathematical or statistical preparation for the statistics course that they are required to complete (Johnson & Kuennen, 2006), and it is not uncommon for undergraduate students to adhere to misconceptions about statistical concepts and faulty statistical reasoning (Garfield & Ahlgren, 1988; delMas & Garfield, 1999; Kahneman, Slovic, & Tversky (1982) as cited in Garfield, 1995; Konold, 1995; Hirsch & O'Donnell, 2001; Castro Sotos, Vanhoof, Van den Noortgate, & Onghena, 2009). Accompanying a lack of academic readiness for post-secondary coursework in applied statistics is often some degree of math-related anxiety or phobia, negative attitudes toward the content or discipline, an overall lack of interest in statistics, or the perception that the course is irrelevant to their chosen major (Gal & Ginsburg 1994; Gal, Ginsburg, & Schau, 1997). Ultimately, these attitudes have been found to be correlated with course performance (Finney & Schraw, 2003; Dempster & McCorry, 2009). Unfortunately for statistics educators, attitudes toward the course and course content, in general, have been found to be difficult to change (Garfield & Ahlgren, 1988; Garfield & Ben-Zvi, 2007). These challenges, both academic and attitudinal, may be particularly pronounced for those educators who teach courses for non-statistics or mathematics majors who enrolled in the

[1] Computer and Information Sciences Department, University of St. Thomas, O'Shaughnessy Science Hall, 2115 Summit Ave., St. Paul, MN 55105, ecurran@stthomas.edu.
[2] Biology Department, University of St. Thomas, Owens Science Hall, 2115 Summit Ave., St. Paul, MN 55105, kerri.carlson@stthomas.edu.
[3] Biology Department, University of St. Thomas, Owens Science Hall, 2115 Summit Ave., St. Paul, MN 55105, dlturvold@stthomas.edu.

course chiefly to fulfill a general graduation requirement or a specific requirement for their majors.

Given the challenges encountered in introductory statistics courses offered at the undergraduate level, educators have looked for ways to improve statistics education. Research on the use of collaborative learning strategies in undergraduate STEM education, in general, suggests that collaborative learning can be a highly effective means for promoting the kind of understanding, problem solving, and social competency that is expected of 21st century professionals (Roseth, Garfield, & Ben-zvi, 2008; Mastascusa, Snyder, & Hoyt 2011). The Guidelines for Assessment and Instruction in Statistics Education (GAISE, 2005) state that statistics educators, specifically, should foster active learning in the classroom through techniques such as group problem solving, hands-on activities and discussion (GAISE, 2005). The aim of this study is to explore the impact of regular, peer-facilitated collaborative problem solving in small-group settings on the learning and attitudes of students taking an introductory course in applied statistics at a medium-sized, liberal arts university in the Midwest.

I. Review of Literature.

A. Collaborative Learning.

One of the major theoretical foundations of collaborative learning is that of social constructivism and the works of psychologist Vygotsky (Cracolice & Tautmann, 2001; Barkely, Cross, & Majro, 2005). Social constructivism is a student-centered view of learning. In constructivist theories, the student (based on past knowledge, experiences, and context of the learning experience) will assimilate new knowledge and construct their own understanding of the information. Social constructivists argue that an important part of constructing this understanding comes from interactions/dialogue with other learners (Svinicki, 2004). While several variations of collaborative learning have been described, at a broad level collaborative learning can be defined as working together to achieve common learning goals.

Elizabeth Barkley and her colleagues (2005) identified three essential elements of collaborative learning: co-labor, intentional design, and meaningful learning. During collaborative learning, all students in the group must be actively engaged (co-labor) in a structured activity designed to complement the course learning objectives (intentional design). This activity results in an increase in a student's knowledge and understanding of course material (meaningful learning).

In the literature, the terms collaborative and cooperative learning are both used to describe situations where students work together to learn, but important distinctions between the two can be made. Bruffee (1995) defines the goal of collaborative learning as shifting the classroom authority from faculty to students and asserts that the goal of cooperative learning is to hold students accountable for learning together. While subtle differences exist, many researchers use the term cooperative and collaborative learning interchangeably. In the review of the literature presented here, the term collaborative learning is used to refer to studies on both collaborative and cooperative learning.

It is well established that in many disciplines and at different education levels collaborative learning has a positive impact on student achievement and student perceptions of both themselves (improved self esteem) and their relationships with others (social support) (Johnson, Maruyama, Johnson, Nelson, & Skon, 1981; Johnson & Johnson, 1994; Johnson,

Johnson, & Smith, 1998). Specifically in STEM disciplines at the post-secondary level, two large, meta-analyses of research show that collaborative learning has substantial positive impacts on student achievement (Springer, Stanne, & Donovan, 1999; Ruiz-Primo, Briggs, Iverson, Talbot, & Shepard, 2011), attitudes towards material, and self-esteem (Springer et al., 1999).

First, in 1999 Springer and colleagues retrieved 383 collaborative learning reports from undergraduate STEM disciplines published between the years of 1980 and 1999. Of these reports, 39 met the strict inclusion criteria for the analysis. Their results show that collaborative learning positively impacts student achievement (effect size (*d)* = 0.51) and improves student attitudes towards the STEM disciplines (*d* = 0.55). When they examined student attitudes further they found collaborative learning specifically increased self-esteem (*d* = 0.61) and student attitudes towards learning the material (*d* = 0.56). There did not appear to be a relationship between collaborative learning and student motivation to achieve (*d* = 0.18) (Springer et al., 1999). A more recent meta-analysis focused on research in undergraduate STEM disciplines from 1990-2007 supported the results of the Springer study (Ruiz-Primo et al., 2011). It examined the impact of multiple innovative science teaching techniques (including collaborative learning) on student achievement. The authors found that collaborative learning improved student achievement when used alone or in conjunction with other innovated technologies such as the introduction of conceptually oriented tasks (mean *d* ranged from 0.46-0.68). (Ruiz-Primo et al., 2011).

Peer-Led Team Learning. Peer-Led Team Learning (PLTL) is a specific form of collaborative learning initially created by educators to improve student interest and success in chemistry courses (Woodward, Weiner, & Gosser, 1993; Gosser et al., 1996; Gosser & Roth 1998). From its inception in the 1990s, PLTL has gained widespread use in STEM (Science, Technology, Engineering and Mathematics) disciplines. To date, at least twenty-six different colleges and universities have been involved in published PLTL research (Gosser, 2011).

PLTL involves a small group of six to eight students that meet weekly with a peer-leader for approximately two hours to tackle difficult course concepts and problems with their peers. Multiple variations of the PLTL model have been reported. They differ in terms of when the PLTL workshops are held (in or out of class time) and whether participation is voluntary or mandatory for all students (Gafney, 2001). Despite these differences six critical components of PLTL have been identified (Gosser, 2001):

- The PLTL group is an essential component of the course.
- The course professors are involved in the PLTL process both in material design and interactions with the peer leaders.
- Peer-leaders have previously been successful in the course and are selected and trained to be skilled in student learning and group facilitation.
- PLTL group materials are designed for small group work, appropriately challenging and directly related to course material.
- The PLTL group is the correct size and the group is held in a space that encourages small group learning.
- Institutional and departmental support for the approach are available.

The first published report of positive student achievement gains in PLTL came from the Workshop Chemistry Project where an average 15% increase in students receiving A, B, or C grades in PLTL-based courses compared to non-PLTL based courses was reported. (Gafney, 2001). Since this initial account, many research groups have reported significant PLTL derived gains in student achievement in chemistry courses as measured by course grades (Hockings,

DeAngelis, & Frey, 2008; Baez-Galib, Colon-Cruz, & Resto, 2005; Lyon & Lagowski, 2008; Quitadamo, Brahler, & Crouch, 2009; Wamser, 2006) and exam scores (Lyon & Lagowski, 2008; Popejoy & Asala, 2013). Additionally, a significant increase in critical thinking skills was reported for students in a PLTL-based general chemistry course compared to their non-PLTL counterparts (Quitadamo et al., 2009).

Consistent with PLTL research in chemistry, significant gains in student learning have also been associated with PLTL in other STEM disciplines including biology (Tenney & Houck, 2003; Petroy-Kelly, 2007; Preszler, 2009), mathematics (Quitadamo et al., 2009; Liou-Mark, Dreyfuss, & Younge, 2010), computer science (Horwitz & Rodger, 2009), and engineering (Loui & Robbins, 2008). In a 2012 review of PLTL research, Gosser reported a 16% average increase in students receiving A, B and C grades in STEM disciplines when exposed to PLTL-based models (Gosser, 2011). Peteroy-Kelly (2007) identified a significant increase in students' conceptual reasoning skills between the beginning and end of the semester in an introductory biology course that incorporated PLTL. While these studies were short-term, focusing on student achievement/learning in the same course where students are exposed to PLTL, a recent study from Penn State-Schuylikk suggests that PLTL benefits might persist beyond the original PLTL-based course. Students who enrolled in a PLTL-modeled one-credit problem solving course accompanying their first semester general chemistry course received higher grades in later chemistry classes than their non-PLTL peers (Eberlein, 2012).

B. Collaborative Learning and PLTL in Applied Statistics.

Like other STEM disciplines, statistics educators have recognized the benefits of collaborative learning for students in their classrooms. Studies have demonstrated greater achievement (as measured by exam scores or total course points earned) in students exposed to collaborative learning techniques compared to their counterparts in introductory statistics courses at both the undergraduate (Keeler & Steinhorst, 1994; Magel, 1998; Ghani, 2009; Perkins & Saris 2001; Giraud, 1997; Potthast 1999; Borreson, 1990) and graduate level (Enders & Diener-West, 2006). In addition to achievement gains, formal and informal assessments suggest that collaborative learning in statistics is viewed positively by most students participating in the experience (Borresen, 1990; Keeler & Steinhorst, 1994).

Published studies on collaborative learning in statistics classrooms focus mostly on in-class collaborative experiences such as small group problem solving (Borresen, 1990; Giraud, 1997; Ghani, 2009) and the jigsaw technique (Perkins & Saris, 2001). These methods differ widely from the PLTL approach described above. In a study most similar to the PLTL method, graduate students in an introductory biostatistics course participated in a one-hour, bi-weekly, project-based learning sessions led by a teaching assistant. In this study, students who participated in learning sessions performed better on course exams than their counterparts (Enders & Diener-West, 2006).

While collaborative learning has been shown to have a positive impact on student attitudes in STEM disciplines (Springer et al., 1999) and is viewed favorably by many statistics students exposed to this form of learning (Borresen, 1990; Perkins & Saris, 2001; Keeler & Steinhorst, 1994), specific studies evaluating the impact of collaborative learning on changes in student attitudes towards statistics are lacking. No published research could be identified that focused primarily on changes in student attitudes with regard to collaborative learning. However, a constructivist learning environment (which includes small group work and other active learning

strategies) has been found to result in a significant improvement in student attitudes towards statistics and a significant decrease in perceived difficulty of the statistics course (Tsao, 2006) as measured by the Survey of Attitudes Toward Statistics (Schau, Stevens, Dauphinee, & Del Vecchio, 1995). While this result may suggest that constructivist learning environments which promote collaborative learning may increase positive attitudes in students taking introductory statistics courses, these results are not well supported. Carnell (2008) failed to find a significant difference between student attitudes in a small observational study when comparing a constructivistically-framed introductory statistics class that included a student-designed project and a class where the student-designed project was omitted.

Researchers examining the impacts of collaborative, problem-based learning in undergraduate STEM education have found results that are largely consistent with the theory underpinning the use of such learning strategies. However, the impact of collaborative, problem-based learning on the undergraduate student and their experience have been found to depend heavily on programmatic structure and implementation. Moreover, very little research has been conducted with students participating in undergraduate applied statistics courses. The emphasis of this research is to investigate the impact of a structured, collaborative learning program incorporating peer-led team-learning on both content mastery (as measured by exam scores) and the dispositions (as measured by the Survey of Attitudes Toward Statistics) of undergraduate students taking an introductory course in applied statistics.

II. Program Description.

A. Background.

In the fall of 2010, the STEM Learning Community Program (LCP) was introduced to undergraduate students enrolled in science and mathematics courses at a private, medium sized, liberal arts university in the Midwest. The objectives of the STEM LCP were to improve the depth and breadth of student learning in introductory science and mathematics courses, facilitate positive attitudes toward course content and discipline, strengthen student capacity for collaborative engagement, and increase the retention of students in STEM disciplines.

STEM Learning Communities (LCs) have been offered every semester since the fall of 2010 with financial support from the Dean of Arts and Sciences, as well as individual departments. To date, LCs have been offered to students taking 100- and 200-level courses in biology, chemistry, mathematics, and applied statistics. Each STEM LC is course-specific, but not necessarily instructor-specific, and provides an opportunity for students to work together (with a peer facilitator), weekly, to enhance their conceptual understanding of course material and develop collaborative problem solving skills. These stable groups of twelve or fewer participants are completely voluntary; however, enrolled students are expected to maintain active and consistent attendance. Between 180 and 240 students participate in the STEM LCs each semester at this university.

B. Program Structure.

The support structure of the STEM LC program consists of a LC coordinator, four faculty liaisons, and 16-20 peer facilitators that are responsible for up to two LCs each semester.

The STEM LC coordinator oversees day-to-day operations of the STEM LC program (i.e., payroll, reserving meeting spaces, etc.). Additionally, this individual develops and delivers ongoing peer facilitator training, oversees LC enrollment and participation, coordinates regular meetings with faculty liaisons to share LC progress, strategies, activities and issues that arise, manages peer facilitator and LC evaluations, and communicates with the broader faculty, department chairs and academic administrators about the STEM LC program. The LC coordinator is compensated through a course release each semester for the time and effort that is expended on this program.

In this particular program, there is a designated faculty liaison for each discipline in which STEM LCs are offered. The liaisons meet with their peer facilitators on a weekly basis to review group progress and to seek input and feedback regarding the development of collaborative problem solving LC activities. The bulk of each faculty liaisons' time is spent working with their peer facilitators to develop appropriate collaborative problem solving activities for use in LC sessions. Liaisons also meet with the STEM LC coordinator regularly and communicate about the STEM LC program to students, faculty and administration. Each faculty liaison has received formal training in collaborative learning theory and practice and receives a small stipend for their involvement with this program.

Peer facilitators are typically upper-level students who have successfully completed the course and have developed a strong foundation in course content. Moreover, these students have exhibited strong peer-to-peer communication skills in the classroom and appear able to foster a positive and productive learning atmosphere. These individuals are hand-picked and invited by the faculty liaisons to participate as facilitators in the STEM LC program. Each semester, these students participate in approximately six hours of training on group facilitation techniques and collaborative learning strategies provided by the LC coordinator. Peer facilitators work with faculty liaisons to develop discipline-specific collaborative, problem-solving activities. They guide the problem solving process in the context of the LC, but do not tutor or teach. These individuals spend roughly four hours each week on STEM LC commitments: they meet for up to an hour with their faculty liaison, devote approximately one hour to individual preparation for their LC, and, finally, spend one and one-half to two hours facilitating their LC. LC facilitators are paid an hourly wage, similar to that of research assistant, for their time.

III. Methods.

A. Participants.

The participants in this study were 46 students enrolled in one of three sections of an undergraduate course in applied statistics. This 200-level course provided students with an opportunity to learn about probability and random variables in an applied context as well as the application, analysis, interpretation and presentation of descriptive and inferential statistics. The vast majority of students taking this applied statistics course enrolled to fulfill a requirement for their majors or for graduation. Half (n = 23) of these students participated in one of three optional collaborative learning community experiences offered throughout the semester. The other half (n = 23) were selected to be members of the control or comparison group based primarily upon first exam scores. When several potential matches were identified, attempts were made to match students by course section and sex.

The participating students closely reflected the demographics of the students taking this course in applied statistics at this institution (Table 1). Most participating students identified as Caucasian between the ages of 18 and 25. Twenty-nine (63.04%) participants were male and 17 (36.96%) were female. Prior to taking this course in applied statistics, nearly all indicated having completed at least one college-level mathematics course (n = 43, 93.48%) and four years of high school math (n = 39, 84.78%). Nearly half (n = 21, 45.65%) of the students participating in the study reported a major in business; the other half of students reported majors in arts/humanities (n = 2, 4.35%), education (n = 3, 6.52%), pre-medicine (n = 3, 6.52%), biology (n = 4, 8.70%), psychology (n = 6, 13.04%), and other (n = 7, 15.22%). The vast majority of students participating in this study, both learning community members and non-members, expected to earn grades of A or A- in this course (n = 37, 80.43%).

Table 1. Demographics of study participants.

Characteristic	Learning Community Participants (n = 23)	Matched Pairs, Control Group (n = 23)
Race		
Caucasian	78.26%	100.0%
Minority	21.74%	0.0%
Sex		
Male	65.22%	60.87%
Female	34.78%	39.13%
Percent who completed AP math or stats in high school	39.13%	43.48%
Percent who completed at least one college math course	86.96%	100.0%
Percent who *expected* to earn final grade of A / B / C / D or F as indicated on SATS – 36 pre-assessment	82.61 / 13.04 / 4.35 / 0.0	78.26 / 21.74 / 0.0 / 0.0
Average age	19.78 (s = 1.59)	20.22 (s = 1.59)
Average reported GPA	3.37 (s = .31)	3.19 (s = .39)

B. Procedure.

All of the students participating in this study were members of one of three sections of an undergraduate course in applied statistics offered in the spring of 2011. The course sections ran for 65 minutes and were taught back-to-back on a Monday, Wednesday, Friday schedule by the same professor. All classes were held prior to noon. Each class used the same syllabus and textbook, worked from the same set of course notes, covered the same lecture material, and were given the same assignments and exams.

Learning community participation was an optional and completely voluntary component of the course and was discussed with students, in tandem with the syllabus, throughout the first week of the course. To allow for the assessment of students' attitudinal change throughout the semester, the SATS-36 pre-assessment was administered to all consenting students during a regularly scheduled class period in the first week of the course. The assessment was administered using procedures similar to those typically used to administer course evaluations: the instrument

and its purpose were briefly introduced to students, students were read the instructions, and the survey administrator left the room while the instrument was being completed by students.

Learning community registration, handled through an online course management system, took place during week two of the semester. Thirty-six learning community spots were available and thirty-two were taken during registration. Learning community sessions began at the start of week three of the spring 2011 semester and continued for 12 weeks through the week of final exams.

Each applied statistics learning community session was structured to reflect the tenants of collaborative learning, providing students with guided opportunities to solve problems together and investigate and apply course material within a friendly, informal context. Applied statistics LC sessions typically began with an icebreaker chosen by the learning community facilitator to meet the specific needs and interests of the participating students. The icebreaker was followed by a carefully constructed collaborative learning or problem solving activity, designed to engage the students with the material and with each other. The session usually concluded with students working together with the peer facilitator to address any areas of confusion or concern from the previous week's lecture, assigned readings, or assigned problems.

The SATS-36 post assessment was administered during class in the last week of the Spring 2011 semester, prior to the week of final exams, to all consenting students. Again, procedures similar to those used to administer course evaluations to students were used.

C. Academic Outcome Measures.

Learning progress and outcomes were assessed using the four exams that are typically administered as a part of this course. These exams consisted primarily of short answer and "multiple justification" questions. Short answer questions allowed the student to structure their own response to the question provided. Multiple justification questions are structured like traditional multiple choice questions; however, in addition to finding the correct answer amongst the four response options provided, students must provide a reasonable justification, or defense, for their chosen response option. A justification may consist of an explanation of a concept, a graph or other diagram, or the mathematical calculations used to arrive at the answer.

Exams covered approximately one-quarter of the course content and, as such, were administered approximately every four weeks throughout the 15-week semester during regularly scheduled class periods. The first exam was given in week four, just following the first learning community session in week three. (Because the introductory learning community session was offered only one week prior to the first exam, it is unlikely to have impacted exam outcomes in any significant or meaningful way.) All exams were graded in a blind fashion by the instructor of the course; additionally, exams from all three course sections were cross-checked for consistency in grading practices.

D. Attitudinal Outcome Measures.

Attitudinal aspects of the student perspective and experience were measured using the Survey of Attitudes Toward Statistics, commonly referred to as the SATS (Schau et. al., 1995). This inventory contains 36 statements, worded either positively or negatively, to which students respond on a 7-point likert scale (1 = Strongly Disagree, 4 = Neutral, 7 = Strongly Agree). The 36 statements can be grouped according to six highly-reliable attitudinal subscales: affect,

cognitive competence, value, difficulty, interest and effort. Cronbach's alpha values, which were calculated for each SATS – 36 subscale at pre- and post-assessment, indicate that the attitudinal constructs were measured with a high degree of internal consistency in this study (Table 2).

Table 2. Survey of Attitudes Towards Statistics – 36.

Subscale	Definition; interpretation	No. of Items	Cronbach's Alpha Pre - Post
Affect	Students' feelings concerning statistics; higher scores indicate more positive feelings about statistics.	6	.834 - .831
Cognitive Competence	Students' attitudes about their intellectual knowledge and skills when applied to statistics; higher scores indicate greater feelings of cognitive competence in statistics.	6	.817 - .867
Value	Students' attitudes about the usefulness, relevance and worth of statistics in personal and professional life; higher scores indicate greater perceptions of the utility, relevance and worth of statistics.	9	.855 - .887
Difficulty	Students' attitudes about the difficulty of statistics as a subject; higher scores indicate lowered perceptions of the difficulty of statistics.	7	.752 - .711
Interest	Students' level of individual interest in statistics; higher scores indicate higher levels of interest in statistics.	4	.829 - .902
Effort	Amount of work the student expends to learn statistics; higher scores indicate greater levels of work intended to learn statistics.	4	.696 - .698

III. Data Analysis and Results.

Learning community participants were matched with non-learning community students based on the score earned on the first exam in this applied statistics course. Additionally, when multiple potential matches were located, efforts were made to match learning community participants (LCP) with non-learning community participants (NonP) by course section and sex. In total, there were 23 LCP and NonP pairs that consented to participate in this study and who completed both the SATS – 36 pre- and post-assessments.

A. Academic Outcomes.

To assess the impact of the learning community, if any, on student learning, paired samples t-tests were used to compare the scores earned on course exams by LCPs and NonPs. The assumptions for paired samples *t*-tests were assessed and satisfactorily met for all comparisons. Due to the high number of significance tests being conducted, a Bonferroni adjustment was applied to alpha for each set of comparisons being made, thereby reducing the chance of a Type I error. Additionally, descriptive statistical analysis was used to examine the trends observed via inferential analysis in greater detail.

For the first set of comparisons, the LCP were compared to NonP on exams I, II, III and IV. Alpha was adjusted relative to the number of pairwise comparisons being made and was set at .0125 for this set of comparisons. Results indicate that, as expected, exam I scores were nearly identical for LCP and NonP students ($t(22) = 1.00$, $p = 0.328$). On exam II, average scores for LCP and NonP were not significantly different ($t(22) = -0.727$, $p = 0.475$). However, LCPs earned significantly higher scores on exam III than their NonP counterparts ($t(22) = -3.12$, $p = 0.005$). On the final exam, which is non-cumulative, LCPs again scored significantly higher than the matched pairs control group ($t(22) = -3.30$, $p = 0.003$) (Table 3).

Table 3. LCP and NonPs exam score results.

	LCP Mean (sd)	NonP Mean (sd)
Exam I	38.33 (4.48)	38.43 (4.27)
Exam II	36.15 (5.11)	35.20 (4.65)
Exam III**	36.22 (5.66)	31.67 (6.21)
Exam IV**	52.32 (5.45)	48.15 (6.81)

** Statistically significant result at $p \leq .01$

In order to examine these patterns more closely, pairs of participants in this study were broken down into groups based upon the letter grade earned on exam I (Figure 1). On exam I, there were two pairs that earned a grade of D/F, five pairs that earned a grade of C, ten pairs that earned a grade of B, and six pairs that earned a grade of A. It is for those students who earned a grade of C or D/F on exam I that the learning community appeared to have the greatest impact on learning outcomes. The average exam score for LCPs who earned a C or D/F on exam I did not fall below 72.6% for the remainder of the semester; on the other hand, the average score for NonPs who earned a grade of C or D/F on exam I fell as low as 53.8%.

B. Attitudinal Outcomes.

Student attitudes toward statistics were assessed using the SATS – 36 pre-assessment in week 1 of the semester and in week 15 using the SATS – 36 post-assessment. The pre- and post-assessment responses of the LCPs and NonPs were compared using paired samples *t*-tests. Additionally, changes in attitudes over the course of the semester were examined by comparing pre- to post-assessment responses for each group (LCPs and NonPs) using paired samples t-tests. Due to the number of significance tests being conducted, a Bonferroni adjustment was applied to the overall alpha for each set of comparisons. Finally, one global attitudinal item on the SATS - 36, "I will like/like statistics" was examined from pre-assessment to post-assessment for LCPs and NonPs separately using Pearson's correlation (*r*).

SATS – 36, Pre-Assessment Comparison. LCPs and NonPs were compared on the six attitudinal subscales of the SATS – 36: affect, cognitive competence, value, difficulty, interest, and effort. It was anticipated that the groups would be relatively similar in respect to their course-related dispositions at the start of the course and this was largely demonstrated through the data analysis, using an adjusted alpha = .0083. There were no statistically significant differences between LCPs and NonPs on the aforementioned subscales at the beginning of the course. However, the NonPs average rating on difficulty indicated that they expected the course to be slightly but significantly less difficult than LCPs ($t(22) = 3.482$, $p = .002$) (Table 4).

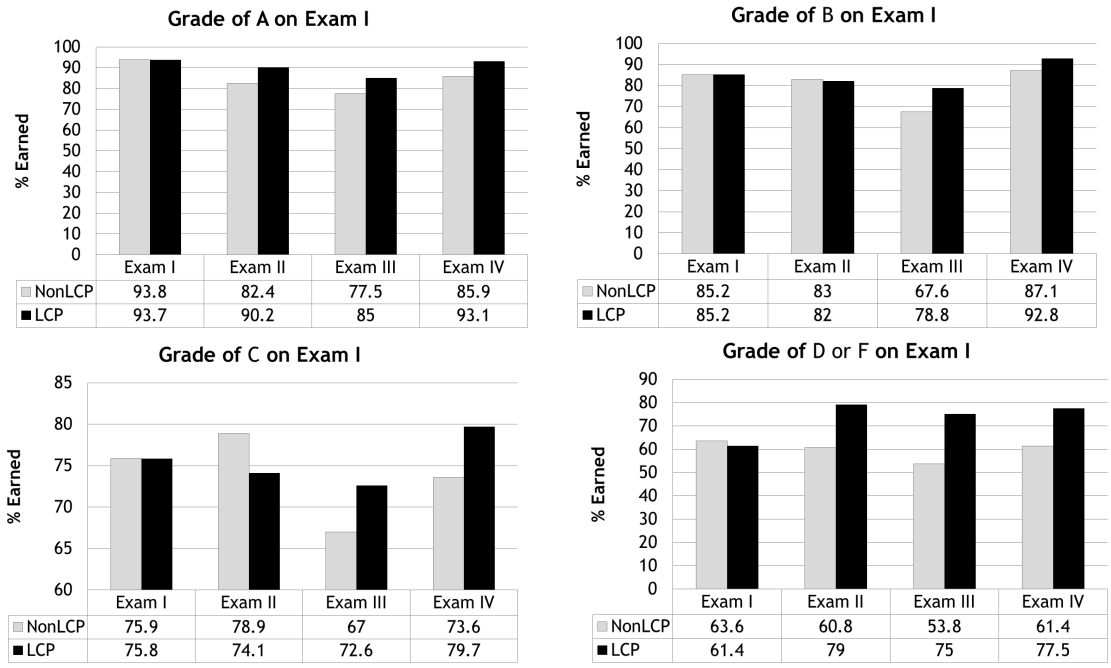


Figure I. Descriptive learning outcomes for LCP and NonPs by grade earned (A, B, C or D/F) on exam I.

Table 4. LCP and NonP means and standard deviations, SATS – 36 pre-assessment.

	LCP Mean (*s*)	NonP Mean (*s*)
Affect	4.30 (0.92)	4.92 (1.05)
Cognitive Competence	5.07 (0.98)	5.49 (0.87)
Value	5.19 (0.80)	5.32 (0.75)
Difficulty**	3.32 (0.40)	3.74 (0.53)
Interest	4.75 (0.88)	4.75 (0.87)
Effort	6.29 (0.61)	6.01 (0.81)

** Statistically significant result at $p \leq .01$

SATS – 36, Post-Assessment Comparison. In order to assess whether there were any statistically significant differences between the LCPs and NonPs on their attitudes toward statistics at the end of the course, these groups were compared on the subscales of the SATS – 36 post-assessment. An adjusted alpha of .0083 was again used as the threshold for statistical significance. Paired-samples t-tests indicate that there were no statistically significant differences between the LCPs and NonPs at the conclusion of the semester in terms of affect, cognitive competence, value, difficulty, interest or effort (Table 5).

Table 5. LCP and NonP means and standard deviations, SATS – 36 post-assessment.

	LCP Mean (*s*)	NonP Mean (*s*)
Affect	4.14 (1.34)	3.73 (1.19)
Cognitive Competence	4.94 (1.20)	4.79 (1.30)
Value	4.73 (0.86)	4.84 (0.97)
Difficulty	3.97 (0.82)	3.80 (0.68)
Interest	3.83 (1.37)	3.93 (1.27)
Effort	5.37 (1.30)	5.29 (1.19)

SATS – 36, Pre- to Post-Assessment, NonPs. Changes in attitude over the course of the semester for NonPs were assessed using paired samples t-tests. An adjusted alpha of .0083 was used as the threshold for detecting statistically significant differences. Average scores on several of the SATS – 36 subscales for NonPs changed from pre-assessment to post-assessment. NonPs were found to have a less positive affect toward statistics at the end of the semester ($t(22) = 6.881$, $p < .001$), to feel less cognitively competent in regard to statistical material and applications ($t(22) = 3.418$, $p = .002$), to express less interested in the subject of statistics ($t(22) = 3.560$, $p = .002$), and to have given less effort than originally intended in the context of the course ($t(22) = 3.418$, $p = .002$). The overall value placed on statistics and perception of the level of difficulty was relatively unchanged for NonPs (Table 6).

Table 6. NonP means and standard deviations, SATS – 36 pre- and post-assessment.

	NonP Pre-Assessment Mean (*s*)	NonP Post-Assessment Mean (s)
Affect***	4.92 (1.05)	3.73 (1.19)
Cognitive Competence**	5.49 (0.87)	4.79 (1.30)
Value	5.32 (0.75)	4.84 (0.97)
Difficulty	3.74 (0.53)	3.80 (0.68)
Interest**	4.75 (0.87)	3.93 (1.27)
Effort**	6.01 (0.81)	5.29 (1.19)

** Statistically significant result, $p \leq .01$
*** Statistically significant result at $p \leq .001$

NonPs attitudes towards statistics at the beginning of the semester, as measured by a global indicator of attitudes toward statistics, were significantly and positively correlated with attitudes toward statistics at the end of the semester ($r = .642$, $p = .001$; $r^2 = .412$). NonP students who anticipated liking statistics at the start of the semester tended to express more favorable attitudes toward statistics at the end. Likewise, NonP students who did not anticipate liking the subject matter at the start of the semester typically held more negative views about statistics at the end of the semester.

SATS-36 Pre- to Post-Assessment, LCPs. The LCPs were also evaluated for changes in attitude over time using the SATs – 36 pre- and post-assessments. The affect LCPs demonstrated toward statistics remained stable over the course of the semester ($p = .543$). Similarly, the LCPs cognitive competence did not change significantly ($p = .517$), nor did LCPs' perceptions of the value of statistics ($p = .024$). Perhaps surprisingly, LCPs found statistics material to be

significantly less difficult at the conclusion of the semester ($t(22) = -3.958$, $p = .001$) than anticipated at the beginning of the semester. Interest ($t(22) = 3.118$, $p = .005$), and actual effort ($t(22) = 3.981$, $p = .001$), were reported at significantly lower levels, however, at the end of the semester than at the beginning (Table 7).

Table 7. LCP means and standard deviations, SATS – 36 pre- and post-assessment.

	LCP Pre-Assessment Mean (*s*)	LCP Post-Assessment Mean (*s*)
Affect	4.30 (0.92)	4.15 (1.34)
Cognitive Competence	5.07 (0.98)	4.94 (1.20)
Value	5.19 (0.80)	4.73 (0.86)
Difficulty***	3.32 (0.40)	3.97 (0.82)
Interest**	4.75 (0.88)	3.83 (1.37)
Effort***	6.29 (0.61)	5.34 (1.30)

** Statistically significant result at $p \leq .01$
*** Statistically significant result at $p \leq .001$

Unlike the moderately strong correlation found between pre- and post-attitudes towards statistics for NonPs, virtually no correlation was identified for LCPs ($r = .040$, $p = .856$; $r^2 = .002$). LCPs who anticipated liking statistics at the start of the semester did not systematically express more favorable or more negative attitudes toward statistics at the conclusion of the semester. Similarly, LCPs who did not anticipate liking the subject matter at the start of the semester did not predictably maintain or change those negative views at the end of the semester.

IV. Discussion.

A. Learning Outcomes.

Inferential Analysis. Inferential analyses indicate that students who participated in the applied statistics LCs performed no differently than their matched pairs on exam II, but significantly better on exams III and IV. Exam II was administered to students just one month after the initiation of the LC. While LC participants certainly had additional structured opportunities to work with the material beyond what was available to NonPs in the month leading up to exam II, it did not result in higher exam scores on exam II for LCPs. This result seems to indicate that a few additional opportunities for collaborative investigation and problem solving are not enough to enhance learning outcomes for undergraduate students taking introductory courses in applied statistics; instead, a longer-term experience with peer-facilitated collaborative learning may be necessary to enhance learning objectives.

Exams III and IV were administered two and three months following initiation of the LC experience and LCPs did in fact earn significantly higher grades than NonPs. Such a result indicates that a LC program may improve the depth and breadth of student learning in the context of an undergraduate course in applied statistics. The fact that exam results of LC participants showed significantly better performance than their matched pairs, but only after an extended period of time working together in stable groups, suggests that the social, collaborative,

and ultimately the community-related aspects of the LC experience are important contributors to learning outcomes.

The gradual improvement of student performance throughout the semester has been noted in other PLTL/PLTL-like research (Hockings et al., 2008, Lewis & Lewis, 2005; Peteroy-Kelly, 2007). Like the students in the applied statistics class, students in these studies having been involved in a PLTL experience showed improvement on achievement measures from the beginning to end of the course. Lewis & Lewis (2005) attributed this result to the students' adjustment to the learning method, but these results are also consistent with research on the nature of team development and learning. Varma-Nelson & Coppola (2005) describe the PLTL model of team learning as groups of undergraduates being transformed into a high performing team by another trained undergraduate student. In this model the peer leader plays a significant role in impacting how quickly this transformation takes place (Varma-Nelson & Coppola, 2005). Based on this idea of team learning, it may take time for a team to develop and mature before the positive impact of PLTL on student learning may be seen.

Descriptive Analysis. An examination of the learning outcomes for LCPs and NonPs who earned grades of A, B, C and D/F on exam I lends further insight into the impact of the LC experience. While inferential analysis on the exam results broken down by grade was not conducted in this study due to the small group sizes, descriptive analysis was used to understand the effect of the LC experience on learning outcomes for students earning C or D/F on exam I. Exam I might be considered largely a review exam in that it covers material and concepts with which students should already be familiar prior to enrolling in a college-level course in applied statistics (e.g., measures of center, measures of spread, and graphical and tabular representations of data). Students who perform poorly on such an exam might be identified by the course instructor as at risk for poor performance throughout, and possible failure of, the course. Not incidentally, it is precisely this group of students who need additional concern and investment from the post-secondary instructor of applied statistics.

The LCPs who earned a grade of C or D/F on exam I passed with an average of 72.6% or better on all of their subsequent exams in this course; exam averages for NonPs, more than half the time, fell below 70.0%. Moreover, the LCPs who earned grades of C, or D/F on exam I outperformed their matched pairs counterparts by particularly wide margins on exams III and IV which, in this particular course, include what is generally perceived by students to be most difficult aspects of the course (e.g., the central limit theorem, sampling distributions, and applications of hypothesis testing in a wide array of contexts). For the students participating in this study, the LC experience was of particular benefit to those who might have been identified as at-risk based on exam I performance. This result suggests that further empirical research on the benefits of collaborative learning, and PLTL in particular, for lower-achieving or at-risk students in post-secondary applied statistics is warranted.

B. Attitudinal Outcomes.

Student attitudes are an important consideration in introductory statistics courses. Attitudes can impact student learning and a student's ability/motivation to apply statistics outside of the course context (Gal et al., 1997) and should be viewed as an important course outcome (Schau & Emmioglu, 2012). Emmioglu & Capa-Aydin (2012) conducted a meta-analysis to examine the relationship between attitudes and statistics achievement in undergraduate classrooms around the world. Focusing on studies that used the SATS-28, they identified 17 studies (published between

1998 and 2011) that met their inclusion criteria. Collectively they report that affect (r = 0.30), cognitive competence (r = 0.30), value (r = 0.21) and difficulty (r = 0.20) are all moderately and significantly correlated with course achievement.

Pre-Assessment. At the start of the semester, the LCPs and NonPs were very similar in terms of their pre-conceived attitudes toward statistics as measured by the six subscales of the SATS – 36. The averages for both groups were similarly high on the subscales of cognitive competence, value and effort during the first week of the course. Such results indicate that LCPs and NonPs alike felt that applied statistics is a relevant and useful discipline worthy of investigation, they felt well prepared to master the material, and they held high expectations for their study and work habits. In fact, a preponderance of LCPs and NonPs expected to earn grades of A or A- (82.61% and 78.26%, respectively) as a result of their efforts in the course. However, the exams scores and final grades that followed were not consistent with such a high degree of initial perceived cognitive competence, anticipated effort and follow-through, or their expected grade outcome.

The initial attitudes towards statistics held by both the LCPs and NonPs in the applied statistics course are consistent with results of past studies (Evans, 2007; Schau & Emmioglu, 2012). One particular study conducted by Schau and Emmioglu (2012) examined the attitudes of 2200 students enrolled in post-secondary introductory statistics courses at multiple institutions in the United States using the SATS - 36. On average, at the start of the semester the students in this large study felt the discipline and study of statistics was relevant ($\bar{x}_{Value}$ = 5.05, n = 2,186), they were confident in their abilities to master the subject ($\bar{x}_{Cognitive\ Competence}$ = 4.94, n = 2,192) and expected to put in a large amount of effort ($\bar{x}_{Effort}$ = 6.32, n = 2,246). Consistent with the results of this study, students enter statistics courses confident they will perform well; however, confidence is not generally indicative of performance in the introductory statistics course. This inconsistency may suggest of a widespread misconception about the contents and expectations of an introductory course in applied statistics at the post-secondary level.

Attitudes between LCPs and NonPs were similarly neutral on the subscales of affect and interest. Students expressed neither overtly positive nor negative attitudes toward personal interest in the discipline of applied statistics. Given that none of these students were mathematics or statistics majors and all were taking this class solely as a requirement for their majors or for graduation, such neutral attitudes are not surprising.

Only in terms of the expected difficulty of the course were the LCPs and NonPs found to differ at the time of pre-assessment. As indicated by their means, both groups perceived the field of applied statistics to be relatively difficult (as opposed to relatively easy). On average, however, the LCPs expected the course to be slightly, but significantly, more difficult than the NonPs. This finding might be expected as those who perceive the course to be more difficult at its outset would be more likely to enroll in a voluntary program such as the one offered.

Post-Assessment. At the conclusion of the semester, responses to the SATS - 36 post-assessment indicate that there were no statistically significant differences between LCPs and NonPs on the six attitudinal subscales of affect, cognitive competence, value, difficulty, interest and effort. The only subscale that maintained its positive average was effort, whereby both LCPs and NonPs agreed that they put forth a high degree of effort throughout the course. At post assessment, relatively neutral attitudes were expressed by LCPs and NonPs on the subscales of affect, cognitive competence, and value. Finally, LCPs and NonPs alike found the study of applied statistics to be relatively difficult and expressed a rather low degree of personal interest in the discipline at the conclusion of the semester. Perhaps not surprisingly, neither of the student

groups who participated in this study came away from the course with overwhelmingly positive attitudes toward the course content or discipline of applied statistics. It would appear that the Learning Community Program does not effectively improve the attitudes of undergraduate students taking an introductory course in applied statistics to fulfill a program or graduation requirement; this is further evidence to support the assertion that attitudes toward statistics are exceptionally difficult to change in the context of a course (Evans, 2007; Sizemore & Lewandowski, 2009; Schau & Emmioglu, 2012).

Change Over Time. However, an examination of attitudes at the conclusion of the course only tells a portion of the story. To more thoroughly investigate the impact of the LC program on participants relative to non-participants, changes in the attitudinal subscales were examined over the course of the semester. The results of inferential analysis indicate that NonPs experienced a significant decline in their attitudes. Significant changes were found for NonPs over the course of the semester on four of the six SATS-36 subscales: affect, cognitive competence, interest and effort. More specifically, NonPs reported feeling less positive toward the discipline, less competent with the material, less interested in the discipline and material, and expending less effort at the end of the semester than what was anticipated at the beginning. The fact that the attitudes toward the course and discipline by NonPs at the end of the semester were consistently more negative at the end of the semester than at the beginning is troubling in that negative attitudes only hinder a student's ability to learn and appropriately generalize the material to relevant contexts. Only the NonPs' expressions of the utility and relevance (value) of the course/discipline and difficulty of the material remained unchanged from the beginning to the end of the semester.

The pre-to post assessment results for LCPs were different from those of NonPs in several areas. While LCPs also expressed significant decreases in interest and effort at the conclusion of the semester, perceived cognitive competence remained stable from the beginning to the end of the semester for LCPs. In other words, the LCPs were able to maintain their relatively strong feelings of cognitive competence in the context of applied statistics. Such a result may have been bolstered by the fact that LCPs, in general, tended to perform better than their matched-pairs counterparts on the exams throughout the semester. Furthermore, the attitude of LCPs towards statistics (e.g., the affect subscale) was stable from pre- to post-assessment. This means that the level of positivity for statistics expressed by LCPs at the beginning of the semester was unchanged at the end of the semester. The difference in results for LCPs and NonPs on perceived cognitive competence and affect may signify the presence of a buffering effect whereby the LC experience, due to the provision of a stable, safe and supportive atmosphere to explore course content and applications, protects participants from a significant decrease in perceived statistics-related competence and affinity. This result is important in that students who do not feel overly negative towards the course or discipline have one less barrier to learning the material, achieving success in the course, and applying course material outside of the context of the course. Moreover, students who feel that they are capable of mastering the material are likely to try to accomplish that goal (Gal et al., 1997).

The final area of contrast for LCPs and NonPs was in terms of perceived difficulty of the course and discipline. While NonPs did not change their perception of the course and discipline from pre- to post-assessment, LCPs reported a lower degree of perceived difficulty at the end of the course than at the beginning. In other words, LCPs' experience in the course helped them to find the material (and its applications) to be more manageable and easily mastered than anticipated. This result can be linked directly to one of the primary objectives of the STEM

Learning Community Program: to strengthen the capacity for collaborative engagement and problem solving. The perception that course content was less difficult than anticipated is likely due to the LCPs enhanced capacity to solve statistics-related problems and apply course content to unfamiliar contexts. In respect to perceived difficulty of the subject, the impact of the LC may extend beyond a buffering effect to actually promoting the kinds of attitudes that foster achievement in post-secondary education.

The importance of buffers on student attitudes and outcomes has been reviewed by Gal and Ginsburg (1994). They suggest that statistics students may experience a series of events similar to the one that math students encounter when course material is difficult and understanding of the material is challenged. When a failure to understand course material takes place, it is usually followed by a failure to receive adequate explanations from the instructor or even peers trying to help. This leads to decreased confidence and panic over a perceived lack of control in the learning process. The student may become bored and disengage entirely. Further, when also considering the negative views of both the course and its content, the student may experience a lifetime of frustration towards the discipline. Programs that foster opportunities for collaborative learning, however, may offer a buffer from these frustrations thus decreasing the likelihood of having an overall negative experience. This could be attributed from support of others encountering similar experiences, thus increasing individuals' confidence and eventually promoting success and achievement.

In addition to the possible buffer that these learning communities may offer, it could also be suggested that they may increase motivation of the enrolled students as well. Bude et al. (2007) built a model of motivation specifically for statistics based on the motivational theories of learned helplessness and attribution. They suggest that educators need to build a learning environment that allows students to tackle feasible tasks. By doing so, students may experience the feeling of success through mastering a certain task. They may then feel more compelled to study and their motivation is heightened because the task/subject at hand now seems attainable. It is believed that these learning communities may offer students the platform on which to accomplish tasks assigned in lecture in a safe and supportive environment outside of the classroom. As a result, the students are able to build a foundation for their studies, solve problems and get feedback from others about the process. This may increase a student's accountability to the group as well as to themselves. Additionally, their self worth and motivation to study increases and leads to improved performance in the course.

Global Assessment – Attitudinal Item. As noted, LCPs and NonPs differed in terms of the change in their overall attitudes toward statistics over the course of the semester. For NonPs, their attitudes toward statistics became significantly more negative as a result of their experience in the class. For LCPs, there was no change in attitude over the course of the semester. Results of correlational analysis on a single global attitudinal item, "I like/will like statistics", yield insight into these pre- and post-assessment attitudes for LCPs and NonPs. As might be expected, attitudes toward statistics as measured by the global attitudinal item at the time of pre-assessment were strongly and positively correlated with the attitudes as measured at the time of post-assessment for NonPs. More specifically, there was a tendency for NonPs who came into the course with relatively favorable responses to the question to leave the course the same way; NonPs who entered the course with relatively negative responses to the question were likely to complete the semester holding on to their negative views. The attitudes of NonPs toward statistics, no matter how accurate or misconceived, were not impacted through regular course participation.

LCPs had very different results on the global attitudinal item, however. The attitudes toward statistics held by LCPs at the beginning of the semester, as measured by the global attitudinal item, were completely uncorrelated to their responses on this item at the semester's end. In other words, there was no systematic tendency for LCPs to leave the course with similar (or opposite) attitudes to what they entered with. For better or for worse, LCPs' attitudes toward the study and discipline of applied statistics were changed, but on an individual basis; the change that one LCP experienced was not necessarily indicative of the chance that another LCP experienced. This statistical result may provide evidence for the potential for a learning community program, like the one investigated here, to provide students with the level of engagement necessary for a highly personalized and meaningful educational experience in the context of an undergraduate course in applied statistics.

V. Conclusion.

Statistics educators face substantial hurdles to engaging their undergraduate non-majors. Whether these students lack the appropriate academic preparation, hold conceptual misconceptions about the material, or simply possess a general lack of interest in the subject, academic and attitudinal barriers can negatively impact students' experiences, attitudes, and course outcomes. As is evidenced by the accumulating research on engaging the hearts and minds of undergraduate statistics students, this is a topic of concern for many. Because it is important that students learn the content offered in an applied statistics course to further the depth and breadth of understanding in their chosen majors, statistics educators must continually strive for ways to stimulate and engage their intellects, while enhancing the accessibility of the material. As demonstrated in this study, the depth and breadth of learning by students in an introductory statistics classroom can be significantly and meaningfully enhanced through a structured peer-led collaborative learning program. While further research is necessary to establish the generalizability of this outcome, peer-led collaborative learning was particularly impactful for the students in this study who were most at risk for failing to meet course learning objectives.

Because student attitudes have been found to be correlated with course performance, statistics educators must also be concerned about the attitudes with which students enter the classroom, the expected trajectory of these attitudes, and how to ultimately minimize the impact of negative attitudes. In general, the students in this study entered the undergraduate statistics classroom with positive attitudes toward the course and discipline and very high expectations for their own competence and performance. However, when the students discovered that the discipline of applied statistics and real-world problem solving is much more complex than calculating batting averages and opinion poll results, many developed feelings of apathy, frustration and discouragement. This trajectory is illuminated by examination of the pre- to post-assessment results of NonPs. LCPs, by contrast, thought that course content actually became *less* difficult over the course of the semester and their sense of cognitive competence remained constant. Moreover, it appears as if the LCPs were able to engage with the class and course material on a very personal and individual level. Unlike the NonPs, the LCPs did not leave this class holding the same attitudes with which they entered and their attitudes did not change in any systematic or predictable way. Such a result may be evocative of a relationship between collaborative learning and change that is possibly transformative in nature; further research is warranted to explore this potential.

Although this study did not employ randomization or utilize strict experimental controls, and the study was implemented throughout just one semester at a singular institution, preliminary results are promising. Frequent, regularly scheduled encounters of guided, collaborative inquiry appear to be an effective strategy for improving learning outcomes for undergraduate students in applied statistics classrooms. Additionally, structured, peer-facilitated collaborative learning may help students cope with the demands of a class for which they have little context and help them to develop an effective buffer against the development of attitudes that can obstruct learning and negatively impact the classroom experience for students and teachers alike.

References

Baez-Galib, R. Colon-Cruz, H, & Resto, W. (2005). Chem-2-Chem: A one-to one supportive learning environment for chemistry. *Journal of Chemical Education, 82*(12), 1859-1863.

Barkley, E.F., Cross, K.P., & Majro, C.H. (2005). *Collaborative learning techniques: A handbook for college faculty*. San Francisco, CA: Jossey-Bass.

Borresen, C.R. (1990). Success in introductory statistics with small groups. *College Teaching, 38(1),* 26-28.

Bruffee, K.A. (1995). Sharing our toys: Cooperative learning versus collaborative learning. *Change, 27*(1). 12-18.

Budé, L., Van De Wiel, M. W. J., Imbos, T., Candel, M. J. J. M., Broers, N. J., and Berger, M. P. F. (2007). Students' achievements in a statistics course in relation to motivational aspects and study behaviour. *Statistics Education Research Journal*, 6(1), 5 – 21. Retrieved from http://www.stat.auckland.ac.nz/~iase/serj/SERJ6(1)

Carnell, L.J. (2008). The effect of a student-designed data collection project on attitudes toward statistics. *Journal of Statistics Education, 16*(1). Retrieved from www.amstat.org/publications/jse/v16n1/carnell.html

Castro Sotos, A. E., Vanhoof, S., Van den Noortgate, W., Onghena, P. (2009). How confident are students in their misconceptions about hypothesis tests? *Journal of Statistics Education, 17*(2). Retrieved from www.amstat.org/publications/jse/v17n2/castrosotos.html

Cracolice, M.S. & Trautmann, J.A. (2001). Vygotsky's theories of education: Theory bases for peer-led team learning. In D.K. Gosser, M.S. Cracolice, J.A. Kampmeier, V. Roth, V.S. Strozak, P. Varma-Nelson (Eds), *Peer-Led Team Learning: A Guidebook* (pp. 94-102). Lower Saddle River, NJ: Prentice Hall.

delMas, R.C. & Garfield, J. (1999). A model of classroom research in action: Developing simulation activities to improve students' statistical reasoning. *Journal of Statistics Education, 7*(3). Retrieved from http://www.amstat.org/publications/jse/secure/v7n3/delmas.cfm

Dempster, M. & McCorry, N.K. (2009). The role of previous experience and attitudes toward statistics in statistics assessment outcomes among undergraduate psychology students. *Journal of Statistics Education, 17*(2). Retrieved from www.amstat.org/publications/jse/v17n2/dempster.html

Eberlein, T. (2012). Evaluation: Beneficial effects of PLTL persist to later courses at Penn State Schukylkill. Retrieved from www.pltlis.org.

Emmioglu, E & Capa-Aydin, Y. (2012). Attitudes and achievement in statistics: A meta-analysis study. *Statistics Education Research Journal, 11*(2), 95-102. Retrieved from: http://www.stat.auckland.ac.nz/serj

Enders, F.B. & Diener-West, M. (2006). Methods of learning in statistical education: A randomized trial of public health graduate students. *Statistics Education Research Journal, 5(1)*, 5-19.

Evans, B. R. (2007). Student attitudes, conceptions, and achievement in introductory undergraduate college statistics. *The Mathematics Educator, 17*(2), 24-30.

Gafney, L. (2001). Workshop Evalution. In D.K. Gosser, M.S. Cracolice, J.A. Kampmeier, V. Roth, V.S. Strozak, P. Varma-Nelson (Eds), *Peer-Led Team Learning: A Guidebook* (pp. 75-93). Lower Saddle River, NJ: Prentice Hall.

GAISE (2005). Guidelines for assessment and instruction in statistics education (GAISE) college report. The American Statistical Association (ASA). Retrieved from http://www.amstat.org/education/gaise/.

Gal, I. & Ginsburg, L. (1994). The role of beliefs and attitudes in learning statistics: Towards an assessment framework. *Journal of Statistics Education, 2*(2). Retrieved from http://www.amstat.org/publications/jse/v2n2/gal.html

Gal, I., Ginsburg, L., & Schau, C. (1997). Monitoring attitudes and beliefs in statistics education. In I. Gal & J.B. Garfield (Eds), *The assessment challenge in statistics* education (pp. 37-51). Burke, VA: IOS Press.

Garfield, J. (1995). How students learn statistics. *International Statistical Review, 63*(1), 25-34.

Garfield, J. & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education, 19*(1), 44 – 63.

Garfield, J. & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review, 75*(3), 372-396.

Ghani, S.A. (2009). Cooperative learning versus the lecture method of instruction in an introductory statistics course. *Jurnal Sains dan Matematik*, *1*(1), 59-71.

Giraud, G. (1997). Cooperative learning and statistics instruction. *Journal of Statistics Education, 5*(3). Retrieved from http://www.amstat.org/publications/jse/v5n3/giraud.html

Gosser, D.K, Roth, V., Gafney, L., Kampmeier, J., Strozak, V., Varma-Nelson, P., Radel, S. & Weiner, M. (1996). Workshop Chemistry: Overcoming the Barriers to Student Success. *The Chemical Educator, 1*(1). doi: 10.1333/s00897960002a

Gosser, D. K., & Roth, V. (1998). The workshop chemistry project: peer-led team learning. *Journal of Chemical Education*, *75*(2), 185-187.

Gosser, D (2001). The peer-led team learning workshop model. In D.K. Gosser, M.S. Cracolice, J.A. Kampmeier, V. Roth, V.S. Strozak, P. Varma-Nelson (Eds), *Peer-Led Team Learning: A Guidebook* (pp. 1-12). Lower Saddle River, NJ: Prentice Hall.

Gosser, D. K. (2011). The PLTL boost: A critical review of research. *Progressions the PLTL Project Newsletter, 14*(1). Retrieved from http://www.pltl.org

Hirsch, L. S. & O'Donnell, A. M (2001). Representativeness in statistical reasoning: Identifying and assessing misconceptions. *Journal of Statistics Education, 9*(2). Retrieved from http://amstat.org/publications/jse/v9n2/hirsch.html.

Hockings, S.C., DeAngelis, K.J., & Frey, R.F. (2008). Peer-led team learning in general chemistry: Implementation and evaluation. *Journal of Chemical Education, 85*(7), 990-996.

Horwitz, S. & Rodger, S.H. (2009). Using peer-led team learning to increase participation and success of under-represented groups in introductory computer science. *SIGCSE-09*, 163-167.

Johnson, D.W., Maruyama, G., Johnson, R., Nelson, D., & Skon, L. (1981). Effects of cooperative, competitive, and individualistic goal structures on achievement: A Meta-Analysis. *Psychological Bulletin, 89*(1), 47-62.

Johnson, R. T. & Johnson, D.W. (1994). An overview of cooperative learning. In: J. Thousand, A. Villa & A. Nevin (Eds), *Creativity and collaborative learning: A practical guide to empowering students and teachers*. Baltimore, MD: Brooks Press.

Johnson, D.W., Johnson, R. T. & Smith, K.A. (1998). Cooperative learning returns to college: What evidence is there that it works? *Change: The Magazine of Higher Learning, 30*(4), 26-35.

Johnson, M. & Kuennen, E. (2006). Basic math skills and performance in an introductory statistics course. *Journal of Statistics Education, 14*(2). Retrieved from www.amstat.org/publications/jse/v14n2/johnson.html.

Keeler, C.M. & Steinhorst, R. K. (1994). Cooperative learning in statistics. *Teaching Statistics, 16*(3), 81-84.

Konold, C. (1995). Issues in assessing conceptual understanding in probability and statistics. *Journal of Statistics Education 3*(1). Retrieved from http://www.amstat.org/publications/jse/v3n1/konold.html.

Lewis, S.E. & Lewis, J.E. (2005). Departing from lectures: An evaluation of a peer-led guided inquiry alternative. *Journal of Chemical Education, 82*(1), 135-139.

Loui M.C. & Robbins B.A. (2008). Work-In progress-assessment of peer-led team learning in an engineering course for freshmen. Presented at the 38th ASEE/IEEE Frontiers in Education Conference. Oct 22-25th Saratoga Springs, NY.

Lyon, D.C. & Lagowski, J.J. (2008). Effectiveness of facilitating small-group learning in large lecture classes: A general chemistry case study. *Journal of Chemical Education, 85*(11), 1571-1576.

Liou-Mark, J., Dreyfuss, A. E., & Younge, L. (2010) Peer assisted learning workshops in precalculus: An approach to increasing student success. *Mathematics and Computer Education*, *44*(3), 249-259.

Magel, R.C. (1998). Using cooperative learning in a large introductory statistics class. *Journal of Statistics Education, 6*(3). Retrieved from http://www.amstat.org/publications/jse/v6n3/magel.html

Mastascusa, E. J., Snyder, W. J., & Hoyt, B. S. (2011). *Effective instruction for STEM disciplines: From learning theory to college teaching*. San Francisco, CA: Jossey-Bass.

Perkins, D.V. & Saris R. N. (2001). A "jigsaw classroom" technique for undergraduate statistics courses. *Teaching of Psychology, 28*(2), 111-113.

Peteroy-Kelly, M.A. (2007). A discussion group program enhances the conceptual reasoning skills of students enrolled in a large lecture-format introductory biology course. *Journal of Microbiology and Biology Education, 8,* 13-21.

Popejoy, K. & Asala K.S. (2013). A team approach to successful learning: Peer learning coaches in chemistry. *Journal of College Science Teaching, 42*(3), 18-23.

Potthast, M.J. (1999). Outcomes of using small-group cooperative learning experiences in introductory statistics courses. *College Student Journal, 33*(1), 34-42.

Preszler, R.W. (2009). Replacing lecture with peer-led workshops improves student learning. *CBE-Life Science Education, 8,* 182-192.

Quitadamo, I.J, Brahler C. J., & Crouch, G.J. (2009). Peer-led team learning: A prospective method for increasing critical thinking in undergraduate courses. *Science Educator, 18*(1), 29-39.

Roseth, C.J., Garfield, J.B., & Ben-Zvi, D. (2008). Collaboration in learning and teaching statistics. *Journal of Statistics Education, 16*(1). Retrieved from www.amstat.org/publications/jse/v16n1/roseth.html

Ruiz-Primo, M.A., Briggs, D., Iverson, H., Talbot, R. & Shepard, L.A. (2011). Impact of undergraduate science course innovations on learning. *Science 331,* 1269-1270.

Schau, C., Stevens, J., Dauphinee, T. L., & Del Vecchio, A. (1995). The development and validation of the Survey of Attitudes Toward Statistics. *Educational and Psychological Measurement, 55,* 868-875.

Schau, C. & Emmioglu, E. (2012). Do introductory statistics courses in the United States improve students' attitudes? *Statistics Education Research Journal, 11*(2), 86-94.

Sizemore, O.J. & Lewandowski Jr, G.W. (2009). Learning might not equal liking: Research methods course changes knowledge but not attitudes. *Teaching of Psychology, 36*, 90-95.

Springer, L., Stanne, M.E., & Donovan, S.S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta analysis. *Review of Educational Research, 69*(1), 21-51.

Svinicki, M.D. (2004). Learning and motivation in the postsecondary classroom. San Francisco, CA, Anker Publishing Company. (pp. 242-243)

Tenney, A., & Houck, B. (2003). Peer-led team learning in introductory biology and chemistry courses: A parallel approach. *The Journal of Mathematics and Science: Collaborative Explorations*, *6*, 11-20.

Tsao, Y. (2006). Teaching statistics with constructivist-based learning method to describe student attitudes toward statistics. *Journal of College Teaching and Learning, 3*(4), 59-64.

Varma-Nelson, P. & Coppola, B.P. (2005). Team Learning. In Pienta, N., Cooper M.M.,Greenbowe, T. (Eds), *Chemist's Guide to Effecive Teaching* (pp. 155-169). Saddle River, NJ: Pearson.

Wamser, C.C. (2006). Peer-led team learning in organic chemistry: Effects on student performance, success and persistence in the course. *Journal of Chemical Education, 83*(1), 1562-1566.

Woodward, A.E, Weiner, M., & Gosser, D. (1993). Problem-solving workshops in general chemistry. *Journal of Chemical Education, 70*(8), 651.

Journal of the Scholarship of Teaching and Learning, Vol. 13, No. 2, May 2013, pp. 72 – 83.

Connecting social psychology to the experience of others through a nonfiction book analysis: New wine in an old bottle

Gregory S. Preuss[1], D. Ryan Schurtz[2], Caitlin A. J. Powell[3], David J. Y. Combs[4], and Richard H. Smith[5]

Abstract: This article evaluates a writing assignment in which students read a non-fiction book that they chose from a list provided by their instructor, identified examples of social psychological phenomena, and fully explained how those examples fit social psychology concepts. This novel twist on a traditional assignment yielded surprisingly robust benefits. Across four samples from two universities and two instructors, students indicated that the assignment furthered their learning beyond other aspects of the course by helping them apply social psychology to "real life" situations that were beyond their own particular experiences. The results suggested that allowing students to choose the book that they would read promoted enjoyment of the assignment. Informal discussion with students, including those who rarely read books for pleasure, indicated that many students took pride in reading a book of their own choosing that they actually enjoyed. Almost all students recommended the continued use of the assignment for future courses. Variations on the assignment that could be utilized by instructors in other psychology courses and other academic disciplines are discussed.

Keywords: active learning, application, book review, choice, psychology

The goal of most educational innovations and active learning approaches is to get students focused, involved, and actively engaged in the learning process (Leamnson, 2000; Prince, 2004). When students are actively engaged and personally invested in the learning process, deep and meaningful learning occurs (Laird, Shoup, Kuh, & Schwarz, 2008). Examples of active learning approaches include small group discussions, experiential activities, and writing assignments. Research has repeatedly demonstrated that active learning promotes a variety of positive outcomes including improved student attitudes (Bleske-Rechek, 2002), increased motivation (Watson, Kessler, Kalla, Kam, & Ueki, 1996), greater engagement with the course material (Smith & Cardaciotto, 2011), enhancements in writing and thinking (Bonwell & Eison, 1991), improvements in retention of the course material (Cherney, 2008; Svinicki & McKeachie, 2011), and higher scores on exams (Yoder & Hechevar, 2005). By providing students with opportunities to discover the relevance and applicability of the course material to everyday life (Bloom, 1956), active learning approaches help students learn the course material at a deeper, personal level (King, 1993; Svinicki & McKeachie, 2011). In addition, involving students in learning concepts through multiple types of exposure (lecture, reflection and application, discussion) can increase

[1] Department of Psychology, Washburn University, 1700 SW College Avenue, Topeka, Kansas 66621, greg.preuss@washburn.edu.
[2] Stevenson University
[3] Georgia College and State University
[4] United States Navy
[5] University of Kentucky

the ability to remember these concepts due to the complexity and variety of retrieval cues available (Craik & Lockhart, 1972; Roediger, 2000; Roediger & Guynn, 1996).

I. The Importance of Course Application and Relevance.

Teachers in numerous academic disciplines want their students to demonstrate a firm understanding of the course material by applying it to new situations that were not faced in the textbook or lecture (Bloom, 1956). Consequently, in order to promote deep comprehension, active learning, and critical thinking, teachers create assignments that encourage students to recognize the many ways in which the course material is relevant to their lives and the world around them (Bloom, 1956; Graham, 2006; Kowalski & Lakey, 2004; Lakin & Wichman, 2005; Svinicki & McKeachie, 2011; Weimer, 2002; Young & Fulwiler, 1986). In this article, we describe a written assignment in which students read a non-fiction book that they chose from a list provided by their instructor, identified examples of social psychological phenomena, and fully explained how those examples fit social psychology concepts. Before discussing our assignment in detail, a brief overview of the common writing assignments that help students apply principles from psychology is in order.

A. Application-oriented writing assignments in psychology.

One of the most common application-oriented writing assignments in psychology courses requires students to keep journals in which they reflect on the relevance and applicability of the course material to their personal experiences, the experiences of others, and/or their observations of real-world events (e.g., Connor-Greene, 2000; Graham, 2006; Hettich, 1990; Miller, 1997; Weber, 1984). For example, in her course on personality theory, Connor-Greene (2000) required students to keep a journal in which they applied the course material to characters from television shows, books, videos, songs, news events, politics, themselves, friends, or family. When writing about each example, students were instructed to identify the relevant theorist, to identify a specific concept from that theory, and to clearly and fully explain how the example illustrated the concept. To assess whether the journal assignments had a beneficial impact on student learning, test grades in two classes that were assigned to write weekly journals were compared to test grades in a class that was not assigned to write weekly journals. As predicted, test scores in the two classes that wrote journals were higher than the test scores of students in the control group.

Writing assignments that require students to apply the course material to events outside the classroom can also be utilized for final papers. For example, in a final paper assigned in their social psychology course, Lakin and Wichman (2005) instructed students to gather examples relevant to the course material from sources such as articles, comic strips, advertisements, advice columns, commercials, television shows, movie clips, or audio recordings. Compared to students in a control group, students who completed the target assignment reported that they found it easier to apply the course material to real-world events. The results of a follow-up survey indicated that this effect persisted nine months after the completion of the assignment.

Previous studies have explored how works of literature can be used to help students find examples of psychological principles in action (e.g., Boyatzis, 1992; Carlson, 1992; Cavanaugh, 1999; Lips, 1990; Osborn, 1990; Williams, 1986). For instance, Boyatzis (1992) had his students reflect on developmental psychology concepts through reading Angelou's (1969) book, *I Know Why the Caged Bird Sings*. Students reported that they found the book to be of high educational

value, and found that they were able to relate specific concepts from the course lecture to Angelou's narrative.

B. Description of the assignment.

In this article, we describe an assignment that takes good advantage of a traditional way of learning and show its benefit for a psychology class. Students read a non-fiction book that was not written by a social psychologist. In each of these books, social psychological principles and theories implicitly permeated the authors' explanations and descriptions of their own, as well as other people's life experiences. This approach is especially relevant because social psychology research draws on the human experience (such as helping behavior, stereotyping and prejudice, aggression, attraction), and frequently, social psychologists utilize real-world events as inspiration for theories of behavior. For example, the story of Kitty Genovese's murder prompted exploration of the situational factors that decrease the likelihood of helping behaviors in emergency situations (Darley & Latané, 1968).

The primary objective for this assignment was to facilitate students' self-discovery of knowledge (Mathie et al., 1993) by providing them with opportunities to apply some of the social psychological concepts and theories to actual real-world events beyond their own personal experience. In addition, the assignment gave students the opportunity to recognize the satisfactions that can come from selecting and reading a good nonfiction book. A final objective of this assignment was to give students practice at expressing themselves in writing. We asked our students a variety of questions about their impressions regarding this assignment. We hypothesized that students would find the assignment beneficial when it came to learning social psychology concepts, that they would find the assignment enjoyable, and that they would appreciate the flexibility they had to choose a book that interested them personally.

Initially, some students were less than enthusiastic about the idea of a book review assignment. We suspect that lukewarm experiences with previous book review assignments made some students resistant to the idea of a teacher "making" them read a book in addition to their textbook. In hopes of overcoming this potential initial resistance, we introduced a novel twist to traditional notions of a book review assignment. In contrast to previous research on book reviews in which the same book was assigned to the entire class (e.g., Carlson, 1992; Cavanaugh, 1999; Lips, 1990; Osborn, 1990; Williams, 1986), we adopted a learner-centered approach (Weimer, 2002) that gave students a substantial amount of freedom in choosing the book that they would review. Students chose from a list of ten books that covered a wide variety of topics of potential interest to students (see Table 1). We believed that providing students with this freedom and personal control over the direction of their learning experience would promote intrinsic motivation (Lepper & Hodell, 1989; Svinicki & McKeachie, 2011) and enjoyment of the assignment (Cialdini, 2009) while still setting the limits necessary to grade the assignment effectively.[6]

Our assignment differed from most teaching-related scholarship in psychology on book reviews in another important way. Most previous research on this topic has focused primarily on the application of psychology to works of fiction (e.g., Carlson, 1992; Cavanaugh, 1999; Lips, 1990; Osborn, 1990; Williams, 1986; for an exception, see Boyatzis, 1992). Although reviews of works of fiction promote the application of course material, they do not always facilitate the

[6] Three students in Sample 1 and two students in Sample 2 asked if they could read books not on the list. In each case, these books fit the criteria and were allowed.

application of course material to contemporary events in the real world. In contrast, by focusing exclusively on works of nonfiction, our assignment provided students with a unique opportunity to apply the course material to actual real-world events beyond their own experiences.

In addition to providing a general summary of the book, students were required to isolate at least three examples of how course-related social psychological concepts were illustrated in the book. Students were instructed to identify the concepts and to describe fully those aspects that were relevant to the examples. Grading of the assignment was based on: 1) the degree to which the examples fit the social psychological concepts, 2) the degree to which the student clearly and fully explained how these examples fit the social psychological concepts, 3) the general structure and organization of the student's points, 4) the general clarity of writing, and 5) the degree to which there was evidence that the student thought deeply about the book. The recommended length was three typed and double-spaced pages. The assignment was worth 20% of the overall course grade in Samples 1 and 2 and 15% in Samples 3 and 4.

Table 1. Relevant concepts and theories identified by students for each book on the list.

Book Title	Concepts and Theories Mentioned by Students
Brother Ray: Ray Charles Own Story	Stereotypes, Discrimination, Reciprocity Norm
Among the Thugs	Deindividuation, Prejudice, Discrimination, Normative Social Influence, Cognitive Dissonance Theory, Social Learning Theory
True Believers: The Tragic Inner Life of Sports Fans	Basking in Reflected Glory, Cognitive Dissonance Theory and Justification of Effort, Stereotypes, Self-Serving Bias, Proximity and Liking, Deindividuation, Normative Social Influence
Rammer Jammer Yellow Hammer: A Journey into the Heart of Fan Mania	Basking in Reflected Glory, Normative Social Influence, Schadenfreude, Investment and Comparison Level of Alternatives, Deindividuation, Stereotypes
How to Close Every Sale	Reciprocity, Overcoming Resistance to Persuasion, Nonverbal Behavior, Scarcity Principle, Commitment and Consistency Principle
The Tipping Point	Descriptive Norms, Peripheral Route to Persuasion, Social Loafing
Under the Banner of Heaven	Social identity, Informational Social Influence, Cognitive Dissonance Theory, Obedience to Authority, Ingroup/Outgroup, Social Exchange Theory, Minority Influence
Halfway Heaven: Diary of a Harvard Murder	Self-presentation, Self-monitoring, Self-esteem, Normative Social Influence, Loneliness, Individualism and Collectivism, Social Comparison, Authority and Social Influence, Attachment, Helping, Close Relationships, Rejection
Makes Me Wanna Holler: A Young Black Man in America	Stereotypes, Prejudice, Discrimination, Cognitive Dissonance Theory
Ghost Soldiers: The Epic Account of World War II's Greatest Rescue Mission	Obedience to Authority, Social Identity, Prejudice, Social Comparison

II. Method.

A. Participants.

Participating students were in one of four sections (n = 46, n = 26, n = 33, and n = 26, respectively) of introduction to social psychology courses taught at two medium-sized state universities. One author taught two sections, and a second author taught the other two sections. The majority of students who take this level of coursework are either psychology majors or minors, and have already taken introductory psychology as a prerequisite. Therefore, they tend to be at the sophomore level of college or higher.

B. Procedure.

During the final week of the academic term and on the day the assignment was turned in, the course instructor left the classroom and a graduate student unconnected to the course asked students if they would be willing to participate in a study involving one of their assignments for the course. If so, they were given a consent form which described the nature of the study and which emphasized both the anonymity of participating and the separation of the study from any aspect of their course grade. Then, students completed an anonymous questionnaire consisting of both Likert items (also using a 5-point scale ranging from *not at all* = 1 to *very much* = 5) and open-ended questions assessing their views on the value of the assignment (see Likert items in Table 2).[7] To further assess the value of the assignment, four questions that provided a standard of comparison were added in Sample 4. Specifically, participants in Sample 4 completed Likert scale items about the educational usefulness of the assignment relative to other final papers that they had completed (using a 10-point scale in which the endpoints and neutral point were labeled; 1 = *much less useful*, 5 = *about the same,* 10 = *much more useful*).

III. Results.

A. Students' Assessment of Assignment.

To examine the extent to which the assignment was linked to students' learning experience, for each sample separately, we conducted single sample *t*-tests in which the mean for each item was compared to the point at or near the middle of the scale on each of the items assessing students' views on the value of the assignment. For the 5-point scale, the mean for each item was compared to the score of 3, which was the midpoint on the scale. Likewise, for the 10-point scale, items were compared to the score of 5, which was labeled as the neutral value. We reasoned that a significant difference by this standard would indicate that the assignment was a positive learning experience for the student (Simmons & Prentice, 2006). As shown in Table 2, only one item in Sample 2 and Sample 4 failed to meet the standard. In further evidence that the assignment was linked to students' learning experience, students in Sample 4 evaluated the assignment more favorably than most final papers that they had previously completed (See Table 3). Furthermore, 88% of the students in Sample 1, 85% of the students in Sample 2, 94% of the

[7] Prior to completing the dependent measures, students in Sample 3 made ungraded informal presentations in which they described the examples of social psychological principles and theories that they found in the book that they chose.

students in Sample 3, and 84% of the students in Sample 4 recommended that the teacher continue to use the book review assignment in future social psychology courses.

In response to an open-ended question about how the assignment furthered their learning experience beyond other aspects of the course, numerous students commented that the assignment helped them apply social psychological principles to events in the real world. Two examples of such comments were: "While I was reading the book, I constantly was applying things I learned in class to the book," and "The assignment helped me see other examples of social psychology besides ones from my own life." Taken together, the results of both quantitative and qualitative analyses suggest that, across samples, students found the assignment to be a valuable learning experience. Specifically, students indicated that this assignment was effective in helping them apply the course material to "real life" events beyond their own particular experiences.

Table 2. Students' Assessment of Assignment.

Question	Sample 1 (n = 46)	Sample 2 (n = 26)	Sample 3 (n = 33)	Sample 4 (n = 25)
To what extent did the book review help you find ways to apply some of the social psychological concepts and/or theories that you learned about in this course?	3.79** (1.04)	3.96** (1.08)	4.33** (.65)	3.64** (.82)
How much did the book review assignment help you to closely examine the relevance of the course material to events beyond your own particular experience?	3.36* (1.13)	3.77** (.99)	4.15** (.76)	3.44* (.88)
How much did your having a choice in what book to read add value to the assignment for you?	4.33** (1.12)	3.96** (1.0)	4.64** (.55)	4.23** (.91)
How much did this assignment encourage you to think critically about the course material?	3.49** (1.11)	3.88** (.91)	4.03** (.77)	3.69** (.62)
To what extent did your book review assignment make a significant contribution to your learning in this course?	3.19** (1.09)	3.19 (.98)	3.79** (.60)	3.21 (.96)

Note. In Samples 1-3, judgments were made on 5-point scales ranging from 1 (not at all) to 5 (very much). Labels were not provided for values 2, 3, and 4 on the scales. In Sample 4, all values on the 5-point scales were labeled (1 = *not at all*, 2 = *slight extent*, 3 = *moderate extent*, 4 = *great extent*, 5 = *very great extent*). Standard deviations are in parentheses. *$p < .05$ **$p < .01$.

Table 3. Students' Relative Assessment of Assignment in Sample 4.

Question	Mean
In comparison to other writing assignments of similar length that required you to read an assigned journal article, did you find the book review assignment to be enjoyable?	6.73** (1.73)
In comparison to most final papers that you have completed, how useful did you find the book review assignment to be in helping you to closely examine the relevance of the course material to events beyond your own particular experience?	6.39** (1.83)
In comparison to most final papers that you have completed, how useful did you find the book review assignment in helping you to find ways to apply the concepts that you learned about in the course?	6.85** (1.89)
In comparison to other writing assignments of similar length that required you to read an assigned journal article, how useful did you find the assignment to be in helping you to think critically about the course material?	6.73** (1.76)

Note. Judgments were made on 10-point scales (1 = *much less*, 5 = *about the same*, 10 = *much more*). Standard deviations are in parentheses. $n = 25$. ** $p < .01$.

B. The Importance of Choice.

As noted earlier, we expected that students would find the assignment to be more enjoyable if they were allowed to choose the book they read. For each sample separately, single sample *t*-tests using the item asking about the value of choice were consistent with this expectation (See Table 2). Sample 4 yielded additional supportive evidence when a strong positive correlation was found between the value of choice and the degree to which students enjoyed the book review assignment relative to other final papers they had completed, $r = .70$, $p < .001$. Multiple possible interpretations of this correlation merit additional discussion. On the one hand, students' greater sense of autonomy associated with choosing the book that they would read might have led to greater enjoyment of the assignment (Deci & Ryan, 1985). On the other hand, the combination of making the commitment to read their book choice and the subsequent time-consuming reading behaviors that were consistent with their commitment may have led to increased perceptions of the value of the assignment (Cialdini, 2009). Stated another way, perhaps it was difficult for students to report disliking an assignment involving a book that they personally chose and spent time reading. Regardless of which interpretation the reader favors, the implication for teachers is still the same. When teachers provide students with multiple options for fulfilling a course requirement, students are more likely to enjoy (or convince themselves that they enjoyed) the assignment.

When students were asked to share their opinions regarding the specific aspects of the exercise that they liked, 25% of the students in Sample 2 and 50% in Sample 3 spontaneously mentioned that they liked being able to choose the book.[8] Based on our discussions with students and the tenor of the responses to this open-ended question, it appears that giving students a voice in charting the direction of their learning experience added value to the assignment not only for students who already enjoy reading (e.g., "I liked that we got to choose what book we read and the choices given were a wide variety."), but also for students who typically do not enjoy reading (e.g., "I enjoyed it only because I found a book on the list that highly related to me. Had I not, I

[8] Open-ended questions were not included in the survey for Sample 1 and Sample 4.

would have hated this assignment. Mainly because book reviews take so long (you have to read the book) and my time is very valuable.")

IV. Discussion.

We designed an assignment that highlighted the relevance of social psychology to events beyond students' own experiences by exploiting a traditional learning tool of reading a book. Students chose a book from a list provided by the instructor that focused on phenomena of a social psychological nature. Their task was to identify three examples of social psychological concepts in the book and to clearly and fully explain how these examples fit the concepts.

Overall, students' attitudes toward the assignment were clearly favorable. Across the four samples from two universities and two instructors, students indicated that this assignment was effective in helping them apply the course material to "real life" situations and to see the relevance of the course material to events beyond their own particular experiences. Almost all students recommended the continued use of the assignment, a noteworthy finding given that many students, in all the authors' experiences, tend to grumble about additional reading assignments. Moreover, the results of both quantitative and qualitative analyses suggested that allowing students to choose the book that they would read promoted enjoyment of the assignment.

One limitation of our research was that we did not have a control condition in which we could compare students' experiences with other students who did not have the assignment. However, given the consistency of responses across the four samples and the preponderance of students who endorsed the assignment, there appears to be solid evidence for its value.

Variations of this assignment may be useful not only for social psychology instructors, but also for instructors of other psychology courses. For example, instructors teaching other seminars such as personality, developmental psychology, or abnormal psychology could easily create a list of biographies or autobiographies and ask the students to identify different personality traits, developmental themes, or clinical symptoms (e.g., Mueller, 1985). Using non-fiction books to provide students with a real-world application of concepts can expand beyond the realm of psychology as well; a powerful personal narrative provides a primary source that could serve to enhance a moment in history, add perspective on sociological and political issues, promote awareness of the struggles faced by people from other cultures and backgrounds, address tough ethical issues, or enhance context for a scientific or medical breakthrough.

In another variation on this assignment, some instructors might choose to incorporate collaborative learning into this exercise by encouraging small groups to meet during the academic term to discuss their reactions to the book with one another. Additionally, we encourage instructors to have students give informal individual or group presentations so that they can tell their classmates about the ways in which social psychological concepts and theories were relevant to the book that they chose. In Sample 3 and Sample 4, the lively discussions that ensued enabled the class to review the material covered throughout the course. Given that some students become quite anxious about public speaking, in future semesters, the first author plans to conduct a "speed dating" activity in which students quickly discuss the book that they read with one classmate at a time.

After using the same list of book choices for several years in our classes, we wish to offer a suggestion that we hope will maximize the excitement that this assignment generates not only for students, but also for instructors. The first author is revitalizing the assignment by adding a

book that he is currently reading, *Ghost in the Wires: My Adventures as the World's Most Wanted Hacker* (Mitnick*,* 2011) to the list of book choices. Given that one of the goals for this assignment is to help students discover the joys of reading a book of their own choosing during their leisure time, it only makes sense for the instructor to practice what he or she preaches. Hopefully, when the teacher enthusiastically does the same academic work that the students are doing, a cooperative atmosphere will arise in which students and their instructor become partners in the learning process.

In closing, we would like to highlight some of the most important benefits of this assignment. First, as facilitators of student learning and thinking, we were delighted to witness students taking pride in reading a book of their own choosing that they actually enjoyed. Indeed, some of our fondest memories involved glowing endorsements of this assignment from students who readily admitted that they rarely read books for pleasure. For example, in response to an open-ended question about the assignment, one student wrote, "I'm glad I read the book. I can honestly say I haven't read a book since high school." In an informal individual presentation, another student who readily acknowledged that she rarely spends her leisure time reading books confessed that she enjoyed the book *Makes Me Wanna Holler: A Young Black Man in America* so much that she later convinced her mother to read it. Second, by applying the course material to the experiences of the characters in the book that they chose, students were able to see the value of what they had been studying. For example, in response to an open-ended question about how the assignment furthered student learning beyond other aspects of the course, one student wrote, "The assignment showed me that there are people in everyday life that use the principles of social psychology." Third, based on informal discussions with many of our former students, we suspect that this assignment allows more of the course material to "stick" (e.g. Roediger, 2000). In the months following the course, some students might actually remember the social psychological aspects of the book that they chose more than certain parts of the textbook.

References

Angelou, M. (1969). *I know why the caged bird sings*. New York: Bantam.

Bleske-Rechek, A. L. (2002). Obedience, conformity, and social roles: Active learning in a large introductory class. *Teaching of Psychology, 28*(4)*,* 260-262.

Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals*. New York: Longman.

Bonwell, C. C., & Eison, J. A. (1991). *Active learning: Creating excitement in the classroom*. (Rep. No. 1). Washington, DC: George Washington University, ASHEERIC Higher Education.

Boyatzis, C. J. (1992). Let the caged birds sing: Using literature to teach developmental psychology. *Teaching of Psychology, 19*(4)*,* 221-222.

*Buford, B. (2011). *Among the thugs*. New York: Random House.

Carlson, J. F. (1992). From Metropolis to Never-neverland: Analyzing Fictional Characters in a Personality Theory Course. *Teaching of Psychology*, *19*(3), 153-155.

Cavanaugh, J. C. (1999). From Metropolis to Never-neverland: Analyzing fictional characters in a personality theory course. *Teaching of Psychology*, *26*(1), 51-53.

*Charles, R., & Ritz, D. (2003). *Brother Ray: Ray Charles' own story*. (2003). New York: De Capo Press.

Cherney, I. (2008). The effects of active learning on students' memories for course content. *Active Learning in Higher Education, 9*(2)*,* 152-171.

Cialdini, R. B. (2009). *Influence: Science and Practice* (*5th ed.*). Boston, MA: Pearson.

Connor-Greene, P. A. (2000). Making connections: Evaluating the effectiveness of journal writing in enhancing student learning. *Teaching of Psychology, 27*(1)*,* 44-46.

Craik. F. I. M., & Lockhart, R. S. (1973). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior, 11*(6), 671-684.

Darley, J. M., & Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology, 8*, 377-383.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

*Gladwell, M. (2000). *The tipping point.* New York: Malcolm Gladwell.

Graham, S. M. (2006). Understanding the applicability of social psychology to everyday life: The benefits of a semiweekly journal assignment. *Teaching of Psychology, 33*(1)*,* 54-55.

Hettich, P. (1990). Journal writing: Old fare or nouvelle cuisine? *Teaching of Psychology, 17*(1)*,* 36-39.

*Jirard, J. (1989). *How to close every sale.* New York: Warner Books.

King, A. (1993). From sage on the stage to guide on the side. *College Teaching, 41*(1), 30-35.

Kowalski, R. M., & Lakey, C. E. (2004). Teaching application and personal relevance through writing in courses on gender. *Teaching of Psychology, 31*(4), 279-281.

*Krakauer, J. (2003). *Under the banner of heaven*: *A story of violent faith*. New York: Anchor Books.

Laird, T. F. N., Shoup, R., Kuh, G. D., & Schwarz, M. J. (2008). The effects of discipline on deep approaches to student learning and college outcomes. *Research in Higher Education, 49,* 469-494.

Lakin, J. L., & Wichman, A. L. (2005). Applying social psychological concepts outside the classroom. *Teaching of Psychology, 32*(2)*,* 110-113.

Leamnson, R. (2000). Learning as biological brain change. *Change, 32*(6)*,* 34-40.

Lepper, M. R., & Hodell, M. (1989). Intrinsic motivation in the classroom. In C. Ames & R. Ames (Eds.), *Research on motivation in education* (Vol. 3, pp. 73-105). San Diego: Academic Press.

Lips, H. M. (1990). Using science fiction to teach the psychology of sex and gender. *Teaching of Psychology, 17*(3)*,* 197-198.

Mathie, V. A., Beins, B., Benjamin, L. T., Jr., Ewing, M. M., Hall, C. C. I., Henderson, B., …, & Smith, R. A. (1993). Promoting active learning in psychology courses. In T. V. McGovern (Ed.), *Handbook for enhancing undergraduate education in psychology* (pp. 183–214). Washington, DC: American Psychological Association.

*McCall, N. (1995). *Makes me wanna holler: A young black man in America*. New York: Vintage.

Miller, S. (1997). Self-knowledge as an outcome of application journal keeping in social psychology. *Teaching of Psychology, 24*(2)*,* 124-125.

*Mitnick, K. (2011). *Ghosts in the wires: My adventures as the world's most dangerous hacker*. New York: Back Bay Books.

Mueller, S.C. (1985). Persons in the personality theory course: Student papers based on biographies. *Teaching of Psychology, 12*(2)*,* 74-78.

Osborn, D. (1990). Samuel Butler's *Erewhon* as social psychology. *Teaching of Psychology, 17*(2), 115-117.

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education, 93*(3), 223-231.

Roediger, H. L., III (2000). Why retrieval is the key process to understanding human memory. In E. Tulving (Ed.), *Memory, consciousness, and the brain: The Tallinn Conference* (pp. 52-75). Philadelphia: Psychology Press

Roediger, H. L., III, & Guynn, M. J. (1996). Retrieval processes. In E. L. Bjork & R. A. Bjork (Eds.), *Memory* (pp. 197-236). New York: Academic Press

*Queenan, J. (2003). *True believers: The tragic inner life of sports fans*. New York: Picador.

Sides, H. (2002). *Ghost soldiers: The epic account of World War II's greatest rescue mission*. New York: First Anchor Books.

Simmons, J. P., & Prentice, D. A. (2006). Pay attention! Attention to the primes increases attitude assessment accuracy. *Journal of Experimental Social Psychology, 42*(6)*,* 784-791.

Smith, C. V., & Cardaciotto, L. (2011). 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.

*St. John, W. (2004). *Rammer jammer yellow hammer: A journey into the heart of fan mania*. New York: Three Rivers Press.

Svinicki, M., & McKeachie, W. J. (2011). *Teaching tips: Strategies, research, and theory for college and university teachers (13th ed.)*. Belmont, CA: Wadsworth.

*Thernstrom, M. (1997). *Halfway heaven: Diary of a Harvard murder*. New York: Doubleday.

Watson, D. L., Kessler, D. A., Kalla, S., Kam, C. M., & Ueki, K. (1996). Active learning exercises are more motivating than quizzes for underachieving college students. *Psychological Reports, 78*(1)*,* 131-134.

Weber, A. (1984). Teaching social psychology. *Contemporary Social Psychology, 10*(3), 9-10.

Weimer, M. (2002). *Learner-centered teaching: Five key changes to practice.* San Francisco, CA: Wiley.

Williams, K. G. (1986). Psychology and literature: An interdisciplinary approach to the liberal curriculum. *Teaching of Psychology, 13*(2)*,* 59-61.

Yoder, J. D., & Hochevar, C. M. (2005). Encouraging active learning can improve students' performance on examinations. *Teaching of Psychology, 32*(2)*,* 91-95.

Young, A., & Fulwiler, T. (1986). *Writing across the disciplines*. Upper Montclair, NJ: Boynton/Cook.

* Indicates a book that students had the option of choosing to read and analyze for the assignment.

Journal of the Scholarship of Teaching and Learning, Vol. 13, No. 2, May 2013, pp. 84 – 100.

Enhancing curriculum through service learning in the social determinants of health course

Ronica Nicole Rooks[1] and Christine Tagliaferri Rael

Abstract: Service learning bridges classroom learning and community volunteerism and is anchored in the curriculum, classroom discussion, and community. We incorporated service learning projects (SLP) into three Social Determinants of Health courses (2008-2010) to promote: experiential learning; undergraduate scholarship; faculty career development through the scholarship of teaching and learning; and collaborative university-community research to reduce social inequalities in health. We examined whether SLP facilitated student learning of course concepts. We used mixed methods analyzing students' (n=25) pre-/post-test surveys, research papers, and site supervisors' (n=17) interviews. Despite positive survey ratings, results showed decreased student agreement about SLP facilitating student learning. Content analysis revealed specific student themes: finding SLP rewarding for future public health careers; aligning student interests with community-based organizations (CBOs)' goals; and valuing interactive experiences with CBOs' clients. Students gained beneficial career development skills with CBOs but needed better preparation for their SLP by increased discussion of their and CBOs' expectations.

Keywords: service learning, social determinants of health, social inequalities, undergraduates, public health

I. Introduction.

Service learning bridges classroom learning and community volunteerism, where service is anchored in the curriculum, classroom discussion, and community (Ballantine & Phelps, 2002), i.e. the laboratory to apply public health and social science knowledge and skills to social issues. It also promotes undergraduate scholarship through research, professional networking, and service. In recognition of these benefits the Institute of Medicine recommends that all undergraduates should have access to public health education, including an understanding of how the social and physical environments shape health through an ecological model (Cashman & Seifer, 2008; Gebbie, Rosenstock, & Hernandez, 2003). Service learning can be a vehicle to fulfill this mandate by applying public health and social and behavioral sciences to health-focused social problems.

We are particularly interested in the benefit of improving student outcomes through problem-based learning with a community-based organization (CBO). The literature on health-focused service learning identifies various benefits to students and CBOs. Students' benefits include: improved academic outcomes such as better grades, critical thinking, and problem-solving skills; connecting theory to practice; increased capacity to view phenomenon from multiple perspectives and a deepened sensitivity to diversity; increased social awareness and

[1] Department of Health and Behavioral Sciences, ronica.rooks@ucdenver.edu

social justice beliefs; increased confidence to make community service contributions (e.g., self-efficacy); leadership; and enhanced understanding of their professional development and career paths (Cashman & Seifer, 2008; Peterson & Yockey, 2006; Reeb, 2010). CBOs' benefits include: volunteer hours to meet the organization's goals addressing social problems; freeing the organization's time and resources for other projects; gaining new ideas from outsider perspectives; programmatic and infrastructure improvements; opportunities to contribute to students' education; and developing positive relationships with the university (Blouin & Perry, 2009; Brown et al., 2006; Cashman & Seifer, 2008; Peterson & Yockey, 2006).

However, little research exists on whether universities are successful in improving student learning outcomes as they address health-focused social problems and social inequalities through community outreach (Kenny & Gallagher, 2002; Mobley, 2007). As a public health student Toboada (2011) critiqued community service learning models designed as charity-oriented activities, traditionally pairing White, middle-class students with low income and communities of color. Based on her experiences, students need proper curriculum training in racism, power, and privilege before and during their interactions with communities so that their experiences do not reinforce bias.

Addressing the prior issues Loewenson and Hunt (2011) used social determinants of health (SDOH), health disparities, and social justice frameworks, as well as theory on working with underserved populations, to positively transform public health nursing students' attitudes toward the homeless. By the end of the semester students reported stronger beliefs related to structural causes for homelessness and more comfort associating with homeless people. This research shares the theme of adequate curriculum preparation on social inequalities in health prior to and concurrent with students' service learning experiences. Thus, as instructors a critical first step to enhancing the health of vulnerable populations is improving students' attitudes towards working with them throughout the semester (Daiski, 2007).

To facilitate student learning, theory on student engagement supports using service learning projects (SLP) to foster community social change. Giles and Eyler (1994) developed service-learning theory from Dewey's principles of experience, inquiry, and reflection within his theory of knowing. Dewey's educational philosophy connects theory to practice in communities for social justice. He presents problems that awaken students' curiosity and demand for information. Students build on their prior experiences through interaction and reflective thinking, resulting in learning. Similarly, Kolb's (1984) experiential learning theory links education, paid work, and personal development to a lifelong active process of learning through direct interactions with phenomena. Kolb believes that learning is an active process where knowledge is created through transforming experiences.

Considering these guidelines, we incorporated SLP into three SDOH courses (fall 2008-2010) with the goals of promoting: experiential learning at the intersection of public health and the social and behavioral sciences; undergraduate scholarship and professional development; faculty career development through the scholarship of teaching and learning; and collaborative university-community research to reduce social inequalities in health. The major objective of this course was for students to analyze the relationships between the individual pursuit of health and the social structural contexts in which this happens in our society, with some cross-national discussions. We examined issues related to the social, psychological, behavioral, economic, political, cultural, and environmental variations in health and disease, particularly focusing on social inequalities in health, social stress, health behaviors, illness experiences, relations between

providers and patients, the structure and processes of healthcare organizations, financial and other barriers to accessing healthcare, health policy, and social change.

Students were required to do SLP, engaging in civic responsibility while connecting course concepts to their service. The assignment specified that students do their SLP with a local health- or healthcare-focused CBO. The specific SDOH topics to focus on in their service learning experiences and research papers were left to students, but we suggested the following general topics: disease (e.g., biomedical perspective) versus illness (e.g., psychosocial perspective) in society, health behaviors, social stressors, experiencing illness and disability, interactions with healthcare professionals, the healthcare system, healthcare delivery, treatment, healthcare policy, etc. from a patient's and/or an organization's perspective. Students had to:

1) Provide a service learning contract describing their SLP.
2) Work at least 20 hours in a CBO.
3) Keep journal notes on their experiences, observations, and reflections.
4) Participate in two in-class discussions, reflecting on their SLP.
5) Complete a final research and evaluation paper of their SLP.
6) Provide a letter verifying their completion of their SLP and hours signed by their site supervisor.

We provided students with SLP learning goals, a rubric, and on-going course discussions about the SLP throughout the semester to model successful student learning. We included the rubric in Appendix 1.

This paper addresses whether SLP facilitate applied student learning of SDOH course concepts using mixed methods. Our intent was to help students gain an awareness of social inequalities in health, understand their own biases, and illustrate how these prior issues could affect health-related problems.

II. Methods.

A. Study Design.

With Institutional Review Board approval from the University of Colorado Denver, we examined the service learning experiences of SDOH students and their site supervisors. The sample included university students from three semesters in 2008 (undergraduate), 2009 (undergraduate and master's level), and 2010 (undergraduate and co-taught course) and their site supervisors, primarily working in non-profit, health-related CBOs in Denver, Colorado. We used mixed methods, descriptively analyzing students' pre- and post-test quantitative surveys and their research papers and site supervisors' interviews using qualitative analysis. We divided students (n=25) and supervisors (n=17) into two mutually exclusive groups and asked them to respond to surveys or interviews unique to their group.

Student Sample. A research assistant proctored student consent forms 4-5 weeks into each course. Students either immediately turned their consents in or by semester's end. Students' involvement in service learning data collection was voluntary, and they were able to refuse participation at any time, for any reason. We asked students to respond to pre- and post-test surveys to assess their service-learning experiences, with open-ended comments at the end (Gelmon, Holland, Driscoll, Spring, & Kerrigan, 2001). Gelmon and colleagues present a long and revised, shorter version of this survey in their research, developed for use with undergraduate students. We used questions from the longer version, but substituted in a few

questions from the revised version. For the pre-test, students were asked to respond to whether they believed their SLP would benefit their learning of SDOH course material (prior to beginning their SLP). The post-test was the same as the pre-test, except we changed questions to past tense (post completing their SLP). We administered the pre-test within the second week and the post-test during the last week of the course. We instructed students to choose a unique, random number to anonymously match their pre- and post-tests. Additionally, we examined students' SLP papers with reflections about their experiences.

Site Supervisors/CBO Sample. We conducted follow-up interviews after final grade submissions so that site supervisors' comments did not factor into students' grades. Prior to conducting mostly telephone interviews a research assistant obtained signed consents from site supervisors via email or fax. Site supervisors' survey participation was completely voluntary; they were allowed to stop the interview at any time. The interview guide included questions about students' responsibilities, whether students had sufficient knowledge to volunteer at the site, the quality of their work, whether students had professional characteristics (e.g., showed initiative, communicated effectively, and reliability), and whether students' work at the sites was beneficial to the CBO, other positive or negative aspects associated with this experience, and suggestions for improvement to the SLP.

Dependent variables. To assess students' learning we used quantitative surveys and qualitative paper reflections focused on the value of service learning as an applied tool for learning course concepts. Student surveys included twenty-four items on whether their attitudes toward community service in the course changed during their SLP experiences (shown in Table 2). Each item indicated students' level of agreement via a 5-point Likert scale (1=strongly disagree to 5=strongly agree). Survey items reflected a Cronbach's alpha of 0.884, indicating high internal question consistency. Students were also asked to qualitatively reflect on their SLP experiences in their papers' research/evaluation and conclusion sections, particularly focusing on any applied learning gained in the course or for their future professional development.

Independent variables. We included five socio-demographic survey items to understand any variation among our students. Race/ethnicity included the following categories: Caucasian/White, African American/Black, Hispanic/Latino, Native American, Asian/Asian American, or Other. Age included: 18-24, 25-34, 35-44, or 45-54. Sex included: male or female. Class level included: freshman, sophomore, junior, senior, post-baccalaureate, graduate, or other. And, paid work/hours per week included: 1-10, 11-20, 21-30, 31-40, 41+ hours, or "I do not have a job."

Statistical Analysis. For our quantitative analysis we used a paired sample t-test to determine whether students' mean level of agreement regarding their learning through community service varied significantly between the pre- and post-tests. Due to the small sample size, we set significance at 0.1 to increase the likelihood of attaining meaningful results. For our qualitative analysis we used content analysis to examine students' survey comments and research papers and site supervisors' interviews in Microsoft Excel (Swallow, Newton, & Van Lottum, 2003). We developed a priori codes based on the surveys and interview questions, as well as deductive codes via emerging themes. To increase inter-rater reliability, both co-authors reviewed all documents. Combining both types of analyses as mixed methods is beneficial for data triangulation, providing generalizability and contextual-depth to our analyses (Borkan, 2004).

III. Results.

A. Quantitative Analyses.

Table 1. shows summary statistics of students' socio-demographic characteristics from the pre- and post-test surveys. The majority of students were female, Caucasian/White, seniors in class level, and working between 21-30 hours a week.

Table 3 shows paired sample t-test mean responses for students' pre- and post-test surveys. Students generally rated the pre-test items between neutral to agree, with the majority of items slightly decreasing in agreement, a negative change, by the time of their post-tests. The only significant items included: Q.1, Q.8, Q.20, and Q.22, showing that students felt less likely to believe: their service learning work could be used in their everyday lives; it benefitted the community; it made them aware of their own biases and prejudices; and it helped them enhance their leadership skills. Students' ratings on eight questions (e.g., Q.4, Q.6, Q.10-11, Q.17-19, and Q.21) slightly increased, ranging from neutral to agree or remaining the same over time, but none of these findings were significant. While the final question (Q.3) did not show a significant change over time, students slightly increased their disagreement that they could have learned more from this course if more time was spent in the classroom instead of doing community work. The highest student agreement items at the beginning and end of the course focused on them being comfortable working with other cultures and being able to make a difference in their communities.

We also showed significant change items by course year (Table 3), where most significant items did not overlap in each year. In 2008, two items showed improvement, with students changing from neutral to agree regarding their interactions with their community partner enhancing their course learning and the community work assisting them in defining their profession. But, the remaining two items decreased from agree to neutral, regarding the community work making students more aware of their own biases and feeling they could make a difference in their communities. In 2009, students increased their disagreement with the item that they could learn more from the course by spending more time in the classroom rather than the community and increased their agreement with the item that performing community work helped them clarify their majors. Students decreased their agreement with the items concerning community work helping them become more aware of their communities' needs; it assisting them in defining their profession; it making them more marketable in their profession; and it helping them develop problem-solving skills. In 2010, students were less likely to believe that community participation helped them to see how the subject matter could be used in everyday life; more likely to agree about volunteering before the course; and more likely to agree that they have a responsibility to serve their community. Despite the items remaining positive, about half of the significant changes across each year occurred in a negative direction.

B. Qualitative Analyses.

Students' comments revealed that those who intended to pursue a public health career found this experience especially relevant to their course learning and career development. Students stated:

> "…The most valuable thing this course did for me was to classify the practical need our individual people, counties, communities, states, and nations as a whole all have for public health."
>
> "I think that the service learning project is so great because it gives us real-world experience and allows us to see the class-content outside of the classroom."
>
> "I would rate this experience as substantial in regards to a research or advocacy position I might employ one day. I feel this way because I assume there will be many jobs within the human and health services sector i.e. public health that require a large amount of research…"
>
> "This project also helped solidify that my future educational endeavors will be focused around HIV/AIDS."
>
> "The lasting effect this project has is that I know how much more research needs to be put into prenatal epidemiology and how important public health issues in maternal and child health need to be addressed… this experience impacted me on a greater level than simply scholastic achievement. I felt like I gained real world knowledge…"

Table 1. *Summary Statistics of Students' Demographic Characteristics from the Pre- and Post-test Surveys.*

		Pre-test	Post-test
		Percent	Percent
Gender	Male	9.4	12.5
	Female	84.4	68.8
Race/Ethnicity	Caucasian/White	53.1	50.0
	African American/Black	12.5	3.1
	Hispanic/Latino	15.6	12.5
	Asian/Asian American	3.1	3.1
	Other	6.2	6.2
	Multiracial	6.2	6.2
Class Level	Sophomore	9.4	3.1
	Junior	28.1	31.2
	Senior	50.0	37.5
	Graduate	9.4	9.4
Hours Worked per Week	1-10 Hrs/Wk	9.4	15.6
	11-20 Hrs/Wk	6.2	3.1
	21-30 Hrs/Wk	28.1	31.2
	31-40 Hrs/Wk	28.1	21.9
	41+ Hrs/Wk	6.2	0
	I do not have a job	18.8	6.2

Source: Social Determinants of Health students in 2008-2010.

Notes. Our original pre-test (n=31) and post-test (n=27) sample sizes varied due to: 1) differences in those students who began the course, prior to the drop/add period, and those who finished the course and 2) absences on the days the surveys were administered. After we matched the pre- and post-test surveys a sample size of 25 students remained. We coded the unmatched

surveys as missing data.

Table 3. *Paired Sample T-test Mean Results for Students' Pre- and Post-Test Surveys.*

Item Statement	Pre-test Mean	Post-test Mean	Sign. (2-tailed)
Q.1) The community participation aspect of this course will help me to see how the subject matter I learn can be used in everyday life.	4.24	3.44	0.011+
(Q.2) The community work I will do through this course will help me to better understand the lectures and readings in this course.	3.76	3.48	0.283
(Q.3) I feel I could learned more from this course if more time was spent in the classroom instead of doing community work.	2.76	2.60	0.491
(Q.4) The idea of combining work in the community with university coursework should be practiced in more classes at this university.	3.88	3.96	0.723
(Q.5) I am responsible for the quantity and the quality of knowledge that I obtain from this course.	4.28	4.00	0.183
(Q.6) I was already volunteering in my community before taking this course.	3.32	3.60	0.230
(Q.7) The community participation aspect of this course will show me how I can become more involved in my community.	4.08	3.96	0.560
(Q.8) I feel that the community work I will do through this course will benefit the community.	4.16	3.56	0.036+
(Q.9) The community work involved in this course will help me to become more aware of the needs in my community.	4.04	3.84	0.446
(Q.10) I have a responsibility to serve my community.	4.1	4.29	0.382
(Q.11) My interactions with the community partner will enhance my learning for this course.	3.72	3.72	1.000
(Q.12) Doing work in the community will help me to define my personal strengths and weaknesses.	3.76	3.64	0.641
(Q.13) Performing work in the community will help me clarify which major I will pursue.	3.46	3.17	0.418
(Q.14) The community work in this course will assist me in defining which profession I want to enter.	3.71	3.38	0.224
(Q.15) The work I will accomplish in this course will make me more marketable in my chosen profession when I graduate.	4.04	3.92	0.450
(Q.16) The community aspect of this course will help me to develop my problem-solving skills.	3.68	3.56	0.574
(Q.17) The syllabus provided for this course outlined the objectives of the community work in relation to the course objectives.	3.96	4.00	0.857
(Q.18) Most people can make a difference in their community.	4.36	4.36	1.000
(Q.19) I am comfortable working with cultures other than my own.	4.48	4.48	1.000
(Q.20) The community work involved in this course will make me aware of some of my own biases and	4.04	3.60	0.013+

prejudices.			
(Q.21) The work I will perform in this course will help me learn how to plan and complete a project.	3.80	3.80	1.000
(Q.22) Participating in the community will help me enhance my leadership skills.	4.04	3.64	0.047[+]
(Q.23) The work I will perform in the community will enhance my ability to communicate my ideas in a real world context.	4.08	3.84	0.110
(Q.24) I can make a difference in my community.	4.56	4.40	0.256

Source: Survey questions from Gelmon et al. (2001) for the Social Determinants of Health students in 2008-2010.

[+] Our significance level is set at α=0.1

Table 3. *Significant Mean Results for Students' Pre-and Post-test Surveys by Course Year.*

Item #	2008				2009				2010			
	N	Mean	t-test	Sig (2-tailed)	N	Mean	t-test	Sig (2-tailed)	N	Mean	t-test	Sig (2-tailed)
1									14	4.29	2.242	.043
									14	3.50		
3					6	2.17	-2.236	.076				
					6	2.67						
6									14	3.00	-2.223	.045
									14	3.64		
9					6	4.33	2.712	.042				
					6	3.50						
10									14	3.86	-2.876	.013
									14	4.36		
11	5	3.60	-2.138	.099								
	5	4.40										
13					6	4.00	3.841	.012				
					6	4.17						
14	5	3.40	-2.138	.099	6	4.00	2.697	.043				
	5	4.20			6	2.67						
15					6	4.17	2.236	.076				
					6	3.67						
16					6	4.00	2.236	.076				
					6	3.00						
20	5	4.20	2.449	.070								
	5	3.60										
24	5	4.80	4.000	.016								
	5	4.00										

Source: Social Determinants of Health students in 2008-2010.

[+] Our significance level is set at α=0.1.

> "The project was a great addition to the Social Determinants of Health course as well as to my personal exposure and experience…I gained valuable knowledge…including skills like data collection methods…I plan to continue my career path focusing on Public Health, Sociology, and Demography so obtaining these skills hands on has been a great asset to my academic and professional life…this was my first time [in college] with the opportunity to

apply my knowledge from inside the classroom, to something that is occurring right now, in the real world."

A second theme was the importance of aligning individual student interests with the organization's goals. Students who provided negative feedback often indicated that had their personal objectives been more compatible with those of their CBOs, they would have had a more enriching learning experience. One student commented, "If I had chosen a better site – many of my 'disagrees' [in the post-test] would have probably been 'agrees.'" Participants in volunteer opportunities stressing labor-oriented tasks, such as restocking shelves, cleaning, or running errands, were less likely to positively rate their experience than students who participated in research or outreach activities. In addition, students who strictly performed labor jobs were more likely to perceive their volunteering as a burden rather than as learning.

A third theme was some students' desires to be more interactive with their CBOs' clients. These students initiated their SLP believing they would have more "hands-on" learning experiences with people, but some CBOs worked at a macro-structural level needing students to do background research for educational materials, designing surveys, and influencing policy. One student who worked with a preventive health center CBO stated, "The weakness of the project was the lack of direct interaction with the students. I never interviewed the transgender students, therefore my poster could have been off or the flyer might have been irrelevant to the

Table 4. *Major Themes Identified from Students' and CBOs' Perceptions*.

Student Themes
1. SLP provided career development.
2. SLP were more beneficial when students' interests aligned with CBOs' goals.
3. SLP were more beneficial when there was interactive, "hands-on" learning with CBOs' clients.
CBOs' Themes
4. Students should understand the CBO's mission and have a clear idea of what they want to accomplish with their SLP.
5. Student SLP volunteers were beneficial to CBOs, who often had limited budgets.
6. Having student volunteers provided publicity for CBOs, raising public awareness about their causes.

Source: Social Determinants of Health students and their site supervisors in 2008-2010.

needs of those particular students." A second student found that her initial expectations of working on child abuse and neglect issues were disappointingly different than her experiences stating, "I thought that by working with organizations whose emphasis was on child wellbeing, that a person-to-person interaction with the children was part of my description…I thought it would be more a community center type atmosphere and not offices pushing out paperwork?…I wanted a more hands on experience...However, when taking a step back and analyzing the situation objectively I felt that what I was trying to accomplish…was exactly what they needed."

While some students expressed these concerns, many others were satisfied with their direct, interactive learning and experienced positive improvement in their early expectations to real experiences by semester's end. One student stated, "While it was an accident that I ended up in the mobile clinic, I think that it was the best place I could have been because of the direct contact with patients and because that is where the need was and continues to be." A student

working in a long-term care facility was initially terrified, "…expecting to find men and women who were on their death beds and getting ready to die…Instead my views and expectations were altered after my time at [anonymous CBO], and I now have a renewed understanding and appreciation for quality long term healthcare." Another student stated, "I would recommend to future students who want to volunteer with this organization to interact more with those who are homeless and spend as much time as you can with this organization because there are a lot of wonderful people to meet and interesting things to learn about the homeless in Denver."

Students' themes were consistent with many made by site supervisors, where the first supervisor theme was that students and CBOs could experience greater benefits if students had an introductory-level background in their CBOs' mission topic and a clear idea of what they wanted to accomplish with their SLP. Site supervisors were willing to train students in all relevant technical skills, but they agreed that incoming students should have some familiarity with the CBO. One supervisor reported, "Students should learn about the background of the organizations…checking the website, being knowledgeable about what the organization does. They should come with a few ideas about what they would like to learn." Once students became proficient with their acquired skills, demonstrating a positive learning transition, supervisors relied on students to provide their CBOs with meaningful assistance.

Despite the prior concern, at least two SLP turned into part-time paid employment for students. These service learners were described as "enthusiastic" and "self-starters" by their site supervisors. Both students forged strong bonds with the populations in which they worked.

Positive responses by site supervisors were overwhelming. This second theme was nearly uniform across all sites. CBOs reported that working with students was a great experience and provided lasting benefits to their organizations. Most CBOs operate on a limited budget and many supervisors reported that having volunteers was extremely helpful.

A third theme was publicity. Many supervisors viewed SLP as a way to raise awareness about their causes. One supervisor required students to shadow her before beginning their individual projects on homelessness. She stated, "A lot of people don't have a background on this and it's nice to be able to educate the community and to get our name and message out there." Often university students had no previous contact with the populations in which they worked. By exposing students to these groups, CBOs were able to promote greater public understanding of their clients.

IV. Discussion.

Perhaps the most striking finding on whether SLP facilitated student learning was the mean decrease in student agreement levels on all significant measures from the pre- to post-tests, despite ratings still occurring between neutral and strongly agree. These findings were not consistent with most service-learning research, where students' attitudes generally showed improvement in social responsibility, awareness, and social justice (Bach & Weinzimmer, 2011; Long et al., 2011; Ottenritter, 2004; Reeb, 2010). However, some researchers have encountered difficulties with decreasing student ratings (Parker-Gwin & Mabry, 1998; Villanueva, Hovinga, & Cass, 2011).

While Gelmon, et al. (2001) state that few students will demonstrate dramatic changes in their pre- and post-tests during a one semester course, we offer possible reasons for students' decreased agreement about their SLP. Two explanations relate to earlier themes of some students not having compatible interests with their CBOs and differences in students' initial expectations

versus later realities regarding client interactions. Third, the higher percentage of students working greater than 20 hours a week while taking this course and doing their SLP could have been an added stressor, dampening their experiences. From post-course student interviews, Madsen and Turnbull (2006) found that students' paid work was the biggest challenge for their service learning experiences. Fourth, while students were advised to avoid organizations that did not include applied research as a part of their SLP, some students did administrative tasks not relevant to their SLP. One future strategy to fix this problem is to channel students toward the faculty pre-approved CBOs who present their possible SLP in the course. Fifth, students' self-selection into this course and their initial naiveté or idealism about doing applied research and helping others could have negatively changed by semester's end. Students' attitude changes were possibly prompted by SLP difficulties, including: site selection, site interactions, or meeting CBOs' project goals.

To address the prior difficulties, we believe that instructor awareness of challenges is warranted to better prepare students and faculty for interaction with CBOs. Various authors reference challenges associated with students' service learning interactions with CBOs (Blouin & Perry, 2009; Cashman & Seifer, 2008; Peterson & Yockey, 2006). They mention unprofessional and/or unreliable students who risk CBOs' invested resources and the populations they serve. Sometimes students misrepresent CBOs in their writing assignments, when CBOs' missions are not learned and/or understood. Another problem is poor communication between faculty and CBOs, when CBOs' roles are not clearly defined in assignments. These challenges can reduce CBOs' enthusiasm to work with university students and faculty and must be addressed to sustain university-community partnerships for service learning.

We addressed these issues by becoming more proactive and inviting CBOs and the university's Experiential Learning Center to present their potential projects to help students find sites. We reminded students early in the semester about being professionals as they interact with CBOs, making sure they dress appropriately, show up on time, and learn the missions of their CBOs. In addition, for CBOs that we recommended to students, we emailed them the assignment and goals within the first two weeks of the semester so that they could ask us any questions if needed. Similarly, when students chose their own CBOs we told them to disclose the assignment, its goals, and instructor contact information at the beginning of their service.

But, we were not able to resolve all difficulties. In some CBOs students worked with the coordinator of volunteers rather than direct supervisors, who were often unaware of details pertaining to students' later volunteer activities. Subsequently, these direct supervisors were unable to provide feedback beyond their initial students' training experiences or they gave

Table 5. *Challenges Identified by Faculty.*

Student Challenges
1. Some students lacked compatible interests with their CBOs.
2. There were differences in some students' initial expectations versus their later realities regarding client interactions.
3. Some students worked more than 20 hours a week, creating an additional stressor on their service learning experiences.
4. Some students did administrative tasks not relevant to their SLP.
5. Some students' initial naivete or idealism about doing applied research and helping others could have negatively changed by semester's end.
CBO challenges
6. Some students worked with volunteer coordinators rather than direct supervisors, where the latter were often unaware of details pertaining to these students' volunteer activities.
Other research challenges
7. We had a small sample size.
8. Selection bias may exist in the students who took the course as an elective, prior to it becoming a requirement for undergraduate public health majors.
9. The survey we used measured attitudinal changes due to students' SLP over a short-term, but we do not know if these changes were maintained in the long-term.
10. We did not adjust for social desirability.

Source: Social Determinants of Health courses in 2008-2010.

secondary reports of student performance. This phenomenon was more common in larger organizations.

This research had other challenges. First, we had a small sample size from a single institution, despite collecting data from three courses over three years. Our sample size was smaller than expected in part due to difficulties with matching pre- and post-test surveys. We asked students to write a three-digit number on their pre-test and write it in their notes to use for the later post-test and anonymous matching. However, many students did not keep or remember the correct numbers to match the surveys. Second, selection bias was a concern in the type of students taking this elective course, prior to it becoming a requirement for undergraduate public health majors in our last course. Because the course work load was high, it likely attracted above average, upper-class level students at the university (Villanueva, Hovinga, & Cass, 2011). Third, the survey design measured short-term attitudinal changes (Gelmon et al., 2001); however, some research has shown long-term changes in student attitudes over years (Fenzel & Peyrot, 2005). Finally, this study did not adjust for social desirability; thus, students' responses could be artificially more positive than negative about their SLP.

Despite students' reduced, but still positive, interest in some aspects of community service over time and our prior limitations, we view service learning as a valuable learning experience for undergraduate public health students. By pairing service learning with undergraduate public health education students gain public health literacy through the application of social determinants to individual and community health (Cashman & Seifer, 2008). Service learning emphasizes reciprocal learning between students, faculty, and CBOs, as

well as reflection that connects practice to theory and critical thinking. It also develops citizenship skills to achieve social change. Our undergraduate public health program, situated within a college of liberal arts and sciences, is an interdisciplinary, pre-professional degree that teaches critical analysis, information synthesis, and problem solving (Cashman & Seifer, 2008; Riegelman, Teitelbaum, & Persily, 2002). This combination of knowledge and skills will better prepare undergraduate students to meet the challenges of and contribute to our nation's health.

In addition, service learning is valuable to undergraduate public health students because it provides practical field experience in a discipline where there are few internship opportunities for undergraduates. Schools of public health have typically favored research institute structures for student development rather than the work-based learning models in professional schools, but even these opportunities are limited for undergraduates (Madsen & Turnbull, 2006; Potter & Eggleston, 2003). Consequently, undergraduate students may lack exposure to important community issues. But, by creating service learning opportunities with CBOs as part of undergraduate public health matriculation, students can gain access to organizations, populations, and problems of interest to their career development.

Based on our experiences with this course we realize that future improvements to students' experiences means that we better prepare them for their SLP through early and on-going class discussions about their and CBOs' expectations. The following student comment continues to inspire us to improve students' service learning experiences—"At the end of my time with [anonymous CBO] for this project I can rate it as one of, if not the, best experiences that I have had as a student so far. It expanded my knowledge of the subject matter in such a way that I'm afraid any future class experiences may fall a little short."

Appendix 1. *Service Learning Project (SLP) Assignment Learning Objectives and Grading Rubric.*

1) Have knowledge of the social determinants of health perspective (SDOH), an interdisciplinary conceptual model for understanding health and healthcare problems. A list of possible issues to address within this model include: social inequalities in health by race, ethnicity, social class, gender, sexuality, disability, etc.; social stress; health behaviors; the life course (in utero to older adults); social and physical environments (social support, social networks, neighborhoods, housing, transportation, working conditions, etc.); doctor-patient relationships; access to and quality of healthcare; etc.

Inadequate	Adequate	Advanced
Student demonstrated poor knowledge of the SDOH perspective on exams, in journal and final paper discussions with no examples from the SDOH conceptual model, and no change in knowledge on pre- to post-test surveys.	Student demonstrated good knowledge of the SDOH perspective on exams, in journal and final paper discussions with multiple examples from the SDOH conceptual model, and change in knowledge on pre- to post-test surveys.	Student demonstrated excellent knowledge of the SDOH perspective on exams, in journal and final paper discussions with multiple examples from the SDOH conceptual model and high initial knowledge or significant change in knowledge from pre- to post-test surveys.

2) Be able to explain and provide examples of macro- and micro-level contributions to health and healthcare problems and offer solutions to address them.

Inadequate	Adequate	Advanced
Student did not discuss or provide examples for macro- & micro-level contributions to health and healthcare through his/her journal and final paper.	Student described macro- & micro-level contributions to health and healthcare problems with multiple examples through his/her journal and final paper.	Student described macro- & micro-level contributions to health and healthcare problems with multiple examples and discussed innovative solutions to health and healthcare problems through his/her journal and final paper.

3) Demonstrate their awareness of social inequalities related to health and healthcare, whether in terms of culture, race, ethnicity, gender, socioeconomic status, sexual orientation, immigration status, age, etc., and understand their own biases and how these could affect health and healthcare problems.

Inadequate	Adequate	Advanced
Student demonstrated poor knowledge of diversity/ social inequality issues in journal and final paper discussions with no examples of how these affect health and healthcare and little knowledge of other cultures or own biases on pre- to post-test surveys.	Student demonstrated good knowledge of diversity/social inequality issues in journal and final paper discussions with multiple examples of how these affect health and healthcare and knowledge of other cultures or own biases on pre- to post-test surveys.	Student demonstrated excellent knowledge of diversity/social inequality issues in journal and final paper discussions with multiple examples of how these affect health and healthcare and knowledge of other cultures or own biases on pre- to post-test surveys. Student demonstrated a deep understanding of structural barriers based on social inequalities in health and healthcare through their journals, final papers, site supervisor's interview report, and having high initial knowledge or significant change in knowledge from pre- to post-test surveys.

4) Demonstrate how their service learning projects can contribute to the organization's growth, meet community needs, and enhance community members' and students' lives.

Inadequate	Adequate	Advanced
Student did not demonstrate both a descriptive and analytical knowledge of the organization's or community's need, how this project meaningfully affected or changed community members' lives, and self-reflection on what was learned from this experience to enhance his/her education or career skills in his/her journal, final paper, and employer's interview report.	Student demonstrated good descriptive and analytical knowledge of the organization's or community's need, how this project meaningfully affected or changed community members' lives, and self-reflection on what was learned from this experience to enhance his/her education or career skills, with one or two examples in each of the prior categories from his/her journal, final paper, and site supervisor's interview report.	Student demonstrated excellent descriptive and analytical knowledge of the organization's or community's need, how this project meaningfully affected or changed community members' lives, and self-reflection on what was learned from this experience to enhance his/her education or career skills, with multiple examples in each of the prior categories from his/her journal, final paper, and site supervisor's interview report.

References

Bach, R., & Weinzimmer, J. (2011). Exploring the benefits of community-based research in a sociology of sexualities course. *Teaching Sociology, 39*, 57-72.

Ballantine, J., & Phelps, K. (2002). What can I do? Applying classroom knowledge to service work, or doing practice sociology. *Sociological Practice, 4*, 41-52.

Blouin, D. D., & Perry, E. M. (2009). Whom does service learning really serve? Community-based organizations' perspectives on service learning. *Teaching Sociology, 37*, 120-135.

Borkan, J. M. (2004). Mixed methods studies: A foundation for primary care research. *Annals of Family Medicine, 2*, 4-6.

Brown, J. D., Bone, L., Gillis, L., Treherne, L., Lindamood, K., & Marsden, L. (2006). Service learning to impact homelessness: The result of academic and community collaboration. *Public Health Reports, 121*, 343-348.

Cashman, S. B., & Seifer, S. D. (2008). Service-learning: An integral part of undergraduate public health. *American Journal of Preventive Medicine, 35*, 273–278.

Daiski, I. (2007). Perspectives of homeless people on their health and health needs priorities. *Journal of Advanced Nursing, 58*, 273–281.

Fenzel, L. M., & Peyrot, M. (2005). Comparing college community participation and future service behaviors and attitudes. *Michigan Journal of Community Service Learning, 12*, 23-31.

Gebbie, K., Rosenstock, L., & Hernandez, L. M. (eds). (2003). *Who will keep the public healthy? Educating public health professionals for the 21st century*. Washington, D.C.: Institute of Medicine of the National Academies, National Academy Press.

Gelmon, S. B., Holland, B. A., Driscoll, A., Spring, A., & Kerrigan, S. (2001). *Assessing Service-Learning and Civic Engagement: Principles and Techniques*. Campus Compact. Providence, RI: Brown University.

Giles, D. E., & Eyler, J. (1994). The theoretical roots of service-learning in John Dewey: Toward a theory of service-learning. *Michigan Journal of Community Service Learning, 1*, 77-85.

Kenny, M. E., & Gallagher, L. A. (2002). Service learning: A history of systems. In Kenny, M. E., Simon, L. A. K., Kiley-Brabeck, K., & Lerner, R. M. (eds), *Learning to Serve: Promoting Civil Society Through Service Learning*. Boston, MA: Kluwer Academic Publishers.

Kolb, D. A. (1984). *Experiential Learning: Experience as the Source of Learning and Development*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Loewenson, K. M., & Hunt, R. J. (2011). Transforming attitudes of nursing students: Evaluating a service-learning experience. *Journal of Nursing Education, 50*, 345-349.

Long, J. A., Lee, R. S., Federico, S., Battaglia, C., Wong, S., & Earnest, M. (2011). Developing leadership and advocacy skills in medical students through service learning. *Journal of Public Health Management & Practice, 17*, 369-372.

Madsen, S. R., & Turnbull, O. (2006). The academic service learning experiences of students in a compensation and benefits course. *Journal of Management Education, 30*, 724-742.

Mobley, C. (2007). Breaking ground: Engaging undergraduates in social change through service learning. *Teaching Sociology, 35*, 125-137.

Ottenritter, N. W. (2004). Service learning, social justice, and campus health. *Journal of American College Health, 52*, 189-191.

Parker-Gwin, R., & Mabry, J. B. (1998). Service learning as pedagogy and civic education: Comparing outcomes for three models. *Teaching Sociology, 26*, 276-291.

Peterson, B. A., & Yockey, J. (2006). Service-learning projects: Meeting community needs. *Home Health Care Management & Practice, 18*, 315-322.

Potter, M. A., & Eggleston, M. M. (2003). Supporting academic public health practice: A survey of organizational structures in public health schools. *Journal of Public Health Management & Practice, 9*, 165-170.

Reeb, R. N. (2010). Service-learning in community action research: Introduction to the special section. *American Journal of Community Psychology, 46*, 413-417.

Riegelman, R. K., Teitelbaum, J. B., & Persily, N. A. (2002). Public health degrees—not just for graduate students. *Public Health Reports, 117*, 485-488.

Swallow, V., Newton, J., & Van Lottum, C. (2003). Research in brief: How to manage and display qualitative data using 'Framework' and Microsoft Excel. *Journal of Clinical Nursing, 12,* 610-612.

Toboada, A. (2011). Privilege, power, and public health programs: a student perspective on deconstructing institutional racism in community service learning. *Journal of Public Health Management Practice, 17*, 376-380.

Villanueva, A. M., Hovinga, M. E., & Cass, J. L. (2011). Master of public health community-based practicum: Students' and preceptors' experiences. *Journal of Public Health Management & Practice, 17*, 337-343.

Journal of the Scholarship of Teaching and Learning, Vol. 13, No. 2, May 2013, pp. 101 – 115.

Demystifying instructional innovation: The case of teaching with case studies

Lina D. Kantar[1]

Abstract: Issues emerging from instructional innovation are inevitable, yet basing any curriculum shift on a theoretical framework is paramount. This paper grounds the case-based pedagogy in three learning theories: behaviorism, cognitivism, and constructivism. The three theories are described and situated in relation to the case study method. An in-depth exploration of the assumptions of each theory helped to identify and analyze several issues that emerged upon the implementation of the case-based pedagogy in a nursing curriculum. In line with the three pedagogical standpoints, and after an extensive literature review, measures are proposed to improve the quality of student learning in a case-based curriculum, and principles are derived to support educators in their teaching with case studies. The application of the three learning theories may be especially useful to educators and instructional leaders when shifting paradigms. By describing the key challenges that educators may face with instructional innovation and the usefulness of the cognitivist, behaviorist, and constructivist perspectives in providing explanations and recommendations, provides a beginning research base for improving pedagogies. Failure to employ theories of learning in similar educational shifts may hinder the progress of any intended curriculum transformation.

Keywords: case studies, CBL, learning theories, nursing

I. Introduction.

How learning occurs remains a perplexing issue for educators and facilitators of learning. Much has been written about learning theories that serve to describe the intricate learning process in all its subtleties. Amid a variety of learning models, behaviorism, cognitivism, and constructivism have gained wide spread interest in higher education (Cicciarelli, 2007; Hemming, 2012; Warin, Kolski, & Sagar, 2011). These models provide the basis for understanding learning behaviors and for designing instruction (Yilmaz, 2011). The intent of this paper is to explore the challenges that educators may face when teaching with cases, to develop theoretical understandings regarding this instructional strategy, and to propose counteractive measures based in relevant theories of learning.

The Case Study Experience. The faculty at a college of nursing in the United Arab Emirates embarked on a curriculum transformation exercise; a shift from the traditional curriculum to the case-based pedagogy. The idea of changing the curriculum was not met with a great deal of enthusiasm by most of the faculty and students simultaneously. I was among the faculty members who felt insecure regarding the demands of the new curriculum and all that it entailed.

[1] American University of Beirut, la67@aub.edu.lb

The student population was quite diverse. One thing they had in common was that they all came from traditional secondary backgrounds with ages ranging between 18 and 28. I was assigned to teach the Adult Health Nursing course for 25 students using case studies. The course needed four hours of theoretical work on a weekly basis over a semester, and was designed to introduce students to the prevalent medical-surgical conditions in the country. The case studies were real-life situations that focused on the analytical approach to learning rather than problem-solving, and were sequenced in the course based on complexity. Students enrolled in the course were expected to complete assignments and prepare tasks prior to the session.

Although the course's educational outcomes were intended to enhance self-directed learning, problem-solving, engagement, and discussion teaching, I felt I was not achieving any. It was three weeks into the beginning of the course and every time I had to check on student preparation, few would have done the required readings. Most of the students had difficulty answering higher order thinking questions, and the majority manifested weaknesses in transferring knowledge to new situations.

Group work was not improving. Although students were divided into groups, they lacked team learning skills and synergy. That was not the worse part; discussion was a complete failure. There were a few students who kept disrupting the discussion, thus deterring the instructional plan and course progress. Given the above challenges, and since the overarching goal of teaching with cases is developing learners' professional attributes for the real world of practice, it is deemed essential to ground any change in teaching practice in appropriate theories of learning.

Teaching and Learning with Cases. Recent calls for transforming higher education aim at implementing radical shifts in curriculum, from top down to bottom up, and from traditional preparation of graduates to the development of self-directed learners (Benner, Sutphen, Leonard, & Day, 2010). Teaching with cases has gained wide spread interest in higher education, and more so in nursing, since the basic tenet of this pedagogy is contextualizing knowledge (Benner et al., 2010; Sankar, Varma, & Raju, 2008). Upon future encounter with like-situations, information can be retrieved for immediate application (Costa, Rensburg, & Rushton, 2007).

The case study educational format emulates the nursing practice environment, thus enables students to think like nurses (Tanner, 2009). Thinking like a nurse may be developed in the classroom by construing real-life situations (Tanner, 2009), or in the clinical laboratory by using simulation (Lasater, 2007). The case method strategy provides the appropriate medium for discussion and for solving problems that students may encounter in practice (Delpier, 2006; Walker, 2009).

Discussion teaching promotes deep learning (Ramaekers, Keulen, Kremer, Pilot, & Beukelen, 2011; Walker, 2009); that is, by analysis and reasoning, learners can be directed to the core of the learning process. From the cognitivist perspective, teaching with cases enhances internalization and processing of knowledge (Henry, 2006), whereas the constructivists believe that engagement of learners in constructing their own meanings helps to promote higher order thinking (Sankar et al., 2008). Accordingly, the quality and nature of discussion influence learning (Barnes, Christensen, & Hansen, 1994).

The overarching goal of case-based teaching is developing learners' higher order thinking dispositions such as conceptualizing the significance of the data, interpreting the information, and creating ideas. Through discussion, learners are prompted to find solutions and determine the best means of implementing the solutions. While learners voice their thinking, teachers can assess learners' thinking processes (Facione, Facione, & Giancarlo, 1997). In this regard, assessment may be geared toward capturing learner's "orderliness" in working with the problem,

"diligence" in searching for relevant data, "reasonableness" in the selection of the actions, "persistence" through encountered difficulties, and "precision" in implementing the actions (Facione et al., 1997). Therefore, it is essential that discussion be structured and skillfully led.
Real-life situations provide the context for learning. The format used to structure the situations in this case study experience is known as 'ground breaking', which, according to Harling and Akridge (1998), such type of cases have an exploratory nature. The focus is analysis rather than problem-solving. Each case study introduces the content to be learned using behavioral terms. Integrated in each case are fundamental concepts such as pathophysiology, pharmacology, professional practice, and social behavior.

To promote learning with cases, students must be responsible and self-directed and must value cooperation and collaboration in learning (Barnes et al., 1994). Students develop these attributes as soon as they take part in the educational process. In their search, students either work individually or in groups.

The effectiveness of case-based teaching is highly contingent on the educator's knowledge and skills in leading discussions and fostering engagement (Barnes et al., 1994). Apart from drafting the cases, educators decide on groups, determine learning activities, monitor group interaction, guide and lead the discussion, and assess and evaluate the learning process. Without seeking insights into issues and concerns of case-based learning, application and effectiveness of the approach will remain unclear (Lauver, West, Campbell, Herrold, & Wood, 2009; Tanner, 2009), and curriculum transformation will be an ordeal.

II. Learning Theories to Guide Instructional Innovation.

Given the concerns discussed earlier in this paper, it is imperative that educators base any shift in teaching practice on a theoretical framework (Chikotas, 2008; Yilmaz, 2011). At present, unfolding cases is touted as a potential strategy to prepare nurses for practice (Benner et al., 2010; Tanner, 2009). Educators often fail to base pedagogical shifts on learning models (Warin et al., 2011), even though models could serve as a guide to conceptualize new roles and responsibilities (Yilmaz, 2011). But which learning theories best describe and support how teachers should teach and learners learn using case studies? How can the applied learning theories help teachers and learners understand the learning process?

Amid the various learning theories, behaviorism, cognitivism, and constructivism have helped to unfold the processes of case method teaching and to predict the impact on teachers and learners. Without understanding the case-based process from the behaviorist, cognitivist, and constructivist perspectives, it would be difficult to analyze the pedagogy's challenges and would be impossible to identify effective measures. To set the framework for the exploration process, deemed essential in this paper, the key features characterizing behaviorism, cognitivism, and constructivism are compared (see Table 1).

Table 1. Behaviorism, Cognitivism, and Constructivism Compared.

Concepts	Behaviorism	Cognitivism	Constructivism
Assumptions	Observable behaviors are indicators for learning	Emphasis on role of mental processes	How learning should happen
Principles	Thorndike's laws of learning behavior; Skinner's operant conditioning for acquiring new behavioral pattern	Gagne's nine instructional events; Piaget's developmental nature of reasoning; Bruner's insights and prior knowledge	Bruner's discovery learning; Piaget's assimilation and accommodation principles of constructing new knowledge
Inputs	Environmental events serve as discriminative stimuli; cueing appropriate behaviors	Processing information in a way that is purposeful to solve problems and form answers	Real-life situations; knowledge resources to influence problem solving; self-regulation
Modes of knowledge acquisition	Connection of the three components of learning: Discriminative stimulus, response, and reinforcing stimulus S^{D} – R – $S^{reinf.}$	Information processing; long term memory; concept linkage; schema formation	Construing situations; constructing own knowledge; portraying useful meanings
Role of learner	Active in the environment; consequences of behavior affect probability of reoccurrence	Perceiving the information; interpreting knowledge in relation to prior ones; reorganizing information into new insights or understanding	Active seeker of knowledge; center of attention; constructor of viable knowledge
Role of teacher	Knowledge expert; designer of behavioral objectives; determinant of contingencies of reinforcement	Organizer of meaningful experiences; activation of learner's mental states; guidance; feedback; assessment	Facilitator learning; an advocate of self-regulated and life-long learning; discussion leader
Learning	Absorber of transmitted knowledge; knowledge is constructed by teacher	Learning is sequential and placed in context	Making meaning of the knowledge; knowledge transfer to solve real-world problems
Nature of educational experience	Dominated by the teacher; learner is conditioned under teacher's expertise	Bridging the gap between what learners know and what they need to know	Engagement; cooperation; exchange of expertise or experiences ; self-direction

Behaviorism in Case-Based Learning. Behaviorism is believed to cause a change in teaching and learning behavior. Behaviorists believe that learning is contingent on what goes on in the environment and on the association between a stimulus and a response. The relation between behavioral change and the environment can be predicted if the expected behavior of learners and the situation in which learning takes place are defined (Gredler, 2005; Jackson, 1996). Because behaviorism focuses on modifying learning behaviors (Warin et al., 2011), and since this paper aims at improving teaching and learning with cases, Thorndike's laws and Skinner's principles of learning are explored and assumptions situated in relation to this pedagogy.

Transforming the Classroom through Thorndike's Laws. Thorndike proposed three laws of learning: readiness, exercise, and transfer (Schunk, 2004). The law of readiness reflects the person's ability and willingness to perform the task. In the case method approach, the readiness disposition is essential to garner the effectiveness of the shift. Therefore, better preparation of educators and learners is prerogative to commencing with this strategy. For educators, familiarity with the new pedagogy may be enhanced through professional development initiatives, during which the skills of discussion teaching are emulated, observed, and taught (Çelik, Çevik, & Haşlaman, 2012).

As for students, abandoning their traditional learning behaviors poses the chief reason for their non-compliance with the requirements of the new pedagogy. Students must be adequately prepared on the principles of case-based learning, as well as the dispositions that garner their success in the learning process. Since learning is developmental in nature, an intended change in learners' behaviors would require the joint efforts of both educators and learners. Learning happens only in terms of what is observed. Educators can assess whether learning is taking place or not by evaluating the nature of student participation in the discussion.

Thorndike's laws of exercise and transfer constitute the practice aspect of learning. According to his theory, to learn a behavior, students must repeat the task until mastery. Transferring learning to new situations is contingent on how similar the new tasks are to the previously learned ones. Although Thorndike's law of transfer is not compatible with the anchored instructional approach of case-based learning, exercise in how students must prepare, discuss, and engage in the case pedagogy would help them embrace the change.

Transforming the Classroom through Skinner's Learning Principles. The overarching principle of behaviorism is the learning behavior. Continuing on with Thorndike's laws of learning, Skinner introduced his principles of operant conditioning and programmed instruction. The principles provide a framework on which to base the change in learning behaviors (Wang, 2012).

In this case study experience, the students presented with problems regarding preparation for the session, group work, and progression of the discussion. Through Skinner's parameters of shaping, reinforcement, and contingency, learners may acquire complex behaviors. The role of educators is vital in this behaviorist approach, yet complex. For example, it is known that reinforcement strengthens desired behaviors; therefore, it is the responsibility of the educator to identify potential reinforcing agents.

Skinner's principles of extinction, generalization, discrimination, and self-regulation, could also contribute to modifying learners' behaviors. For example, the principle of extinction best applies to those students in the case study scenario who demonstrated tardiness in preparing for the session or to those who sought to disrupt classroom discussion. Whereas generalization, discrimination, and self-regulation are appropriate for students with non-specific difficulties,

such as difficulties that directly relate to the shift. An example of shift-related issues is student acquisition of new roles and responsibilities of the case-based learning pedagogy. The instructional application of Skinner's learning principles is described in the following section.

Extinction. According to Skinner, the strength of a behavior can be eliminated or reduced by reinforcement (Schunk, 2004). How would extinction help change the behavior of learners who come to class unprepared? For these learners, it is obvious that preparation for the session is perceived a low-priority activity. At first place, extinction of this behavior requires the identification of effective reinforcers, which, if identified, they can be coupled with any low-priority activity. Eventually, learners will be prompted to complete the requirements.

In the case-based scenario, extinction may be applied to learners who demonstrate tardiness in preparing the assigned tasks by giving them the choice to select between three types of learning activities: watching pertinent electronic-based instructional material, reading assigned material, or going to the simulation laboratory. The educator observes what learners will do with these choices, and orders them in priority of preference. Eventually, the learner's most valued activities will be paired with the perceived low-priority task. Although arduous, educators may employ extinction to modify learning behaviors.

Generalization. Skinner's principle of generalization has substantial benefits to the students. With this principle, replication of learned behaviors may be maintained with repeated reinforcement (Schunk, 2004). Every case study has specific content to be covered; yet learning behaviors, such as preparing the assigned tasks, participating in the discussion, and cooperating in groups, are similar. With repeated reinforcement, there is a high tendency that learners generalize the case-based learning behaviors to new cases and courses across the curriculum (Skinner & Daly, 2010).

Discrimination. The Adult Health Nursing course examined in this paper has seven case studies of varying complexity, focus, and purpose. Through this course, case-based learning skills can be generalized to other courses in the curriculum. However, the situation is different with discrimination, since the emphasis of this principle is on enabling students to cue learning behaviors in terms of appropriateness to the situation. The discriminative stimulus, inherent in each case study, develops from the case purpose, outcomes, content, and context. Eventually, the stimulus will identify behaviors that are needed for learning to take place.

Self-regulation. From a reinforcement perspective, students regulate their behavior by carefully assessing their learning needs, determining the discriminative stimulus, evaluating the outcomes, and reinforcing self. In the case study experience, students experienced a sudden shift in their learning; from traditional receivers of knowledge to regular observers of their own learning. The tradeoff for using this approach is one of reducing rote learning for self-regulation. Self-regulation is a major principle of case-based learning; its skills include: self-assessment, self-instruction, and self-reinforcement (Loyens, Rikers, & Schmidt, 2007). The acquisition of these skills is the responsibility of educators who must exercise every effort to render the shift in learners' roles and responsibilities a gradual one, not abrupt.

Self-regulation inherently links to Weiner's attribution theory. Attribution focuses on how individuals interpret events (Demetriou, 2011). Events may have the power to motivate prospective learning. In this respect, Weiner proposed four factors, "ability, effort, task difficulty, and luck" (Demetriou, 2011, p. 16), which, if integrated in case-based teaching, will promote self-regulation in learners.

A general note that applies to Weiner's assumptions is that the integration of case studies in the nursing curriculum has underscored many key features essential to move learners toward a

case-based learning paradigm. Initially, educators must design the tasks by taking into account the ability of learners. Failure to do so may lead to an undesired state of complexity resulting in frustration of educators and learners, and a state of incompatibility between learners and the nature of the tasks. Securing a match will allow the enhancement of self-concept and the disclosure of fear related to the approach.

Another factor that deserves consideration on part of the educator is assessing the magnitude of effort that learners need to prepare the tasks. Instead of frustrating learners with difficult tasks, educators may need to revisit the structure of each case and the purpose of each task. A good case study is constructed to provide learners with some helpful pedagogical goals such as knowing how to set their goals, develop action plans, monitor achievement, and derive meaningful attribution (Poulton, 2009). As for Weiner's difficulty factor, learners may be provided with various approaches that facilitate the accomplishment of the assigned tasks.

Observing learning behaviors remains the chief focus of behaviorism. Although most behaviorists incorporate mental processes in learning, they do not elaborate much on these processes. Cognitive learning theory helps to explain what behaviorism has failed to achieve so far. However, both theories converge on behavioral changes which are observable in behaviorism, yet implicit in cognitivism (Huitt, 2006).

Cognitivism in Case-Based Learning. The focus of teaching with case studies is developing the cognitive abilities of learners (Yilmaz, 2011). Case-based learning helps to prepare nurses to be adept at making sound clinical judgment (Tanner, 2009). Cognitive learning capitalizes on learner's abilities to attain highest level of thinking, with less emphasis on factors within the environment to influence behavior (Huitt, 2006). Learning activates internal mental processes, which include: perception, rehearsal, problem-solving, memory, imaging, and processing and structuring of knowledge (Schunk, 2004; Warin et al., 2011).

Case method teaching has advantages over traditional teaching; it promotes theoretical understanding and develops insights (Loghmani, Bayliss, Strunk, & Altenburger, 2011). Building of new knowledge is contingent on classroom discussion and contextual relevance of learning. Case studies provide special teaching environments in which content can be linked to cognitive processes of perception, interpretation, and information processing. How learning situations are presented to learners will in due course affect learners' attempts to perceive, interpret, and store the information. Therefore, the case studies must be truly engaging and have all factual data adequate for analysis and reasoning (Poulton, 2009; Walker, 2009).

According to Bruner, cognitive development takes place in three stages: "enactive, iconic, and symbolic representation" (Gredler, 2005, p. 76). These stages represent the mechanisms by which information is processed using one's mental capabilities and intellectual and psychomotor efforts (Jackson, 1996). Along with the information processing theory, the construct of cognitive learning in case-based teaching is explained through Gagne's perspective of learning and Piaget's developmental epistemology.

Information Processing Theory Applied. There are a plethora of cognitive paradigms used to describe how learning takes place, yet information processing is recognized as the backbone of cognitive science that explains mental processes (Huitt, 2006; Warin et al., 2011). Case studies provide an appropriate pedagogy to activate these processes. With this approach, learners are encouraged to retrieve prior knowledge and focus attention on meaningful ones. When all aspects of a real situation are integrated in the case study, learners will, ultimately, make sense of the information. Contextual learning ensures retention and transfer since such

executive functions recommend more indulgence of the learner than when doing a routine learning chore (Huitt, 2003).

The case study approach can serve as the means for integrative teaching (Tanner, 2009). Integration provokes learners to call to mind existing information. To optimize the number of concepts that can be processed by each case study, educators may opt to incorporate learning resources and instructional material that are deemed essential to create the robust for cognitive learning.

The students in the case study scenario demonstrated difficulty in processing the information. The difficulty was diagnosed from the quality of answers to higher order thinking questions and the inability of learners to transfer knowledge to new situations. From the cognitivist perspective, cognitive states and processes may be developed and explored through instruction (Alutu, 2006). Therefore, educators may assume a vital role in helping learners perceive the importance of the information, encode it, and connect it to existing knowledge. The basic principles of information processing lie in the work of Robert Gagne' (Gredler, 2005).

In his theory, Gagne' viewed human learning as developmental, generalizable, and contingent on instructional effectiveness (Gredler, 2005). Accordingly, he proposed nine instructional principles that help students process the information: (1) learners' attention, (2) learning outcomes, (3) prior knowledge, (4) content, (5) guidance, (6) performance, (7) feedback, (8) assessment, and (9) diversity in practice. From a practical perspective, the principles guide educators in how to start the session, engage learners, link new information with prior knowledge, guide the organization of information, check on ability to demonstrate understanding, provide prompts, and create triggers to check on transference of knowledge. These principles resonate with the general aims of teaching with cases: (a) knowledge acquisition, (b) development of psychomotor and affective skills, (c) transfer of knowledge to new settings (Çelik et al., 2012), and (d) deep learning (Walker, 2009). Once again, it is the responsibility of the educator to convert case studies into meaningful learning experiences.

Piaget's Developmental Epistemology Applied. Piaget's cognitivism focuses on the developmental nature of reasoning governed by a variety of factors, with social factors having a major impact on learning. According to Piaget, interactive teaching enhances student inquiry about previously accepted assumptions and promotes the development of problem-solving and life-long learning skills (Young & Paterson, 2007). His theory sheds light on the role of interactive teaching in developing cognitive abilities. In the course of pursuing information about the educator's experience with case-based teaching in this paper, many challenges surfaced; most of these challenges reflect the effectiveness of dialogue and the teacher student interaction.

Since interaction enhances the exchange of cognitive processes, educators need to focus efforts on improving dialogue between and among various partners in the educational process, and on creating an educational milieu. However, effective dialogue, known as reciprocal teaching, can be promoted through the incorporation of four cognitive-based techniques: summarization, question generation, explication, and prediction (Yilmaz, 2011). Given the discussion difficulties endorsed earlier in the case study pedagogy, educators are prompted to consider these techniques in the discussion process.

Constructivism in Case-Based Learning. Case-based learning has its roots in constructivism (Hartfield, 2010); a learning theory that capitalizes on learner's abilities to construct viable knowledge and an education paradigm that fosters discovery learning (Hartfield, 2010). As advanced by Jerome Bruner, discovery learning is perceived "a necessary condition for learning the variety of techniques of problem-solving, of transferring information for better

use …" (1961, p. 60). Effective discussion and reflection create a state of dissonance in the learner, thus stimulating higher order thinking (Hmelo-Silver, Duncan, & Chinn, 2007). Through reflection on prior knowledge and exposure to new information, case studies provide the means for scaffolding learning and for building a repertoire of nursing knowledge prior to commencing with practice (Delpier, 2006; Walker, 2009). However, knowledge making primarily depends on efforts of learners to generate meaningful ideas (Nikitina, 2010).

Nursing students must be prepared to solve problems of the workplace (Institute of Medicine, 2010). Learning with case studies trains learners how to question assumptions, explore options, base new knowledge on context, and consider the constant transformation of knowledge. These skills represent the rationale for using real-life situations in the program.

Although constructivism is a learning theory, it is considered a mode to improve instruction (Clark, 2000). Stemming from the notion that engaging learners in the meaning-making process is crucial for constructing viable knowledge, constructivists base learning on prior knowledge, which transforms at a later stage into dynamic, concrete, and lasting knowledge (Young & Paterson, 2007).

How does constructivism relate to teaching with case-studies? It ties in numerous ways. Case studies create the tool for active learning; its parameters include: engagement, self-regulation, and motivation (Friedlander et al., 2011). The following section will describe each parameter as it relates to constructivism and the case-based pedagogy.

Engagement Applied. A fundamental curriculum strand intended to enhance student engagement is discussion teaching (Barnes et al., 1994). Although discussion empowers students to generate reality, two principles dominate: prior knowledge to guide learning and social collaboration to enhance learning. The principles are implemented with the use of cooperative learning, self-observation, and reflection. Educators must ensure that the tasks in the case study build on prior knowledge, invoke problem-solving processes, and promote schematic representations of knowledge. Eventually, knowledge acquisition will be enhanced and easily transferred.

Knowledge construction is activated with small-group interaction and discussion tied to the clinical situation (Dzerviniks & Poplavskis, 2012). Interactive discussion best takes place over propitious learning situations (Richards & Inglehart, 2006); situations that induce "discovery learning" (Bruner, 1961). It is expected that case studies invoke the learner's internal processing of information. As a result, knowledge becomes viable when learners actively engage in connecting the processed information with classroom discussion.

The significance of encouraging learners to evaluate the learning process must not be ignored. With effective questioning, learners will be enabled to identify strategies that foster the educational process (Brandon & All, 2010). Therefore, educators must seek every opportunity to promote learner engagement in the case-based paradigm.

Discussion teaching is believed to foster the notion of a learning community created by learners, educators, and subject matter. Through team learning, learners are persuaded to compare their analytic modes, interact with one another, further achievement, and improve attitudes toward learning. Team learning contributes to building team ethos and synergy (Barnes et al., 1994); thus motivating learners when recognizing the value of their input.

Constructivist educators seek to embark on strategies that promote learner engagement, with emphasis on the learning environment and team learning. Since the environment capitalizes on "optimal arousal" (Kiger, 2004), educators are prompted to combine learning with context

(Brandon & All, 2010). A movement like this leans on the role of educators in capturing learners' attention on the intricate aspects of a learning situation.

Self-Regulation Applied. Self-regulated learning has gained widespread interest in higher education (Loyens et al., 2007) as well as general education (Kistner, Rakoczy, Otto, Ewijk, Büttner, & Klieme, 2010). Being a premise in the preparation of life-long learners (Kistner et al., 2010), the means to develop self-regulation skills have captured the attention of educational theorists (Zimmerman, 2008). Among the recognized self-regulation skills are setting goals, deciding on actions to achieve expected learning outcomes, selecting appropriate learning means, and monitoring and evaluating achievement (Kistner et al., 2010). Self-regulation skills can be taught and promoted using direct (implicit and explicit) and indirect instructional methods (Otto, 2010).

Direct implicit teaching of self-regulation skills prompts constructivist educators to model the skills by voicing their thought processes using 'elaboration', 'organization', and 'problem-solving' (Kistner et al., 2010, p. 163), or by engaging learners in an inquiry process. In explicit promotion, educators request from learners to demonstrate self-regulation by engaging them in activities designed for this purpose. For indirect promotion, the emphasis is generating a learning environment, guided by content, tasks, and instructional strategies (Kistner et al., 2010). Therefore, the development of self-regulation skills is contingent on how educators lead discussions and trigger curiosity and interest of learners. Good case studies can contribute to self-regulation.

Case-based learning stands as an approach to promote self-directed learning. Being an expansion of self-regulation (Candy, 1991), the key features of self-directed learning include: autonomy, responsibility, independent inquiry, and self-teaching (Loyens, Magda, & Rikers, 2008). Although all the features influence the learning process, autonomy and responsibility are core attributes. Upon the basis of this analysis, it is essential that educators specify activities or tasks that involve learners in building their knowledge (autonomy in construction) and in expanding this knowledge outside the narrow confines of the classroom (responsibility for contextual relevance).

Conscious and voluntary involvement motivates the learner to participate in the learning process, which in turn activates intrinsic interest and motivation (Sivan, 1986). Constructivism applied to case-based learning commends that discovery and self-regulated learning be introduced to educators and learners prior to commencing with this approach, thus ensuring effectiveness of the shift (Loyens et al., 2007). Along the same vein, students and educators must be aware of their new roles and responsibilities in the case-based pedagogy.

III. Conclusion.

Learning is the conscious engagement of students in situations and experiences that help in changing attitudes and behaviors, developing higher order thinking skills, improving perceptions, and constructing knowledge. Case-based learning is a promising pedagogy for developing the aforementioned attributes. Since not a single theory can represent the teaching learning process in its entirety, teaching with case studies requires a strong command of learning theories.

While case-based learning (CBL) is grounded in behaviorism, cognitivism, and constructivism, educators are expected to teach their courses in accordance with the principles of these theories. The scarcity of literature on philosophical perspectives related to the case study pedagogy necessitated the in-depth exploration of each learning theory as related to the key

challenges arising from the pedagogy. A better understanding of each theory led to the identification of corrective measures and transformative actions; all of which converged on five principles for teaching with cases (see Table 2). Educators are bound to the educational process through their continuous analysis of the process. These principles create the framework to help analyze and assess the effectiveness of the case method pedagogy.

Table 2. Theoretical Underpinnings for and Principles of Case-Based Teaching and Learning.

Emergent Issues	Theoretical Assumptions	Transformation in Perspective	Principles of Case Method Approach
▪ Readiness disposition for both educators and students ▪ Non-compliance with requirements of case-based learning	**Behaviorism** ▪ Thorndike's laws of readiness, exercise, and transference ▪ Skinner's extinction, generalization, and discrimination ▪ Weiner's attribution theory	▪ Educators: challenge assumptions; facilitator role ▪ Students: exploration of abilities; self-regulation role ▪ Linking low-priority activities with reinforcers ▪ Discriminative stimulus ▪ Attribution-based factors: ability, effort, task difficulty, and luck	1. Conceptualize the new roles and responsibilities of educators and students 2. Construct/review case studies for specific pedagogical goals, considering complexity of tasks, learners' capability, and potential for engagement
▪ Transfer of knowledge ▪ Higher order thinking skills ▪ Discussion ▪ Teacher student interaction	**Cognitivism** ▪ Information processing: Gagne's cognitive apprenticeship ▪ Piaget's interactive pedagogy	▪ Prior knowledge ▪ Sequential approach ▪ Group interaction ▪ Building team ethos, synergy, and cooperation	3. Alternate individual learning with team learning 4. Foster the notion of a learning community
▪ Engagement ▪ Responsibility and autonomy for learning ▪ Group dynamics	**Constructivism** ▪ Knowledge generation: prior and social collaboration ▪ Bruner's discovery learning	▪ Cooperative learning, self-observation, and reflection ▪ Small group interaction and discussion ▪ Schematic representation of knowledge	5. Analyze the educational process for effectiveness

References

Alutu, A. (2006). The guidance role of the instructor in the teaching and learning process. Journal of Instructional Psychology, 33(1), 44-49.

Barnes, L.B., Christensen, C.R., & Hansen, A.J. (1994). Teaching and the Case Method (3rd ed.). Boston: Harvard Business School Press.

Benner, P., Sutphen, M., Leonard, V., & Day, L. (2010). Educating Nurses: A Call for Radical Transformation. San Francisco, CA: Jossey-Bass.

Brandon, A.F., & All, A.C. (2010). Constructivism theory analysis and application curricula. Nursing Education Perspectives, 31(2), 89-92.

Bruner, J. (1961). The act of discovery. Harvard Educational Review, 31, 21-32.

Candy, P. (1991). Self-direction for Lifelong Learning. San Francisco, CA: Jossey-Bass.

Çelik, S., Çevik, Y.D., & Haşlaman, T. (2012). Reflections of prospective teachers regarding case-based learning. Turkish Journal of Qualitative Inquiry, 3(4), 68-82.

Chikotas, N.E. (2008). Theoretical links supporting the use of problem-based learning in the education of the nurse practitioner. Nursing Education Perspectives, 29(6), 359-362.

Cicciarelli, M. (2007). Behavioral, cognitive, and humanistic theories: Which theories do online instructors utilize? International Journal of Information and Communication Technology Education, 3(4), 1-12.

Clark, D. (2000). Developing instruction or instructional design. Retrieved from http://www.nwlink.com/~donclark/hrd/learning/development.html

Costa, M.L., Rensburg, L.V., & Rushton, N. (2007). Does teaching style matter? A randomized trial of group discussion versus lectures in orthopedic undergraduate teaching. Medical Education, 41, 214-217.

Delpier, T. (2006). Cases 101: Learning to teach with cases. Nursing Education Perspective, 27(4), 204-209.

Demetriou, C. (2011). The attribution theory of learning and advising students on academic probation. NACADA Journal, 31(2), 16-21.

Dzerviniks, J., & Poplavskis, J. (2012). Acquisition of physics in comprehensive school: Accents of constructivism approach. Problems of Education in the 21st Century, 41, 10-17.

Facione, P., Facione, N., & Giancarlo, C.A. (1997). The Motivation to Think in Working and Learning. Millbrae, CA: The California Academic Press, 1-12.

Friedlander, M.J., Andrews, L., Armstrong, E.G., Aschenbrenner, C., Kass, J.S., Ogden, P., et al. (2011). What can medical education learn from the neurobiology of learning? Academic Medicine, 86(4), 415-420. DOI: 10.1097/ACM.0b013e31820dc197.

Gredler, M.E. (2005). Learning and Instruction: Theory into Practice (5th ed.). Upper Saddle River: Pearson Merrill Prentice Hall.

Harling, K.F., & Akridge, J. (1998). Using the case method of teaching. Agribusiness, 14(1), 1-14.

Hartfield, P.J. (2010). Reinforcing constructivist teaching in advanced level biochemistry through the introduction of case-based learning activities. Journal of Learning Design, 3(3), 20-31.

Hemming, T. (2012). Simulation methodology in nursing education and adult learning theory. Adult Learning, 23(3), 129-137. DOI: 10.1177/1045159512452848.

Henry, P. (2006). Making groups work in the classroom. Nurse Educator, 31, 26-30.

Hmelo-Silver, C., Duncan, R.G., & Chinn, C.A., (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Sweller, Kirschner and Clark. Educational Psychologist, 42, 99-107.

Huitt, W. (2003). The information processing approach to cognition. Educational Psychology Interactive. Valdosta, GA: Valdosta State University. Retrieved from http://chiron.valdosta.edu/whuitt/col/cogsys/infoproc.html

Huitt, W. (2006). The cognitive system. Educational Psychology Interactive. Valdosta, GA: Valdosta State University. Retrieved from http://chiron.valdosta.edu/whuitt/col/cogsys/cogsys.html

Institute of Medicine. (2010). A summary of the February 2010 forum on the future of nursing. IOM, Washington, DC.

Jackson, W.H. (1996). Survey of an adult learner. Retrieved from http://internet.cybermesa.com/bjackson/Papers/Adultlearner.html

Kiger, A.M. (2004). Teaching for Health (3rd ed.). Edinburgh: Churchill Livingstone.

Kistner, S., Rakoczy, K., Otto, B., Ewijk, C., Büttner, G., & Klieme, E. (2010). Promotion of self-regulated learning in classrooms: Investigating frequency, quality, and consequences for student performance. Metacognition Learning, 5, 157-171. DOI 10.1007/s11409-010-9055-3.

Lasater, K. (2007). Clinical judgment development: Using simulation to create an assessment rubric. Journal of Nursing Education, 46(11), 496-503.

Lauver, L.S., West, M.M., Campbell, T.B., Herrold, J., & Wood, G.C. (2009). Toward evidence-based teaching: Evaluating the effectiveness of two teaching strategies in an associate degree nursing program. Teaching and Learning in Nursing, 4, 133-138.

Loghmani, M.T., Bayliss, A.J., Strunk, V., & Altenburger, P. (2011). An integrative, longitudinal case-based learning model as a curriculum strategy to enhance teaching and learning. Journal of Physical Therapy Education, 25(2), 42-50.

Loyens, S.M., Rikers, R., & Schmidt, H.G. (2007). Students' conceptions of distinct constructivist assumptions. European Journal of Psychology of Education, XXII(2), 179-199.

Loyens, S.M., Magda, J., & Rikers, R. (2008). Self-directed learning in problem-based learning and its relationships with self-regulated learning. Educational Psychology Review, 20, 411-427.

Nikitina, L. (2010). Addressing pedagogical dilemmas in a constructivist language learning experience. Journal of Scholarship of Teaching and Learning, 10(2), 90-106.

Otto, B. (2010). How can motivated self-regulated learning be improved? In A. Mourad & J. de la Fuente Arias (Eds.), International Perspectives on Applying Self-regulated Learning in Different Settings. Frankfurt am Main: Lang.

Poulton, M.S. (2009). Undergraduate business ethics pedagogy: Writing constructed narratives. Journal of Business Ethics Education, 6, 93-102.

Ramaekers, S., Keulen, H., Kremer, W., Pilot, A., & Beukelen, P. (2011). Effective teaching in case-based education: Patterns in teacher behavior and their impact on the students' clinical problem-solving and learning. International Journal of Teaching and Learning in Higher Education, 23(3), 303-313.

Richards, P.S., & Inglehart, M.R. (2006). An interdisciplinary approach to case-based teaching: Does it create patient-centered and culturally sensitive providers? Journal of Dental Education, 70(3), 284-291.

Sankar, C.S., Varma, V., & Raju, P.K. (2008). Use of case studies in engineering education: Assessment of changes of cognitive skills. Journal of Professional Issues in Engineering Education and Practice, 287-296. DOI: 10.1061/(ASCE) 1052-3928(2008) 134:3(287).

Schunk, D.H. (2004). Learning Theories: An Educational Perspective (4th ed.). Upper Saddle River: Pearson Merrill Prentice Hall.

Sivan, E. (1986). Motivation in social constructivist psychology. Educational Psychologist, 21(3), 209-233.

Skinner, C.H., & Daly III, E.J. (2010). Improving generalization of academic skills: Commentary on the special issue. Journal of Behavioral Education, 19, 106-115. DOI: 10.1007/s10864-010-9100-y

Tanner, C.A. (2009). The case for cases: A pedagogy for developing habits of thought. Journal of Nursing Education, 48(6), 299-300.

Walker, C. (2009). Teaching policy theory and its application to practice using long structured case studies: An approach that deeply engages undergraduate students. International Journal of Teaching and Learning in Higher Education, 20(2), 214-225.

Wang, V.C.X. (2012). Understanding and promoting learning theories. International Forum of Teaching and Studies, 8(2), 5-11.

Warin, B., Kolski, C., & Sagar, M. (2011). Framework for the evolution of acquiring knowledge modules to integrate the acquisition of high-level cognitive skills and professional competencies: Principles and case studies. Computers and Education, 57, 1595-1614. DOI: 10.1016/j.compedu.2011.02.013.

Yilmaz, K. (2011). The cognitive perspective on learning: Its theoretical underpinnings and implications for classroom practices. The Clearing House, 84(5), 204-212. DOI: 10.1080/00098655.2011.568989

Young, L.E., & Paterson, B.L. (2007). Teaching Nursing: Developing a Student-centered Learning Environment. Philadelphia: Lippincott Williams & Wilkin

Zimmerman, B.J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. American Educational Research Journal, 45(1), 166-183.

Journal of the Scholarship of Teaching and Learning, Vol. 13, No. 2, May 2013, pp. 116 – 132.

Family context predictors of math self-concept among undergraduate STEM majors: An analysis of gender differences

Anne N. Rinn[1], Kathi Miner[2] and Aaron B. Taylor[3]

Abstract: The purpose of the current study was to examine four family context variables (socioeconomic status, mother's level of education, father's level of education, and perceived family social support) as predictors of math self-concept among undergraduate STEM majors to better understand the gender differential in math self-concept. Participants included 499 undergraduates (75% of whom were female) at a large research university in the southwestern United States. Results indicated that males had higher math self-concepts than females and that social support predicted math self-concept, particularly for males.

Keywords: STEM, math self-concept, college students, undergraduate major, gender

I. Introduction.

There has been great concern, both from academics and the general public, about the small number of women in the fields of science, technology, engineering, and mathematics (STEM). Despite an increase in the percentage of women earning STEM degrees, the number is still relatively small (National Science Foundation, 2007); and women who initially decide to pursue degrees in STEM fields leave them at a significantly higher rate than men. This higher rate of leaving STEM fields is especially seen among women of color (Kohlstedt, 2004). Even intellectually advanced women drop out of STEM fields at a higher rate than intellectually advanced men (Johnsen & Kendrick, 2005).

One possibility for the low number of women in STEM is that females consistently report lower perceptions than males of their abilities in math, or lower math self-concepts, across all age and grade levels, often grossly under representing their actual ability levels (Else-Quest, Hyde, & Linn, 2010; Furnham, 2001; Furnham, Hosoe, & Tang, 2002; Heller & Ziegler, 1996; Juang & Silbereisen, 2002; Marsh & Yeung, 1998; Meece & Jones, 1996; Pajares, 1996; Sax, 1994; Skaalvik & Skaalvik, 2004; Steinmayr & Spinathm, 2009; Watt, 2005).

This gender difference in math self-concept may lead to differences in later math achievement, which may in turn lead to differences in STEM undergraduate degree attainment, STEM graduate degree attainment, and STEM career aspirations. Indeed, math self-concept is a known predictor of persistence in math courses, performance on standardized tests measuring math ability, achievement in math, and advanced math coursework selection (Lee, 2009; Marsh & Yeung, 1998). All of these differences could easily affect STEM-related outcomes in undergraduate or graduate study.

[1] Anne N. Rinn, Department of Educational Psychology, University of North Texas, 316 Matthews Hall, 1155 Union Circle #311335, Denton, TX 76203. E-mail: anne.rinn@unt.edu Research for this study was supported by Grant #0936780 from the National Science Foundation.
[2] Department of Psychology and Women's and Gender Studies Program, Texas A&M University
[3] Department of Psychology, Texas A&M University

Numerous educational and institutional factors contribute to the educational experiences of women majoring in STEM fields in college. These factors likely include such things as having female role models (Lockwood, 2006; Sonnert, Fox, & Adkins, 2007) and the overall educational climate (Cabrera, Nora, Terenzini, Pascarella, & Hagedorn, 1999; Eimers & Pike, 1997; Graham & Gisi, 2000). Researchers have typically examined educational and institutional factors such as these when studying predictors of persistence in the STEM fields, such as a high math self-concept. Few research studies, though, have considered the influence of one's family in predicting STEM degree persistence and attainment, and research has not adequately addressed family context predictors of math self-concept. As family context variables greatly influence an individual's educational opportunities, and families are among the first to expose children to gender stereotypes and gender roles, the purpose of the present study is to examine multiple family context variables as predictors of math self-concept among male and female undergraduate STEM majors. Economic resources, parent education levels, and parental social support, in particular, have the potential to influence student's experiences in higher education (Marable, 2003; Perna & Titus, 2004) and are therefore examined in the present study.

Math Self-Concept. Shavelson, Hubner, and Stanton (1976) were among the first to assert that self-concept is both multifaceted and hierarchical. Self-concept is multifaceted "in that people categorize the vast amount of information they have about themselves and relate these categories to one another" and is hierarchically arranged "with perceptions of behavior at the base moving to inferences about self in sub areas (e.g., academic-English, science, history, math), then to inferences about self in general" (Marsh & Shavelson, 1985, p. 107). Thus, self-concept can be defined as "a person's perceptions of him- or herself … formed through experience with and interpretations of one's environment" (Marsh & Shavelson, p. 107). Math self-concept, then, pertains to one's perceptions of his or her ability related to math.

Self-concept becomes more differentiated with age. As individuals get older, they are better able to determine their strengths and weaknesses in various areas (e.g., math, English, social skills). Math self-concept is a facet of self-concept that typically declines throughout the college years for both males and females (Astin, 1993; Pascarella & Terenzini, 2005; Sax, 1994). The rate of decline is not equal, though. Gender differences in STEM related self-concept appear to be larger among high school and college students than among younger students (Hyde, Fennema, Ryan, Frost, & Hopp, 1990). Hence, while both males and females experience a decrease in math self-concept throughout college, females experience that decline at a greater rate. In fact, women undergraduates may enter a STEM major relatively confident in their STEM-related abilities, but will likely experience a significant drop in confidence during their first year of STEM coursework (Brainard & Carlin, 2001; Seymour, 1995), larger than what is experienced by their male STEM major classmates.

Many researchers have also suggested academic self-concept and other academic self-beliefs are important factors in choosing a career path (e.g., Bandura, 1986; Farmer, 1985; Gottfredson, 2002; Rottinghaus, Lindley, Green, & Borgen, 2002), and the direct relationship between academic self-concept and aspirations is supported by several meta-analyses (e.g., Lent, Brown, & Hackett, 1994; Multon, Brown, & Lent, 1991). This relationship is likely to occur because "people evaluate themselves in terms of their ability to be successful in certain occupational roles" (Quilter, 1995, p. 40).

In particular, a positive relationship between math self-concept and math achievement has been documented extensively (Watt, 2005). Students are generally more likely to select STEM related courses and careers when they have high math, science, and/or technology self-

concepts (Jacobs, 2005) and are more likely to persist with STEM related aspirations when they have high math self-beliefs (Mau, 2003) and science self-beliefs (Larose, Ratelle, Guay, Senécal, & Harvey, 2006). Gainor and Lent (1998) found math self-beliefs predicted math interests, which predicted math career choice intentions among a sample of undergraduate students. Byars-Winston, Estrada, Howard, Davis, and Zalapa (2010) found similar results with a sample of racial minority undergraduate students.

Because of the importance of math self-concept in predicting later achievement related outcomes, and because of the known discrepancy between males and females with regard to math self-concept, it is critical that research examine factors that may contribute to decreases in math self-concept. If low math self-concept is a factor that prevents women from pursuing a degree in a STEM field, finding ways to address this issue can aid researchers and administrators in directly addressing a cause for women's attrition from STEM fields and begin to create educational environments where women can intellectually thrive and progress toward their educational and career goals. In so doing, women will be more apt to continue in STEM, strive to reach the highest levels in their field, and ultimately become role models for other women. Retaining women in STEM via positive educational climates will also equip women with the skills and knowledge to compete in the global economy with its increasing focus on technological advancements and scientific complexity.

For these reasons, then, the examination of potential predictors of math self-concept is a necessity. The following section will examine what research has already shown with regard to family context variables and math self-concept before discussing the current study.

Family Context Variables and Math Self-Concept. While researchers have already documented several known predictors of math self-concept, such as math achievement and the number of math and science courses taken (Sax, 1994), several family context variables also have the potential to impact an individual's math self-concept, including socioeconomic status, parents' level of education, and family social support.

Socioeconomic Status. The term *socioeconomic status* (SES) is used by researchers to designate a family's rank in an economic hierarchy and is usually measured by some combination of parents' level of education, family income, and parents' occupational prestige (Sirin, 2005). A family with a low SES, for example, might be characterized by parents who did not attend college, a limited amount of monthly income, and blue-collar or hourly wage positions. A family with high SES, then, would be characterized by college educated parents, disposable income, and white-collar positions or professional employment.

Students from lower SES backgrounds show disadvantages in educational settings. For example, at the elementary and secondary school levels, a medium to strong correlation exists between SES and academic achievement (Sirin, 2005). In addition, students from low SES backgrounds, and in particular those whose parents did not attend college, are far less likely to score at an average or above average level on the SAT, as compared to those students from higher SES backgrounds and whose parents graduated from college (Bowen, Kurzweil, & Tobin, 2005). Further, only about one-fourth of academically qualified students from low SES backgrounds even attend college (Marable, 2003). Students from low SES backgrounds are less likely to persist to a bachelor's degree and have graduate degree aspirations than students from higher SES backgrounds (Pascarella & Terenzini, 1991; Walpole, 2003).

Family SES likely influences one's math self-concept, either directly, or indirectly with academic achievement as a mediating variable. For example, SES has been shown to have a direct effect on math and science test scores (Yavuz, 2009), and because the relationship between

achievement and self-concept is reciprocal (Hamachek, 1995; House, 2000), SES likely indirectly affects academic and math self-concept. The direct relationship between SES and math self-concept is also supported. In a study examining math self-efficacy (a construct conceptually related to math self-concept) among adolescents across 33 countries, SES had a direct positive effect on adolescents' math self-efficacy in 19 of the countries (Williams & Williams, 2010). In addition, middle and high SES parents are more likely to have higher educational aspirations for their children than low SES parents (Khattri, Riley, & Kane, 1997), which may also impact children's sense of self. Further, research suggests that SES may be an important predictor of math self-concept for white men, but not for white women, black men, or black women (Pascarella, Smart, Ethington, & Nettles, 1987).

It is important to note, though, that there is no "automatic privilege" that comes with being raised in a middle or high SES family (Ozturk & Singh, 2006, p. 31). Parental involvement, through modeling, expectations, and support, likely remains essential for children to benefit from a middle or high SES home.

Parents' Level of Education. Parents' level of education is often included as a control variable in studies of children's achievement, aspirations, and other educational outcome variables. Parents' education levels are known to impact children's test scores, particularly math and science test scores (Yavuz, 2009), as well as their levels of academic achievement (Orr & Dinur, 1995). Among women, in particular, having parents who attended college is associated with entering a high-paying, male-dominated occupation (Pascarella & Terenzini, 2005).

Research findings are somewhat mixed with regard to the effects of parents' level of education on children's academic self-concepts. Fathers' level of education has been shown to have a direct effect on children's perceptions of academic efficacy, or sense of competence (Hortaçsu, 1995). However findings also differ by gender of the parent. Having a highly educated father has been found to have a negative effect on females' math self-concepts and having a highly educated mother has been found to have a negative effect on males' math self-concepts (Sax, 1994).

Family Social Support. Social support can be defined as an individual's perception of "general support or specific supportive behaviors (available or enacted upon) from people in their social network, which enhances their functioning and/or may buffer them from adverse outcomes" (Malecki, Demaray, & Elliott, 2000, pg. 3). Social support systems may include people such as friends, family members, and peers, as well as groups of individuals in specific institutions or circumstances (e.g., schools, classmates, athletic team members). Utilizing social support systems may help college students cope with the diverse academic and social demands they face (Azmitia & Cooper, 2001; Budny & Paul, 2003; Jacobs, 2005; Maccoby & Martin, 1983).

Perceived family support, in particular, has been shown to affect a variety of educational and academic outcomes among college students. For example, research consistently indicates parents are among the most important influences on late adolescents' and young adults' career aspirations and educational persistence (Flores & O'Brien, 2002; Otto, 2000; Whiston & Keller, 2004). Multiple studies have also reported that supportive environments in general lead to persistence in STEM fields (Bonous-Hammarth, 2000; Grandy, 1998).

Especially relevant to the present study, Ahmed, Minnaert, Werf, and Kuyper (2010) found perceived social support (in this case, from parents, peers, and teachers) influenced math achievement among early adolescents, as mediated by a combination of competence beliefs (i.e., math self-concept), value of the subject, anxiety, and enjoyment. In other words, perceived social

support influenced math self-concept (and other variables), which in turn influenced one's math achievement. Among a sample of Mexican American adolescents, perceived parent social support significantly predicted math and science self-efficacy for both males and females (Navarro, Flores, & Worthington, 2010).

II. The Current Study.

The purpose of the current study was to examine family context variables, including SES, mother's level of education, father's level of education, and perceived family social support, as predictors of math self-concept among undergraduate STEM majors while including gender as a moderator. We predict that the family context variables will show direct relationships with math self-concept. Research and theory, however, suggest that gender may also moderate these relationships. Such variables as SES (Pascarella et al., 1987) and father's level of education (Sax, 1994), for example, have been shown to differentially affect females' and males' math self-concepts.

This study is important for both theoretical and practical reasons. Males' and females' math self-concepts have been examined at the elementary and secondary levels, but less so at the university level, and particularly not among declared STEM majors. This study thus adds valuable information to the available research literature. Further, but few research studies have examined the influence of one's family in predicting factors that might lead to STEM degree persistence and attainment. By including family context variables as predictors of math self-concept, the findings from this study might provide insight regarding the causes of lower and steeper declining math self-concept among women. Using this information, researchers, administrators, educators, and parents could work to potentially improve the math self-concepts of women, and men, and to increase retention rates within STEM undergraduate majors.

Specifically, we make the following hypotheses about declared STEM majors:

1) Females will report lower math self-concepts than males.
2) Higher SES will predict higher math self-concept.
3) Mothers' higher level of education will predict higher math self-concept.
4) Fathers' higher level of education will predict higher math self-concept.
5) Higher perceived family social support will predict higher math self-concept.
6) Because we examine gender exploratory, we make no formal hypotheses regarding gender as a moderator of Hypotheses 2-5.

III. Method.

A. Participants.

Participants for this study were recruited from a large, research university located in the southwestern United States. This university serves approximately 40,000 undergraduate students (nearly half female) and is White-dominated (more than 75% White). Of these, approximately 12,000 students are declared STEM majors. Of the 499 students who participated in the current study, 373 were female and 126 were male. Of these, 25 identified as African American/Black, two as American Indian/Alaska Native, 49 as Asian, 107 as Hispanic/Latino, two as Native Hawaiian/Pacific Islander, 298 as White, and 16 as Bi- or Multi-racial. Regarding the students' year in school, 388 were freshmen, 76 were sophomores, 17 were juniors, and 12 were seniors

(with the remaining six labeling themselves as "other" or not responding to this question). Regarding students' mothers' levels of education, 33 indicated their mothers did not graduate from high school, 76 had mothers with a high school diploma, 114 had mothers with some college experience, 184 had mothers who were college graduates, 23 had mothers with some graduate level training, and 69 had mothers with a Master's degree or higher. Regarding students' fathers' levels of education, 32 indicated their fathers did not graduate from high school, 64 had fathers with a high school diploma, 83 had fathers with some college experience, 178 had fathers who were college graduates, 24 had fathers with some graduate level training, and 114 had fathers with a Master's degree or higher. Regarding students' family finance levels, or SES, one student indicated his/her family was "very poor/not enough to get by", 46 students had "barely enough to get by", 173 had "enough to get by, but not many extras", 174 had "more than enough to get by", 87 were "well-to-do", 17 were "extremely well-to-do", and one student did not answer this question. Finally, about half of the students (49.5%) reported taking between seven and nine high school math and science classes.

B. Materials.

Data for this study were gathered as part of a larger study. Only the variables used in the current research are discussed here.

Demographic information. A demographic questionnaire was administered to assess participants' major, age, gender, race, number of science and math courses taken during high school, and other information.

Math self-concept. The Self Description Questionnaire III (SDQ-III) was designed to measure the self-concepts of late adolescents, and was created in response to a need for high quality measurement instruments with a strong theoretical basis that provide support for the multidimensionality of self-concept (Marsh, 1989). The SDQ-III is based on the Shavelson model of self-concept, which assumes that self-concept is both multifaceted and hierarchical, as previously mentioned (Shavelson et al., 1976).

The SDQ-III contains 136 items and measures 13 facets of self-concept (Marsh & O'Neill, 1984). Each facet is measured by 10 to 12 items, for which responses range from 1 (definitely false) to 8 (definitely true). Half of the items are negatively worded (Marsh, 1989). The SDQ-III assesses four areas of academic self-concept (math, verbal, general academic, and problem solving), eight areas of nonacademic self-concept (physical ability, physical appearance, relations with the same sex, relations with the opposite sex, relations with parents, spiritual values/religion, honesty/trustworthiness, and emotional stability), and general self-concept.

For the purposes of this study, we only used data from the math self-concept subscale. Sample items include, "I have hesitated to take courses that involve mathematics", "I have generally done better in mathematics courses than other courses", and "I am quite good at mathematics." Marsh (1989) reports strong psychometric support for the SDQ-III, based on scores of reliability, correlations with external criteria, and self-other agreement. Scores of internal consistency range from $\alpha = .74$ to $\alpha = .95$ for the subscales of the SDQ-III. Studies of discriminant validity are extensive. For example, the relationship between one's math self-concept and math achievement score has been noted as $r = .58$, $p < .01$ (Marsh & O'Neill, 1984).

Social support. Perceived social support was measured by the Child and Adolescent Social Support Scale (CASSS; Malecki et al., 2000) The CASSS is designed to measure social support as perceived by children and adolescents. It is made up of 60 items, evenly divided

between five subscales (Parent, Teacher, Classmate, Close Friend, and School). Responses range from 1 (never) to 6 (always).

For the purposes of this study, we only used data from the Parent subscale. Sample items include, "My parents show they are proud of me", "My parents listen to me when I need to talk", and "My parents make suggestions when I don't know what to do." Internal consistency scores on subscales of the CASSS range from .92 to .96 (Malecki et al., 2000). Validity evidence for the CASSS has been documented through correlations with other measures of social support including the Social Support Scale for Children (r = .55) and with the Social Support Appraisals Scale (r = .56; as cited in Malecki et al.).

C. Procedure.

A random sample of female students who were declared STEM majors at a large research university in the southwestern United States (N = 758) were invited to participate in a study examining "women's experiences in math and science." A subset of male students (N = 300) was also invited to participate to serve as a comparison group. From the 1058 students who were invited to participate, a total of 499 students participated in the current study resulting in a convenience sample. Students completed a series of online questionnaires after electronically signing an informed consent document, and were compensated with a gift card.

IV. Results.

Means, standard deviations, and intercorrelations for all variables in the present study are presented in Table 1. Most noteworthy, math self-concept is correlated with family social support (r = .17, p < .001), mother's level of education (r = .13, p < 0.01), father's level of education (r = .10, p < 0.05), and SES (r = .13, p < 0.01). An independent samples t-test shows a significant difference between the math self-concepts of females (M = 6.22, SD = 1.38) and males (M = 6.68, SD = 1.02), t(498) = 3.44, p =.001, but not between the family social support of females (M = 4.85, SD = .86) and males (M = 4.69, SD = .88), t(498) = -1.80, p = .07.

Table 1. Means, Standard Deviations, and Intercorrelations of Variables of Interest.

Variable	***Females***		***Males***		**1**	**2**	**3**	**4**	**5**
	M	*SD*	*M*	*SD*					
1. Math SC	6.22	1.37	6.69	1.01	--				
2. Family SS	4.85	.86	4.70	.88	.17***	--			
3. Mother E					.13**	.08	--		
4. Father E					.10*	.04	.56***	--	
5. SES					.13**	.11*	.33***	.43***	--

Note. The correlation matrix includes both females and males. **1** Math SC = Math Self-Concept. **2** Family SS = Family Social Support. **3** Mother E = Mother's Level of Education. **4** Father E = Father's Level of Education. **5** SES = Family Socioeconomic Status. * p < .05, ** p < .01, ***p < .001

Hypotheses were tested via moderated regression. Gender, family income, mother's education, father's education, family support and the interactions between gender and each of the family variables were entered simultaneously as predictors of math self-concept. Race (entered

as five dummy codes, with Black as the reference group) and numbers of science and math classes taken in high school were entered as covariates in the analysis. All quantitative predictors were standardized, as was the quantitative covariate (number of science and math classes) and the outcome variable (math self-concept). Categorical variables entered using dummy codes (gender and race) were not standardized.

Results from the moderated regression appear in Table 2. It is worth noting that several of the covariates were significant predictors of math self-concept. Both Asians and Whites had significantly higher math self-concepts than did Blacks (Asians: $B = .499$, $p = .040$; Whites: $B = .554$, $p = .007$). In addition, respondents who had taken more science and math classes in high school had significantly higher math self-concepts than did those who had taken fewer such classes ($B = .157$, $p < .001$).

Table 2. Math Self-Concept Regression Results.

Predictor	B	SE	*t*	*p*
Intercept	-.138	.209	-0.663	.508
American Indian	.077	.701	0.110	.913
Asian	.499	.242	2.057	.040
Hispanic	.330	.216	1.526	.128
Pacific Islander	1.092	.702	1.555	.121
White	.554	.206	2.695	.007
Multiracial	.124	.315	0.393	.695
Science/math classes	.157	.044	3.538	< .001
Family finance	.062	.105	0.591	.555
Mother's education	-.105	.101	-1.042	.298
Father's education	.031	.112	0.274	.785
Family support	.351	.087	4.030	< .001
Gender	-.415	.100	-4.155	< .001
Gender × Family finance	.005	.119	0.042	.966
Gender × Mother's education	.219	.119	1.844	.066
Gender × Father's education	-.056	.129	-0.436	.663
Gender × Family support	-.229	.101	-2.269	.024

Hypothesis 1 predicted that women would report a lower math self-concept compared to men. Supporting our hypothesis, results showed that after controlling for race and number of math/science classes, gender significantly predicted math self-concept in that women reported a lower math self-concept than did men ($B = -.415$, $p < .001$). There was also a main effect for family support on math self-concept; students who felt more supported by their family reported a higher math self-concept ($B = .351$, $p < .001$). These main effects were qualified by an interaction between gender and family support on math self-concept. To examine the nature of the relationships, an analysis of simple slopes was conducted (Aiken & West, 1991). Supporting Hypothesis 2, results showed that higher family social support predicted higher math self-concept for both women ($B = .123$, $p = .017$) and men ($B = .351$, $p < .001$), but the association was more than twice as strong for men as for women. Figure 1 displays these relationships.

The main effect for gender was also nearly qualified by an interaction between gender and mother's education on math self-concept ($p = .066$), thus somewhat supporting Hypothesis 3. An analysis of simple slopes revealed a positive association between mother's education and math

self-concept for women (B = .114, p = .071), but a negative association for men (B = -.105, p = .298). The other hypothesized relationships were not significant.

Figure 1. **Interaction between Gender and Family Support of Math Self-Concept**.

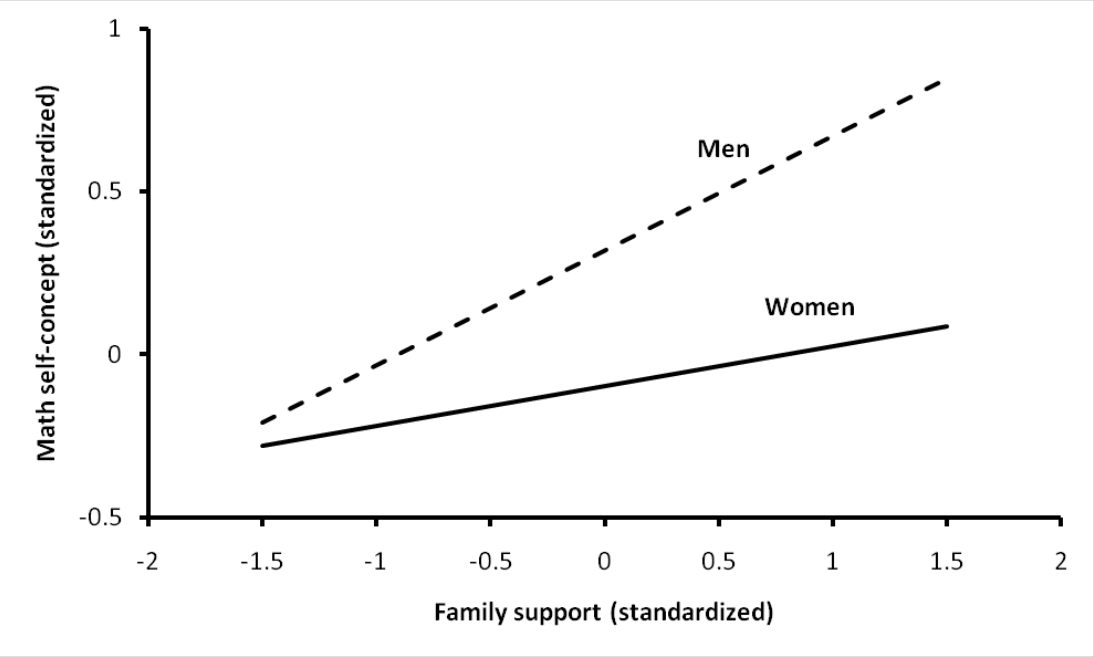
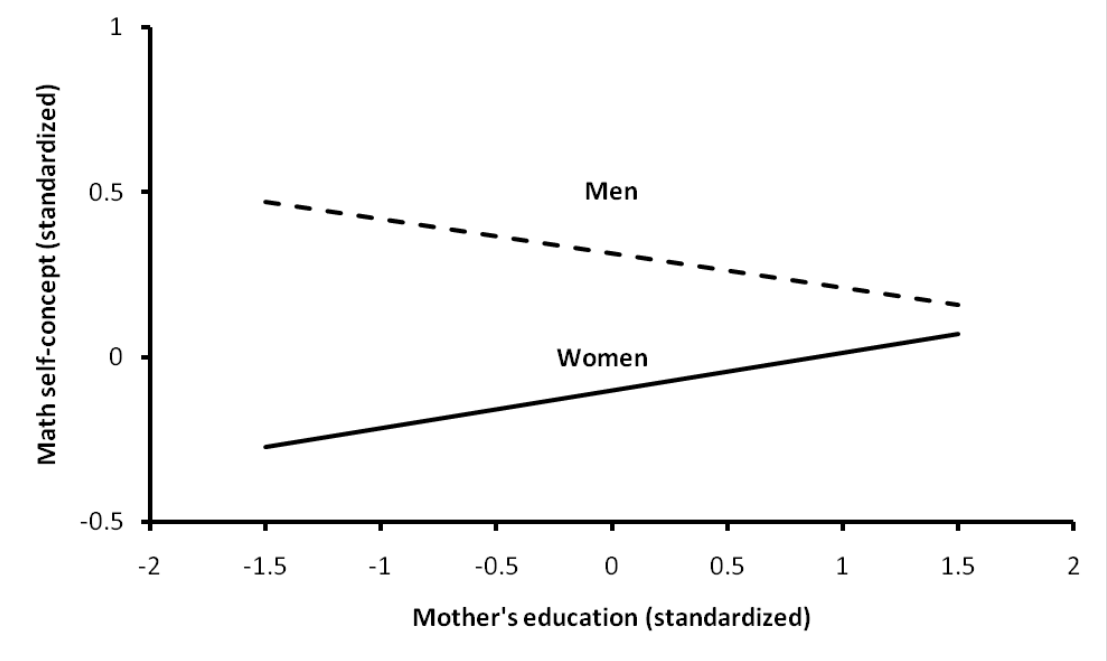


Figure 2. **Interaction between Gender and Mother's Education of Math Self-Concept.**



V. Discussion.

The purpose of the current study was to examine family context variables, including SES, mother's level of education, father's level of education, and perceived family social support, as predictors of math self-concept among undergraduate male and female STEM majors. Results indicate that Asians, Whites, and those who took more science/math classes while in high school had the highest math self-concepts. Males had higher math self-concepts than females, even after controlling for race and the number of science/math classes taken while in high school. Social support predicted math self-concept for males and females, but more than twice as strongly for males as for females. Finally, mothers' level of education nearly predicted math self-concept among females.

African American and Hispanic students, both males and females, are less likely to have access to advanced math and science courses while in high school (May & Chubin, 2003; Tyson, Lee, Borman, & Hanson, 2007), so it is not surprising to find, in the current study, that Asians and Whites had higher math self-concepts than other racial/ethnic groups.

Even though all participants surveyed in this study were declared STEM majors, males in the current sample had higher math self-concepts than females, even after controlling for race and the number of science/math classes taken while in high school. This finding is consistent with previous findings that have shown females report lower math self-concepts than males, regardless of actual ability levels in math (Else-Quest et al., 2010; Furnham, 2001; Furnham et al., 2002; Heller & Ziegler, 1996; Juang & Silbereisen, 2002; Marsh & Yeung, 1998; Meece & Jones, 1996; Pajares, 1996; Sax, 1994; Skaalvik & Skaalvik, 2004; Steinmayr & Spinathm, 2009; Watt, 2005). This gender difference in math self-concept is important and worth following longitudinally to determine whether math self-concept significantly predicts attrition from an undergraduate STEM major, STEM undergraduate degree attainment, STEM graduate degree attainment, or STEM career aspirations.

In the current study, higher family social support predicted math self-concept for both men and women, but this relationship was more than twice as strong for men as for women. This finding is in line with previous research that has shown perceived social support from parents has an impact on one's math self-concept and math self-efficacy (Ahmed et al., 2010; Navarro et al., 2007). Considering there was no significant difference in the amount of perceived parental social support between males and females in the current study, we are left to wonder why parental social support has such a stronger effect on males' math self-concepts than on females' math self-concepts. We did not ask participants to elaborate about the social support messages they are receiving from their parents, particularly with regard to their STEM major, but future research should consider this as an option. If parents are providing support based on their perceptions of their children's abilities in STEM areas, or based on the perceived appropriateness of majoring in a STEM area, this type of support could differentially impact a male versus a female's math self-concept. Generic parental support related to college (i.e., financial support, praise, positive feedback), then, might not affect math self-concept in the same way.

Mother's education nearly predicted math self-concept among females, but there would have been a negative trend among males, such that more educated mothers would have resulted in lower math self-concepts among males. Father's education did not significantly predict math self-concept for either males or females. This trend partially replicates Sax's (1994) finding that having a highly educated mother has a negative effect on males' math self-concepts. Sax provides the following explanation for this seemingly contradictory finding: "…perhaps men

with highly educated mothers would not be as overconfident as other men, because they have contact with, and are influenced by, highly educated, intelligent women… although men with highly educated mothers have overall greater confidence in math, these men might be less likely to *overestimate* their math abilities" (p.155-157). As Sax's findings might now be outdated, more research is certainly needed to tease apart the relationship between parents' level of education and children's math self-concepts.

Although the two variables were correlated, SES was not a predictor of math self-concept in the current study. As there was little variation among participants with regard to SES, this finding is not entirely surprising. Replicating this study on a more diverse university campus, or across multiple institutions, might provide further conclusive evidence regarding the relationship between SES and math self-concept.

Limitations and Directions for Future Research. The current sample was obtained at a single institution and by using a convenience sampling method, both of which naturally limit the generalizability of the findings. Future research should replicate the current study across multiple institutions and multiple groups of undergraduates. The following limitations should be understood with the lack of generalizability of the current study in mind.

Future research examining math self-concept among those undergraduates majoring in a STEM field should include undergraduates' particular STEM fields of study (e.g., math, physics, chemistry, engineering) as a covariate or predictor. Researchers, administrators, and educators assume math self-concept is predictive of persistence in a STEM field, but it might not be for each individual STEM major. For example, math self-concept might not be as important or predictive of success in less mathematically demanding STEM fields such as biology, as it is in more mathematically demanding STEM fields such as physics.

While we included the number of math and science classes taken during high school as a covariate in the current study to account for academic preparation, future research should include undergraduates' total SAT or ACT scores, or their SAT or ACT math subscale scores, as a covariate to account for the potential relationship between aptitude and self-concept.

Finally, in order to gain more conclusive evidence regarding the predictive aspects of family context variables on math self-concept, the experiences of undergraduate STEM majors should be compared to undergraduate non-STEM majors.

Conclusion. Even among declared STEM majors, females are still reporting lower math self-concepts than males. Among the family context variables examined in the current study, only parental social support played a significant role in predicting math self-concept among female undergraduate STEM majors. Future research should continue the examination of math self-concept among female and male STEM majors, as the gender difference in math self-concept may lead to differences in later math- and STEM-related achievement, which may lead to differences in STEM undergraduate and graduate degree attainment.

References

Ahmed, W., Minnaert, A., Werf, G., & Kuyper, H. (2010). Perceived social support and early adolescents' achievement: The mediational roles of motivational beliefs and emotions. *Journal of Youth and Adolescence, 39*, 36-46.

Aiken, L. S., & West, S. G. (1991). *Multiple Regression: Testing and Interpreting Interactions.* Newbury Park, CA: Sage.

Astin, A. W. (1993). *What Matters in College? Four Critical Years Revisited.* San Francisco: Jossey-Bass.

Azmitia, M., & Cooper, C. R. (2001). Good or bad? Peer influences on Latino and European American adolescents' pathways through school. *Journal of Education for Students Placed at Risk, 6*, 45-71.

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Approach.* Englewood Cliffs, NJ: Prentice Hall.

Bonous-Hammarth, M. (2000). Pathways to success: Affirming opportunities for science, mathematics, and engineering majors. *Journal of Negro Education, 69*, 92-111.

Bowen, W. G., Kurzweil, M. A., & Tobin, E. M. (2005). *Equity and Excellence in American Higher Education.* Charlottesville, VA: University of Virginia Press.

Brainard, S. G., & Carlin, L. (2001). A six-year longitudinal study of undergraduate women in engineering and science. In M. Lederman & I. Bartsch (Eds.), *The Gender and Science Reader.* New York: Routledge.

Budny, D. D., & Paul, C. A. (2003). Working with students and parents to improve the freshman retention. *Journal of STEM Education: Innovations and Research, 4*, 1-9.

Byars-Winston, A., Estrada, Y., Howard, C., Davis, D., & Zalapa, J. (2010). Influence of social cognitive and ethnic variables on academic goals of underrepresented students in science and engineering: A multiple-groups analysis. *Journal of Counseling Psychology, 57*, 205-218.

Cabrera, A. F., Nora, A., Terenzini, P. T., Pascarella, E., & Hagedorn, L. S. (1999). Campus racial climate and adjustment of students in college: A comparison between White and African-American students. *The Journal of Higher Education, 70,* 134-155.

Eimers, M. T., & Pike, G. R. (1997). Minority and nonminority adjustment to college: Differences or similarities? *Research in Higher Education, 38,* 77-97.

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136*, 103-127.

Farmer, H. S. (1985). Model of career achievement motivation for women and men. *Journal of Counseling Psychology, 32*, 363-390.

Flores, L. Y., & O'Brien, K. M. (2002). The career development of Mexican American adolescent women: A test of social cognitive career theory. *Journal of Counseling Psychology, 49*, 14-27.

Furnham, A. (2001). Self-estimates of intelligence: Culture and gender difference in self and other estimates of both general (g) and multiple intelligences. *Personality and Individual Differences, 31,* 1381-1405.

Furnham, A., Hosow, T., & Tang, T. L.-P. (2002). Male hubris and female humility? A cross-cultural study of ratings of self, parental, and sibling multiple intelligence in America, Britain, and Japan. *Intelligence, 30,* 101-115.

Gainor, K. A., & Lent, R. W. (1998). Social cognitive expectations and racial identity attitudes in predicting the math choice intentions of black college students. *Journal of Counseling Psychology, 45*, 403-413.

Gottfredson, G. D. (2002). Interests, aspirations, self-estimates, and the self-directed search. *Journal of Career Assessment, 10*, 200-208.

Graham, S. W., & Gisi, S.L. (2000). The effects of instructional climate and student affairs services on college outcomes and satisfaction. *Journal of College Student Development, 41,* 279-291.

Grandy, J. (1998). Persistence in science of high-ability minority students: Results of a longitudinal study. *The Journal of Higher Education, 69*, 589-620.

Hamachek, D. (1995). Self-concept and school achievement: Interaction dynamics and a tool for assessing the self-concept component. *Journal of Counseling and Development, 73*, 419-425.

Heller, K. A., & Ziegler, A. (1996). Gender differences in mathematics and the sciences: Can attributional retraining improve the performance of gifted females? *Gifted Child Quarterly, 40,* 200-210.

Hortaçsu, N. (1995). Parents' education levels, parents' beliefs, and child outcomes. *The Journal of Genetic Psychology: Research and Theory on Human Development, 156,* 373-383.
House, J. D. (2000). The effect of student involvement on the development of academic self-concept. *Journal of Social Psychology, 140*, 261-263.

Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A., & Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect: A meta-analysis. *Psychology of Women Quarterly, 14*, 299-324.

Jacobs, J. E. (2005). Twenty-five years of research on gender and ethnic differences in math and science career choices: What have we learned? *New Directions for Child and Adolescent Development, 110*, 85-94.

Johnsen, S. K., & Kendrick, J. (Eds.). (2005). *Teaching and Counseling Gifted Girls.* Waco, TX: Prufrock Press.

Juang, L. P., & Silbereisen, R. K. (2002). The relationship between adolescent academic capability beliefs, parenting, and school grades. *Journal of Adolescence, 25,* 3-18.

Khattri, N., Riley, K. W., & Kane, M. B. (1997). Students at risk in poor, rural areas: A review of the research. *Journal of Research in Rural Education, 13*, 79-100.

Kohlstedt, S. G. (2004). Sustaining gains: Reflections on women in science and technology in 20th-century United States. *NWSA Journal, 16,* 1-26.

Larose, S., Ratelle, C. F., Guay, F., Senécal, C., & Harvey, M. (2006). Trajectories of science self-efficacy beliefs during the college transition and academic and vocational adjustment in science and technology programs. *Educational Research and Evaluation, 12*, 373-393.

Lee, J. (2009). Universals and specifics of math self-concept, math self-efficacy, and math anxiety across 41 PISA 2003 participating countries. *Learning and Individual Differences, 19*, 355-365.

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior, 45*, 79-122.

Lockwood, P. (2006). "Someone like me can be successful": Do college students need same-gender role models? *Psychology of Women Quarterly, 30*, 36-46.

Maccoby, E. E., & Martin, J. A. (1983). Socialization in the context of the family: Parent-child interaction. In E. M. Hetherington (Ed.) & P. H. Mussen (Series Ed.), *Handbook of child psychology:* Vol. 4. *Socialization, personality, and social development* (pp. 1-101). New York, NY: Free Press.

Malecki, K. C., & Demaray, M. K. (2003). Perceptions of the frequency and importance of social support by students classified as bullies, victims, and bully/victims in an urban middle school. *School Psychology Review, 32,* 471-489.

Malecki, C. K., Demaray, M. K., & Elliott, S. N. (2000). *A working manual on the development of the child and adolescent social support scale.* DeKalb, IL: Northern Illinois University.

Marable, M. (2003). *The great wells of democracy: The meaning of race in American life.* Cambridge, MA: BasicCivitas Books.

Marsh, H. W. (1989). *Self Description Questionnaire (SDQ) III: A Theoretical and Empirical Basis for the Measurement of Multiple Dimensions of Late Adolescent Self-Concept: An Interim Test Manual and a Research Monograph.* New South Wales, Australia: University of Western Sydney.

Marsh, H. W., & O'Neill, R. (1984). Self Description Questionnaire III: The construct validity of multidimensional self-concept ratings by late adolescents. *Journal of Educational Measurement, 21,* 153-174.

Marsh, H. W., & Shavelson, R. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist, 20*, 107-123.

Marsh, H. W., & Yeung, A. S. (1998). Longitudinal structural equation models of academic self-concept and achievement: Gender differences in the development of math and English constructs. *American Educational Research Journal, 35*, 705-738.

Mau, W. C. (2003). Factors that influence persistence in science and engineering career aspirations. *The Career Development Quarterly, 51*, 234-242.

May, G. S., & Chubin, D. E. (2003). A retrospective on undergraduate engineering success for underrepresented minority students. *Journal of Engineering Education, 92*, 27-40.

Meece, J. L., & Jones, M. G. (1996). Gender differences in motivation and strategy use in science: Are girls rote learners? *Journal of Research in Science and Teaching, 33*, 407-431.

Multon, K. D., Brown, S. D, & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology, 38*, 30-38.
National Science Foundation. (2007). *Women, minorities, and persons with disabilities in science and engineering: NSF 07-315.* Division of Science Resource Statistics: Arlington, VA.

Navarro, R. L., Flores, L. Y., & Worthington, R. L. (2007). Mexican American middle school students' goal intentions in mathematics and science: A test of social cognitive career theory. *Journal of Counseling Psychology, 54*, 320-335.

Orr, E., & Dinur, B. (1995). Actual and perceived parental social status: Effects on adolescent self-concept. *Adolescence, 30*, 603-616.

Otto, L. B. (2000). Youth perspectives on parental career influence. *Journal of Career Development, 27*, 111-117.

Ozturk, M. A., & Singh, K. (2006). Direct and indirect effects of socioeconomic status and previous mathematics achievement on high school advanced mathematics course taking. *The Math Educator, 16*, 25-34.

Pajares, F. (1996). Self-efficacy beliefs and mathematical problem-solving of gifted students. *Contemporary Educational Psychology, 21,* 325-344.

Pascarella, E. T., Smart, J. C., Ethington, C. A., & Nettles, M. T. (1987). The influence of college on self-concept: A consideration of race and gender differences. *American Educational Research Journal, 24*, 49-77.

Pascarella, E., & Terenzini, P. (1991). *How College Affects Students.* San Francisco: Jossey-Bass.

Pascarella, E. T., & Terenzini, P. T. (2005). *How College Affects Students. Volume 2: A Third Decade of Research.* San Francisco, CA: Jossey-Bass.

Perna, L. W., & Titus, M. A. (2004). Understanding differences in the choice of college attended. *The Review of Higher Education, 27*, 501-525.

Quilter, S. M. (1995). Academic self-concept and the first-year college student: A snapshot. *Journal of the Freshman Year Experience, 7*, 39-52.

Rottinghaus, P. J., Lindley, L. D., Green, M. A., & Borgen, F. H. (2002). Educational aspirations: The contribution of personality, self-efficacy, and interests. *Journal of Vocational Behavior, 61*, 1-19.

Sax, L. J. (1994). Mathematical self-concept: How college reinforces the gender gap. *Research in Higher Education, 35*, 141-166.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Validation of construct interpretations. *Review of Educational Research, 46*, 407-441.

Seymour, E. (1995). The loss of women from science, mathematics, and engineering under-graduate majors: An explanatory account. *Science Education, 79*, 437-473.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*, 417-453.

Skaalvik, S., & Skaalvik, E. M. (2004). Gender differences in math and verbal self-concept, performance expectations, and motivation. *Sex Roles, 50,* 241-252.

Sonnert, G., Fox, M. F., & Adkins, K. (2007). Undergraduate women in science and engineering: Effects of faculty, fields, and institutions over time. *Social Science Quarterly, 88,* 1333-1356.

Steinmayr, R., & Spinath, B. (2009). What explains boys' stronger confidence in their intelligence? *Sex Roles, 61*, 736-749.

Tyson,W., Lee, R., Borman, K. M., & Hanson, M. A. (2007). Science, technology, engineering, and mathematics (STEM) pathways: High school science and math coursework and postsecondary degree attainment. *Journal of Education for Students Placed at Risk, 12*, 243-270.

Walpole, M. (2003). Socioeconomic status and college: How SES affects college experiences and outcomes. *The Review of Higher Education, 27*, 45-73.

Watt, H. M. G. (2005). Explaining gendered math enrollments for NSW Australian secondary school students. *New Directions for Child and Adolescent Development, 110*, 15-29.

Whiston, S. C., & Keller, B. K. (2004). The influences of the family of origin on career development: A review and analysis. *The Counseling Psychologist, 32*, 493-568.

Williams, T., & Williams, K. (2010). Self-efficacy and performance in mathematics: Reciprocal determinism in 33 nations. *Journal of Educational Psychology, 102*, 453-466.

Yavuz, M. (2009). Factors that affect mathematics-science (MS) scores in the secondary education institutional exam: An application of structural equation modeling. *Educational Sciences: Theory and Practice, 9*, 1557-1572.

Journal of the Scholarship of Teaching and Learning, Vol. 13, No. 2, May 2013, pp. 133 – 147.

Management of classroom behaviors: Perceived readiness of education interns

Dennis Garland[1], Krista Vince Garland[2], and Eleazar Vasquez, III[3]

Abstract: Education students at a large research university participated in internships during their final semesters as part of their respective programs of study as a capstone experience. Qualitative and quantitative methods were used to collect data on the perceptions of interns' readiness and knowledge of evidence-based practices to manage classroom behaviors for students with exceptionalities in inclusive settings. Emergent themes include general feeling of readiness to manage classroom behaviors. However, a desire for earlier access to actual classroom experiences was also expressed, as was the desire for greater instruction in evidence-based practices to manage behaviors for students with exceptionalities.

Keywords: teacher preparation, behavior management, evidence-based practices, student internships

I. Purpose.

Teacher quality is the single most accurate indicator of students' academic success, and teachers who leave the profession often cite a lack of adequate preparation as one of the reasons for their departure (Darling-Hammond, 2010; McKinney, Haberman, Stafford- Johnson, & Robinson, 2008). Reschly and Holdheide (2008) found that teachers who are skilled in scientifically based instruction, classroom organization, and behavior management had the competencies to establish classroom environments conducive to learning and improved academic performance for all students.

In a literature review of evidence-based practices in classroom management, Simonsen, Fairbank, Briesch, Myers, and Sugai (2008) identified five critical features of effective classroom management: (a) maximize structure; (b) post, teach, review, monitor, and reinforce expectations; (c) actively engage student in observable ways; (d) use a continuum of strategies for responding to appropriate behaviors; and (e) use a continuum of strategies to respond to inappropriate behaviors. Proactive, evidence-based programs are currently being implemented in school districts nationally (Sugai & Horner, 2006) and disseminated through resources such as pbis.org to support teachers in managing behaviors. However, early career teachers have frequently stated that they are unprepared to address problematic behaviors (Cooper, Kurtts, Baber, & Vallecorsa, 2008), especially among students with exceptionalities in inclusive settings (Billingsley, Israel, & Smith, 2011; Regan & Michaud, 2011).

[1] Department of Child, Family, and Community Sciences, University of Central Florida, Orlando, FL 32816, dpgarland@knights.ucf.edu

[2] Department of Exceptional Education, Buffalo State College, Buffalo, NY 14222, garlankm@buffalostate.edu

[3] Department of Child, Family, and Community Sciences, University of Central Florida, Orlando, FL 32816, eleazar.vasquez@ucf.edu

Traditionally teacher preparation programs provide a culminating, capstone experience (Backhus & Thompson, 2006; Fernandez & Erbilgin, 2009; Kenny, 1998) in the form of a semester-long internship during a senior student's final semester of baccalaureate study. These internships provide an opportunity for students to implement and refine teaching strategies learned through their coursework in actual classrooms with K-12 students.

The purpose of this study was to investigate undergraduate pre-service teachers' levels of preparedness when managing student behaviors in inclusive settings. Specifically, we asked (a) what are the perceptions of readiness among college of education interns for managing classroom behaviors for students with exceptionalities when they begin teaching, and (b) to what extent do interns identify best practices for managing behaviors of students with and without exceptionalities?

II. Background.

The No Child Left Behind Act of 2001 (NCLB) has dramatically increased the demands that all teachers encounter in the classroom. New teachers need a broad continuum of abilities to teach more complex curriculum to the growing number of public school students who have limited educational resources at home, those whose primary language is not English, and those who have special needs (Darling-Hammond, 2010). Moreover, established research on teacher development has shown that early career teachers have long had feelings of "inadequacy and unpreparedness" (Katz, 1972, p. 51.) as well as concerns about classroom management (Burden, 1979; Fuller, 1969; Fuller & Brown, 1975; Katz, 1972). These factors emphasize the need for Institutions of Higher Education (IHE) to not only evaluate the outcomes of their teacher preparation programs, but assess processes that lead to those outcomes in the name of high quality education for all teachers and their students (Slavin, 2007). This is particularly important with respect to teachers' classroom management readiness (Cooper et al., 2008).

Objectives of the reauthorization of the Elementary and Secondary Education Act (ESEA; 1965) include (a) improving teacher and principal effectiveness, (b) providing information to families to help them improve their children's schools and to educators to help them improve their students' learning, (c) implementing college and career-ready standards and development of improved assessments aligned with those standards, and (d) providing support and interventions to improve student learning and achievement in the nation's lowest performing schools. Above all, the reauthorization emphasizes the goal to meet the needs of diverse learners (Department of Education [ED], 2010). An emphasis on preparation in content knowledge that applies to special education teachers has been explicated in the Individuals with Disabilities Education Improvement Act (IDEIA) of 2004 (Boe, Shin, & Cook, 2007).

For their part, general education teachers have indicated the need for ongoing professional development in the management of student behavior (Cooper et al., 2008). Many new special educators conveyed that when it comes to behavior management, they faced comparable challenges to their general educator colleagues (Keller, Brady, & Taylor, 2005; White & Mason, 2006). In a meta-analysis of studies concerning teacher induction programs, Billingsley et al. (2011) found that new teachers focus more on behavioral challenges than any other area of their jobs. The impetus for the current research came from the legislative context of increased teacher accountability and the increased momentum toward including students with exceptionalities in general education settings (Darling-Hammond, 2010; Forlin, Loreman, Sharma, & Earle, 2009), thus informing programming among colleges of education, and

establishing an engaging and safe environment for all students (Lane & Carter, 2006; Oblinger, 2008; Regan, 2009; Jukes & McCain, 2011).

III. Methodology.

A. Setting.

To address the findings of previous research, a survey was conducted at a large research university in the southeastern United States, having an enrollment of 58,587 students. At the time of the study, the college of education had an enrollment of 5,590 students. In an attempt to triangulate the data and increase reliability, subsequent semi-structured interviews were held on the main campus of the university in a medium-sized student lounge setting.

B. Sample.

A convenience sample was obtained in collaboration with the director of the office of clinical experiences at the college of education of all students enrolled in a teaching internship (N=891). An invitation to complete an anonymous online survey (Appendix A) was distributed by the director to the interns, and the response rate was 34% (see Table 1).

C. Materials.

Participant response to survey questions was conducted using a free online survey platform. Design and deployment of the survey followed protocols set forth by Dillman and Bowker (2000) on four types of errors- sampling, coverage, measurement, and nonresponse. The online survey was composed of a total of 18 questions with 15 questions on a Likert scale ranging from one to five (1= strongly disagree to 5 = strongly agree) and three open-response questions. Questions 1-15 were sectored into three factors: (a) preparedness; (b) accommodations for students with exceptionalities; and (c) communication. Three open-response questions were included at the end of the survey that asked respondents (a) their age; (b) what strategies they were planning to use to manage behaviors in inclusive classrooms; and (c) their program of study.

Strategies for classroom management and program of study were included because they could be indicators of how particular programs of study prepared their students to teach in inclusive settings. Kaplowitz, Hadlock, and Levine (2004) determined mean response rates for web-based surveys to be 21% compared to surveys delivered by mail (31.5%). The survey used in this study included a small number of items with the intention of increasing the likelihood of participant response. Survey response rate was 34%, higher than the average, as reported by Kapliwitz et al. (2004).

Validation of survey items 1-15 was accomplished through factor analysis (Dillman & Bowker, 2000). Items that correlated the highest with a factor defined the meaning of the factor as judged by what conceptually tied the items together, thereby providing internal structure evidence (Hair, Anderson, Tatham, & Black, 2009). The factors, ascertained by the magnitude of the coefficients are identified in Table 2 by the shading, where shaded coefficients are the largest coefficients for a factor. Names for the factors are as follows: Factor 1, Preparedness

Considerations; Factor 2, Accommodations Considerations; and Factor 3, Communication Considerations.

Table 1. Survey participants.

College of Education Major	N	Percent of N
Elementary Ed. BS	163	53.98%
English L.A. Ed. BS	21	6.96%
Exceptional Ed. BS	20	6.63%
Social Science Ed. BS	28	9.27%
Art Ed. BS	11	3.64%
Social Science Ed. MAT	5	1.66%
Mathematics Ed BS	14	4.64%
Secondary Ed. BS Biology	2	.66%
No Answer	7	2.32%
Elementary Education MA	1	.33%
Early Childhood Education BS	6	1.99%
Exceptional Education MA	8	2.65%
Middle School Mathematics MAT	1	.33%
Science Education BS-Chemistry	1	.33%
Foreign Language Education BS	4	1.31%
Science Education BS – Physics	2	.66%
English Language Arts Education with ESOL Endorsement MAT	2	.66%
Counselor Education MA	1	.33%
Counselor Education Med	2	.66%
Foreign Language Education BS Spanish	3	.99%
Total	302	100%

Descriptive statistics were performed to examine the frequency of responses to items 1-15 of the survey. Cross tabs were examined to determine if there existed an influence of the respondents' program of study or their age on their response to items 1-15. Open-response items from the questionnaire were analyzed using a grounded theory approach (Glaser & Strauss, 1967; Cresswell, 2007; Glesne, 2011), which consisted of coding the survey responses and aggregating the codes to identify themes. The themes that emerged served as foundation for the development of semi-structured interview questions.

The director of the Office of Clinical Experiences at the College of Education facilitated the distribution of invitations via email to the respondents of the survey to participate in semi-structured interviews. The resulting sample size for the semi-structured interviews was five participants from four educational programs of study (See Table 3).

Table 2. Structure matrix.

	Factor		
	1	2	3
i5	.886		
i6	.848		
i9	.755	.510	
i1	.732	.444	
i10	.605		.319
i8		.846	.151
i4		.802	.111
i2	.307	.790	
i7		.642	
i13		.602	
i14		.501	
i11	.344		.764
i3	.264		.599
i12		.126	.392
i15		.119	.382

Extraction Method: Maximum Likelihood.
Rotation Method: Promax with Kaiser Normalization.

Table 3. Interview participants.

College of Education Major	**N**
Exceptional Education B.A.	2
Elementary Education B.A.	1
Mathemetics Education M.A.T.	1
Early Childhood Education B.A.	1

D. Data Collection.

All the interviews used a standard protocol of questions and were conducted according to Bogdan's and Biklen's (2007) recommended approach. The primary author conducted each of the sessions, served as facilitator, and audio taped the interviews. Interview questions were developed based upon analysis of survey results and the protocol of interview questions followed the structure of the overall research questions regarding perceptions of readiness for behavioral management when working with students with special needs in inclusive settings, knowledge of best practices for behavioral management when working with students with exceptionalities and

curriculum coursework that student participants had completed. Recommendations for future curriculum preparation and in-service programs were also solicited.

E. Data Analysis.

Grounded theory (Cresswell, 2007) procedures were utilized to examine responses to question 17 of the survey, "What strategies are you planning for managing classroom behaviors for students with exceptionalities in inclusive settings?" These procedures consist of "developing categories of information (open coding), interconnecting the categories (axial coding), building a 'story' that connects the categories (selective coding), and ending with a discursive set of theoretical propositions" (Creswell, 2007, p. 150). Due to the open-ended nature of these questions, a participant could have identified several concerns within one answer. As such, more than one theme could have been identified and coded for the question. To establish inter-rater reliability, the primary and secondary authors independently coded the first 50 responses into categories whereupon themes emerged, then compared each response and corresponding theme category. Response codings that were not agreed upon were discussed and agreement was sought. Final inter-rater reliability on the question was higher than 95%. Themes from the qualitative analysis of the survey were fundamental to the creation of interview questions.

Subsequent interview data were analyzed using Krueger and Casey's (2000) recommended methods to identify themes and their prevalence within and across individuals and to contrast the views of interns from different education programs. Audiotapes of the sessions were transcribed and combined with affiliated notes for analysis. Member checks were conducted throughout the duration of the study (Merriam, 1998). Lincoln and Guba (1985) describe member checks as "the most crucial technique for establishing credibility" (p. 314) in a study and consist of taking data and interpretations back to the participants in the study to confirm credibility of information and narrative account.

Participants were emailed transcripts of their interview and the researchers' interpretations of the interviews and were invited to comment on accuracy of interview transcripts and interpretations prior to subsequent interviews. None of the participants objected to the interview transcripts or the researchers' interpretations. Related themes and patterns emerged in the course of the conversations. Results of the data analysis were then examined in the light of current research and literature about the preparation of general and special educators for behavioral management.

IV. Results.

A regression analysis (see Table 4) of survey items 1-15 on items 16 and 17 resulted showed the following areas of significance:

Table 4. Survey item regression.

Nominal item	Interval item	Significance
Age	The physical environment of my classroom is an important consideration toward the learning of all students.	.025
Program of Study	I am confident that I will be able to cope with the pressures of classroom behavior management.	.043
Program of Study	I am confident in my ability to manage classroom behaviors	.050
Age	Seating arrangements of students can promote positive behaviors or negative behaviors.	.023
Age	I should be aware if any of my students take medication or not.	.031

Among the responses of the survey, nine predominant themes emerged: (a) positive reinforcement; (b) seating arrangements; (c) collaborative development of class rules; (d) posting class rules and consequences; (e) rewards systems; (f) consistency; (g) smooth instructional transitions; (h) use of behavior charts; and (i) not sure. The survey results provided salient concerns of teaching interns. In order to establish deeper context to those concerns (Krueger & Casey, 2000), semi-structured interviews were conducted. Three main themes emerged from the analysis of the transcribed participant interviews. Emergent themes include: (a) a feeling of uncertainty to manage classroom behaviors; (b) a desire for earlier access to actual classroom experiences; and (c) a need for greater instruction in evidence-based practices to manage behavior for students with exceptionalities.

A. A Feeling of Uncertainty to Manage Classroom Behaviors.

The first major theme to emerge from the structured interviews was consistent with the literature as it relates to the feelings of teachers regarding their readiness to manage inappropriate classroom behaviors (Brownell, Ross, Colon, & McCallum, 2005; Burden, 1979; Fuller, 1969; Fuller & Brown, 1975; Katz, 1972; Melnick & Meister, 2008; Westling, 2010). Clarissa (pseudonym) stated:

> As teachers, we are asked to go into classrooms and be highly qualified teachers to both students that want to learn and students that don't want to learn or only want to disrupt the flow of a lesson; but principals expect us to be both experts in our content area and experts at controlling behaviors, we are just not prepared or trained to do that.

A comment from another student from a different education program was:

> I feel that I am progressively getting better and learning new techniques to manage behaviors in the classroom, but I'm not sure how well I could implement them if I were in my own classroom and in a real situation.

Although the respondents were uncertain of their readiness, there was a sense of optimism among some. "I am planning to use my personality and love for teaching as tools to enthuse my students about learning, and I hope that will make a big difference in terms of how they behave", wrote one intern.

B. A Desire for Earlier Access to Actual Classroom Experience.

The second theme that emerged among participants was the desire for earlier access to actual classroom experiences. As one student expressed:

> We come into our programs thinking that we are made for this, and we spend a lot of money to become teachers. But we don't even get a chance to spend any real time in a classroom until our last two semesters. What if we got all that way and realized that we really weren't cut out for teaching?

Another intern had similar concerns:

> I really wish that my program gave us exposure to a range of classroom experiences sooner than at the end, when we are slapped with the realities of juggling lesson plans, classroom management, state exam prep, and everything else that can overwhelm us. That way, we could have a better chance to reflect on what we've learned and prepare better.

One student framed her response from another point of view:

> I had the opportunity to visit a friend of mine who teaches students with special needs early in my program to see what it would be like to really teach. That experience solidified my motivation to get into the classroom and teach. I know it will be tough, but every kid deserves an opportunity to have a great education and if they have the right teacher it will make all the difference.

C. A Need for Greater Instruction in Evidence-Based Practices to Manage Behavior.

The final theme that emerged clearly during interviews with the interns was the need for greater instruction in evidence-based practices to manage behaviors in the classroom. When asked, "what kind of strategies will you use to manage classroom behaviors?" a range of positive responses were given among the participants, including:

> You need to allow for student differences. For example, say you may have a child who has ADHD and is fidgety. I would allow him to take a calm lap around the classroom and return to his studies. This way, he uses up his energy and can focus better on the task at hand.

Although the participants largely had well-meaning plans for the management of classroom behaviors, there were a number of responses that did not align with evidence-based practices for students with exceptionalities. One such example being:

> I would just nip it in the bud straight off the bat. Students are just kids, and they need to know that the adult is in charge of the classroom. I mean, I am not going to be a tyrant, but my students will be well aware that the classroom is mine, and that misbehavior will not be tolerated. That way, they will be able to anticipate what I want from them.

Despite their lack of experience in implementing evidence-based practices, there were responses that indicated a degree of willingness to learn more about the how to put them into practice. As one participant said, " I think I have learned a lot of great strategies for managing behavior and working with diverse students. I just need some coaching and practice in using them in real teaching situations".

In summary, analysis of data from both the survey results and the semi-structured interviews during this study indicated that students at this university had feelings of uncertainty about managing classroom behaviors and possessed a desire for earlier access to actual classroom experiences. They also expressed a need for greater instruction in evidence-based practices for behavior management when working with students in inclusive settings. However, there was a sense of optimism about having the opportunity to teach that coincided with the feelings of reticence.

V. Discussion.

This study sought insight into pre-service education interns' readiness to manage behaviors of students with exceptionalities in inclusive classrooms. Upon analysis of the results, there was evidence that some teaching interns felt uncertain of their ability to manage classroom behaviors. Many respondents expressed a desire access to real classrooms earlier in their educational program and felt that having a more initial exposure could solidify foundational pedagogies of best practices in behavior management.

Lastly, both survey and interview responses indicated that students recognized the importance of honing a skill set in behavioral management, and evidenced that they grasped some of the concepts integral to best practices when working with students who have special needs. However, many respondents also stated that they required more in-depth instruction before becoming the teacher of record themselves. These findings mirror previous study findings which report that the incorporation of increased instruction on the implementation of evidence-based practices for behavior management when working with students who have exceptional needs is a predominant factor to student success and teacher retention (Billingsley et al., 2011; McKinney et al., 2008; Reschly & Holdheide, 2008), and should occur within teaching curricula.

In looking toward future studies on perceived readiness of pre-service education interns and given the state of available technologies, institutes of higher education may seek to develop and evaluate teacher preparation curricula that harnesses cloud technology. By utilizing a web-based learning and supervision platform that allows student teachers to progress through clinical teaching coursework, intern supervision can take place online with opportunities for immediate feedback and coaching (Rock, Gregg, Gable, & Zigmond, 2009; Scheeler, McKinnon, & Stout, 2012). Possibilities may include synchronous online tutoring (Vasquez & Slocum, 2011) and experiences in simulated learning environments (Hughes, Stapleton, & O'Connor, 2006; Vince Garland, Vasquez, & Pearl, 2012).

Such learning opportunities could be infused into teaching curricula to enhance traditional face to face instruction (Billingsley et al., 2011) and increase opportunities for earlier access to classroom experiences among teaching interns prior to their final semesters of programming. It is clear that targeted instruction of evidence-based practices for behavior management when working with students with exceptionalities is recognized to be important by the pre-service teachers in this study, and necessary to ensure the retention of teachers who serve even our most vulnerable students.

It is recognized that the sample for this study was limited to one university, and generalization to other programs in the U.S. is slight. Future research should include a larger sample size from several IHEs in different geographical areas. It is important to note that findings echo results from previous studies across the years and emphasize a continued ongoing desire among teaching interns for additional preparation in behavior management (Burden, 1979; Fuller, 1969; Fuller & Brown, 1975; Katz, 1972).

Acknowledgements

We wish to thank Dr. Bryan Zugelder, Director of Clinical Experiences at the University of Central Florida College of Education for distributing all correspondence concerned with this study to the respective participants.

Appendix 1. Survey.

A Survey of Pre-Service Teachers' Perceptions of Readiness for Behavior Management	Strongly Disagree	Disagree	Neither	Agree	Strongly Agree	Not Applicable
Instructions: Please circle one answer for each statement below.	SD	D	NA/D	A	SA	N/A
START HERE						
1. I know that students with disabilities have legal rights.	1	2	3	4	5	N/A
2. Some students need accommodations for their behavior.	1	2	3	4	5	N/A
3. It is a good idea to consult with parents at least once per grading period.	1	2	3	4	5	N/A
4. The physical environment of my classroom is an important consideration toward the learning of all students.	1	2	3	4	5	N/A
5. I am confident that I will be able to cope with the pressures of classroom behavior management.	1	2	3	4	5	N/A
6. I am confident in my ability to manage classroom behaviors	1	2	3	4	5	N/A
7. There should be opportunities for individual and group work on assignments.	1	2	3	4	5	N/A

8.	Seating arrangements of students can promote positive behaviors or negative behaviors.	1	2	3	4	5	N/A
9.	I should know if any of my students has a behavioral disability.	1	2	3	4	5	N/A
10.	I should be aware if any of my students take medication or not.	1	2	3	4	5	N/A
11.	I should communicate with my students' other teachers about their behaviors in other settings.	1	2	3	4	5	N/A
12.	I will allow for students to communicate their emotional needs.	1	2	3	4	5	N/A
13.	I will allow for students to have a "break" when they appear to be stressed or anxious.	1	2	3	4	5	N/A
14.	It is okay to send a student out of the room when they are being disruptive.	1	2	3	4	5	N/A
15.	Using physical restraint is used as a last resort and only when a child is a threat to herself or others.	1	2	3	4	5	N/A

16. What is your program of study (your major)?

17. What is your age?

18. What strategies are you planning for managing classroom behaviors for students with exceptionalities in inclusive settings?

****Thank you for your time to complete this survey!****

References

Backhus, D. A., & Thompson, K. W. (2006). Addressing the nature of science in preservice science teacher preparation programs: Science educator perceptions. *Journal of Science Teacher Education, 17*(1), 65-81.

Billingsley, B., Israel, M., & Smith, S. (2011). Supporting new special education teachers: How online resources and web 2.0 technologies can help. *TEACHING Exceptional Children, 43*(5), 20-29.

Boe, E. E., Shin, S., & Cook, L. H. (2007). Does teacher preparation matter for beginning teachers in either special or general education? *Journal of Special Education, 41*(3), 158 170.

Bogdan, R. C., & Biklen, S. K. (2007). *Qualitative Research for Education: An Introduction to Theory and Methods.* (5th ed.). Boston, MA: Allyn and Bacon.

Brownell, M. T., Ross, D. D., Colon, E. P., & McCallum, C. L. (2005). Critical features of special education teacher preparation: A comparison with general teacher education. *The Journal of Special Education, 38*, 242–252.

Colvin, G. (2007). *7 Steps for Developing a Proactive Schoolwide Discipline Plan: A Guide for Principals and Leaderships Teams*. Thousand Oaks, CA: Corwin Press.

Cooper, J. E., Kurtts, S., Baber, C. R., & Vallecorsa, A. (2008). A model for examining teacher preparation curricula for inclusion. *Teacher Education Quarterly, 35*(4), 155-176.

Creswell, J.W. (2007). *Qualitative Inquiry and Research Design: Choosing among Five Traditions* (2nd ed). Thousand Oaks,CA: Sage Publications.

Darling-Hammond, L. (2010). *Evaluating teacher effectiveness: How teacher performance assessments can measure and improve teaching*. Retrieved from the Center for American Progress website:http://www.americanprogress.org/issues/2010/10/pdf/ teacher_effectiveness.pdf

Department of Education (ED), Office of Planning, Evaluation and, Policy Development. (2010). *A blueprint for reform. the reauthorization of the elementary and secondary education act.* US Department of Education. Retrieved from http://ezproxy.lib.ucf.edu/login?URL=http://search.ebscohost.com.ezproxy.lib.ucf.edu/login.aspx?direct=true&db=eric&AN=ED508795&site=ehost-live

Dillman, D.A., & Bowker, D. K. (2000). The web questionnaire challenge to survey methodologists. In B. Batinic, U.D.Reips, M. Bosnjak & A. Werner (Eds.), *Online Social Sciences* (pp. 53-72). Seattle, WA: Hogrefe & Huber.

Education of All Handicapped Children Act of 1975, Public L. No. 94-142 (1975).

Elementary and Secondary Education Act, 20 U.S.C. § 1400 et seq. (1965).

Fernandez, M. L., & Erbilgin, E. (2009). Examining the supervision of mathematics student teachers through analysis of conference communications. *Educational Studies in Mathematics, 72*(1), 93-110.

Forlin, C., Loreman, T., Sharma, U., & Earle, C. (2009). Demographic differences in changing pre-service teachers' attitudes, sentiments and concerns about inclusive education. *International Journal of Inclusive Education, 13*(2), 195-209.

Fuller, F. F. (1969). Concerns of teachers: A developmental conceptualization. *American Education Research Journal, 6,* 207-226.

Fuller, F. F., & Brown, O. (1975). Becoming a teacher. In K. Ryan (Ed.), *Teacher Education: Seventy-fourth Yearbook of the National Society for the Study of Education.* Chicago: University of Chicago Press.

Glaser, B., & Strauss, A. (1967). *The Discovery of Grounded Theory*: *Strategies for Qualitative Research.* Hawthorne, NY: Aldine Transaction.

Glesne, C. (2011). *Becoming Qualitative Researchers: An Introduction* (4th ed.). Boston, MA: Pearson.

Guskey, T.R. (2000). *Evaluating Professional Development.* Thousand Oaks, CA: Corwin Press.

Hair, F. H., Anderson, R. E., Tatham, R. L., & Black, W. C. (2009). *Multivariate Data Analysis* (7th ed.). Englewood Cliffs, NJ: Prentice Hall.

Hughes, C. E., Stapleton, C. B., & O'Connor, M. (2006). *The Evolution of a Framework for Mixed Reality Experiences in Emerging Technologies of Augmented Reality: Interfaces and Design*. Idea Group, Inc., Hershey: PA.

Jukes, I., & McCain, T. (2011). *Understanding the digital generation.* Retrieved from http://www.21stcenturyfluency.com/perspectives/UDG_perspective.pdf

Kaplowitz, M. D., Hadlock, T. D., & Levine, R. (2004). A comparison of web and mail survey response rates. *Public Opinion Quarterly, 68*(1), 94-101.

Katz, L. G. (1972). Developmental stages of preschool teachers. *Elementary School Journal, 73*(1), 50-54.

Keller, C. L., Brady, M. P., & Taylor, R. L. (2005). Using self evaluation to improve student teacher interns' use of specific praise. *Education and Training in Developmental Disabilities, 40*(4), 368-376.

Kenny, R. W. (1998). Reinventing undergraduate education: A blueprint for America's research universities. The Boyer Commission on Educating Undergraduates in the Research University. Retrieved from http://naples.cc.sunysb.edu/Pres/boyer.nsf/673918d46fbf653e852565ec0056ff3e/955b61ffddd590a852565ec005717ae/$FILE/boyer.pdf

Krueger, R., & Casey, M. A. (2000). *Focus Groups.* Thousand Oaks, CA: Sage.

Lane, K. L., & Carter, E. W. (2006). Supporting transition-age youth with and at risk for emotional and behavioral disorders at the secondary level: A need for further inquiry. *Journal of Emotional and Behavioral Disorders, 14*(2), 66-70.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry.* Newbury Park, CA: Sage.

McKinney, S. E., Haberman, M., Stafford-Johnson, D., & Robinson, J. (2008). Developing teachers for high-poverty schools. *Urban Education, 43*(1), 68-82.

Melnick, S. A., & Meister, D. G. (2008). A comparison of beginning and experienced teachers' concerns. *Educational Research Quarterly, 31*(3), 39-56.

Merriam, S. B. (1998). *Qualitative Research and Case Study Applications in Education.* San Francisco: Jossey-Bass Publishers.

National Center for, L. D. (2007). *Commission on no child left behind summary of key recommendations affecting students with learning disabilities.* National Center for Learning Disabilities.

National Council for Accreditation of Teacher Education Blue Ribbon Panel (2010). *Transforming teacher education through clinical practice: A national strategy to prepare effective teachers.* Retrieved from the National Council for Accreditation of Teacher Education website: http://www.ncate.org/LinkClick.aspx?fileticket= zzeiB1OoqPk%3d&tabid=715

No Child Left Behind Act of 2001, 115 U.S.C. § 1425 et seq. (2002).

Oblinger, D. G. (2008). Games and Learning: Digital Games have the potential to bring play back to the learning experience. *Educause Quarterly, 3*, 5-7.

Regan, K. S. (2009). Improving the way we think about students with emotional and/or behavioral disorders. *TEACHING Exceptional Children, 41*(5), 60-65.

Regan, K. S., & Michaud, K. M. (2011). Best practices to support student behavior. *Beyond Behavior, 20*(2), 40-47.

Reschly, D. J., & Holdheide, L. (2008). Innovation configurations: Implementing evidence based practices in teacher preparation (Report for National Comprehensive Center for Teacher Quality). Washington, DC: NCCTQ.

Reschly, D., Oliver, R., Schumaker, J., Smartt, S., & Holdheide, L. (2009). *Innovation configurations: Implementing evidence-based practices in teacher preparation*. Retrieved from the National Comprehensive Center for Teacher Quality website: http://www.tqsource.org/webcasts/innovationConfigurations_CriticalAreas/ DanReschly.pdf

Rock, M. L., Gregg, M., Gable, R. A., & Zigmond, N. P. (2009). Virtual coaching for novice teachers. *Phi Delta Kappan, 91*(2), 36-41.

Scheeler, M. C., McKinnon, K., & Stout, J. (2012). Effects of immediate feedback delivered via webcam and bug-in-ear technology on preservice teacher performance. *Teacher Education and Special Education, 35*(1), 77-90.

Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education and Treatment of Children, 31*(1), 351-380.

Slavin, R. E. (2007). *Educational Research in an Age of Accountability.* Boston, MA: Pearson.

Sugai, G., & Horner, R. H. (2006). A promising approach for expanding and sustaining the school-wide positive behavior support. *School Psychology Review, 35*, 245–259.

Vasquez, E., & Slocum, T. A. (2011). Evaluation of synchronous online tutoring for students at risk of reading failure. *Exceptional Children, 78*(2), 221-235.

Vince Garland, K. M., Vasquez, E., & Pearl, C. (2012). Efficacy of individualized clinical coaching in a virtual reality classroom for increasing teachers' fidelity of implementation of discrete trial teaching. *Education and Training in Autism and Developmental Disabilities, 47(*4), 502-515.

Westling, D. L. (2010). Teachers and challenging behavior: Knowledge, views, and practices. *Remedial and Special Education, 31*(1), 48-63.

White, M., & Mason, C. Y. (2006). Components of a successful mentoring program for beginning special education teachers: Perspectives from new teachers and mentors. *Teacher Education and Special Education, 29*(3), 191-201.

JOSOTL
JOURNAL OF THE SCHOLARSHIP OF TEACHING AND LEARNING

Mission

Founded in 2001, the Journal of the Scholarship of Teaching and Learning (JoSoTL) is a forum for the dissemination of the Scholarship of Teaching and Learning in higher education for the community of teacher-scholars. Our peer reviewed Journal promotes SoTL investigations that are theory-based and supported by evidence. JoSoTL's objective is to publish articles that promote effective practices in teaching and learning and add to the knowledge base.

The themes of the Journal reflect the breadth of interest in the pedagogy forum. The themes of articles include:

1. Data-driven studies: formal research projects with appropriate statistical analysis, formal hypotheses and their testing, etc. These studies are either with a quantitative or qualitative emphasis and authors should indicate the appropriate domain. Acceptable articles establish a research rigor that leads to significant new understanding in pedagogy
2. Reviews: Literature reviews illuminating new relationships and understanding, meta-analysis, analytical and integrated reviews, etc
3. Case studies: To be considered a case study, a manuscript should focus on an intense analysis of a specific teaching situation or problem that led to a solution. Case studies should have the following components: description of the teaching situation or problem, solution or solutions attempted, quantitative or qualitative analysis of the effectiveness of the solution, reflection on the implications and possible generalization to other settings or populations
4. Invited Comments: Comments about previously published manuscripts
5. Invited Essays: Discussion of a topic not tied to a previously published manuscript

Submissions

Authors are encouraged to submit work in one of the following categories:

- Data-driven studies: formal research projects with appropriate statistical analysis, formal hypotheses and their testing, etc. These studies are either with a quantitative or qualitative emphasis and authors should indicate the appropriate domain. Acceptable articles establish a research rigor that leads to significant new understanding in pedagogy.
- Reviews: Literature reviews illuminating new relationships and understanding, meta-analysis, analytical and integrated reviews, etc.
- Case studies: To be considered a case study, a manuscript should focus on an intense analysis of a specific teaching situation or problem that led to a solution. Case studies should have the following components: description of the teaching situation or problem, solution or solutions attempted, quantitative or qualitative analysis of the effectiveness of the solution, reflection on the implications and possible generalization to other settings or populations.
- Invited Comments: Comments about previously published manuscripts
- Invited Essays: Discussion of a topic not tied to a previously published manuscript.

All submissions for JoSoTL should be submitted using the online submission process on our website beginning on July 1, 2012.

Please follow this link to get more detailed information on the submission process for the Journal of the Scholarship of Teaching and Learning:
http://josotl.indiana.edu/about/submissions#onlineSubmissions

If you have additional questions or run into problems please contact the journal at josotl@iu.edu.

Journal of the Scholarship of Teaching and Learning, Vol. 13, No. 2, May 2013, pp. 150 – 152.

Style Sheet for the *Journal of the Scholarship of Teaching and Learning*

John Dewey[1] and Marie Curie[2]

Abstract: This paper provides the style sheet for the Journal of the Scholarship of Teaching and Learning. Manuscripts submitted for publication should adhere to these guidelines.

Keywords: radiation, metacognition, identity theory, constructivism, educational philosophy.

I. General Guidelines for the Manuscript.

The final manuscript should be prepared in 12-point, Times New Roman, and single-spaced. Submissions should be double-spaced. All margins should be 1 inch. The text should be fully left- and right-justified. The title (in 16 point bold) and author's name (in 12 pt. bold) should be at the top of the first page. The author's name should be followed by a footnote reference that provides the author's institutional affiliation and address. The abstract should be indented 0.5" left and right from the margins, and should be in italics.

Except the first paragraph in a section subsequent paragraphs should have a 0.5" first line indent. Use only one space after the period of a sentence (word processors automatically adjust for the additional character spacing between sentences). The keywords should be formatted identically to the abstract with one line space between the abstract and the keywords. Authors should use keywords that are helpful in the description of their articles. Common words found in the journal name or their title article are not helpful.

Pages should be unnumbered since they will be entered by the Journal editorial staff. We will also insert a header on the first page of the article, as above.

References should be incorporated in the text as authors name and date of publication (Coffin, 1993), with a reference section at the end of the manuscript (see below for the desired format for the references). Titles of articles should be included in the references in sentence case. Unless instructed otherwise in this Style Sheet, please use APA style formatting. Footnotes should incorporate material that is relevant, but not in the main text.

A. Plagiarism.

It is essential that authors refrain from plagiarism. Plagiarism is a violation of ethics and, in serious cases, will lead to a manuscript being rejected by this journal. No future manuscripts will be accepted from authors who have submitted a plagiarized manuscript.

[1]Department of Educational Philosophy, Indiana University Northwest, 3400 Broadway, Gary, IN 46408, jdewey@iun.edu.
[2]Institut Pasteur, University of Paris, 75015 Paris, France.

B. Unique work.

This journal does not accept previously published work. We also do not accept work that is being considered for publication by another journal. If your manuscript is accepted, you will be required to sign a form stating that your manuscript has not been previously published.

II. Section and Sub-Section Headings.

A. Major Sections.

Major section headings should be flush-left, bold-faced, and Roman numeral numbered. Major section headings should have one-line space before and after. The first paragraph(s) of the article do not require a major heading.

B. Sub-Sections.

Sub-section headings should also be flush-left, in italics, and alphabetically numbered. Sub-section headings should have a one-line space before and after. Sub-sub-sections should appear at the beginning of a paragraph (i.e., with an 0.5" indent, followed immediately by the text of the sub-sub-section), with the heading also in italics.

III. Tables and Figures.

Tables and figures should be inserted in the text where the author believes they best fit. They may be moved around a little to better correspond to the space requirements of the Journal. If necessary, tables and figures may occupy an entire page to ensure readability and may be in either portrait or landscape orientation. Insofar as possible, tables should fit onto a single page. All tables and figures should be germane to the paper. Tables should be labeled as follows with the title at the beginning (in bold), with data entries single-spaced, and numbered. Column labels should be half-line spacing above data.

Table 1. The title of the table.

Unit	Length, inches
Point	1/12
Pica	1/6

Figures should have their captions follow the image. Captions should be single-spaced, with title in bold. Additional text should not be in bold. The Editorial staff may adjust layout to allow optimal use of space.

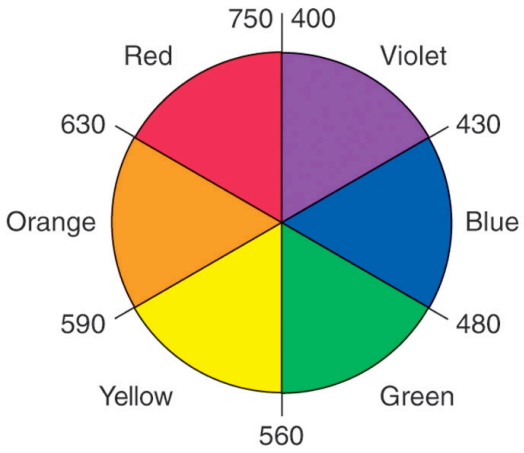


Figure 1. Color wheel with wavelengths indicated in millimicrons. Opposite colors are complementary.

Acknowledgements

Acknowledgements should identify grants or other financial support for this research by agency (source) and number (if appropriate). You may also acknowledge colleagues that have played a significant role in this research.

Appendix

Please insert any appendices after the acknowledgments. They should be labeled as follows:

Appendix 1. The Title of the Appendix.

References

Coffin, D.A. (1993). Using the competitive edge. *Journal of Economic Education*, *24* (1), 62-69.

Garcia, J. and Rodriguez, P. (2002). The determinants of football match attendance revisited: Empirical evidence from the Spanish football league. *Journal of Sports Economics*, *3* (1), 18-38.

Hamilton, S. J. (1995). My name's not Susie: A life transformed by literacy. Portsmouth, NH: Boynton/Cook Publishers.

Pappas, D. (2004). Fixing the fan cost index: A more realistic view. Retrieved April 21, 2004, from http://www.baseballprospectus.com/article.php?articleid=2790.

Contact Info for the Journal

JoSoTL Editorial Office

Indiana University Purdue University Indianapolis
755 W. Michigan St, UL 1180D
Indianapolis, IN 46202

josotl@iupui.edu

ISSN: 1527-9316