LUMAT 1(5), 2013

Modeling Active Engagement Pedagogy through Classroom Response Systems in a Physics Teacher Education Course

Marina Milner-Bolotin
Department of Curriculum and Pedagogy, Faculty of Education, The University of British Columbia • marina.milner-bolotin@ubc.ca

Heather Fisher
Department of Curriculum and Pedagogy, Faculty of Education, The University of British Columbia

Alexandra MacDonald
Department of Curriculum and Pedagogy, Faculty of Education, The University of British Columbia

Abstract One of the most commonly explored technologies in Science, Technology, Engineering, and Mathematics (STEM) education is Classroom Response Systems (clickers). Clickers help instructors generate in-class discussion by soliciting student responses to multiple-choice conceptual questions and sharing the distribution of these responses with the class. The potential benefits of clicker-enhanced pedagogy include: increased student engagement, reduced anxiety, continuous formative assessment, and enhanced conceptual understanding. Most studies, however, investigate the effects of clicker-enhanced instruction in large undergraduate STEM courses. The impact of this pedagogy on learning in small secondary or post-secondary classrooms is still relatively unexplored. The context of this study is a secondary physics methods course in a Teacher Education Program at a large Canadian university. One of the course assignments required future teachers to develop multiple-choice conceptual questions relevant to the secondary physics curriculum. This study investigates the impact of modeling clicker-enhanced active engagement pedagogy on future teachers' Content Knowledge, Pedagogical Knowledge, and Pedagogical Content Knowledge, as revealed by this assignment. The results of the study indicate that: (1) modeling clicker-enhanced pedagogy in a physics methods course increases future teachers' interest in active learning; (2) clicker-enhanced pedagogy is a powerful vehicle for developing Pedagogical Content Knowledge of future physics teachers; (3) clicker-enhanced pedagogy is a useful tool for teacher educators for identifying and addressing the gaps in the Content Knowledge of future teachers. This study sheds light on developing future teachers' capacities to design and implement instruction that is driven by conceptual questions in the presence or absence of technology and the impact of this process on their Pedagogical Content Knowledge and attitudes about conceptual STEM learning.

Keywords STEM teacher education, Technological Pedagogical Content Knowledge, Classroom Response Systems, conceptual science learning, misconceptions

1 Introduction

Science education has undergone dramatic changes in North American post-secondary institutions over recent decades (Deslauriers, Schelew, & Wieman, 2011; Mazur, 2009). There has been a continuous growth in the number of post-secondary Science, Technology, Engineering, and Mathematics (STEM) courses offering student-centered active learning environments. This shift towards the active engagement of students was prompted by a number of factors: increased knowledge of how students learn (Bransford, Brown, &

Cocking, 2002); availability of reliable instruments measuring conceptual understanding (Hestenes & Wells, 1992; Hestenes, Wells, & Swackhamer, 1992; R. Chabay, 2006); improved access to computerized ways to administer these instruments to obtain quantitative data on student achievement; the realization that traditional teacher-centered pedagogies are largely ineffective for promoting conceptual understanding (Hake, 1998) and positive attitudes about science (Adams & Wieman, 2011); the recognition that engaging all students in STEM learning is important (Hazari, Sonnert, Sadler, & Shanahan, 2010; Let's Talk Science, 2013); and lastly, the availability of new educational technologies to promote the active engagement of students (Milner-Bolotin, 2012; Milner-Bolotin, Kotlicki, & Rieger, 2007; Wieman, Adams, Loeblein, & Perkins, 2010).

One of the most common active engagement pedagogies in postsecondary STEM classrooms is Peer Instruction (PI) (Lasry, Mazur, & Watkins, 2008; Mazur, 1997). It utilizes Classroom Response Systems (clickers) to engage students in interactive activities and discussions through conceptual multiple-choice questions that target student difficulties, often referred to as misconceptions. PI has been found to be very effective in college STEM classrooms when students used either clickers (Hake, 1998; Milner-Bolotin, Antimirova, & Petrov, 2010) or flashcards (Lasry, 2008).

Despite its success at the college level, PI has not yet penetrated secondary school walls. One reason might be that PI pedagogy requires teachers to possess deep conceptual understanding of the content, and an awareness of potential student difficulties and multiple ways of addressing them. Student-centered pedagogies that promote active student engagement are much more demanding of teachers than their teacher-centered counterparts, where the teachers are in control of the classroom discourse at all times. Student-centered pedagogies require a significant intellectual investment from teachers, especially if they have not experienced them as learners. As STEM teacher educators, we believe that it is important to engage future teachers (FTs) in active learning during their Teacher Education Program (TEP). Science methods courses – in which FTs acquire subject-specific pedagogical knowledge – are a perfect opportunity for this. These courses allow the instructors to model active engagement pedagogies so FTs can experience them as both learners and teachers. Moreover, they provide opportunities for FTs to practice designing and modifying pre-existing subject-specific educational materials, such as conceptual STEM questions for PI pedagogy.

It might appear that the abundance of freely available STEM teaching materials online should reduce the pressure on teachers. Yet experienced educators know that in order to teach successfully, teachers must have pedagogical ownership of the materials they are using. This means that they have to be comfortable with these materials in terms of the required content, pedagogical, and technological knowledge. STEM educators who want to incorporate PI into their teaching should learn how to ask meaningful multiple-choice conceptual questions (Beatty, Gerace, Leonard, & Dufresne, 2006). In this paper, *conceptual science questions* are questions that involve little calculation, yet require a high level of critical thinking. Teachers should also be able to evaluate materials and questions designed by others in terms of accuracy and potential to uncover and address student difficulties, and then modify these to fit their own educational goals. In order to design and modify powerful questions, a teacher must be aware of potential student difficulties and ways of addressing them. Through experiencing PI in their own science methods courses, FTs have an opportunity to acquire these skills with the purpose of promoting active student engagement in their future classrooms. The following literature review provides a theoretical background for the current study.

2 Literature Review

This study was inspired by Shulman's 1986 paper, provocatively titled *Those Who Understand: Knowledge Growth in Teaching*. In this paper he asks a number of important questions that stemmed from a comparison of teacher education standards of the mid-80s to those of more than a century ago: "Where did the subject matter go? What happened to the content?" (p. 5) "What pedagogical prices are paid when the teacher's subject matter competence is itself compromised by deficiencies of prior education and ability? How does learning for teaching occur?" (p. 8). He was so concerned with the lack of research focus on the process of how subject matter transforms from the knowledge of the teacher into the content of instruction that he dubbed it the "missing paradigm". In Shulman's words, this missing paradigm was reflected in continuing to treat teaching "more or less generically, or at least as if the content of instruction was relatively unimportant" (p. 6).

In order to address the "missing paradigm" problem, he suggested a new perspective on teacher education: to consider teacher's knowledge as the combination of content and pedagogy to develop a content-embedded pedagogy, or Pedagogical Content Knowledge (PCK). According to Shulman, PCK is knowledge *for teaching specific content*: it is the knowledge of content, multiple ways of representing it, potential student difficulties, available resources, helpful illustrations, demonstrations, experiments, and examples. In the context of science education, it is also the knowledge of how to ask meaningful questions that promote conceptual science understanding, facilitate independent thinking, and encourage student interest in science. Almost three decades later, the concept of PCK (updated to Technological Pedagogical Content Knowledge (TPCK), to emphasise the knowledge of educational technologies (Koehler & Mishra, 2009)) is widely used in teacher education, yet the problem still persists. Few North American TEP graduates have acquired solid TPCK in their disciplines, necessary for becoming effective STEM teachers (Meltzer, Plisch, & Vokos, 2012). There are a few reasons for that in Canada. First, most TEPs are post-baccalaureate professional programs that accept students who have already earned a Bachelor's degree in their teachable subject. These TEPs are therefore relatively short, lasting between eight months and two years.

Second, since most FTs have earned a Bachelor's degree prior to enrolling into TEP, it is often assumed that they have mastered the necessary Content Knowledge and only lack the Pedagogical Knowledge required to teach it. Extensive research evidence indicates that science knowledge of many B.Sc. graduates in North America is very limited (Deslauriers et al., 2011; Wieman, 2012). This is exacerbated by the low entrance requirements of many TEPs. For example, at the University of British Columbia, the minimum Grade Point Average for entering into the TEP is 65%. Yet time constraints prompted many TEPs to reduce time devoted to methods courses where PCK is acquired (Milner-Bolotin, 2014). As a result subject-specific methods courses represent less than 10% of the total time devoted to teacher education. Third, TEPs currently offer an increasing number of general education courses, divorced from the subject-matter content, that are intended to equip FTs with "applied-to-all" pedagogical strategies. However, educational research shows that transfer of general pedagogical strategies into context specific teaching is one of the hardest tasks for novice teachers to master (Fensham, Corrigan, Dillon, & Gunstone, 2011). Fourth, in Canada, unlike some other countries, there are no institutionalized mentorship programs for new teachers during their formative years that could help new teachers enhance their TPCK.

The lack of time devoted by TEPs to helping FTs develop a solid PCK base posts significant challenges to STEM teacher education. TEPs are places where subject-specific PCK should be acquired. As the international tests of student STEM achievement repeatedly show, failing to help teachers acquire solid PCK has dire consequences for the students (Alphonso & Morrow, 2013; OECD, 2009a, 2009b, 2013). The lack of PCK also threatens the success of educational reforms that aim at promoting inquiry and active engagement in STEM (Crippen, Biesenger, & Ebert, 2010; Feder, 2010).

These are only a few reasons why we have no choice but to re-consider Shulman's vision for the role of PCK in teacher education. If STEM teaching is to be considered a serious profession in N. America (Tobias & Baffert, 2009), the process of educating FTs should be considered as seriously, as the education of future doctors, engineers and other professionals.

3 Methods

3.1 Goals and Research Questions

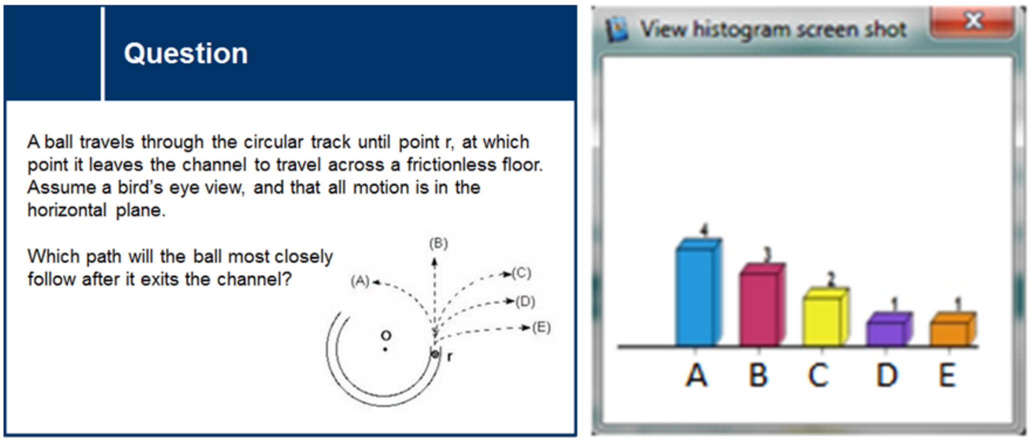
The goal of this study is to investigate the impact of modeling PI pedagogy in a small physics methods course on FTs' PCK, attitudes about the active engagement of students, and pedagogical decisions. In order to achieve this goal, we have developed the following research questions:

1. How does modeling active engagement pedagogy in the form of PI in a secondary physics methods course influence classroom dynamics and FTs' learning?

2. What is the evidence that modeling PI affects FTs' pedagogical decisions in their own classrooms?
3. How well are FTs able to:
    a. Critique, modify, and develop conceptual questions?
    b. Justify their content and pedagogical decisions?
4. How does PI pedagogy:
    a. Impact the instructor's ability to diagnose gaps and misconceptions in FTs' Content Knowledge?
    b. Allow the instructor to address these challenges?

3.2. Study Context

This study took place in the TEP at a large research university in Western Canada. The secondary physics methods course involved in the study was dedicated to helping FTs acquire physics-related TPCK. It was designed to model active engagement to ensure FTs experience it as learners and as teachers. This was accomplished in two ways. First, FTs participated in bi-monthly lab sessions, in which they engaged in hands-on physics activities, such labs, both low and high-tech (Milner-Bolotin & Moll, 2008; Moll & Milner-Bolotin, 2009), demonstrations, and problem-solving. Second, PI pedagogy was modeled in each class to facilitate discussions about: concepts related to relevant physics content, how students might approach it, where they might experience difficulties, and how a teacher could help them overcome these challenges. An example of a PI question is seen in Figure 1.



Question

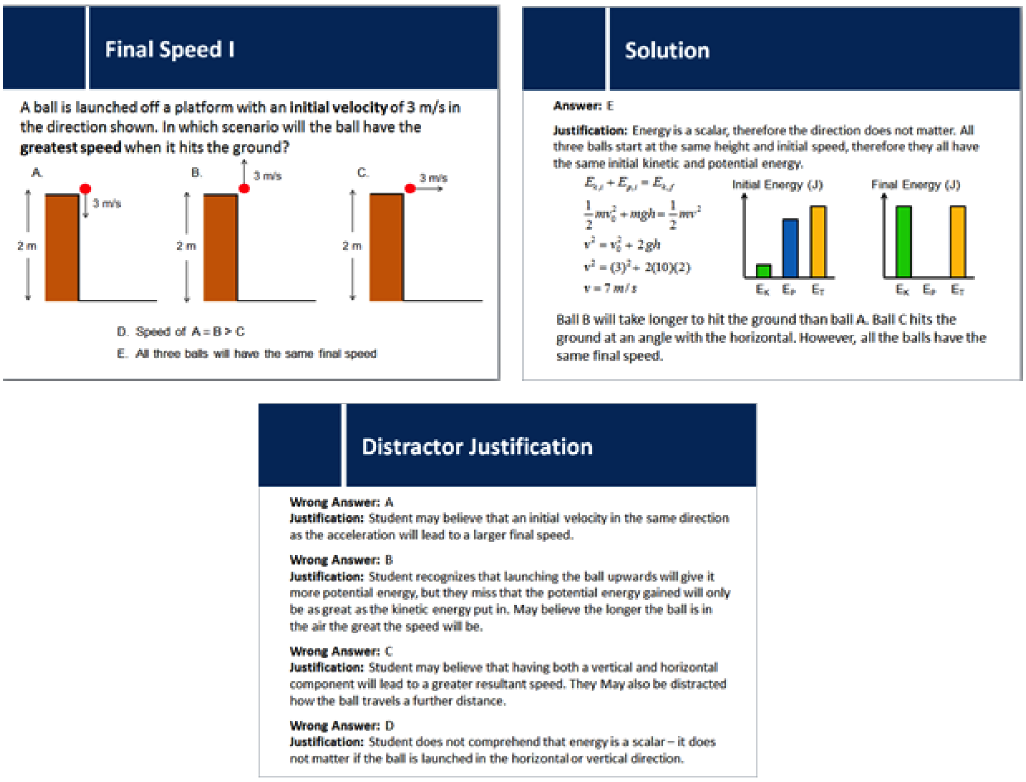
A ball travels through the circular track until point r, at which point it leaves the channel to travel across a frictionless floor. Assume a bird's eye view, and that all motion is in the horizontal plane.

Which path will the ball most closely follow after it exits the channel?

Figure 1: Example of a conceptual science question and the related histogram of student responses.

Questions were presented using PI, allowing FTs to answer them in a risk-free anonymous manner. The following sequence illustrates one version of this pedagogy: (a) the instructor poses a question and asks FTs to respond individually using clickers, (b) a histogram of their responses is revealed (Figure 1), (c) FTs discuss their responses in small groups focusing on how secondary students might respond to the question, (d) FTs answer the same question individually a second time(Kalman, Milner-Bolotin, & Antimirova, 2010). Other methods of implementing PI pedagogy were also modeled and discussed. For example, an instructor might decide not to reveal the response histogram after the first vote, but ask students to discuss their responses with each other in small groups. Another version of PI asks students to discuss the question in groups right away, before individual responses. Alternatively, the dynamic histogram can be revealed to the students during voting. While the traditional PI pedagogy was used most of the time, FTs were exposed to different versions of PI pedagogy in order to encourage them to be open to adjusting available pedagogies to their own teaching style and pedagogical goals.

As a part of the course, FTs were required to complete an assignment where they designed (or substantially modified) at least five conceptual multiple-choice physics questions. FTs were asked to include meaningful distractors (incorrect responses to multiple-choice questions), an explanation of the correct answer, and the logic behind each distractor. Some of the FTs also chose to include notes clarifying their pedagogical choices (Figure 2).



Final Speed I

A ball is launched off a platform with an **initial velocity** of 3 m/s in the direction shown. In which scenario will the ball have the **greatest speed** when it hits the ground?

D. Speed of A = B > C

E. All three balls will have the same final speed

Solution

Answer: E

Justification: Energy is a scalar, therefore the direction does not matter. All three balls start at the same height and initial speed, therefore they all have the same initial kinetic and potential energy.

$$E_{k,i} + E_{p,i} = E_{k,f}$$

$$\frac{1}{2}mv_0^2 + mgh = \frac{1}{2}mv^2$$

$$v^2 = v_0^2 + 2gh$$

$$v^2 = (3)^2 + 2(10)(2)$$

$$v = 7\ m/s$$

Ball B will take longer to hit the ground than ball A. Ball C hits the ground at an angle with the horizontal. However, all the balls have the same final speed.

Distractor Justification

Wrong Answer: A
Justification: Student may believe that an initial velocity in the same direction as the acceleration will lead to a larger final speed.

Wrong Answer: B
Justification: Student recognizes that launching the ball upwards will give it more potential energy, but they miss that the potential energy gained will only be as great as the kinetic energy put in. May believe the longer the ball is in the air the great the speed will be.

Wrong Answer: C
Justification: Student may believe that having both a vertical and horizontal component will lead to a greater resultant speed. They May also be distracted how the ball travels a further distance.

Wrong Answer: D
Justification: Student does not comprehend that energy is a scalar – it does not matter if the ball is launched in the horizontal or vertical direction.

Figure 2 An example of a multiple-choice question, its solution and justification of the distractors

3.3. Participants

Thirteen future secondary physics teachers were enrolled in the required physics methods course in the fall of 2012. In Canada, physics is taught in grades 8-10 as part of the science studies, and in grades 11 and 12 as a separate subject. Among the study participants, 11 FTs were enrolled in a single-year TEP and completed the course in their first term of study. Of the remaining two, one was completing course requirements for recertification from a previous TEP, and the second was in the fourth year of a concurrent Education and Physics program. Twelve FTs had secondary physics as their primary teachable subject, and one had physics as their secondary teachable subject. Nine did not have a secondary teachable subject.

The FTs had diverse undergraduate backgrounds, but all had obtained at least a B.Sc. degree or its equivalent. Nine had Physics or combined Physics and Astronomy B.Sc. or M.Sc. degrees. Three had equivalent engineering degrees, and one had a Chemistry B.Sc. degree. Only one of them earned their undergraduate B.Sc. degree outside of Canada (in India).

3.4. Quantitative Measures: Evaluation of the Conceptual Science Questions' Quality

Quantitative data in this study came from the analysis of each conceptual science question FTs submitted for this course assignment. A quantitative assessment rubric was designed to evaluate their Content and Pedagogical Knowledge as demonstrated by each conceptual question (Table 1). Each of these categories was explored by evaluating multiple dimensions of that knowledge and rating it on a relevant scale. The rubric, developed specifically for this study, draws from Shulman's PCK model (1986).

The three researchers in this study, the course instructor and two Research Assistants, served as evaluators. The instructor, a physics educator with 20 years of secondary and post-secondary physics teaching experience and 5 years of teaching FTs was the leader of the study. She also has more than 20 years of physics education research experience. One of the Research Assistants holds a B.Sc. degree in Physics and had some prior experience in science education research. She was also the Teaching Assistant for this course. The other Research Assistant holds a B.Sc. in biological and political sciences and possesses expertise in conceptual question design and science education research. Due to her limited physics expertise, she focused on general pedagogical aspects of the questions.

The researchers independently rated each of the questions using the rubric (Table 1). After all questions had been evaluated, the ratings were compared. Each case where discrepancies existed was discussed and resolved. This happened in fewer than 10% of the total ratings and, most of the time, the ratings were off by only 1 value.

Due to the diverse backgrounds of the research team, a complete agreement of the ratings was not expected. When this happened, each evaluator presented their reasons for the rating. The rating was then discussed and a final value was assigned. When making these decisions it was especially important to consider the backgrounds of the researchers, as each brought different experience and insight to the discussion.

Table 1: A rubric for evaluating pedagogical and content effectiveness of conceptual questions (CA – Correct Answer and IA – Incorrect Answer)

Level	Content Knowledge measures				Pedagogical Knowledge measures				
	Cognitive level	Targeting student difficulties	Science accuracy	Distractors' quality	Answer justification	Question clarity	Multiple representations MRs	Part of a sequence	Originality
1	Knowledge	Doesn't target any concept. difficulty	Major mistakes in the question and in solutions	All irrelevant distractors	No answer justification	Unclear, misleading and inaccurate	1 MR	No	Exactly copied from a known source
2	Comprehension	Targets a minor concept ineffectively	An accurate question but an inaccurate and unclear solution		Incomplete CA justification; no IA justification		2 MRs	Yes	Copied with minor modifications
3	Application	Targets a minor concept effectively	The question is clear; the solution is accurate but unclear	Half of distractors are meaningful	Incomplete CA and IA justification	Minor problems in the question or in the solution	3 MRs		Copied with some modifications
4	Analysis	Targets a few conceptual difficulties	Both question and solution are somewhat clear and accurate		Complete and accurate CA justification only		4 MRs		Copied with interesting modifications
5	Synthesis/ evaluation	Clearly targets major conceptual difficulties	Both question and solution are clear and accurate	All distractors are meaningful	Complete and accurate CA and IA justification	Both quest. and solution are very clear and accurate	5 MRs		Original question

3.4.1 Content Knowledge

Four measures were evaluated within the Content Knowledge category, and all were rated on a five-point Likert scale (Table 1). First, each question was rated on its cognitive level, as identified in Bloom's taxonomy (Bloom, 1956). For clarity and consistency between raters, a list of verbs associated with each level of Bloom's Taxonomy, as seen in Morrison and Walsh Free (2001), was utilized (Table 2). Synthesis and evaluation were grouped into a single category, as it is difficult to create high-level multiple-choice questions, and it was therefore unlikely that these would be present.

Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation
Define	Describe	Apply	Analyze	Compose	Appraise
Identify	Differentiate	Calculate	Categorize	Construct	Assess
Know	Discuss	Classify	Compare	Create	Evaluate
List	Explain	Develop	Contrast	Design	Judge
Name	Rephrase	Examine	Distinguish	Formulate	
Recognize	Restate	Solve	Determine	Modify	
State	Reword	Use	Investigate	Plan	

Table 2 Verbs Association with Categories of Cognition for Bloom's Taxonomy of Educational Objectives as seen in Morrison & Walsh Free (2001)

Second, questions were rated on whether they targeted specific student difficulties. This is important, as conceptual questions that are not tailored to address specific conceptual challenges hold limited pedagogical promise (Beatty et al., 2006). In this area, it was particularly important to account for raters' backgrounds, as an instructor with more than 20 years of teaching experience is better equipped to speak to the conceptual difficulties of novice physics learners. Hence, in situations where inter-rater disagreements occurred, deference was given to the instructor and Research Assistant, who had significantly more physics content expertise.

Third, the questions and solutions were rated for their science accuracy. This included two characteristics of the question and its solution: clarity and accuracy. In order for a question to be valuable, the content must not only be accurate, but clearly presented. FTs should be able to clearly convey scientific concepts. For example, if a question was clear but the answer was unclear, even though it accurately dealt with a scientific concept, it would be rated as a 3. One of the team members, who did not have an extensive physics background focused on rating the clarity of the questions, while the other two members rated questions' clarity and accuracy.

Finally, the quality of the distractors was rated in relation to their value to the question. High quality distractors play an important role in multiple-choice questions and should differentiate between those students who understand concepts and those who do not. FTs were not required to include a specific number of distractors. As the number of distractors varied from two to four, it was important to develop this scale in percentages of distractors instead of absolute numbers. This allowed for comparability between questions with differing numbers of distractors.

3.4.2 Pedagogical Knowledge

Five measures were used to evaluate the Pedagogical Knowledge category of conceptual questions designed by FTs. Three used a five-point Likert scale, while two others used a ratio and a binary scale.

First, each solution was rated on the quality of provided justifications. The ratings were determined using a five-point Liker scale. The intention of the assignment was to have solutions that justified the correct and incorrect answers, as both play an important role in constructing student understanding. Providing explanations for both correct and incorrect answers, as well as understanding what role distractors play in student learning, demonstrates FTs' Pedagogical Knowledge. Solutions were rated for completeness and correctness. A solution that included only an explanation of the correct answer would be rated lower than one which also explained why each distractor was incorrect.

Second, the questions were rated for clarity. Clarity is important because an unclear question undermines its pedagogical intent and misleads students. This measure was also rated on a five-point Likert scale, but unlike the other measures, descriptions of the ratings were only provided for the extremes and the center of the scale (i.e., 1, 3, and 5). This allowed raters to interpret the question's clarity and use their judgment to rank questions in the intermediate rating spaces (i.e., 2 and 4).

Third, the questions were rated for the number of multiple representations (MRs) present. They could include, but were not limited to, verbal, graphical, schematic, diagrammatic, and algebraic representations. Demonstrating a concept in multiple ways acknowledges that students learn from different forms of content (Kohl & Finkelstein, 2007). A representation was only counted if it was included in the question, in the solution options or in conjunction with the solution. This ratio scale corresponded with the number of representations present in these areas. Any disagreements between raters were discussed and resolved, as the method of rating this measure was clearly outlined.

Fourth, each question was rated on a binary scale identifying whether or not it was part of a sequence. Sequential questions demonstrate an awareness of how concepts can be broken down into constituent parts, and thus made more approachable for students. This is an important pedagogical skill. Independent questions were rated as a 1, while questions that were part of a sequence were rated as a 2.

Finally, questions were rated for originality using a five-point Liker scale. There are many available resources for conceptual physics questions (i.e. www.compadre.org): from the questions supplied by textbook publishers to questions designed by physics education researchers (Mazur, 1997) or by individual physics teachers who might not be familiar with physics education research. One of these resources is the Mathematics and Science

Teaching and Learning through Technology database of conceptual questions, designed by the research team (Milner-Bolotin, 2013). This resource was used extensively in the course. FTs were encouraged to use its questions, modify them and expand on them. For example, they could have started with a borrowed question and developed a sequence of questions around it. Even if the question wasn't novel, FTs had to develop and justify meaningful distractors and clearly describe its pedagogical purpose. While there is no scarcity of available conceptual questions, not all of these questions are accurate or pedagogically effective. As such, it is important for FTs to be able to evaluate the quality of available resources, and to modify them to meet their needs. This measure was very important, but it was difficult to know prior to rating the questions if the three raters would be able to identify whether questions were original, directly copied from a known source, or copied from a known source, but modified. Thus, this measure relied heavily on the research and teaching experiences of the course instructor. Disagreements between raters were discussed, and deference was given to the instructor's experience and broader awareness of available resources.

3.5 Qualitative Measures

Qualitative data consisted of the semi-structured interviews with FTs, classroom observations, and FTs' and course instructor's reflections.

3.5.1 Interviews with FTs

Before entering and after completing their ten-week school practicum, the FTs were invited to participate in a semi-structured interview with the two Research Assistants. Eight FTs participated. These interviews allowed the Research Assistants to probe the FTs' PCK, as well as their awareness of this knowledge. The interviews were conducted after the completion of their physics methods course, and had no possible impact on FTs' academic performance. The interviews were transcribed verbatim and anonymized, after which the entire research team gained access to them. The interviews were analyzed for themes relating to Content and Pedagogical Knowledge, their interactions, and the development of PCK. FTs were also asked to reflect on their experience with active engagement in the classroom, how it impacted them as learners, and whether or not they intended to take PI pedagogy into their own classrooms. In addition to the individual semi-structured interviews, six FTs participated in a focus group discussion.

3.5.2 Classroom observations

One Research Assistant was present during all class meetings and made observations of classroom dynamics, FTs' participation, their activities and their interactions.

3.5.3 FTs' reflections

FTs reflected on their course experiences twice – at the midterm point and at the end of the term. At the midterm point, they reflected on how PI impacted their behavior as learners in their physics methods course and as teachers in their two-week short practicum that took place during the 8th and 9th weeks of the course. At the end of the term, they reflected on what they learned in the course, what they felt was not explored in detail and what they wish had been included. This open reflection allowed them to express opinions on a variety of topics, not limited to the role of PI pedagogy.

3.5.4 Instructor's reflections

One of the central research questions aimed to explore how PI can impact an instructor's ability to address gaps in FTs' PCK. As such, the instructor reflected on their experiences of implementing PI, including how active engagement influenced classroom dynamics and learning.

4 Results

Since the study followed a mixed-method design, we will report the results for the quantitative part first, followed by the qualitative data analysis. In the following section we will discuss how the quantitative and qualitative findings support each other.

4.1 Quantitative Results

All FTs submitted at least five multiple-choice conceptual questions and corresponding answers: 72 questions in total (Table 3).

	Number of questions submitted per FT	Number of FTs who submitted the indicated number of questions	Percentage of FTs who submitted the questions
	5	8	62%
	6	3	23%
	7	2	15%
Total	**72**	**13**	**100%**

Table 3 Distribution of conceptual questions submitted by FTs

Table 4 shows the results of the analysis that include the average rating value for each of the rubric's categories and the frequency of each of the values in the ratings. In the Content Knowledge part of the data, these results indicate that FTs submitted multiple-choice questions that were at the middle levels of Bloom's taxonomy. The average cognitive level of the questions was 3.2, corresponding to the application level. Moreover, most of the questions specifically targeted student conceptual difficulties that FTs were able to identify and address. Most of the questions were scientifically accurate and had meaningful distractors that FTs were able to justify. The last two statements are very important, as they indicate FTs' growing PCK.

In the Pedagogical Knowledge categories, the data indicate that FTs were able to provide pedagogically sound justifications for their answers, their questions were clear and appropriate for secondary students, FTs used most often 2 or 3 multiple representations, and most of the questions were part of a pedagogical sequence. Based on our knowledge, most of the questions included original elements designed by FTs and were not copied directly from known sources. This was also verified during the interviews, as FTs felt pride and ownership of their questions. In addition, 38% of the FTs submitted more questions that they were asked for.

		Content Knowledge				**Pedagogical Knowledge**				
		Cognitive level (Bloom) (1-5)	Targets student difficulties (1-5)	Science accuracy (1-5)	Distractors' quality (1-5)	Answer justification (1-5)	Question clarity (1-5)	Multiple repre-sentations (1-5)	Part of a sequence (1, 2)	Originality (1-5)
Average		3.04	4.38	4.59	4.06	4.58	4.58	2.47	1.9	1.8
Frequency of ratings	1	2	0	0	2	0	0	9	7	24
	2	9	2	1	1	2	3	23	65	39
	3	45	12	9	2	17	7	38		8
	4	16	15	22	14	27	7	1		1
	5	0	43	40	53	26	55	1		0
	Total	**72**	**72**	**72**	**72**	**72**	**72**	**72**	**72**	**72**

Table 4: Summary of the results of the analysis of multiple-choice conceptual questions developed by FTs: The most frequent rating in each rubric category is shaded

4.2 Qualitative Results

The qualitative data included FTs' responses during the semi-structured interviews, classroom observations during PI-generated discussions, and FTs' anonymous reflections and feedback. Classroom observations focused on how FTs responded to PI pedagogy and how their answer distributions changed as a result of in-class discussions. Since PI was modeled during most lessons (at least one clicker question was asked in every lesson) it was important to observe how FTs engaged with conceptual multiple-choice questions and if they found these questions valuable for their own learning. Figure 1 shows their responses to a conceptual question from a well-known introductory physics instrument, called the

Force Concept Inventory (Hestenes et al., 1992). This question clarifies student understanding of the concept of inertia and does not require any advanced mathematical knowledge. While the correct answer to the question is B, only 3 out of 11 FTs who participated in the poll chose it. This question demonstrates that FTs themselves, despite having earned at least a B.Sc. or its equivalent in physics, face significant conceptual difficulties in understanding basic topics relevant to the secondary physics curriculum. These difficulties would not have been revealed if FTs were not asked to vote on this question and the results of the voting were then shared with the group Moreover, as soon as the answer distribution was revealed, FTs felt compelled to discuss the question and justify their answers. As a result, they were able to appreciate the difficulties their future students might encounter while exploring the concept of inertia and Newton's laws of motion.

FTs' interviews and feedback during and after class also indicated how their engagement with PI had affected their PCK and their views on science teaching. One FT indicated that "clicker questions are a great way to test student understanding of a topic and distractors can test misconceptions", while another one mentioned "the importance of using conceptual questions in assessment of/for/as learning". A third FT pointed out: "I've learnt the benefits of "clicker" conceptual type questions and how to implement them into the class". While not all FTs mentioned PI in their feedback, most of them noted the value of conceptual questions in STEM teaching emphasizing that "physics concepts can continually be reinforced and reimagined".

In the following section we discuss these results and focus on their implications for STEM teacher education.

5 Discussion

The results of the study will be discussed in light of the four research questions the study aimed to address.

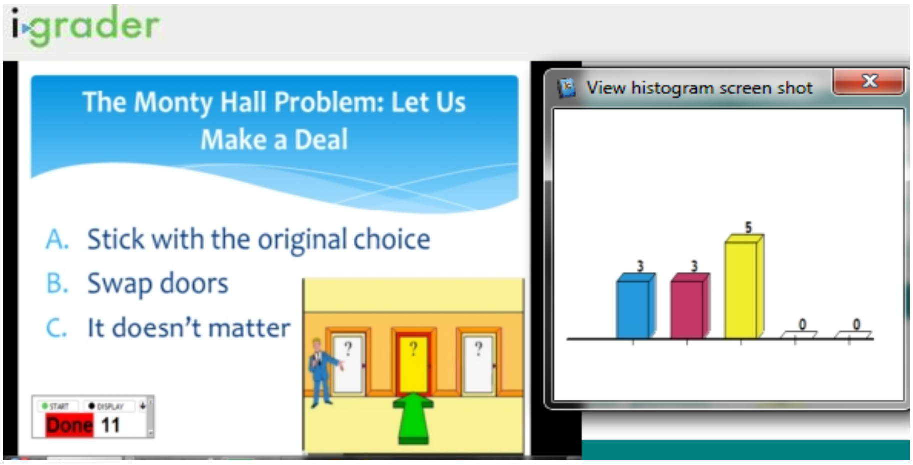
5.1 Research Questions

Research Question 1: How does modeling active engagement in the form of PI in the secondary physics methods course influence classroom dynamics and FTs' learning?

Answer: We have collected overwhelming evidence that PI can be successfully implemented in a physics methods course and it has a potential to facilitate student-centered active engagement pedagogy. Most of the conceptual questions chosen to be discussed during class meetings produced conflicting results (Figure 1). This prompted FTs to articulate and expose their ways of thinking often suggesting additional questions and activities that would be useful in a science classroom. One example of a question was particularly revealing. FTs were asked to respond to a well-known probability puzzle (Figure 3). The problem is based on the American television show "Let's Make a Deal" and named after its original host Monty Hall. The problem is stated as follows:

> Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which is known to conceal a goat. He then says to you, "Do you want to change to door No. 2?" Is it to your advantage to swap your door of choice? (http://en.wikipedia.org/wiki/Monty_Hall_problem)

FTs were asked to vote anonymously on their choice just to confirm their understanding. As it happened, this problem was discussed earlier in their mathematics methods course. Nevertheless, only 3 out of 11 FTs answered it correctly (to switch doors). However, most of them were confident they understood it and they were ready to explain it to their future students.



The Monty Hall Problem: Let Us Make a Deal

A. Stick with the original choice
B. Swap doors
C. It doesn't matter

Figure 3: Monty Hall problem and teacher-candidates' responses to it (B is the correct answer was chosen by only 3 teacher-candidates out of 11 who voted – 27%).

This teachable moment was a powerful reminder of the importance of active student engagement either in secondary or teacher education classrooms: students have to be actively engaged cognitively and emotionally to gain meaningful understanding.

Research Question 2: What is the evidence that modeling active engagement pedagogy affects FTs' pedagogical decisions in their own classrooms?

Research Question 3: How well are FTs able to:

a. Critique, modify, and develop conceptual questions?
b. Justify their content and pedagogical decisions?

Answer: The second and third research questions will be answered below. The quality of in-class discussions, FTs' feedback, as well as the quality of their conceptual questions' (Table 4) can speak to the pedagogical effectiveness of PI in science methods courses and increased PCK skills of FTs to ask pedagogically effective conceptual questions. As the course progressed, FTs became much more focused not only on answering conceptual questions correctly in class, but also on analyzing science concepts targeted by these questions and discussing the distractors. Designing powerful multiple-choice science questions is difficult (Beatty et al., 2006) as it requires teachers to possess deep PCK of the subject. Novice teachers think of multiple-choice questions as belonging solely to the low levels of Bloom's taxonomy (Knowledge and Comprehension), while experts are able to design more sophisticated multiple-choice questions at the higher levels (Application, Analysis, and Synthesis) (Lord & Baviskar, 2007). Most of the questions designed by the FTs (Figure 1) belonged to the Application and Analysis levels (Table 4). More than 90% of them were part of a sequence, thus opening doors to an in-depth investigation and in-class discussion of more complex concepts. About 87% of the questions used more than one representation (words, diagrams, graphs, algebraic expressions) in the question and in the solution. Approximately 86% of the questions were rated high or very high on the scientific accuracy of the question and the solution. In addition, from the FTs' pedagogical notes to their questions, it was clear that most of the questions (~ 95%) targeted very specific science concepts in an attempt to address potential student difficulties. Not surprisingly none of the questions were entirely original, yet most of them were new to FTs. Many of the questions included significant pedagogically justified modifications and improvements. In addition, most of the FTs used PI pedagogy in their own classrooms during the practicum. While only three FTs were able to use clickers (as they were available at the schools they were assigned to teach at), the rest used either flashcards, or facilitated PI-style discussions. We believe that this evidence strongly supports the claim that FTs were able to critique, modify and develop conceptual questions, as well as justify their pedagogical decisions. This supports a conclusion that PI should have a place in science methods courses, as it effectively promotes the development of FTs' PCK and their ability to ask pedagogically effective conceptual multiple-choice questions.

Research Question 4: How does PI pedagogy

a. Impact the instructor's ability to diagnose gaps and misconceptions in FTs' Content Knowledge?
b. Allow the instructor to address these challenges?

*Answer: T*he last question of the study focuses on the instructor teaching a science methods course, rather than on the FTs in the course. This is a very important question, as it emphasizes the value of continuous formative assessment and gauging FTs' Content Knowledge. While all FTs have fulfilled the formal requirements for admissions (B.Sc. degree or its equivalent), PI revealed significant gaps in their Content Knowledge, thus allowing the instructor to develop strategies for addressing the problem. As this science methods course had only 39 hours (13 weeks of 3 hours per week), the instructor did not have an opportunity to address all the topics in the secondary physics curriculum. However, PI helped her to engage FTs in the topics that were the most difficult or confusing for them. This was done in a non-threatening and supportive manner and, as a number of FTs indicated in their feedback, allowed them to improve their PCK without being embarrassed or afraid to reveal their challenges.

5.2 Limitations

STEM education is a complex research field with many challenges. Education, in general, is a difficult area to study because the context of an educational intervention can drastically impact its effects and pedagogical implications. STEM educational researchers have often struggled to replicate findings across different disciplines and contexts. Implementing successful interventions across different disciplines, at different institutions, that involve different instructors is even more challenging (Wieman, 2012). This makes it difficult to expand successful educational interventions beyond the pilot stage, since instructors are often apprehensive and are rarely compelled enough to make significant changes in their own classrooms. What works in one context might not always work in another, discouraging the widespread implementation of educational innovations. This is further exacerbated by the time-intensive process of adapting materials and approaches to new contexts, as every educational context has its own unique characteristics.

Another limitation of investigating technology-enhanced STEM pedagogies is the risk of getting caught up in the technology itself instead of the underlying principles of the intervention. We have outlined in this paper how to apply the principles of technology-enhanced pedagogy in high-tech or low-tech settings; focusing on the role the technology plays to facilitate learning instead of the technology itself. It is important for an ongoing dialogue between instructors interested in implementing similar approaches and educational researchers, to ensure that the results are not mitigated by the allure of focusing on the technology alone.

Finally, it is difficult to conclusively identify the impact of an intervention on FTs. We have developed a rubric to begin dissecting the process describing how future STEM teachers acquire PCK in relevant methods courses, but it is only the beginning. Further research is necessary to validate these findings and to apply them to novel contexts. As demonstrated above, this is an extremely difficult but also a very worthwhile task.

6 Conclusions and Future Directions

Secondary STEM FTs are faced with a variety of challenges during their TEP, exacerbated by the assumption that they have already mastered their content and, therefore, only need to acquire general (not-subject-specific) Pedagogical Knowledge. In practice FTs often lack the content expertise in the area they will be certified to teach (Wieman, 2012). This discrepancy, between the Content Knowledge they were supposed to possess prior to being admitted into the program and the Content Knowledge they actually have, places a large amount of pressure on them and on the methods courses' instructors. TEPs should begin to acknowledge and address this issue in order to enhance FTs' PCK.

Consequently, FTs who are not confident in their Content Knowledge are likely to be unable to apply the general pedagogies to the STEM context. This is also relevant to technology-enhanced STEM teaching. These pedagogies require a much deeper PCK than a traditional lecture-style method. Moreover, it is unreasonable to expect that FTs who have not experienced these pedagogies as students or as FTs will be open to these teaching methods.

FTs are the agents of change of the STEM education of tomorrow. Therefore, a call to change STEM education should begin with FTs. The STEM teachers we educate today should possess a solid PCK foundation coupled with the knowledge of modern technology-enhanced pedagogies, such as PI. These active engagement pedagogies can help teacher educators to close the gap between STEM teachers we certify and STEM teachers we would like to be teaching our children. This paper begins to address this gap by evaluating the pedagogical effectiveness of a clicker-enhanced pedagogy, PI, in a physics methods course. The study also assesses FTs' PCK gains as a result of their active engagement with conceptual science questions in two capacities: as students and as FTs. This research has also demonstrated the value of PI in aiding instructors in identifying gaps in FTs' PCK and in addressing these gaps in a supportive way. Most FTs conveyed their appreciation of how clickers were used in the course and expressed their interest in using PI with their future students. There is a need for a study that will expand upon these findings, for example, by exploring the secondary impact of this intervention on FTs' self-efficacy. Regardless of the intervention, it is essential to consider how changes to TEPs can impact curriculum that graduates from those programs enact in their future classrooms.

References

Adams, Wendy K., & Wieman, Carl E. (2011). Development and validation of instruments to measure learning of expert-like thinking. *International Journal of Science Education, 33*(9), 1-24.

Alphonso, Caroline, & Morrow, Adrian. (2013, August 28). Ontario Teachers Need Better Math Training, Minister Says, *The Globe and Mail*. Retrieved from http://www.theglobeandmail.com/news/national/education/ontario-testing-shows-dip-in-student-math-skills-for-fifth-year-in-a-row/article13996303/

Beatty, Ian D., Gerace, William J., Leonard, William J., & Dufresne, Robert J. (2006). Designing effective questions for classroom response systems teaching. *American Journal of Physics, 74*(1), 31-39.

Bloom, Benjamin S. (1956). *Taxonomy of Educational Objectives: Cognitive Domain*. (Vol. 1). New York: Longman.

Bransford, John D., Brown, Ann L., & Cocking, Rodney R. (2002). *How People Learn: Brain, Mind, Experience, and School*. Washington, DC: The National Academies Press.

Crippen, K. J., Biesenger, K. D., & Ebert, E. E. (2010). Using professional development to achieve classroom reform and science proficiency: An urban success story from Southern Nevada. *Professional Development in Education, 36*(4), 637-661.

Deslauriers, Louis, Schelew, Ellen, & Wieman, Carl. (2011). Improved learning in a large-enrollment physics class. *Science, 332*(6031), 862-864.

Feder, Toni. (2010). Europe reflects on a decade of higher education reforms. *Physics Today, 63*(May), 24-27.

Fensham, Peter J., Corrigan, Deborah, Dillon, Justin, & Gunstone, Richard. (2011). Knowledge to Deal with Challenges to Science Education from Without and Within *The Professional Knowledge Base of Science Teaching* (pp. 295-317): Springer Netherlands.

Hake, Richard R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66*(1), 64-74.

Hazari, Zahra, Sonnert, Gerhard, Sadler, Philip M., & Shanahan, Marie-Clair. (2010). Connecting high school physics experiences, outcome expectations, physics Identity, and physics career choicer: A gender study. *Journal of Research in Science Teaching, 47*(8), 978-1003.

Hestenes, D., & Wells, M. (1992). A Mechanical Baseline Test. *The Physics Teacher, 30*, 159-166.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher, 30*(3), 141-157.

Kalman, Calvin S., Milner-Bolotin, Marina, & Antimirova, Tetyana. (2010). Comparison of the effectiveness of collaborative groups and peer instruction in a large introductory physics course for science majors. *Canadian Journal of Physics, 88*(5), 325-332.

Koehler, Mathew J., & Mishra, Punya. (2009). What is technological pedagogical content knowledge? *Contemporary Issues in Technology and Teacher Education, 9*(1), 60-70.

Kohl, Patrick B., & Finkelstein, Noah D. (2007, August). *Expert and novice use of multiple representations during physics problem solving*. Paper presented at the 2007 Physics Education Research Conference, Greensboro, NC.

Lasry, Nathaniel. (2008). Clickers or flashcards: Is there really a difference? *The Physics Teacher, 46*(May), 242-244.

Lasry, Nathaniel, Mazur, Eric, & Watkins, Jessica. (2008). Peer Instruction: From Harvard to the two-year college. *American Journal of Physics, 76*(11), 1066-1069.

Let's Talk Science. (2013). Spotlight on science learning: The high cost of dropping science and math. Toronto, ON: Let's Talk Science.

Lord, Thomas, & Baviskar, Sandhya. (2007). Moving students from information recitation to information understanding: Exploiting Bloom's Taxonomy in creating science questions. *Journal of College Science Teaching, 37*(March-April), 40-45.

Mazur, Eric. (1997). *Peer Instruction: User's Manual*. Upper Saddle River, NJ: Prentice Hall.

Mazur, Eric. (2009). Farewell, Lecture? *Science, 323*, 50-51.

Meltzer, David E., Plisch, Monica, & Vokos, Stamatis. (2012). Transforming the preparation of physics teachers: A call to action. A Report by the Task Force on Teacher Education in Physics (T-TEP) (pp. 123). College Park, MD: American Physical Society.

Milner-Bolotin, Marina. (2012). Increasing interactivity and authenticity of chemistry instruction through data acquisition systems and other technologies. *Journal of Chemical Education, 89*(4), 477-481.

Milner-Bolotin, Marina. (2013, November, 2013). Mathematics and Science Teaching and Learning through Technologies. 2013, from http://scienceres-edcp-educ.sites.olt.ubc.ca/

Milner-Bolotin, Marina. (2014). Promoting research-based physics teacher education in Canada: Building bridges between theory and practice. *Physics in Canada, In press*.

Milner-Bolotin, Marina, Antimirova, Tetyana, & Petrov, Anna. (2010). Clickers beyond the first year science classroom. *Journal of College Science Teaching, 40*(2), 18-22.

Milner-Bolotin, Marina, Kotlicki, Andrzej, & Rieger, Georg. (2007). Can students learn from lecture demonstrations: The role and place of interactive lecture experiments in large introductory science courses. *Journal of College Science Teaching, 36*(4), 45-49.

Milner-Bolotin, Marina, & Moll, Rachel F. (2008). Physics exam problems reconsidered: Using Logger Pro technology to evaluate student understanding of physics. *The Physics Teacher, 46*(8), 494-500.

Moll, Rachel, & Milner-Bolotin, Marina. (2009). The effect of interactive lecture experiments on student academic achievement and attitudes towards physics. *Canadian Journal of Physics, 87*(8), 917-924.

Morrison, S., & Walsh Free, K. (2001). Writing multiple choice test items to promote and measure critical thinking. *Journal of Nursing Education, 40*(1), 17-24.

OECD. (2009a). Programme for International Student Assessment (PISA).

OECD. (2009b). Top of the class: High performers in science in PISA 2006 (pp. 168): OECD.

OECD. (2013). PISA 2012 Results: OECD.

R. Chabay, B. Sherwood. (2006). Brief Electricity and Magnetism Assessment. *Physical Review Special Topics - Physics Education Research, 2*(1), 7-13.

Shulman, Lee S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Tobias, Sheila, & Baffert, Anne. (2009). *Science Teaching as a Profession: Why It Isn't How It Could Be*. Tucson, AZ: Research Corporation for Science Advancement.

Wieman, Carl E. (2012). Applying new research to improve science education. *Issues in Science and Technology*(Fall), 1-9.

Wieman, Carl E., Adams, Wendy K., Loeblein, Patricia, & Perkins, Katherine K. (2010). Teaching physics using PhET simulations. *The Physics Teacher, 48*(April), 225-227.