ACTIVE LEARNING: THEORETICAL PERSPECTIVES, EMPIRICAL STUDIES AND DESIGN PROFILES

EDITED BY: Robert Cassidy, Elizabeth S. Charles, James D. Slotta and Nathaniel Lasry

PUBLISHED IN: Frontiers in ICT and Frontiers in Education

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88945-885-1
DOI 10.3389/978-2-88945-885-1

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.
By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

ACTIVE LEARNING: THEORETICAL PERSPECTIVES, EMPIRICAL STUDIES AND DESIGN PROFILES

Topic Editors:
Robert Cassidy, Concordia University, Canada
Elizabeth S. Charles, Dawson College, Canada
James D. Slotta, University of Toronto, Canada
Nathaniel Lasry, John Abbott College, Canada

Image: Tom Berthelot, licensed under CC-BY.

This book represents the emerging efforts of a growing international network of researchers and practitioners to promote the development and uptake of evidence-based pedagogies in higher education, at a level approaching large-scale impact. By offering a communication venue that attracts and enhances much needed partnerships among practitioners and researchers in pedagogical innovation, we aim to change the conversation and focus on how we work and learn together – i.e. extending the implementation and knowledge of co–design methods.

In this first edition of our Research Topic on Active Learning, we highlight two (of the three) types of publications we wish to promote. First are studies aimed at understanding the pedagogical designs developed by practitioners in their own practices by bringing to bear the theoretical lenses developed and tested in the education research community. These types of studies constitute the “practice pull” that we see as a necessary counterbalance to “knowledge push” in a more productive pedagogical innovation ecosystem based on research-practitioner partnerships.

Second are studies empirically examining the implementations of evidence-based designs in naturalistic settings and under naturalistic conditions. Interestingly, the teams conducting these studies are already exemplars of partnerships between researchers and practitioners who are uniquely positioned as “in-betweens” straddling the two worlds. As a result, these publications represent both the rigours of research and the pragmatism of reflective practice.

In forthcoming editions, we will add to this collection a third type of publication—design profiles. These will present practitioner-developed pedagogical designs at varying levels of abstraction to be held to scrutiny amongst practitioners, instructional designers and researchers alike.

We hope by bringing these types of studies together in an open access format that we may contribute to the development of new forms of practitioner-researcher interactions that promote co-design in pedagogical innovation.

Citation: Cassidy, R., Charles, E. S., Slotta, J. D., Lasry, N., eds. (2019). Active Learning: Theoretical Perspectives, Empirical Studies and Design Profiles. Lausanne: Frontiers Media. doi: 10.3389/978-2-88945-885-1

Table of Contents

SECTION 4
DESIGN PROFILES

SECTION 5
CHANGES IN PRACTITIONERS

frontiers in ICT

EDITORIAL
published: 07 March 2019
doi: 10.3389/fict.2019.00003

Editorial: Active Learning: Theoretical Perspectives, Empirical Studies, and Design Profiles

Robert Cassidy[1*†], *Elizabeth S. Charles*[2†] *and James D. Slotta*[3†]

[1] Department of Education and Centre for Teaching and Learning, Concordia University, Montreal, QC, Canada, [2] Dawson College, Montreal, QC, Canada, [3] Department of Curriculum, Teaching and Learning, Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, Canada

Keywords: research-practice partnership, pedagogical innovation, evidence-based teaching, active learning, design-based research, learning sciences

OPEN ACCESS

Edited by:
Leman Figen Gul,
Istanbul Technical University, Turkey

Reviewed by:
Aydin Oztoprak,
TOBB University of Economics and Technology, Turkey

Correspondence:
Robert Cassidy
rob.cassidy@concordia.ca

† These authors have contributed equally to this work and are listed in alphabetical order

Specialty section:
This article was submitted to Digital Education, a section of the journal Frontiers in ICT

Received: *21 December 2018*
Accepted: *12 February 2019*
Published: *07 March 2019*

Citation:
Cassidy R, Charles ES and Slotta JD (2019) Editorial: Active Learning: Theoretical Perspectives, Empirical Studies, and Design Profiles. Front. ICT 6:3. doi: 10.3389/fict.2019.00003

Editorial on the Research Topic

Active Learning: Theoretical Perspectives, Empirical Studies, and Design Profiles

Scholars recognize our transition into a "Knowledge Society," where citizens are increasingly engaged in critical thinking, collaborative problem solving and evidence-based reasoning, and the workplace is defined by its complexity and rapid evolution (Hargreaves, 2003; Zuboff and Maxmin, 2004). As technologies like artificial intelligence and automation further affect the nature of work, educators, learning scientists and psychologists are now questioning whether our current educational approaches are adequately preparing students for this transforming landscape. In such a world, it is arguable that education should focus on helping students develop new skills, literacies and learning dispositions—e.g., complex problem-solving, digital literacy, initiative, self-direction and lifelong learning—in addition to basic skills and factual knowledge (Acosta and Slotta).

Educational researchers and practitioners have begun to respond to this challenge, leading to an instructional paradigm at the boundary of theory and practice, known as "active learning" (Bonwell and Eison, 1991). Translating research knowledge into practice, active learning develops and uses modes of instruction grounded in social constructivist theories and technological innovations to engage students and focus more intentionally on learning processes to improve learning outcomes. In the other direction, practitioners build highly effective active learning practices that challenge and inform our theoretical understanding, demonstrate effective principles of design, and are useful to other practitioners. In this productive exchange between research and practice, active learning designs are researched to produce rigorous evidence for what works and what does not for active learning methods.

New areas of research have been spawned by innovative learning technologies and the learning environments that support active learning, for instance, technology-rich classrooms such as SCALE-UP (Foote et al., 2014) and TEAL (Belcher, 2003). Among practitioners, there is a surge of interest in approaches such as the "flipped classroom" where students engage in the "lecture-like" activity at home, watching videos and reading texts, while they enact more active forms of problem solving, small group work, tutorial and recitation during class time (Bens, 2005; Lasry et al., 2014). Enjoying equal attention among instructors are student-centered methods such as peer instruction (Mazur, 1997; Balta et al., 2017; Cormier and Voisard; Fagen et al., 2002; Lasry et al., 2008, Schell and Butler), peer assessment (Panadero et al., 2018), peer annotations (Miller et al.), 2-stage exams (Wieman et al., 2014), to name a few. This movement has begun to generate new knowledge, as practitioners adapt and innovate theoretically driven, evidence-based pedagogies and technologies to make them work in real classroom settings.

Biesta (2015) has outlined two different roles for the way research can be useful to practice: (1) the technical, in which research provides practitioners with knowledge about effective teaching strategies, assessment practices, etc.; and (2) the cultural, in which research helps practitioners to acquire a different understanding of their practice. With respect to a technical role for research in practice, instructors typically find research knowledge inaccessible and irrelevant—removed from the contextual needs and realities of teachers (Hargreaves, 1997). In general, there is a knowledge translation chasm across which practitioners are expected to deconstruct abstract instructional principles and research findings and reconstruct them around their own learning context and content objectives. In a technical sense, educational research is failing practitioners.

Hence, there is an emerging consensus that effective and sustainable implementation of instructional innovations such as active learning can only be achieved through new ways of conceptualizing the transfer of knowledge from research to practice and vice versa (Biesta, 2007; Broekkamp and van Hout-Wolters, 2007; Vanderlinde and van Braak, 2010). Research-Practice partnership (RPP) is one such conceptualization that offers the approach of co-design, where researchers and practitioners learn from each other (Coburn and Penuel, 2016; Fishman et al., 2013). RPP recognizes the importance of the practical expertise of practitioners and their role as front-line designers who consistently innovate and "make the magic happen." Practitioners must understand research findings and apply them in creating tools and methods that serve to implement principled pedagogies. These implementations then serve as a crucial source of insight for our wider community, rather than just implementations of the research. In focusing on "problems of practice," RPP interventions integrate practitioners' expertise and anchor the development of educational solutions in true collaborations.

It is increasingly clear that the transformation of the educational landscape should involve a thoughtful examination of the work of RPPs, which in turn can lead to new approaches to both research and practice. The traditional "knowledge push" approach (i.e., of research into practice) entails a unidirectional movement of generalized theory into practice, placing the researcher as the primary agent and holder of knowledge and the practitioner as the recipient. We seek to add a dynamic of "practice pull." We see practitioners as being useful to researchers, not as a mere testing ground, but as a source of insight from which researchers can reciprocally acquire a different understanding of their research and its objectives. Recognizing the limits of a "knowledge translation" approach to innovation, we seek to develop a social and cultural approach to innovation in which the voice of the practitioner is equal to that of the educational researcher. By capturing and sharing stories from RPPs and the wider active learning community, we seek to move beyond "pushing" and "pulling" into the more complex and recursive relationships of co-design, co-understanding and collaboration.

One example can be seen in our own work, to establish a professional learning community called SALTISE (Supporting Active Learning & Technological Innovation in Studies of Education; see saltise.ca), which features a growing collection of successful implementations of active learning collected from practitioners, analyzed and codified by researchers. One goal of SALTISE is to help shed light on the tensions between the generalizability goals of research and the contextual realities of practitioners. Researchers look for relatively well-defined projects that result in publishable findings at the cusp of what is known or has been previously demonstrated, often relying on methodological traditions of comparative intervention studies. Practitioners are interested in longer-term refinement, experimentation and ongoing optimisation based on experience and feedback. Their primary goal is to improve student learning outcomes. SALTISE thus offers the promise of studying what Penuel (2014, p. 101), describes as interventions that are "developed in practice by participants in that practice, rather than in a controlled laboratory."

In this *Frontiers in ICT Research Topic*, 12 articles touch on these various aspects of mixing researcher and practitioner knowledge to understand and evaluate active learning in action. Three papers demonstrate the value of using a research-based theoretical perspective to examine effective active learning practices developed by practitioners. Schell and Butler analyse Peer Instruction through a cognitive psychology lens to derive principles that can help understand its effectiveness and provide guidance for practitioners who need to adapt it to fit their own implementation contexts. Brewe et al. similarly use a neurobiological lens to examine the efficacy of Modeling Instruction (MI) as a step toward understanding the changes in neural activity consequent to this style of learning. Furthermore, grounded in the theories of epistemological beliefs and conceptual change, Kalman and Lattery describe critical obstacles for learning post-secondary science and advance three instructional design principles for practically working through them.

Another seven papers empirically examine design principles in action. Ehrlick and Slotta present work in which the Knowledge Community and Inquiry model (Slotta and Najafi, 2013) is put sustainably and effectively into a situated practice through the adaptive iterations of design-based research (DBR). Similarly, Cormier and Voisard describe how the abstract "flipped" approach is concretely applied to an organic chemistry context and provide evidence for its effectiveness at driving student learning outcomes Acosta and Slotta present theoretical and practical design principles for the implementation of active learning curricula in grade 12 biology classrooms.

Researcher's theories have helped physics professors hone in on learning outcomes, as Marshman et al. engage the elusive "transfer of learning" problem through the design of a digital tutorial platform. The authors outline the impact and lessons learned from its implementation. Akiha et al. describe a cross-sectional study of the instructional methods students experience as they advance through an educational system, identifying gaps in instructors' understanding of other parts of the system and emphasizing the need for inter-order communication and collaboration.

Miller et al. present the design and implementation of a computer-supported collaborative learning technology developed from practitioner experience, and present evidence for its increase in engagement and learning. Poellhuber et al. describe methods for the functional analysis of active learning spaces to identify the most valued features of these spaces and the relations among learning behaviors and attitudes toward the learning spaces.

Finally, two studies examine the types of changes in instructors that are associated with the adoption of active learning strategies. Fournier St-Laurent and Poellhuber present a case study of the changes that instructors undergo during the early stages of adopting active learning pedagogies; while, Laferrière presents the transformative nature of an ongoing DBR experiment on a sustained community of practice.

This collection of papers will hopefully engage a broad audience of researchers and practitioners as a knowledge community whose goal is to understand such pedagogical approaches. In what ways are they effective, and how do we know if they are effective? What aspects of student and teacher interactions are responsible for their efficacy? What are important principles underlying effective curricular designs? What are the most compelling applications of media and technology? How does active learning vary across different disciplines (e.g., physics, biology) and age levels (elementary, secondary and undergraduate education). We hope to engage both researchers and practitioners from a range of disciplines and contexts, to gather a wealth of evidence demonstrating the efficacy of new, principles and practical approaches that emphasize student inquiry, problem solving and collaboration.

AUTHOR CONTRIBUTIONS

All authors listed contributed equally to the work, and approved it for publication.

REFERENCES

Balta, N., Michinov, N., Balyimez, S., and Ayaz, M. F. (2017). A meta-analysis of the effect of peer instruction on learning gain: Identification of informational and cultural moderators. *Int. J. Educ. Res.* 86, 66–77. doi: 10.1016/j.ijer.2017.08.009

Belcher, J. W. (2003). Improving student understanding with TEAL. *The MIT Faculty Newsletters* 16. Available online at: http://web.mit.edu/fnl/vol/162/fnl162.pdf

Bens, I. (2005). *Facilitating With Ease! Core Skills for Facilitators, Team Leaders and Members, Managers, Consultants, and Trainers,* 2nd Edn. San Francisco, CA: Jossey Bass.

Biesta, G. (2007). Bridging the gap between educational research and educational practice: the need for critical distance. *Educ. Res. Eval.* 13, 295–301. doi: 10.1080/13803610701640227

Biesta, G. (2015). On the two cultures of educational research, and how we might move ahead: Reconsidering the ontology, axiology and praxeology of education. *Eur. Educ. Res. J.* 14, 11–22. doi: 10.1177/1474904114565162

Bonwell, C. C., and Eison, J. A. (1991). "Active learning: Creating excitement in the classroom," in *ASHE-ERIC Higher Education Reports* (Washington, DC: ERIC Clearinghouse on Higher Education).

Broekkamp, H., and van Hout-Wolters, B. (2007). The gap between educational research and practice: A literature review, symposium, and questionnaire. *Educ. Res. Eval.* 13, 203–220. doi: 10.1080/13803610701626127

Coburn, C. E., and Penuel, W. R. (2016). Research–practice partnerships in education: Outcomes, dynamics, and open questions. *Educ. Res.* 45, 48–54. doi: 10.3102/0013189X16631750

Fagen, A. P., Crouch, C. H., and Mazur, E. (2002). Peer instruction: results from a range of classrooms. *Phys. Teach.* 40, 206–209. doi: 10.1119/1.1474140

Fishman, B. J., Penuel, W. R., Allen, A.-R., Cheng, B. H., and Sabelli, N. (2013). Design-based implementation research: an emerging model for transforming the relationship of research and practice. *Natl. Soc. Study Educ.* 112, 136–156. https://www.sri.com/sites/default/files/publications/fishman_penuel_allen_cheng_sabelli_2013.pdf

Foote, K. T., Neumeyer, X., Henderson, C., Dancy, M. H., and Beichner, R. J. (2014). Diffusion of research-based instructional strategies: the case of SCALE-UP. *Int. J. STEM Educ.* 1, 1–18. doi: 10.1186/s40594-014-0010-8

Hargreaves, A. (2003). *Teaching in the Knowledge Society: Education in the Age of Insecurity.* New York, NY: Teachers College Press.

Hargreaves, D. H. (1997). In defence of research for evidence–based teaching: a rejoinder to Martyn Hammersley. *Br. Educ. Res. J.* 23, 405–419. doi: 10.1080/0141192970230402

Lasry, N., Dugdale, M., and Charles, E. (2014). Just in time to flip your classroom. *Phys. Teach.* 52, 34–37. doi: 10.1119/1.4849151

Lasry, N., Mazur, E., and Watkins, J. (2008). Peer instruction: from harvard to the two-year college. *Am. J. Phys.* 76, 1066–1069. doi: 10.1119/1.2978182

Mazur, E. (1997). *Peer instruction*. Upper Saddle River, NJ: Prentice Hall.

Panadero, E., Jonsson, A., and Alqassab, M. (2018). "Providing formative peer feedback: what do we know?" in *The Cambridge Handbook of Instructional Feedback Cambridge Handbooks in Psychology*, eds A. Lipnevich and J. Smith (Cambridge: Cambridge University Press). doi: 10.1017/9781316832134.020

Penuel, W. R. (2014). Emerging forms of formative intervention research in education. *Mind. Culture Act.* 21, 97–117. doi: 10.1080/10749039.2014.884137

Slotta, J. D., and Najafi, H. (2013). "Supporting collaborative knowledge construction with Web 2.0 technologies," in *Emerging Technologies for the Classroom: A Learning Sciences Perspective*, eds C. Mouza and N. Lavigne (New York, NY: Springer). 93–112. doi: 10.1007/978-1-4614-4696-5_7

Vanderlinde, R. and van Braak, J. (2010). The gap between educational research and practice: Views of teachers, school leaders, intermediaries and researchers. *Br. Educ. Res. J.* 36, 299–316. doi: 10.1080/01411920902919257

Wieman, C. E., Rieger, G. W., and Heiner, C. E. (2014). Physics exams that promote collaborative learning. *Phys. Teach.* 52, 51–53. doi: 10.1119/1.4849159

Zuboff, S., and Maxmin, J. (2004). *The Support Economy: Why Corporations Are Failing Individuals and the Next Episode of Capitalism.* New York, NY: Penguin.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Education

REVIEW
published: 28 May 2018
doi: 10.3389/feduc.2018.00033

Insights From the Science of Learning Can Inform Evidence-Based Implementation of Peer Instruction

Julie A. Schell[1,2]* and *Andrew C. Butler*[3,4]

[1] Department of Educational Leadership and Policy, The University of Texas at Austin, Austin, TX, United States, [2] School of Design and Creative Technologies, The University of Texas at Austin, Austin, TX, United States, [3] Department of Education, Washington University in St. Louis, St. Louis, MO, United States, [4] Department of Psychological and Brain Sciences, Washington University in St. Louis, St. Louis, MO, United States

Peer Instruction is a popular pedagogical method developed by Eric Mazur in the 1990s. Educational researchers, administrators, and teachers laud Peer Instruction as an easy-to-use method that fosters active learning in K-12, undergraduate, and graduate classrooms across the globe. Research over the past 25 years has demonstrated that courses that incorporate Peer Instruction produce greater student achievement compared to traditional lecture-based courses. These empirical studies show that Peer Instruction produces a host of valuable learning outcomes, such as better conceptual understanding, more effective problem-solving skills, increased student engagement, and greater retention of students in science majors. The diffusion of Peer Instruction has been widespread among educators because of its effectiveness, simplicity, and flexibility. However, a consequence of its flexibility is wide variability in implementation. Teachers frequently innovate or personalize the method by making modifications, and often such changes are made without research-supported guidelines or awareness of the potential impact on student learning. This article presents a framework for guiding modifications to Peer Instruction based on theory and findings from the science of learning. We analyze the Peer Instruction method with the goal of helping teachers understand why it is effective. We also consider six common modifications made by educators through the lens of retrieval-based learning and offer specific guidelines to aid in evidence-based implementation. Educators must be free to innovate and adapt teaching methods to their classroom and Peer Instruction is a powerful way for educators to encourage active learning. Effective implementation, however, requires making informed decisions about modifications.

Keywords: Peer Instruction, cognitive science, retrieval practice, instructional design, Eric Mazur, research-based instructional strategies, learning science, active learning

OPEN ACCESS

Edited by:
Rob Cassidy,
Concordia University, Canada

Reviewed by:
Tammy M. Long,
Michigan State University,
United States
Lisa McDonnell,
University of California, San Diego,
United States

****Correspondence:***
Julie A. Schell
Juile.schell@austin.utexas.edu

Specialty section:
This article was submitted to Digital Education, a section of the journal Frontiers in Education

Received: *06 December 2017*
Accepted: *27 April 2018*
Published: *28 May 2018*

Citation:
Schell JA and Butler AC (2018) Insights From the Science of Learning Can Inform Evidence-Based Implementation of Peer Instruction. Front. Educ. 3:33. doi: 10.3389/feduc.2018.00033

INTRODUCTION

In today's classrooms, there is great demand for active learning among both students and educators. Calls for active learning are not new (see Eliot, 1909), but a recent surge of interest in this concept is transforming pedagogical practices in higher education. The inspiration for this movement comes in large part from the now well-established benefits for student achievement and motivation produced by active learning environments (Bonwell and Eison, 1991; Braxton et al., 2000; National Research Council, 2000; Ambrose et al., 2010; Freeman et al., 2014). With a growing number of educators keenly aware of the limitations of "transmissionist" teaching methods, many of them are trying out new pedagogical methods that encourage active learning (Dancy et al., 2016).

Despite its popularity and general effectiveness, active learning is a broad concept and it is often vaguely defined, which leads to a great variability in its implementation within formal and informal education environments. We define active learning as a process whereby learners deliberately take control of their own learning and construct knowledge rather than passively receiving it (National Research Council, 2000). Active learning is not necessarily synonymous with liveliness or high levels of engagement, even if classrooms that feature active learning are often dynamic; and it is qualitatively different from more passive learning processes, such as listening to a lecture or reading a text, that primarily involve the transmission of information. Active learners construct meaning by integrating new information with existing knowledge, assess the status of their understanding frequently, and take agency in directing their learning. Even though control over learning ultimately resides with students, educators play a crucial role because they create classroom environments that can either foster or hinder active learning.

In this article, we explore the challenges faced by educators who want to effectively foster active learning using established pedagogical methods while retaining the ability to innovate and adapt those methods to the unique needs of their classroom. One challenge that educators face is that they often must teach themselves to use new methods that are very different from the teaching that they experienced as students. Moreover, graduate and post-doctoral education rarely focus on teaching, so most educators do not have any formal training to draw upon when trying to implement new methods or innovate. In addition, many educators who are trying new methods must do so with little or no feedback on effective implementation from more experienced teachers. Under these conditions, pedagogical improvement is exceedingly difficult, which makes it all the more impressive that the switch to active learning generally produces good results. Nevertheless, changes to pedagogy do not always result in positive effects. Indeed, when educators make modifications to established pedagogical methods, it may have the unintended consequence of limiting, inhibiting, or even preventing active learning. Thus, it is important for educators to understand how omitting or changing aspects of a pedagogical method might affect student learning and motivation.

We chose to focus on an established and popular pedagogical method called Peer Instruction, which researchers have demonstrated encourages active learning in a wide range of classrooms, disciplines, and fields (Mazur, 1997; Crouch and Mazur, 2001; Schell and Mazur, 2015; Vickrey et al., 2015; Müller et al., 2017). Eric Mazur developed Peer Instruction in the early 1990s at Harvard University (Mazur, 1997). The method is well-regarded in the educational research community for its demonstrated ability to stimulate active learning and achieve desired learning outcomes in a variety of educational contexts (Vickrey et al., 2015; Müller et al., 2017). One of the key features of Peer Instruction is its flexibility that enables adaptation to almost any context and instructional design (Mazur, 1997). However, this flexibility comes with a potential cost in that modifications to the method may limit its effectiveness as it relates to active learning. Indeed, when educators modify Peer Instruction, they may be unaware that these modifications can disrupt the benefits of active learning (Dancy et al., 2016).

The primary goal of this paper is to provide Peer Instruction practitioners with an understanding of why the method is effective at fostering active learning so that they can make informed choices about how to innovate and adapt the method to their classroom. A secondary goal of this article is to respond to a need for explicit collaborations between educational researchers and cognitive scientists to help guide the implementation of innovative pedagogical methods (Henderson et al., 2015). Integrating basic principles from the science of learning into the classroom has been shown to increase learning in classrooms in ways that can easily scale and generalize to a variety of subjects (e.g., Butler et al., 2014). Unfortunately, the diffusion of general principles from the science of learning into the classroom has been much slower than innovative pedagogical methods that provide "off-the-shelf" solutions, such as Peer Instruction. Accordingly, analyzing such pedagogical methods to identify the mechanisms and basic principles that make them effective may be beneficial for both implementation in educational practice and scientific research on learning.

By way of providing the reader with an outline, our article begins with an overview of the Peer Instruction method, including a brief history and a description of the advice for implementation from the manual created by the developer (Mazur, 1997). Next, we provide an in-depth analysis of the efficacy of Peer Instruction by drawing upon theory and findings from the science of learning. Finally, we conclude with a discussion about the many common modifications users make to Peer Instruction. In this concluding section, we also provide clear recommendations for modifying Peer Instruction based on findings from the science of learning with a specific focus on a driving mechanism underlying the potent achievement outcomes associated with Mazur's method—retrieval-based learning. Taken as a whole, we believe this article represents a novel, evidence-based approach to guiding Peer Instruction innovation and personalization that is not currently available in the literature.

PEER INSTRUCTION: A POPULAR PEDAGOGICAL METHOD THAT PROMOTES ACTIVE LEARNING

Mazur developed Peer Instruction in 1991 in an attempt to improve his Harvard undergraduates' conceptual understanding of introductory physics (Mazur, 1997). Previously, Mazur's teaching was lecture-based and his instructional design featured passive learning before, during, and after class. The impetus for the change in his teaching method came from David Hestenes and his colleagues who published the Force Concept Inventory (FCI)—a standardized test that evaluated students' abilities to solve problems based on their conceptual understanding of Newton's Laws, which is a foundational topic in introductory physics (Hestenes et al., 1992). In their classroom research using the FCI, Hestenes and colleagues found that most students could state Newton's Laws verbatim, but only a small percentage could solve problems that relied on mastery of the concept. Mazur learned about the FCI and decided to deliver the test to his students. To his surprise, the results were similar to Hestenes. After a brief period of questioning the validity of the test, Mazur became convinced that there was a serious gap in students' learning of physics in introductory college classrooms. The vast majority of physics education at the time was lecture-based. Mazur developed Peer Instruction to target the gap in conceptual understanding because he was convinced that it resulted from passive learning experiences and overreliance on transmission-based models of teaching.

The Peer Instruction Method

In 1997, Mazur published *Peer Instruction: A User's Manual* in which he describes the seven steps that constitute the method (Mazur, 1997, page 10). The seven steps are the following:

1. Question posed (1 min)
2. Students given time to think (1 min)
3. Students record individual answers [*optional*]
4. Students convince their neighbors—peer instruction (1–2 min)
5. Students record revised answers [*optional*]
6. Feedback to teacher: Tally of answers
7. Explanation of correct answer (2+ min)

As can be gleaned from the list, the Peer Instruction method involves a structured series of learning activities. The overall learning objective is the improvement of conceptual understanding, or in Mazur's words: "The basic goals of Peer Instruction are to exploit student interaction during lectures and focus students' attention on underlying concepts" (Mazur, 1997, p. 10). Accordingly, the method begins with the teacher focusing students' attention by posing a conceptual question called a ConcepTest that is generally in a multiple-choice format (but increasingly short answer format is being used), and then the remaining activities build on this question. The method is designed to take between 5 and 15 min depending on the complexity of the concept and whether all of the seven steps are used.

Given the central importance of the ConcepTest to Peer Instruction, it is no surprise that the efficacy of the method depends upon the quality of the question. Although a ConcepTest is a question, not all questions are a ConcepTest—a ConcepTest has specific features that distinguish it from other types of questions. First, as an assessment item, a ConcepTest is designed to test and build students' conceptual understanding rather than factual or procedural knowledge. Another distinct feature of a ConcepTest is the list of multiple choice alternatives. A well-designed, multiple choice ConcepTest will follow published guidelines for designing effective multiple choice questions (Haladyna et al., 2002). In particular, the teacher will construct the responses by including a correct answer and viable distractors that elicit common misconceptions about the concept.

After the teacher poses the ConcepTest (Step 1), she gives students time to think and construct an answer based on their current understanding (Step 2). The teacher then directs students to record and display their answers to the the teacher using a classroom response method (Step 3). The response method can be low-tech (e.g., hand signals, flashcards, or student whiteboards) or high-tech (e.g., clickers, text messages, or cloud-based courseware). The "modality" in which students record and/or display their answer is not critical—the key is that students generate and commit to a response (Lasry, 2008). That said, the higher-tech response systems (clickers, web-based response systems) have benefits to consider. For students, the systems record answers for later review and provide greater anonymity than using hand signals or flashcards. For teachers, the higher-tech systems enable the analysis of student responses that may inform teacher behavior and future assessment planning based on the pattern of answer choices (Schell et al., 2013). For example, students may surprise the teacher if the majority chooses a distractor as the right answer, thereby prompting the teacher to modify her teaching plan.

Once the teacher collects the responses, she reviews them without disclosing, displaying, or sharing the correct answer or the frequency of choices among the students. Next, the teacher cues students to "turn to their neighbor" to use reasoning to convince their peer of their answer (Step 4). If their neighbor has the same answer, Mazur recommends cueing students to find someone with a different answer (Mazur, 2012). Students then engage in a brief discussion in pairs where they have the opportunity to recall their response as well as justify why they responded the way they did. Mazur emphasizes that during the discussion students must defend their answers with reasoning based on what they have previously heard, read, learned, or studied. After the discussion is complete, the teacher gives students time to think about their final answer—whether they want to keep the same answer or change answers. Once they have had a moment to think, students record their final responses (Step 5), which are communicated to the teacher using the same classroom response method (Step 6).

The teacher closes the series of activities by finally revealing and explaining the correct answer (Step 7). Some teachers display the pre-post response frequencies so students can see how their answers changed (often, in the direction of the correct answer) and how many others selected specific answer choices. After revealing the correct answer, some teachers ask for explanations from representatives from each answer choice to explain their reasoning. Students are often willing to explain their reasoning despite the revelation that their response was incorrect. The purpose of this additional exercise is to help students interrogate and resolve any potential misconceptions that led them to select one of the distractors. Hearing the correct answer explained by their peers can be more effective because other novices may be able to better communicate it than the teacher who is an expert (Mazur, 1997).

Finally, it is important to note some of the key features of the method that are critical to the efficacy of Peer Instruction. In a recent article, Dancy et al.(2016, p. 010110-5) analyzed the method in consultation with Mazur and other experienced Peer Instruction practitioners. They identified nine key features of Peer Instruction based on their research:

1. Instructor adapts instruction based on student responses
2. Students are not graded on in-class Peer Instruction activities
3. Students have a dedicated time to think and commit to answers independently
4. The use of conceptual questions
5. Activities draw on student ideas or common difficulties
6. The use of multiple choice questions that have discrete answer options
7. Peer Instruction is interspersed throughout class period
8. Students discuss their ideas with their peers
9. Students commit to an answer after peer discussion

These features, which are present in the original Peer Instruction user manual (Mazur, 1997), have proven to be essential to the success of the method.

Diffusion of Peer Instruction

Over the past quarter-century, the use of Peer Instruction has expanded far beyond Ivy League undergraduate physics education. Educators from wildly diverse contexts have used the method to engage hundreds of thousands of students in active learning. For example, middle school, high school, undergraduate, and graduate students studying Biology, Chemistry, Education, Engineering, English, Geology, US History, Philosophy, Psychology, Statistics, and Computer Science, in a variety of countries in Africa, Australia, Asia, Europe, North America, and South America, have all experienced Mazur's Peer Instruction (Mazur, 1997; Schell and Mazur, 2015; Vickrey et al., 2015; Müller et al., 2017). The widespread adoption of Peer Instruction by a diverse array of educators over the past 25 years has prompted a new area of research and large body of scholarship. Studies that support the efficacy of Peer Instruction run the gamut from applied research in a single classroom (Mazur, 1997) to multi-course, large-sample investigations (Hake, 1998), comparisons across institutional types (Fagen et al., 2002; Lasry et al., 2008), and meta-analyses covering a variety of educational contexts (Vickrey et al., 2015; Müller et al., 2017).

The consensus woven through the fabric of over two and a half decades of scholarship is that when compared to traditional lecture-based pedagogy, Peer Instruction leads to positive outcomes for multiple stakeholders, including teachers, institutions, disciplines, and (most importantly) students. For example, large-sample studies of Peer Instruction report that teachers observe lower failure rates even in challenging courses (Porter et al., 2013). On a more structural level, researchers have also demonstrated that Peer Instruction may offer a high impact solution to stubborn educational problems, such as retention of STEM majors and reduction of the gender gap in academic performance in science (Lorenzo et al., 2006; Watkins and Mazur, 2013). Peer Instruction efficacy is not limited to STEM courses. For example, Draper and Brown (2004) and Stuart et al. (2004) investigated the use of Peer Instruction in the humanities. And Chew (2004, 2005) has studied Peer Instruction use in the social sciences. Both Stuart and Chew observed positive outcomes. Finally, the benefits of Peer Instruction are most notable for students. In particular, research has shown that learners in Peer Instruction courses develop more robust quantitative problem-solving skills, more accurate conceptual knowledge, increased academic self-efficacy, and an increased interest in and enjoyment of their subject (Hake, 1998; Nicol and Boyle, 2003; Porter et al., 2013; Watkins and Mazur, 2013; Vickrey et al., 2015; Müller et al., 2017). However, this literature is limited in the sense that it mainly focuses on educational outcomes that result from the use of Peer Instruction without considering why and how the method produces those outcomes. In the remainder of this article, we contribute such an analysis through the lens of the science of learning.

WHY IS PEER INSTRUCTION EFFECTIVE? PERSPECTIVES FROM THE SCIENCE OF LEARNING

We now turn to analyzing why Peer Instruction is an effective teaching method for fostering active learning by drawing upon theory and findings from the science of learning. As a framework for presenting our analysis, we have grouped the key aspects of Peer Instruction into four general categories of factors that form the context that an educator must consider in order to facilitate student learning in any course (see **Figure 1**). Learner objectives (Course Material and Skill), learner characteristics, learner activities, and learner outcomes.

Objectives

One of the first steps in designing any course should be the development of specific, achievable student learning objectives (Tyler, 1949; Wiggins and McTighe, 2005). Ideally, the process of developing such learning objectives grows out of a careful analysis of the goals of the course material in the context of the broader curriculum and the skills and knowledge that students

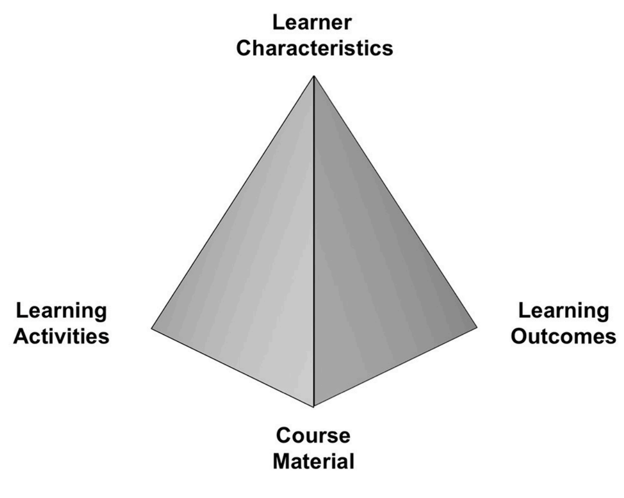

FIGURE 1 | A tetrahedral model of classroom learning (adapted from Jenkins, 1979).

need to acquire to achieve these goals. In large part, the creation of Peer Instruction was born out of a recognition that the skills and knowledge that students acquired in introductory physics courses were qualitatively different from what is needed to progress in physics education. More specifically, the key insight was that students were acquiring procedural skills and knowledge but lacked the conceptual understanding to effectively use them, which is a common issue in many STEM disciplines (e.g., Rittle-Johnson et al., 2015). In addition to content-specific learning objectives, Mazur (1997) also emphasizes the importance of domain-general objectives that active learning can help achieve, such as critical thinking and metacognitive monitoring. Mazur states that Peer Instruction, "forces the students to think through the arguments being developed and provides them (as well as the teacher) with a way to assess their understanding of the concept" (p. 10). As a result, Peer Instruction fosters critical thinking in the domain of study and metacognition. Indeed, these cognitive skills are essential components of active learning – it is impossible to monitor and direct one's own learning without them. When students receive feedback throughout each cycle of Peer Instruction on how well they "understand" the concepts, they can direct their efforts toward learning concepts they are struggling with. In sum, a clear sense of the skills and knowledge that students need to acquire is critical to selecting the learning activities and outcome measures that will be appropriate for any given group of students.

Activities

Educators have a multitude of instructional activities from which to choose in order to facilitate student learning and active learning more specifically (see Hattie, 2009). Importantly, there are substantial differences among this broad array activities in terms of how they affect student learning, and thus selecting an effective learning activity depends upon the learning objective (Koedinger et al., 2012). In addition, the effectiveness of a given learning activity can also differ as a function of where students are in the process of learning, so it is also imperative to consider how to structure and scaffold learning as student knowledge and skills progress. The complexity underlying how learning occurs and the need to align teaching accordingly can seem daunting to educators (Koedinger et al., 2013), which is one reason that Peer Instruction is so useful. That is, Peer Instruction provides educators with a well-structured method that includes a potent mix of effective learning activities that are designed to foster active learners.

One key to understanding the utility of any learning activity is to analyze the types of cognitive processes that are required to perform the task. Although a multitude of basic cognitive processes are engaged during learning, educators are understandably more interested in types of processing that facilitate the construction of meaning from information (Craik and Lockhart, 1972; Craik and Tulving, 1975). At this more complex level, there are many ways in which people can process information (e.g., Packman and Battig, 1978; Hunt and Einstein, 1981). One framework that can inform the analysis of the cognitive processes that are engaged by a particular learning activity is the updated Bloom's taxonomy of educational objectives (see too Bloom, 1956; Anderson et al., 2001). Does the activity involve application, analysis, classification, evaluation, comparison, etc.? The reason that such analysis is important is that the cognitive processes that are used during the activity will dictate what is learned and as such, how students will direct further learning. With this idea in mind, a clear advantage of Peer Instruction is that it provides educators with great flexibility in deciding how students should process the information during learning. For example, the question posed on a ConcepTest, whether in multiple choice or constructed response format, could induce students to engage any one or multiple processes described in Bloom's taxonomy.

While on the topic of cognitive processing, one critical distinction is that learning activities differ in the extent to which they involve perceiving and encoding new information relative to retrieving and using information that has already been stored in memory. In more simplistic terms, this distinction is between how much the activity involves "putting information in" vs. "getting information out." Many of the learning activities traditionally used in college courses predominantly involve perceiving and encoding new information—listening to a lecture, reading a textbook, watching video, etc. One of the key innovations in Peer Instruction is to introduce more activities that require students to retrieve and use information (e.g., ConcepTests), a change that is reflective of a broader movement toward active learning in STEM courses (Freeman et al., 2014). Perceiving and encoding new information is imperative during the initial stages of learning. However, after students have some knowledge to work with, it is often much more effective for them to engage in activities that require them to retrieve and use that knowledge (Roediger and Karpicke, 2006; for review see Dunlosky et al., 2013; Rowland, 2014).

Retrieval practice is a low-threshold instructional activity in that it is simple and easy to implement for teachers. Engaging in retrieval practice has both direct and indirect effects on learning. The direct effect stems from the fact that retrieving information from memory changes memory, and thus causes learning (Roediger and Butler, 2011). Retrieval practice has been shown to improve long-term retention (e.g., Larsen et al., 2013) and transfer of learning to new contexts (e.g., Butler, 2010; for review see Carpenter, 2012). In addition, the indirect effects are numerous—students are incentivized to keep up with material outside of class (Mawhinney et al., 1971) and they become less anxious about assessments (Agarwal et al., 2014), among other benefits. When educators use Peer Instruction following Mazur's protocol, students engage in more than three distinct retrieval practice opportunities in a single cycle (see above section on Peer Instruction Method, Steps 2,4, and 5). In short, retrieval practice is a critical mechanism in the Peer Instruction method that facilitates the development of deeper understanding that enables students to transfer their knowledge to new contexts (e.g., solve problems, analyze new ideas). We discuss retrieval as a mechanism for learning in Peer Instruction in the next section on common modifications.

Having students engage in activities that require retrieving and using recently acquired knowledge also has another important indirect benefit—it provides feedback to both students and educators (Black and Wiliam, 1998; Hattie, 2009). Feedback is one of the most powerful drivers of learning because it enables students to check their understanding and address any potential gaps (Bangert-Drowns et al., 1991; Butler and Winne, 1995; Hattie and Timperley, 2007). In particular, explanation feedback promotes the development of deeper understanding (Butler et al., 2013). Equally important is the information that is provided to educators about the current state of student understanding, which enables them to circle back and address misunderstandings. In comparison to the traditional lecture method, Peer Instruction is rich with opportunities for feedback from student-to-student, teacher-to-student, and student-to-teacher (i.e., in addition to the metacognitive benefits of feedback that results from retrieval practice). The student-to-student feedback may be particularly valuable given the benefits of collaborative learning (Nokes-Malach et al., 2015). As described in the Peer Instruction manual, students can often explain concepts better to each other than their teacher can, providing both valuable feedback and new information (Smith et al., 2009). In addition, the act of explaining to someone else is a powerful learning event as well, so both students benefit.

Finally, it is important to consider how activities are structured in order to continuously facilitate learning as the acquisition of knowledge and skills progresses. Peer Instruction does a good job of scaffolding student learning—pre-class readings and reading quizzes prepare students to learn in class, lectures present new information that extends from the readings, ConcepTests provide further practice and an opportunity for feedback. All of these activities are aligned and build upon each other. This structure also incorporates many basic principles from the science of learning that are known to promote long-term retention and the development of understanding. For example, learning is spaced or distributed over time rather than massed (Dempster, 1989; Cepeda et al., 2006) and variability is introduced during the learning of a particular piece of concept or skill by using different examples, contexts, or activities (Glass, 2009; e.g., Butler et al., 2017). Variation of this sort is particularly useful for honing students' abilities to be active learners who transfer their knowledge across contexts (Butler et al., 2017). A single cycle of Peer Instruction, which could be as short as 2–3 min, is packed with variation in learning activities. For example, students think on their own, retrieve, discuss, retrieve again, and then receive feedback on their responses.

Learner Characteristics

Perhaps the most important set of factors that influence learning in any course are the characteristics of the learners—their individual knowledge and experiences, expectations, interests, goals, etc. Individual differences play a major role in determining student success in STEM disciplines (Gonzalez and Kuenzi, 2012), and yet it is this aspect of the learning context that is so often ignored in large introductory STEM courses. One of the reasons that Peer Instruction is so effective is that it directly addresses this issue in that it is "student-centered." Throughout the Peer Instruction manual there is a consistent focus on the student experience when explaining the methodology. In addition, the rationale for focusing on students is bolstered by insightful anecdotes and observations: "Students' frustration with physics—how boring physics must be when it is reduced to a set of mechanical recipes that do not even work all the time!" (Mazur, 1997, p. 7). Taken as a whole, the manual makes clear that student engagement is essential to the successful implementation of Peer Instruction.

Nevertheless, it is possible for a pedagogy to be "student-centered" and yet ineffective on this front; what sets Peer Instruction apart is that it is consistent with many principles and best practices from research on student motivation. Chapter 3 of the manual, which focuses on student motivation, begins with advice about "setting the tone" that addresses student expectations and beliefs about learning. One theme that emerges is that students should embrace the idea that learning is challenging and requires effort and strategic practice (i.e., a growth mindset; Yeager and Dweck, 2012). Students who adopt such a mindset often show greater resilience and higher achievement (e.g., Blackwell et al., 2007). Another theme that emerges is about the importance of students coming to value what they are learning in the course and the methodology used for learning. People's perceptions about the value of an activity (e.g., self-relevance, interest, importance, etc.) can have a strong effect on their motivation to engage in that activity (Harackiewicz and Hulleman, 2010; Cohen and Sherman, 2014). Examples from the manual include Mazur's introductory questionnaire that probes student goals and interests and the explanation provided on the first day of class about why the course is being taught in this manner. A third theme that emerges is the benefit of creating a cooperative learning environment rather than a competitive one. Classrooms that foster cooperation lead students to adopt mastery learning goals (i.e., rather than performance goals) and produce greater achievement relative to classrooms that foster competition (Johnson et al., 1981; Ames, 1992). Numerous aspects of Peer Instruction help produce a cooperative environment, from the student-to-student peer instruction at the core of the pedagogy to the use of an absolute grading scale that enables everyone to succeed.

Outcomes

The purpose of any course is to facilitate learning that will endure and transfer to new situations. In education, summative assessment provides a proximal measure of learning that is assumed to predict future performance (Black, 2013). As such, it is imperative that the nature of the assessment used reflect such future performance to the extent that it is possible. The assessment tools used within Peer Instruction and afterwards to evaluate its effectiveness are derived from a careful analysis of what students must know and do in future courses. The result of this analysis is mix of different types of assessments each designed to measure a different aspect of the knowledge and skills that students need to acquire. The use of one or more diagnostic tests that tap fundamental concepts in the discipline are recommended (e.g., the FCI and the Mechanics Baseline Test in physics). Course exams are meant to feature different types of questions, such as conceptual essays and conventional problems, that engage students in types of cognitive processing (see discussion of learning activities above; Anderson et al., 2001).

Importantly, the assessment tools used in Peer Instruction are not only aligned with the future, but also with the activities that are used to facilitate student learning. As discussed above, the cognitive processes that students engage during activity determines what is learned; however, a student's ability to demonstrate that learning depends upon the nature of the assessment task. Performance tends to be optimized when the processes engaged during learning match the processes required for the assessment, a concept known as *transfer-appropriate processing* (Morris et al., 1977; for a review see Roediger and Challis, 1989). When there is a mismatch in cognitive processing (e.g., learning involved application but the test requires evaluation), then assessment can fail to accurately measure student learning.

Finally, it is critical to remember that every assessment provides the opportunity to both measure learning and facilitate learning. Every question that a student answers, regardless of whether it is in the context of a low-stakes ConcepTest or a high-stakes exam, provides summative information (i.e., measuring learning up until that point), formative information (feedback for the student and teacher), and an opportunity to retrieve and use knowledge that directly causes learning. Thus, assessment is learning and learning is assessment, and this inherent relationship makes it even more imperative that assessment reflect what students must be able to know and do in future.

In summary, Peer Instruction is an effective pedagogy because it utilizes many principles and best practices from the science of learning, while also allowing flexibility with respect to implementation. No laws of learning exist (McKeachie, 1974; Roediger, 2008), and thus facilitating student learning involves considering each category of factors shown in **Figure 1** in the context of the other three categories to optimize learning (see McDaniel and Butler, 2011). By allowing flexibility, Peer Instruction enables educators to foster active learning in ways that are optimal for their particular context. In the next section, we use the insights about Peer Instruction gleaned from the science of learning to evaluate the potential impact of common modifications to the method made by teachers. Our goal is to provide evidence-based guidance for how to make decisions about modifying Peer Instruction in ways that will not undermine student learning and motivation.

IMPLEMENTING PEER INSTRUCTION: RECOMMENDED GUIDELINES ON COMMON MODIFICATIONS BASED ON RETRIEVAL-ENHANCED LEARNING

Teachers commonly modify their use of Peer Instruction (Turpen and Finkelstein, 2007; Dancy et al., 2016; Turpen et al., 2016). In physics education, where Peer Instruction has been most widely practiced, Dancy et al. (2016) found that teachers often make modifications to Mazur's method by omitting one or more of the seven steps outlined in the original user manual. In addition, Dancy et al. found that teachers also modify the nine key features identified through their analysis (see above section on The Peer Instruction Method). Teachers gave variety of reasons, both personal and structural, for their modifications. Some teachers revealed that they modified the method because they did not have a clear understanding of it (e.g., they often confused Peer Instruction with general use of peer-to-peer engagement). Other teachers reported making modifications due to concerns about the limited time to cover content during class time or a perceived difficulty with motivating students to engage in the method. Finally, many teachers stated they modified the Peer Instruction method by omitting key steps and features because they were unaware that eliminating them might negatively affect learning, motivation, or other desired outcomes (Dancy et al., 2016; Turpen et al., 2016). Taken as a whole, studies on teacher implementation of Peer Instruction indicate that the common changes made to the method are not informed by the science of learning, educational research on active learning in the classroom, or even the literature on Peer Instruction itself.

The overwhelmingly positive results produced by Peer Instruction despite the prevalence of relatively uninformed modifications to the method is intriguing. This finding speaks to the robust effectiveness of Peer Instruction because a potent cocktail of mechanisms for learning remain even if one aspect of the method is removed. For example, eliminating one of the many retrieval attempts in the 7-step cycle still leaves many opportunities for retrieval practice. However, it also obscures the possible reductions in effectiveness of the method that such changes might cause. Much of the literature on Peer Instruction is built on studies in which the method is implemented in full fidelity or modified by researchers who have carefully designed the modification. The subset of studies in which modifications have been made to the method usually find positive results, but the magnitude of the observed effects may be lower, indicating an overall reduction in effectiveness. Of course, modifications could also maintain or even improve the effectiveness of the method. However, we argue that the changes to the Peer Instruction most likely to improve the effectiveness of the method are ones that are supported by theory, findings, and evidence from the science of learning and classroom research.

In this final section, we aim to help Peer Instruction practitioners understand how their choices with respect to common modifications could affect active learning in their classroom. More specifically, we provide answers to the following two questions: If a Peer Instruction user wishes to promote active learning in their classroom, what should they understand about common modifications to the method? What are some other modifications teachers can make that would be aligned with the science of learning? We focus on the concept of retrieval-based learning in order to further explicate one of the key mechanisms that drives learning in Peer Instruction. We hone in on retrieval to explain Peer Instruction effectiveness and to guide implementation for two reasons. First, as aforementioned, Peer Instruction is packed with retrieval events. Second, retrieval is one of the most firmly established mechanisms for causing student learning, retention of learning regardless of complexity of the material, and the ability to transfer learning to new contexts (Roediger and Butler, 2011). Many of the modifications made to the method reduce the number of opportunities for students to retrieve and use their knowledge. The reminder of the paper is dedicated to describing six common modifications made to Peer Instruction and discussing the potential effects of these changes. The result is a set of detailed decision-making guidelines supported by the science of learning with clear recommendations for modifying Peer Instruction.

Retrieval-Based Learning: A Key Mechanism in Peer Instruction

As explained above, retrieval practice is one of the most robust and well-established active-learning strategies in the science of learning (for review see Roediger and Karpicke, 2006; Roediger and Butler, 2011; Carpenter, 2012; Dunlosky et al., 2013; Rowland, 2014), and it pervades the Peer Instruction method. Retrieval involves pulling information from long-term memory into working memory so that it can be re-processed along with new information for a variety of purposes. The cue used to prompt a retrieval attempt (e.g., the question, problem, or task) determines in large part what knowledge is retrieved and how it is re-processed. The information can be factual, conceptual, or procedural in nature, among other types and aspects of memory. Thus, depending on the cue, retrieval can be used for anything from rote learning (e.g., the recall of a simple fact) to higher-order learning (e.g., re-construction of a complex set of knowledge in order to analyze a new idea). As people attempt to retrieve a specific piece of information from memory, they also activate related knowledge, making it easier to access this other knowledge if needed and integrate new information into existing knowledge structures. In the foregoing discussion of common modifications, we refer to the act of *attempting* to pull knowledge from memory as a retrieval opportunity. It is important to note that such an attempt to retrieve can be a potent learning event even if retrieval is unsuccessful. Science of learning researchers have demonstrated that even when students fail to generate the correct knowledge or make an error, the mere act of trying to retrieve potentiates (or facilitates) subsequent learning, especially when feedback is provided after the attempt (Metcalfe and Kornell, 2007; Arnold and McDermott, 2013; Hays et al., 2013).

The effectiveness of retrieval-based learning can be enhanced in several ways depending on how retrieval practice is implemented and structured. In our subsequent analysis of modifications to Peer Instruction, we will focus on four specific ways to make instruction that employs retrieval practice more effective:

1) Feedback—Retrieval practice is beneficial to learning even without feedback (e.g., Karpicke and Roediger, 2008), but it becomes even more effective when feedback is provided (Kang et al., 2007; Butler and Roediger, 2008)
2) Repetition—A single retrieval opportunity can be effective, but retrieval practice becomes even more effective when students receive multiple opportunities to pull information from memory and use it (Wheeler and Roediger, 1992; Pyc and Rawson, 2007).
3) Variation—Verbatim repetition of retrieval practice can be useful and effective for memorizing simple pieces of information (e.g., facts, vocabulary, etc.), but introducing variation in how information is retrieved and used can facilitate the development of deeper understanding (Butler et al., 2017).
4) Spacing—When repeated, retrieval practice is more effective when it is spread out or distributed over time, even if the interval between attempts is just a few minutes (Kang et al., 2014).

Of course, these four ways can also be used in various combinations, which creates the potential for even greater effectiveness.

Peer Instruction involves numerous retrieval opportunities that are implemented and structured in a way that would enhance the benefits of such retrieval practice. Many of the common modifications to Peer Instruction involve eliminating opportunities for retrieval practice in ways that might reduce active learning. The simplest recommendations for guiding Peer Instruction modification through the lens of retrieval-based learning are to consider increasing the number of opportunities to engage in retrieval practice, implement and structure retrieval practice in effective ways (e.g., provide feedback), and avoid omitting the retrieval opportunities present in the original method (Mazur, 1997). With that advice in mind, we now turn to analyzing some of the common modifications to Peer Instruction.

Common Modification #1: Skipping Initial Individual Thought and Response

One of the most common modifications to Peer Instruction is skipping the first retrieval event (Steps 2 and 3) and moving right into the peer discussion (Step 4) (Turpen and Finkelstein, 2009; Vickrey et al., 2015). In this modification scenario, teachers pose the question or ConcepTest, but they immediately direct students to turn to their neighbor to discuss instead of giving students time to think and respond on their own. Nicol and Boyle (2003) report that students prefer Peer Instruction when the initial individual think and response steps are included, but there are additional, more important reasons to keep the first response in the Peer Instruction cycle.

Through the lens of retrieval-based learning, skipping the initial opportunity for students to generate a response is problematic for several reasons. First, there is a learning benefit to students from attempting to retrieve information without immediate feedback, even if they are not able to generate the correct response. Second, a prominent finding from the science of learning literature is that repeated retrieval of the same question enhances learning (see Roediger and Butler, 2011). Removing the first response reduces the benefits of engaging in multiple rounds of retrieval practice on the same question throughout the Peer Instruction cycle. Finally, removing the first retrieval attempt eliminates a powerful opportunity for students to engage in metacognitive monitoring about their current understanding of the content being tested. Fostering student metacognition is critical to helping students direct their subsequent learning behavior. In summary, we offer the following guideline for Common Modification #1: Removing the first "think and response" steps eliminates a key retrieval practice opportunity and thereby reduces a key opportunity for active learning produced by Peer Instruction. Avoid this modification unless it is absolutely necessary or if you plan to replace the omitted retrieval with another equally powerful learning activity.

Common Modification #2: Revealing the Frequency of Responses Before Peer Discussion

Another common modification to Peer Instruction is revealing the results of the initial thought and response (Steps 2 and 3) before the peer discussion begins (Step 4) without revealing the correct answer (Vickrey et al., 2015). For example, some educators using clickers or other voting devices will show the results on the screen via a projector; or if using flashcards, they will reveal by verbal description the frequency of student responses after the first round (e.g., 70% of students voted A, 20% voted for B, 5% for C, and 5% for D). Some researchers have found that revealing the results of the vote (but not the correct answer) before peer discussion biases student responses to the most commonly chosen answer even if that answer is incorrect (Perez et al., 2010; Vickrey et al., 2015). However, a smaller study in chemistry education did not find a student bias when the responses were revealed before (see Vickrey et al., 2015). Although the effects of this modification deserve further investigation, we think that it is helpful to consider how it might influence retrieval-based learning in Peer Instruction. By showing the distribution of responses in the class, students may misinterpret this information as feedback and think that the most popular answer choice is correct. Such a misinterpretation could potentially confuse students or even lead them to acquire a misconception. In addition, the benefits of retrieval practice are enhanced when there is a delay between the retrieval attempt and corrective feedback (e.g., Butler and Roediger, 2008). By contrast, providing students with the class response frequencies right after the initial individual thought and response essentially constitutes immediate feedback. In summary, we offer the following guideline for Common Modification #2: Educators who elect to reveal the response frequencies before peer discussion may confuse students and negate the benefits of delaying feedback (e.g., time for students to reflect on their understanding), so we recommend not revealing students' answers after the first response round.

Common Modification #3: Refashioning Question Design

Educators use many different types and formats of questions during Peer Instruction cycles that do not always align with the original conceptualization of a ConcepTest, which is a multiple-choice test designed to build conceptual understanding (Mazur, 1997). Popular modifications include switching from multiple choice to constructed response format and using types of questions that are not necessarily aimed at conceptual understanding (Smith et al., 2009; Vickrey et al., 2015). Routinely, Peer Instruction practitioners also fill class time with administrative questions, such as polling to record attendance, using questions that require recall of basic facts to determine if students completed pre-assigned homework, or to check if students are listening during a lecture. The consensus from reviews of Peer Instruction efficacy is that questions that are challenging and involve higher-order cognition (e.g., application, analysis; see Anderson et al., 2001) are correlated with larger gains in learning than questions that require the recall of basic facts (Vickrey et al., 2015). As such, modification recommendations for Peer Instruction tend to emphasize that ConcepTest questions should tap higher-order cognition and not recall of basic facts. For the most part, theory and findings from the science of learning would agree with these recommendations. However, it is important for educators to consider the learning objectives of the course and each particular class when creating or selecting questions. If mastery of basic knowledge (e.g., vocabulary, facts) is important then giving students retrieval practice through Peer Instruction on such information is useful. Indeed, improving students' basic knowledge can form a strong foundation that enables them to effectively engage in higher-order cognition. Nevertheless, it is probably best that retrieval practice of such basic knowledge be given outside of class time and the use of ConcepTests focused on engaging students in higher-order cognition during class when the teacher and peers are available to aide in understanding.

With respect to format, Peer Instruction researchers emphasize that writing multiple-choice questions with viable distractors is one of the key elements that represent fidelity of implementation, but practitioners often lament that multiple-choice questions are difficult to construct. Although there are clear benefits to the use of multiple-choice format (e.g., ease of grading responses), the type of question being asked is much more important for learning than the format of the question (McDermott et al., 2014; Smith and Karpicke, 2014). In summary, we offer the following guideline for Common Modification #3: Feel free to be creative with the ConcepTest using different formats and types of questions, but it is probably best if ConcepTest questions posed during class time engage students in higher-order cognition. And, because even one act of retrieval can significantly enhance students' knowledge retention, ConcepTests or other Peer Instruction questions should always be aligned with specific learning objectives and not content teachers do not really want students to remember or use in the future.

Common Modification #4: Skipping Peer Discussion

Some Peer Instruction practitioners elect to skip peer discussion and only require a single round for individual thought and response. However, peer discussion represents an important learning opportunity for students because it requires them to engage in many different higher-order cognitive processes. When following Mazur's protocol, students must first retrieve their response to the ConcepTest, which provides another opportunity for retrieval practice. Next, they must discuss it with their partner, a complex interaction which involves explaining the rationale for why their answer is the correct answer, considering another point of view and (potentially) new information, thinking critically about competing explanations, and updating knowledge (if the response was incorrect). Although he does not detail it in the Peer Instruction manual (Mazur, 1997), Mazur now recommends educators to instruct students not to just "turn to your neighbor and convince them you are right" but to "find someone with a different answer and convince them you are right." Note that Smith et al. (2009) found that "peer discussion enhances understanding, even when none of the students in a discussion group originally knows the correct answer" (p. 010104-1). The task of convincing someone else about the correctness of a response may require retrieving other relevant knowledge (e.g., course content, source information about where they learned it), and thus it might be considered additional retrieval practice that is distinct but related to the ConcepTest question itself. Peer discussion also allows students to practice a host of domain-general skills, such as logical reasoning, debating, listening, perspective-taking, metacognitive monitoring, and critical thinking. Removing such a rich opportunity for active learning seems like it would have negative consequences, and indeed it does: Smith et al. (2009) found that the inclusion of peer discussion was related to larger gains in learning relative to its omission. That said, Mazur does endorse skipping peer discussion if during the first "think and respond" rounds, more than 70% of students respond correctly OR less than 30% of students correctly (Mazur, 2012). In summary, we offer the following guideline for Common Modification #4: Eliminating peer discussion removes the central feature of Peer Instruction, one that contains a cocktail of potent mechanisms for learning, especially variation in retrieval practice. Because there are benefits to peer discussion even when students have the wrong answer, we recommend always including peer discussion. In cases where the majority of the students have responded correctly, consider shortening the discussion period.

Common Modification #5: Skipping Final Individual Thought and Response (Step 5)

In Peer Instruction, some teachers may skip the final individual response round (Step 5). In this scenario, teachers deliver the ConcepTest question, solicit individual thinking and responses, engage students in peer discussion, but then move directly to an explanation of the correct answer. Although it is less common than skipping the initial individual "think and response" rounds, some teachers eliminate this step if they need to save time or a large percentage of students are correct on their initial response. Like skipping the peer discussion round when a large percentage (over 70%) of students' initial responses are correct, skipping the final response round in the same situation is endorsed by Mazur (2012). Skipping the final "think and respond" rounds eliminates an opportunity for repeated, spaced retrieval practice. Importantly, retrieval practice is substantively distinct from rote repetition—students have been exposed to new information in the interim between retrieval attempts and thus the second retrieval attempt represents a learning event that can facilitate the updating of knowledge. Such knowledge updating is likely to occur regardless of whether students' responses were correct or incorrect initially because either way they are being exposed to new information during peer discussion. In summary, we offer the following guideline for Common Modification #5: The time saved by skipping the final individual thought and response probably does not outweigh the benefits of repeated spaced retrieval practice, but a potential alternative would be shift its timing by asking students to provide their final answer and an explanation for it after class as homework (i.e., further increasing the spacing between retrieval attempts, which would be beneficial).

Common Modification #6: Skipping the Explanation of the Correct Answer

Occasionally, educators choose to eliminate the final step of the Peer Instruction method—the explanation of the correct answer (Step 7). However, this step is critically important, especially if steps 1–6 reveal that student understanding is poor, because of the powerful effects of explanation feedback on student understanding (for review see Hattie and Timperley, 2007; e.g., Butler et al., 2013). In separate studies on Peer Instruction each in a different discipline, Smith et al. (2011) and Zingaro and Porter (2014) observed larger gains in learning when an explanation was provided relative to when it was not. An ideal implementation of this final step might proceed as follows: Once students have recorded their final response, the teacher reveals the correct answer, provides explanatory feedback, and then potentially engages students in additional learning activities if the desired level of mastery has not been achieved. However, there is ample room for flexibility and customization in how explanatory feedback is provided. When using Peer Instruction, the first author often implements the final step by asking student representatives from each answer choice to again retrieve their answers and explain the rationale for supporting their response. The following script illustrates this version of Step 7:

> Teacher: "The correct answer was C; can I get a volunteer who answered differently to explain their thinking?"
>
> Student: "[Provides one or two explanations for answer choice A]"
>
> Teacher: "[Takes the opportunity to address misconceptions underlying answer choice A]. How about a volunteer who chose B or who can understand why someone else might do so?"
>
> Student: "[Provides one or two explanations for answer choice B]"
>
> Teacher: "[Takes the opportunity to address misconceptions underlying answer choice A]. Thank you, how about answer C? Why did you select C?"
>
> Student: "[Provides one or two explanations for answer choice C]"
>
> Teacher: "[Takes the opportunity to address misconceptions underlying answer choice C and provides the final explanation]"

A script for constructed responses rather than multiple choice questions would be analogous, but the teacher might specify several possible answers generated by students instead of the multiple-choice alternatives (A, B, C, etc.). It is also worth noting that this particular implementation of the final explanatory feedback step adds yet another repeated, spaced retrieval opportunity to the original method. However, students who volunteer to explain their response in front of a large group are engaging in learning event that is somewhat different from the other retrieval attempts that occurred earlier and thus it incorporates valuable variation in retrieval practice as well. In summary, we offer the following guideline for Common Modification #6: The final step of Peer Instruction invites opportunities for innovation and customization, but the one modification that we discourage is the elimination of explanatory feedback. That said, teachers should feel free to customize their approach to this explanation, such as through the above script, demonstrations, discussion, simulations, and more.

CONCLUSION

Teaching is an incredibly personal endeavor. Part of the beauty of teaching is the opportunity it provides an educator to breathe unique life into a subject to which they have dedicated their careers. Thus, it seems both natural and important for teachers to be able to personalize the way they teach so that it fits within their teaching context. Given the desire for personalization in teaching, it is imperative to allow flexibility in the use of instructional methods developed by others. A key characteristic of innovations that scale, pedagogical or otherwise, is the innovation's capacity for reinvention or customization in ways the developer did not anticipate (Rogers, 2003). Indeed, experts who study the uptake of pedagogical innovation report that teachers "rarely use a research-based instructional strategy 'as is.' They almost always use it in ways different from the recommendations of the developer" (Dancy et al., 2016, p. 12; see too Vickrey et al., 2015).

Yet, allowing the flexibility for teachers to modify instructional methods also comes with a potential cost because modifications can reduce the efficacy of the method. If a teacher using a modified version of a method observes limited or no improvement in learning outcomes, their tweaked version may lead to the erroneous conclusion that the method itself does not work; and if teachers sense the new method they have adopted does not work, they may choose to return to more familiar pedagogical habits that encourage passivity in students and yield middling results for learning (Vickrey et al., 2015; Dancy et al., 2016).

The potential for evidence-based pedagogical methods to produce poor results due to modifications creates a tension between the need to personalize teaching and the need to follow protocols that are designed to produce specific learning outcomes. We believe this tension can be resolved if teachers understand why a method is effective at facilitating learning so that they can make informed decisions about potential modifications. To this end, we have provided an analysis of why Peer Instruction is effective through the lens of the science of learning and clear guidelines regarding common modifications of the method. Peer Instruction is a remarkably flexible, easy-to-use, high-impact pedagogy that has been shown to foster active learning in a variety of contexts. By simply following the original method described by Mazur (1997), educators can infuse the state-of-the-art learning science in their classrooms and be assured they are using practices demonstrated to foster active learning. Nevertheless, the personal nature of teaching guarantees that teachers will modify Peer Instruction. We love the spirit of teaching improvement and innovation that educators are embracing, and we encourage them to make their choices by evaluating evidence from the science of learning while also considering their own unique classroom context.

AUTHOR CONTRIBUTIONS

JS and AB co-developed the concept of the paper. JS lead contributions to the Introduction, the first two sections of the paper, the section on common modification, and the conclusion and provided feedback on the remainder of the paper. AB led the section on perspectives from the science of learning and made substantive conceptual and written contributions to all sections of the paper.

ACKNOWLEDGMENTS

JS would like to acknowledge Eric Mazur for consultation on the ConcepTest portion of this paper and for his development of Peer Instruction and Charles Henderson, Melissa Dancy, Chandra Turpen, and Michelle Smith for their scholarship on educator modifications of Peer Instruction. While co-writing this article, the second author (AB) was supported by the James S. McDonnell Foundation Twenty first Century Science Initiative in Understanding Human Cognition—Collaborative Grant No. 220020483.

REFERENCES

Agarwal, P. K., D'Antonio, L., Roediger, H. L., McDermott, K. B., and McDaniel, M. A. (2014). Classroom-based programs of retrieval practice reduce middle school and high school students' test anxiety. *J. Appl. Res. Mem. Cogn.* 3, 131–139. doi: 10.1016/j.jarmac.2014.07.002

Ambrose, S. A., Bridges, M. W., DiPietro, M., Lovett, M. C., and Norman, M. K. (2010). *How Learning Works: Seven Research-Based Principles for Smart Teaching*. San Francisco, CA: John Wiley & Sons.

Ames, C. (1992). Classrooms: goals, structures, and student motivation. *J. Educ. Psychol.* 84, 261–271. doi: 10.1037/0022-0663.84.3.261

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., et al. (Eds.) (2001). *A Taxonomy For Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York, NY: Longman, Inc.

Arnold, K. M., and McDermott, K. B. (2013). Test-potentiated learning: distinguishing between direct and indirect effects of tests. *J. Exp. Psychol. Learn. Mem. Cogn.* 39, 940–945. doi: 10.1037/a0029199

Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., and Morgan, M. (1991). The instructional effect of feedback in test-like events. *Rev. Educ. Res.* 61, 213–238. doi: 10.3102/00346543061002213

Black, P. (2013). "Formative and summative aspects of assessment: theoretical and research foundations in the context of pedagogy," in *Handbook of Research on Classroom Assessment*, eds James H. McMillan Sage. (Thousand Oakes, CA: SAGE Publications, Inc), 167–178.

Black, P., and Wiliam, D. (1998). Assessment and classroom learning. *Assess. Educ. Princ. Pol. Pract.* 5, 7–74. doi: 10.1080/0969595980050102

Blackwell, L. S., Trzesniewski, K. H., and Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: a longitudinal study and an intervention. *Child Dev.* 78, 246–263. doi: 10.1111/j.1467-8624.2007.00995.x

Bloom, B. S. (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals*. Essex, England: Harlow.

Bonwell, C. C., and Eison, J. A. (1991). *Active Learning: Creating Excitement in the Classroom.* 1991 ASHE-ERIC Higher Education Reports. ERIC Clearinghouse on Higher Education, The George Washington University, One Dupont Circle, Suite 630, Washington, DC 20036-1183.

Braxton, J. M., Milem, J. F., and Sullivan, A. S. (2000). The influence of active learning on the college student departure process: toward a revision of Tinto's theory. *J. High. Educ.* 71, 569–590. doi: 10.2307/2649260

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 1118–1133. doi: 10.1037/a0019902

Butler, A. C., Black-Maier, A. C., Raley, N. D., and Marsh, E. J. (2017). Retrieving and applying knowledge to different examples promotes transfer of learning. *J. Exp. Psychol. Appl.* 23, 433–446. doi: 10.1037/xap0000142

Butler, A. C., Godbole, N., and Marsh, E. J. (2013). Explanation feedback is better than correct answer feedback for promoting transfer of learning. *J. Educ. Psychol.* 105, 290–298. doi: 10.1037/a0031026

Butler, A. C., Marsh, E. J., Slavinsky, J. P., and Baraniuk, R. G. (2014). Integrating cognitive science and technology improves learning in a STEM classroom. *Educ. Psychol. Rev.* 26, 331–340. doi: 10.1007/s10648-014-9256-4

Butler, A. C., and Roediger, H. L. III. (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Mem. Cogn.* 36, 604–616. doi: 10.3758/MC.36.3.604

Butler, D. L., and Winne, P. H. (1995). Feedback and self-regulated learning: a theoretical synthesis. *Rev. Educ. Res.* 65, 245–281. doi: 10.3102/00346543065003245

Carpenter, S. K. (2012). Testing enhances the transfer of learning. *Curr. Dir. Psychol. Sci.* 21, 279–283. doi: 10.1177/0963721412452728

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., and Rohrer, D. (2006). Distributed practice in verbal recall tasks: a review and quantitative synthesis. *Psychol. Bull.* 132, 354–380. doi: 10.1037/0033-2909.132.3.354

Chew, S. L. (2004). Using ConcepTests for formative assessment. *Psychol. Teach. Netw.* 14, 10–12. Available online at: http://www.apa.org/ed/precollege/ptn/2004/01/issue.pdf

Chew, S. L. (2005). "Student misperceptions in the psychology classroom," in *Essays from excellence in teaching,* eds B. K. Saville, T. E. Zinn and V. W. Hevern (Available online at: Society for the Teaching of Psychology Web site: https://teachpsych.org/ebooks/tia2006/index.php)

Cohen, G. L., and Sherman, D. K. (2014). The psychology of change: self-affirmation and social psychological intervention. *Annu. Rev. Psychol.* 65, 333–371. doi: 10.1146/annurev-psych-010213-115137

Craik, F. I., and Lockhart, R. S. (1972). Levels of processing: a framework for memory research. *J. Verbal Learn. Verbal Behav.* 11, 671–684. doi: 10.1016/S0022-5371(72)80001-X

Craik, F. I. M., and Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *J. Exp. Psychol. Gen.* 104, 268–294. doi: 10.1037/0096-3445.104.3.268

Crouch, C. H., and Mazur, E. (2001). Peer Instruction: ten years of experience and results. *Am. J. Phys.* 69, 970–977. doi: 10.1119/1.1374249

Dancy, M., Henderson, C., and Turpen, C. (2016). How faculty learn about and implement research-based instructional strategies: the case of Peer Instruction. *Phys. Rev. Phys. Educ. Res.* 12, 010110. doi: 10.1103/PhysRevPhysEducRes.12.010110

Dempster, F. N. (1989). Spacing effects and their implications for theory and practice. *Educ. Psychol. Rev.* 1, 309–330. doi: 10.1007/BF01320097

Draper, S. W., and Brown, M. I. (2004). Increasing interactivity in lectures using an electronic voting system. *J Comput. Assist. Learn.* 20, 81–94. doi: 10.1111/j.1365-2729.2004.00074.x

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., and Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychol. Sci. Public Interest* 14, 4–58. doi: 10.1177/1529100612453266

Eliot, C. W. (1909). *Educational Reform: Essays and Addresses*. New York, NY: Century.

Fagen, A. P., Crouch, C. H., and Mazur, E. (2002). Peer Instruction: results from a range of classrooms. *Phys. Teacher* 40, 206–209. doi: 10.1119/1.1474140

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., et al. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceed. Natl. Acad. Sci.* 111, 8410–8415. doi: 10.1073/pnas.1319030111

Glass, A. L. (2009). The effect of distributed questioning with varied examples on exam performance on inference questions. *Educ. Psychol. Int. J. Exp. Educ. Psychol.* 29, 831–848. doi: 10.1080/01443410903310674

Gonzalez, H. B., and Kuenzi, J. J. (2012). *Science, Technology, Engineering, and Mathematics (STEM) Education: A Primer*. Washington, DC: Congressional Research Service, Library of Congress.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses. *Am. J. Phys.* 66, 64–74. doi: 10.1119/1.18809

Haladyna, T. M., Downing, S. M., and Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Appl. Measure. Educ.* 15, 309–344. doi: 10.1207/S15324818AME1503_5

Harackiewicz, J. M., and Hulleman, C. S. (2010). The importance of interest: the role of achievement goals and task values in promoting the development of interest. *Soc. Personal. Psychol. Compass* 4, 42–52. doi: 10.1111/j.1751-9004.2009.00207.x

Hattie, J. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. Oxford, UK: Routledge.

Hattie, J., and Timperley, H. (2007). The power of feedback. *Rev. Educ. Res.* 77, 81–112. doi: 10.3102/003465430298487

Hays, M. J., Kornell, N., and Bjork, R. A. (2013). When and why a failed test potentiates the effectiveness of subsequent study. *J. Exp. Psychol. Learn. Mem. Cogn.* 39, 290–296. doi: 10.1037/a0028468

Henderson, C., Mestre, J. P., and Slakey, L. L. (2015). Cognitive science research can improve undergraduate STEM instruction: what are the barriers? *Pol. Insights Behav. Brain Sci.* 2, 51–60. doi: 10.1177/2372732215601115

Hestenes, D., Wells, M., and Swackhamer, G. (1992). Force concept inventory. *Phys. Teacher* 30, 141–158.

Hunt, R. R., and Einstein, G. O. (1981). Relational and item-specific information in memory. *J. Verbal Learn. Verbal Behav.* 20, 497–514. doi: 10.1016/S0022-5371(81)90138-9

Jenkins, J. J. (1979). "Four points to remember: a tetrahedral model of memory experiments," in *Levels of Processing in Human Memory,* eds L. S. Cermak and F. I. M. Craik (Hillsdale, NJ: Erlbaum), 429–446.

Johnson, D. W., Maruyama, G., Johnson, R., Nelson, D., and Skon, L. (1981). Effects of cooperative, competitive, and individualistic goal structures on achievement: a meta-analysis. *Psychol. Bull.* 89, 47–62. doi: 10.1037/0033-2909.89.1.47

Kang, S. H., Lindsey, R. V., Mozer, M. C., and Pashler, H. (2014). Retrieval practice over the long term: should spacing be expanding or equal-interval? *Psychon. Bull. Rev.* 21, 1544–1550. doi: 10.3758/s13423-014-0636-z

Kang, S. H., McDermott, K. B., and Roediger, H. L. III (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *Eur. J. Cogn. Psychol.* 19, 528–558. doi: 10.1080/09541440601056620

Karpicke, J. D., and Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science* 319, 966–968. doi: 10.1126/science.1152408

Koedinger, K. R., Booth, J. L., and Klahr, D. (2013). Instructional complexity and the science to constrain it. *Science* 342, 935–937. doi: 10.1126/science.1238056

Koedinger, K. R., Corbett, A. T., and Perfetti, C. (2012). The knowledge-learning-instruction framework: bridging the science-practice chasm to enhance robust student learning. *Cogn. Sci.* 36, 757–798. doi: 10.1111/j.1551-6709.2012.01245.x

Larsen, D. P., Butler, A. C., and Roediger, H. L. III (2013). Comparative effects of test-enhanced learning and self-explanation on long-term retention. *Med. Educ.* 47, 674–682. doi: 10.1111/medu.12141

Lasry, N. (2008). Clickers or flashcards: is there really a difference? *Phys. Teach.* 46, 242–244. doi: 10.1119/1.2895678

Lasry, N., Mazur, E., and Watkins, J. (2008). Peer Instruction: from Harvard to the two-year college. *Am. J. Phys.* 76, 1066–1069. doi: 10.1119/1.2978182

Lorenzo, M., Crouch, C. H., and Mazur, E. (2006). Reducing e Gender Gap In e Physics Classroom. *Am. J. Phys.* 74, 118–122. doi: 10.1119/1.2162549

Mawhinney, V. T., Bostow, D. E., Laws, D. R., Blumenfeld, G. J., and Hopkins, B. L. (1971). A comparison of students studying-behavior produced by daily, weekly, and three-week testing schedules. *J. Appl. Behav. Analysis* 4, 257–264. doi: 10.1901/jaba.1971.4-257

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

Mazur, E. (2012). *Peer Instruction Workshop*. Available online at: https://www.smu.edu/Provost/CTE/Resources/Technology/~/media/63F908C3C1E84A5F93D4ED1C2D41469A.ashx.

McDaniel, M. A., and Butler, A. C. (2011). "A contextual framework for understanding when difficulties are desirable," in *Successful Remembering and Successful Forgetting: Essays in Honor of Robert A. Bjork*, eds A. S. Benjamin (New York: Psychology Press), 175–199.

McDermott, K. B., Agarwal, P. K., D'antonio, L., Roediger, H. L. III, and McDaniel, M. A. (2014). Both multiple-choice and short-answer quizzes enhance later exam performance in middle and high school classes. *J. Exp. Psychol. Appl.* 20, 3–21. doi: 10.1037/xap0000004

McKeachie, W. J. (1974). Instructional psychology. *Annu. Rev. Psychol.* 25, 161–193. doi: 10.1146/annurev.ps.25.020174.001113

Metcalfe, J., and Kornell, N. (2007). Principles of cognitive science in education: the effects of generation, errors, and feedback. *Psychon. Bull. Rev.* 14, 225–229. doi: 10.3758/BF03194056

Morris, C. D., Bransford, J. D., and Franks, J. J. (1977). Levels of processing versus transfer-appropriate processing. *J. Verbal Learn. Verbal Behav.* 16, 519–533. doi: 10.1016/S0022-5371(77)80016-9

Müller, M. G., Araujo, I. S., Veit, E. A., and Schell, J. (2017). Uma revisão da literatura acerca da implementação da metodologia interativa de ensino Peer Instruction (1991 a 2015). *Rev. Bras. Ensino Fís.* 39:e3403. doi: 10.1590/1806-9126-rbef-2017-0012

National Research Council (2000). *How People Learn: Brain, Mind, Experience, and School: Expanded Edition*. Washington, DC: National Academies Press.

Nicol, D. J., and Boyle, J. T. (2003). Peer Instruction versus class-wide discussion in large classes: a comparison of two interaction methods in the wired classroom. *Stud. High. Educ.* 28, 457–473. doi: 10.1080/0307507032000122297

Nokes-Malach, T. J., Richey, J. E., and Gadgil, S. (2015). When is it better to learn together? Insights from research on collaborative learning. *Educ. Psychol. Rev.* 27, 645–656. doi: 10.1007/s10648-015-9312-8

Packman, J. L., and Battig, W. F. (1978). Effects of different kinds of semantic processing on memory for words. *Mem. Cogn.* 6, 502–508. doi: 10.3758/BF03198238

Perez, K. E., Strauss, E. A., Downey, N., Galbraith, A., Jeanne, R., and Cooper, S. (2010). Does displaying the class results affect student discussion during Peer Instruction? *CBE Life Sci. Educ.* 9, 133–140. doi: 10.1187/cbe.09-11-0080

Porter, L., Bailey Lee, C., and Simon, B. (2013). "Halving fail rates using Peer Instruction: a study of four computer science courses," in *Proceeding of the 44th ACM Technical Symposium on Computer Science Education* (Denver, CO: ACM), 177–182.

Pyc, M. A., and Rawson, K. A. (2007). Examining the efficiency of schedules of distributed retrieval practice. *Mem. Cogn.* 35, 1917–1927. doi: 10.3758/BF03192925

Rittle-Johnson, B., Schneider, M., and Star, J. R. (2015). Not a one-way street: bidirectional relations between procedural and conceptual knowledge of mathematics. *Educ. Psychol. Rev.* 27, 587–597. doi: 10.1007/s10648-015-9302-x

Roediger, H. L. III (2008). Relativity of remembering: why the laws of memory vanished. *Annu. Rev. Psychol.* 59, 225–254. doi: 10.1146/annurev.psych.57.102904.190139

Roediger, H. L. III, and Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends Cogn. Sci.* 15, 20–27. doi: 10.1016/j.tics.2010.09.003

Roediger, H. L. III, and Karpicke, J. D. (2006). The power of testing memory: basic research and implications for educational practice. *Perspect. Psychol. Sci.* 1, 181–210. doi: 10.1111/j.1745-6916.2006.00012.x

Roediger, H. L. III, Weldon, M. S., and Challis, B. H. (1989). "Explaining dissociations between implicit and explicit measures of retention: a processing account," in *Varieties of Memory and Consciousness: Essays in Honor of Endel Tulving*, eds H. L. Roediger, III and F. I. M. Craik (Hillsdale, NJ: Erlbaum), 3–41.

Rogers, E. M. (2003). *Diffusion of Innovations 5th Edn*. New York, NY: Free Press.

Rowland, C. A. (2014). The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychol. Bull.* 140, 1432–1463. doi: 10.1037/a0037559

Schell, J., Lukoff, B., and Mazur, E. (2013). "Catalyzing learner engagement using cutting-edge classroom response systems in higher education," in *Increasing Student Engagement and Retention using Classroom Technologies: Classroom Response Systems and Mediated Discourse Technologies*, eds C. Wankel, P. Blessinger (Bingley, UK: Emerald Group Publishing Limited), 233–261.

Schell, J., and Mazur, E. (2015). "Flipped the chemistry classroom with Peer Instruction," in *IChemistry Education: Best Practices, Opportunities and Trends, 1*, eds J. García-Martínex and E. Serrano-Torregrosa (Wiley-VCH), 319–334.

Smith, M. A., and Karpicke, J. D. (2014). Retrieval practice with short-answer, multiple-choice, and hybrid tests. *Memory* 22, 784–802. doi: 10.1080/09658211.2013.831454

Smith, M. K., Wood, W. B., Adams, W. K., Wieman, C., Knight, J. K., Guild, N., et al. (2009). Why peer discussion improves student performance on in-class concept questions. *Science* 323, 122–124. doi: 10.1126/science.1165919

Smith, M. K., Wood, W. B., Krauter, K., and Knight, J. K. (2011). Combining peer discussion with instructor explanation increases student learning from in-class concept questions. *CBE Life Sci. Educ.* 10, 55–63. doi: 10.1187/cbe.10-08-0101

Stuart, S. A., Brown, M. I., and Draper, S. W. (2004). Using an electronic voting system in logic lectures: one practitioner's application. *J. Comput. Assist. Learn.* 20, 95–102. doi: 10.1111/j.1365-2729.2004.00075.x

Turpen, C., Dancy, M., and Henderson, C. (2016). Perceived affordances and constraints regarding instructors' use of Peer Instruction: implications for promoting instructional change. *Phys. Rev. Phys. Educ. Res.* 12:010116. doi: 10.1103/PhysRevPhysEducRes.12.010116

Turpen, C., and Finkelstein, N. D. (2007). "Understanding how physics faculty use Peer Instruction," in *AIP Conference Proceedings* (Melville, NY: AIP), 204–207.

Turpen, C., and Finkelstein, N. (2009). Not all interactive engagement is the same: variations in physics professors' implementation of *Peer Instruction. Phys. Rev. Spec. Top Phys. Educ. Res.* 5:020101. doi: 10.1103/PhysRevSTPER.5.020101

Tyler, R. W. (1949). *Basic Principles of Curriculum and Instruction*. Chicago, IL: The University of Chicago Press.

Vickrey, T., Rosploch, K., Rahmanian, R., Pilarz, M., and Stains, M. (2015). Research based implementation of Peer Instruction: a literature review. *CBE Life Sci. Educ.* 14:es3. doi: 10.1187/cbe.14-11-0198

Watkins, J., and Mazur, E. (2013). Retaining students in science, technology, engineering, and mathematics (STEM) majors. *J. Coll. Sci. Teach.* 42, 36–41. Available online at: https://eric.ed.gov/?id=EJ1011746

Wheeler, M. A., and Roediger, H. L. III (1992). Disparate effects of repeated testing: reconciling Ballard's (1913) and Bartlett's (1932) results. *Psychol. Sci.* 3, 240–246.

Wiggins, G., and McTighe, J. (2005). *Understanding by Design 2nd Edn*. Alexandria, VA: Association for Supervision and Curriculum Development.

Yeager, D. S., and Dweck, C. S. (2012). Mindsets that promote resilience: when students believe that personal characteristics can be developed. *Educ. Psychol.* 47, 302–314. doi: 10.1080/00461520.2012.722805

Zingaro, D., and Porter, L. (2014). Peer instruction in computing: the value of instructor intervention. *Comput. Educ.* 71, 87–96. doi: 10.1016/j.compedu.2013.09.015

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in ICT

ORIGINAL RESEARCH
published: 24 May 2018
doi: 10.3389/fict.2018.00010

Toward a Neurobiological Basis for Understanding Learning in University Modeling Instruction Physics Courses

Eric Brewe[1*†‡], *Jessica E. Bartley*[2†‡], *Michael C. Riedel*[2], *Vashti Sawtelle*[3], *Taylor Salo*[4], *Emily R. Boeving*[4], *Elsa I. Bravo*[4], *Rosalie Odean*[4], *Alina Nazareth*[5], *Katherine L. Bottenhorn*[4], *Robert W. Laird*[2], *Matthew T. Sutherland*[4], *Shannon M. Pruden*[4] and *Angela R. Laird*[2]

[1] Department of Physics, School of Education, Drexel University, Philadelphia, PA, United States, [2] Department of Physics, Florida International University, Miami, FL, United States, [3] Lyman Briggs College, Department of Physics and Astronomy, Michigan State University, Lansing, MI, United States, [4] Department of Psychology, Florida International University, Miami, FL, United States, [5] Department of Psychology, Temple University, Philadelphia, PA, United States

OPEN ACCESS

Edited by:
Nathaniel Lasry,
John Abbott College, Canada

Reviewed by:
Chandralekha Singh,
University of Pittsburgh, United States
Rachel E. Scherr,
Seattle Pacific University,
United States

***Correspondence:**
Eric Brewe
eric.brewe@drexel.edu

†These authors have contributed equally to this work.

‡co-first author.

Specialty section:
This article was submitted to Digital Education, a section of the journal Frontiers in ICT

Received: 29 October 2017
Accepted: 26 April 2018
Published: 24 May 2018

Citation:
Brewe E, Bartley JE, Riedel MC, Sawtelle V, Salo T, Boeving ER, Bravo EI, Odean R, Nazareth A, Bottenhorn KL, Laird RW, Sutherland MT, Pruden SM and Laird AR (2018) Toward a Neurobiological Basis for Understanding Learning in University Modeling Instruction Physics Courses. Front. ICT 5:10. doi: 10.3389/fict.2018.00010

Modeling Instruction (MI) for University Physics is a curricular and pedagogical approach to active learning in introductory physics. A basic tenet of science is that it is a model-driven endeavor that involves building models, then validating, deploying, and ultimately revising them in an iterative fashion. MI was developed to provide students a facsimile in the university classroom of this foundational scientific practice. As a curriculum, MI employs conceptual scientific models as the basis for the course content, and thus learning in a MI classroom involves students appropriating scientific models for their own use. Over the last 10 years, substantial evidence has accumulated supporting MI's efficacy, including gains in conceptual understanding, odds of success, attitudes toward learning, self-efficacy, and social networks centered around physics learning. However, we still do not fully understand the mechanisms of how students learn physics and develop mental models of physical phenomena. Herein, we explore the hypothesis that the MI curriculum and pedagogy promotes student engagement via conceptual model building. This emphasis on conceptual model building, in turn, leads to improved knowledge organization and problem solving abilities that manifest as quantifiable functional brain changes that can be assessed with functional magnetic resonance imaging (fMRI). We conducted a neuroeducation study wherein students completed a physics reasoning task while undergoing fMRI scanning before (pre) and after (post) completing a MI introductory physics course. Preliminary results indicated that performance of the physics reasoning task was linked with increased brain activity notably in lateral prefrontal and parietal cortices that previously have been associated with attention, working memory, and problem solving, and are collectively referred to as the central executive network. Critically, assessment of changes in brain activity during the physics reasoning task from pre- vs. post-instruction identified increased activity after the course notably in the posterior cingulate cortex (a brain region previously linked with episodic memory and self-referential thought) and in the frontal poles (regions linked with learning). These preliminary outcomes highlight brain regions linked with physics reasoning and, critically, suggest that brain activity during physics reasoning is modifiable by thoughtfully designed curriculum and pedagogy.

Keywords: modeling instruction, physics reasoning, mental models, force concept inventory, fMRI, STEM learning, brain network, neuroeducation

INTRODUCTION

Active learning is neither a curriculum nor a pedagogy. Active learning is a class of pedagogies and curriculum materials that strive to more fully engage students and promote critical thinking about course material. Students learn more effectively when they engage in investigations, discussions, model building, problem solving, and other active explorations (National Research Council, 2012; Kober, 2014). However, typical university instruction in physics (and other Science, Technology, Engineering, and Mathematics [STEM] fields) has been lecture-based. While lectures can be interesting, and some students clearly have been trained to become engaged during lectures (Schwartz and Bransford, 1998), for the majority of students, lectures are passive activities. This mismatch between the ways that students learn and the way many classes are taught is the primary motivation for the transformation of STEM instruction. When classrooms are transformed, the evidence is overwhelming; students learn more and are more likely to succeed in active learning settings (Freeman et al., 2014).

Multiple transformative curricula and pedagogical approaches have been developed for introductory physics to promote active learning. For example, *Peer Instruction* emerged to enhance standard lecture-based approaches by incorporating conceptual questions for discussion and, in turn, facilitated development of personal response systems (Crouch and Mazur, 2001). *Tutorials in Physics* were developed to supplement standard lectures through use in recitation sections (McDermott and Shaffer, 2001). Other materials such as *Student Centered Active Learning Environment with Upside-down Pedagogies* [SCALE-UP] (Beichner and Saul, 2003) and *Investigative Science Learning Environments* [ISLE] (Etkina et al., 2006; Etkina and Van Heuvelen, 2007) implement a studio-format that integrates lab and lecture, including greater amounts of conceptual reasoning and greater emphasis on exploration. Modeling Instruction (MI) is an active learning approach (Brewe, 2008) similar to SCALE-UP and ISLE in that it is a complete course transformation integrating lab and lecture components into one studio format class. However, MI is distinct from other reforms in that it was built around an explicit epistemological theory of science, and this foundation is one of the motivations for using functional magnetic resonance imaging (fMRI) to study how learning physics may impact brain network development.

Hestenes (1987) avers that science by its very nature is a modeling endeavor. Science proceeds by developing models that describe and ultimately predict phenomena. As a model is developed, it is validated through the interplay between the predictions generated by the model and the evidence that emerges supporting such predictions. Once a valid model has been developed, the model is deployed to new situations. This is a process which Kuhn (1970) called "normal science," whereby scientists use existing prevalent models to explore the models' limits of applicability and search for places where the models give rise to predictions in contrast with evidence. Ultimately, models reach their limits of applicability and need to be revised or in some cases abandoned entirely, beginning what Kuhn called "revolutionary science." When this happens, a new model is proposed, and the cycle begins anew.

The modeling theory of science is the theoretical and epistemological basis of MI. This, however, is a theory of *science*, not a theory of *science instruction*. It translates to instruction through the premise that, if modeling is how science proceeds and we believe students should be engaged in authentic scientific practices, then instruction should be designed to engage students in the process of modeling. Wells et al. (1995) describe the Modeling Cycle as the recursive process of engaging students in model development, validation, deployment, and revision.

In this paper, we first provide an overview of the theoretical background, development process and critical features behind MI as a transformative curricula and model-building endeavor. This overview serves to motivate why scientific model development in students resulting from university instruction warrants further investigation not only at the academic (e.g., grades) and social level (e.g., social networks) but also at the neurobiological level as a putatively measurable phenomena that occurs within the brain. Then, we shift focus to present results from a fMRI study in which we measured brain activity among students engaged in physics reasoning and model use before and after they completed a MI course. We subsequently discuss the results which show distinctive brain activity related to physics reasoning and that instruction consistent with a Modeling theory of science modifies brain activity from pre to post-course.

Role of Conceptual Models in Introductory Physics Curriculum

Building instruction around modeling necessitates a working understanding of models. To date, research in the MI context has focused on conceptual models, which are instructionally useful, rather than mental models, which have been difficult to directly observe. Herein, we seek to expand upon existing research by adopting neuroimaging techniques to interrogate mental models among students receiving instruction via an explicit conceptual modeling approach (i.e., MI). We operate from the following definition of a conceptual model: conceptual models are purposeful coordinated sets of representations (e.g., graphs, equations, diagrams, or written descriptions) of a particular class of phenomena that exist in the shared social domain of discourse. This definition has several features worth elaborating. First, it fits on a t-shirt. Second, this definition establishes the domain, purpose, and composition of conceptual models, which we expand upon below. Finally, this definition of conceptual models has helped us design research to look for evidence of the modeling process in classrooms. **Figure 1** illustrates the relationship between conceptual and mental models.

Attempting to synthesize the many definitions and descriptions of models is not our purpose. Instead, we aim to highlight some of the features of our definition that were relevant to the development of the MI approach based on building, validating, deploying and revising models. These features (i.e., the composition, purpose, and domain of conceptual models), then will be used to structure the investigations into the nature of student's mental model formation as measured via brain-based fMRI data.

Composition

Conceptual models are composed of representations. Representations are human inventions/constructs that stand in for the phenomena (Morgan and Morrison, 1999; Giere, 2005; Frigg and Hartmann, 2006; Windschitl et al., 2008; Schwarz et al., 2009). In physics, common types of representations include graphs, vector diagrams, equations, simulations, words, and pictures (Krieger, 1987). From the MI perspective, this means that instruction should focus on helping students to identify, use, and interpret representational tools that are useful in describing physical systems. Instruction around model building necessarily focuses on what representations are common to a discipline, how they are used, and how information can be extracted from them. Further, the coordination of these representations helps to build a more robust model, and provide a variety of ways to extract information from the model (Hestenes, 1992; Halloun, 2004).

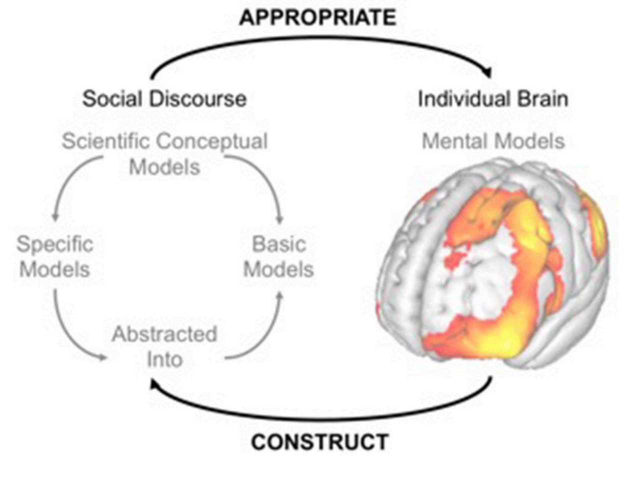

FIGURE 1 | Schematic of the relationship between conceptual and mental models in physics curriculum.

Purpose

Morgan and Morrison (1999) described mental models as mediators of thought, autonomous from, but in correspondence with the system they represent. This mediating function of models establishes the roles that models have within science as the center of thought, explanation, and prediction (Craik, 1943; Johnson-Laird, 1983). For example, Craik (1943) stated, "*If the organism carries a 'small-scale model' of external reality and of its own possible actions within its head, it is able to try out various alternatives...*" Instructionally, if models fill this role of mediators of thought, then models should structure the organization of the curriculum. Models also allow students to address new phenomena (Odenbaugh, 2005; Svoboda and Passmore, 2011; Gouvea and Passmore, 2017). This purpose is built into the instructional modeling cycle where students are encouraged to understand new phenomena by deploying existing models to extract information about and characterize the phenomena. When existing models do not work, students *are expected to adapt or redevelop models that can account for these new phenomena.*

Domain

We propose a distinction between *scientific conceptual* and *mental model* domains and place conceptual models in the shared social domain of discourse. This perspective differs from other conceptualizations where mental models within individuals' minds/brains are implicitly or explicitly the center of focus (Greca and Moreira, 2000, 2001). Specifically, to infer the status of a student's mental model, investigators typically assess students' actions or behaviors, such as writing, speaking, drawing, predicting, or arguing (Halloun, 1996a; Justi and Gilbert, 2000; Lehrer and Schauble, 2006). Thus, evidence of model-based reasoning exists external to the individual and is contingent on an external evaluation. Instructionally, our efforts have been to help students develop models as a distributed cognitive element. Meaning that each individual student will have an instantiation of the shared model, but the visible elements of the model exist external to individuals through writing, speaking, drawing, diagraming, predicting, and/or simulating. This notion of shared models improves team performance and the learning process (Mathieu et al., 2000). As such, the design of the MI curriculum and pedagogy focuses not on mental models *per se*, but on the social construction of a model. In other words, we focus students on using consistent representational tools to build models of phenomena in an interactive team environment. Models are shared among class members and agreed upon before deploying these models to analyze new situations. We provide a more detailed description of the classroom setting in section "Features of MI Learning Environment" but much of class time is spent in small groups developing models of specific phenomena on small portable whiteboards, which are then presented at larger "board meetings." The interplay between smaller and larger groups provides a vehicle for students to use diagrams, equations, or graphs to represent elements of the model.

We do not reject that individuals have internal mental models, or that these mental models include connections between representations and concepts, or interactions between mathematics and intuition, for example. As Rogoff (1990) points out, cognitive functions are essential components of purposeful action. We are aligned with the notion that scientific conceptual models are distributed cognitive elements, which are then appropriated by individuals. During the appropriation, students construct the mental models in correspondence with the scientific conceptual models. Rather our point is that assessing external behaviors speaks to the conceptual model domain and assessing the mental model domain would benefit from directly considering the brain.

Role of Conceptual Models in Instruction

For instructional purposes, models represent an appropriate and accessible level of abstraction (Halloun, 2004). Within a larger context, models occupy the middle level of a conceptual hierarchy (**Table 1**; Halloun, 2004; Matthews, 2007) which is best illustrated by a representative example (Lakoff, 1987). Veterinarians are not likely to study the superordinate category of animals, which is too broad a categorization to be useful. Nor are they likely to study the subordinate category of retrievers; this is too specific to be broadly useful. Instead, dogs are likely to be the level of focus. This level is referred to as the "basic" level and is considered the ideal focus for instruction (Halloun, 2004).

In the MI classroom, building basic conceptual models begins with considering a specific phenomenon to be described. Once a target phenomenon is established, the next step is to characterize the phenomena through relevant representational tools. For example, using velocity vs. time graphs to represent the motion of a moving object. As students create representations of the object's motion, a model of this specific phenomenon is being developed, or what we call a *specific model*. These specific models are not generally applicable, they pertain to the specific details of the situation being considered. By necessity, specific models are predecessors to *basic models*. Specific models are made more robust as additional representational tools are introduced and integrated with existing ones. Introduction of representational tools and the subsequent negotiation of their use and interpretation are motivated by specific phenomena to be modeled, so the models created are always specific models.

However, a desirable scientific skill is to reason based on *general models* (Nersessian, 1995, 2002a,b). As such, the MI curriculum and pedagogy is specifically designed to facilitate the students' transition from specific to basic models. Basic models, which are general and represent entire classes of phenomena (such as a constant acceleration model), are abstracted from a collection of specific models (Halloun, 1996b, 2004). For example, the general features of a basic constant acceleration model can be abstracted from specific models of objects undergoing constant acceleration, such as objects in free fall, or uniformly slowing down. This is achieved in the MI classroom by having students consider a number of specific models, and then identifying the features that are similar to all such models. For example, all constant acceleration models include a linear velocity time graph. These similar features are then compiled into one model that can be used for all situations, a basic model. Basic models are useful because they are not tied to a specific phenomenon, much like the Standard Model is a basic model built up and abstracted from the specific models of atomic collisions, particle interactions, etc. Basic models are essential in science as they promote abstract reasoning about novel phenomena (Nersessian, 1995); when physicists seek to understand interactions of atomic particles they start by using the Standard Model.

TABLE 1 | Conceptual and Categorical Hierarchies.

	HIERARCHY	
Conceptual	**Level**	**Categorical**
Theory	Superordinate	Animal
Model	Basic	Dog
Concept	Subordinate	Retriever

Once a basic model is established, students deploy the model in a variety of settings. This deployment phase is most aligned with the standard problem solving that happens in physics classes. The purpose is to develop skill at adapting the representations that make up the model to new situations and extracting information about the situation from the representations.

The final stage in the MI instructional cycle is revision. Revision of a basic model happens when students encounter a phenomenon that does not fit with the model's assumptions. An example often encountered comes when students attempt to generate a specific model of two-dimensional motion on the basis of a one-dimensional constant acceleration model. The one-dimensional case is inadequate without modification to understand motion in two dimensions, and thus must be revised. In some cases, revision involves a simple modification of the representational tools, and in other cases, it requires starting with an entirely different model.

In summary, the modeling cycle of MI describes the progression of course content. In addition, MI also interweaves social interactions designed to facilitate discourse in the service of building conceptual models. Next, we more fully describe the precise aspects of the MI learning environment that support the development, validation, deployment, and revision of models.

Features of MI Learning Environment

Basic conceptual models are often well-developed for scientists and course instructors, yet these models are not well-developed for the students in introductory physics courses. Accordingly, the first contextual feature of the MI classroom is to support students in re-developing constituent basic models within their own learning environment. The MI instructor's role is thus to guide students through the development of these basic conceptual models by establishing activities and providing scaffolding to manage student discourse and promote model building and deployment. In this way, the MI curriculum and pedagogy can be considered a guided inquiry approach. Students are not expected to discover physical laws without strong instructor guidance who chooses activities, introduces representational tools, and guides students toward their appropriate use and interpretation. In this way, the instructor is a guide to the disciplinary norms and tools.

Student Participation in a Model-Centered Learning Environment

Accomplishing this fundamental re-development of basic conceptual models requires students to be active and engaged participants in the learning environment. Accordingly, there are specific ways MI students are expected to participate in the re-development of basic conceptual models. First, students are expected to be involved in identifying the way that tools such as pictures, diagrams, graphs, and equations are used to represent phenomena. They are not expected to invent or discover these tools, but instead to determine with instructor guidance how these tools are used and how to interpret these representations. For example, how does a vector representation of forces describe interactions the object is involved in, and what do these forces allow us to infer about the current state of the object and its future behavior? Second, students are expected to be involved in the interpretation of these representational tools and drawing inferences from them as they pertain to physical laws. Third, students are expected to then deploy these established basic conceptual models by extending them to novel situations. Finally, students are expected to communicate basic conceptual models. This promotes greater expertise with the models when presenting to others and facilitates competence in scientific communication skills.

Studio Format

MI is designed for implementation in a studio-format classroom. In studio physics classrooms students are able to flexibly engage in various types of activities, which may include labs, conceptual reasoning, or problem-solving activities. At Florida International University (FIU), the MI classroom integrates both the lecture and lab components of the introductory physics course and meets for a total of 6 h per week across 3 days. Typically, students work in small groups of three to complete in-class activities. This small group work is summarized on small portable whiteboards. These whiteboards are then presented in larger group "board meetings" where all students in the class actively participate.

Small Group Participation

During the small group component, students work on model-building activities. In these groups, students begin the process of reaching consensus by creating whiteboards for sharing or "publishing" their lab results and/or solutions to problems. The instructor's role is to circulate through the classroom, asking questions, introducing new content, and examining the whiteboards that are being prepared. This small group work allows students to work together on a model-building activity, generate conceptual models, and practice communicating scientific information in a relatively "low-stakes" setting.

Large Group Participation: The "Board Meeting"

The practice of having students first work in small groups and then present their outcomes to a larger group provides students with multiple opportunities to negotiate the use of conceptual models. The board meetings involve all students in the class gathering in a circle such that every member can see every other member and every groups' boards. During the board meeting, the instructor assumes the role of disciplinary expert and guides the discourse toward a shared conceptual model. Facilitating the discussion involves moderating the groups' whiteboard presentations, addressing student questions, and helping groups clarify their presentations and understanding. The instructor's guidance during the board meetings relies heavily on providing student groups with formative feedback. The explicit goal of these board meetings is to reach consensus regarding the conceptual models. In addition to the explicit goals, tacit goals include establishing the norms of a discourse community and encouraging students to utilize scientific argumentation strategies (Passmore and Svoboda, 2012). These strategies include supporting claims with evidence and reasoning based on the shared conceptual models.

Pairing Large and Small Group Interactions

The combined interaction structure is designed to elicit target conceptual models. The structure of these interactions also mimics the structure of science in general and physics in particular as practiced in a research setting. Students work in small research groups, building up and synthesizing the conceptual model that is subsequently 'published' at the board meeting, much like a scientific meeting. Both the small and large group settings rely on the pedagogical skill of the instructor. In MI-like environments (which are less "instructor-centered" than traditional classrooms), the trajectory of the learning takes varied paths based on the input of the participants. For this reason, the curriculum and pedagogy of MI are less like a script for an actor to follow, and more like a set of guidelines for an improvisational comedienne.

Impact on Student Outcomes

The combination of curriculum materials designed to recursively implement the modeling cycle and a learning environment and pedagogy that are similarly supportive have been shown to be effective at promoting learning. Like other transformed curricula in university physics, MI promotes both conceptual understanding and student success in introductory physics (Brewe et al., 2010b). A survival analysis suggests that the increased success rate in introductory physics is not a result of lowered standards, as students from MI classes showed equivalent likelihood of success in completing a major in physics as students from lecture classes (Rodriguez et al., 2016). MI students also report improved attitudes about learning physics (Brewe et al., 2009, 2013) and these attitudinal shifts are equitable in terms of ethnicity (Traxler and Brewe, 2015). The group interactions in a MI class promote more well-developed classroom networks (Brewe et al., 2010a), and these networks are known to facilitate retention in physics courses (Zwolak et al., 2017). Positive shifts in self-efficacy associated with participating in MI have been documented, (Sawtelle et al., 2010) although not consistently (Dou et al., 2016). We are in the process of studying qualitatively the construction of a conceptual model in MI (Brewe and Sawtelle, 2018) and investigating students' representational choices in problem solving (McPadden and Brewe, 2017). These studies are consistent with students constructing and using conceptual models to solve problems and analyze physical systems. The successes coming from the MI classroom motivate our current research into the neurobiological mechanisms of reasoning in physics.

Investigating Mental Model Development Using Neuroimaging

While prior assessments of MI's impact on students has typically focused on the social construction of conceptual models (Brewe, 2008, 2011; Sawtelle et al., 2012), here we consider MI's potential impact on mental models using brain imaging techniques. This study aimed to investigate brain activation during a physics reasoning task and changes in brain activation after MI course instruction relative to before such instruction. Previous neuroimaging studies have localized brain activity associated with reasoning across various modalities (e.g., mathematics, formal logic, and fluid reasoning; Prabhakaran et al., 1997; Arsalidou and Taylor, 2011; Prado et al., 2011), but no investigations have probed for such brain activity in the field of physics or across physics classroom instruction. Because of this, no standardized tasks have been adapted for the MRI environment to examine such brain activation. Therefore, as a first step, we sought to develop a novel neuroimaging paradigm to probe brain activity during physics reasoning. We focused the development of this task on mental model use during physics reasoning, as previous research has provided evidence that students' use a variety of mental models during conceptual physics reasoning (Nersessian, 1999; Hegarty, 2004). Thus, we adapted items from the well-known *Force Concept Inventory* (FCI; Hestenes et al., 1992) which is known to engage conceptual physics reasoning. FCI questions were modified to fit with the parameters of the MRI data collection, and to investigate physics reasoning, (see section "Physics Reasoning Task" for further details. Simultaneously, to facilitate formation of neuroanatomical hypotheses regarding the brain networks we might observe during physics reasoning, we conducted a neuroimaging meta-analysis (Bartley et al., in press) of fMRI studies that investigated problem solving across a diversity of representation modalities. Briefly, the primary outcome of that meta-analysis was that similar reasoning tasks using mathematical, verbal, and visuospatial stimuli involving attention, working memory, and cognitive control, activated dorsolateral prefrontal and parietal regions. Participants completed this physics reasoning task while undergoing functional magnetic resonance imaging (fMRI) scanning, both before (pre) and after (post) completing a physics course in order to investigate the putative impact of physics instruction on brain function. Driving this neuroeducation project were two main hypotheses: (1) This novel physics reasoning task would induce increased activity in brain regions previously associated with attention, working memory, and problem solving (e.g., lateral prefrontal and parietal regions), and (2) Activation patterns would differ from pre- to post-course, indicating that brain activity can be modified as a result of physics instruction.

A few prior studies have demonstrated that short- and long-term course instruction can impact brain function. Differences in brain function have been observed from pre- to post-course among students enrolled in a 90-day Law School Admission Test preparation course (Mackey et al., 2013). Mason and Just (2015) showed that providing information to research participants about mechanical systems while in the MRI scanner, which they called physics instruction, led to changes in knowledge representation during successive stages of learning. In a separate study, they were also able to use machine learning and factor analysis to identify neural representations of four physics concepts: motion visualization, periodicity, algebraic forms, and energy flow (Mason and Just, 2016). However, to our knowledge, this is the first neuroeducational study to consider the impact of a full, semester-long physics class on the brain.

Brief Primer on Neuroimaging Studies

This manuscript is intended for an educational research audience, with the expectation that readers have not had extensive experience with neuroimaging as a research methodology. As such, this section provides a brief overview of neuroimaging studies, particularly fMRI. In neuroimaging studies, researchers develop an experimental task to isolate mental operations of interest that participants perform lying in a MRI scanner while a series of three-dimensional brain images are acquired. Typically, these brain images are acquired approximately every 2 s and are composed of small volume elements called voxels, which in this study measured 3.4 mm^3. Within each voxel, the blood's changing oxygen levels (known as the blood-oxygenation level-dependent [BOLD] signal) are measured. Task-related changes in the BOLD signal provide an indirect measure of brain activity. In one implementation of fMRI experimental design, brain images are collected in blocks. During 'active task' blocks, participants are presented a stimulus (e.g., a physics question) engendering cognitive processes of interest (e.g., physics reasoning) and are instructed to make a response using a MRI-compatible keypad. During carefully constructed 'control task' blocks, participants are also presented with stimuli and give responses; however, the stimuli presented do not engender the cognitive processes of interest. Contrasting active blocks with control blocks presumably isolates task-related brain activity associated with the cognitive processes of interest and excluding those common to both conditions (e.g., visual processing, word reading, button pressing).

Following data collection, fMRI data are processed to correct for in-scanner head movement and fitted to a standardized brain template to enable averaging over a group of participants. BOLD time series from each voxel are input into a general linear model (GLM) including distinct regressors for various task events (and other known sources of noise) to characterize the degree to which variability in the BOLD signal correlates with those task events. Resulting beta weights from active and control task blocks can then be contrasted and significant differences are interpreted as differences in brain activity between blocks. This procedure is repeated for the BOLD time series across all voxels in the entire brain. Additional multi-level modeling can be performed on these results, as was done in this study, to test for changes in brain activity across repeated measures (i.e., from pre- to post-instruction).

METHODS

Participants

Participants were drawn from MI classes at FIU over the course of 3 years (academic years 2014–2017). We recruited 55 students (33 male, and 22 female) in the age range of 18–25 years old (mean ± SD: 20.1 ± 1.4). All participants were screened to be right-handed, not using psychotropic medications, and free of psychiatric conditions, cognitive or neurological impairments, and MRI contraindications. Volunteers invited to participate had not previously taken a college physics course and met either a GPA (>2.24) or SAT Math (>500) inclusion criteria. These criteria were implemented to minimize between-participant variability that could confound brain measurements associated with the experimental conditions. Written informed consent to a protocol approved by the FIU Institutional Review Board was obtained from all participants. Imaging data were collected on a General Electric 3-Tesla Healthcare Discovery 750 W MRI scanner located in the Neuroimaging Suite (NIS) of the Department of Psychology at the University of Miami (Coral Gables, FL). Each participant completed a 90-min MRI scanning session at both a pre- and post-instruction time point. The pre-session scans were scheduled within the first 4 weeks of the semester and the post-session scans were completed in the first 2 weeks following the semester. All participants were compensated for their time participating in the MRI assessment ($50 for pre- and $100 for post-scans).

Physics Reasoning Task

We adapted a set of questions from the Force Concept Inventory (FCI) for presentation in the MRI scanner (**Figure 2A**). The FCI was chosen given the substantial amount of extant data from students in MI at FIU on this measure (Brewe et al., 2010b), established reliability measures (Lasry et al., 2011), and known time requirements (Lasry et al., 2013). The FCI is a 30 question, multiple choice conceptual survey of students understanding of Newtonian mechanics (Hestenes et al., 1992). Each question has five multiple choice options, one correct and four distractors which were originally generated from student responses to open-ended versions of the same questions. The questions present "every-day scenarios," do not require any mathematical calculations, and are presented as text describing the scenario accompanied by a representational diagram. To ensure that MRI data collection sessions were manageable and well-tolerated by participants, we reduced the number of FCI questions from 30 to nine (FCI 2, 3, 6, 7, 12, 14, 27, and 29). These nine questions were selected to span a range of difficulty levels that were simultaneously challenging enough to tax the mental resources of participants, but not necessarily the most difficult items in the FCI, as determined by item response curves in Morris et al. (2012) (**Table 2**). Additionally, because measurement of brain networks via fMRI require the repeated observations across multiple yet similar experimental trials, we sought to narrow the broad range of physics-related cognition being probed in this task and selected questions that required students to determine the trajectories and motion of objects as resulting from different scenarios and combinations of initial velocities and/or force configurations. Given technical constraints associated with the use of a four-button MRI-compatible keypad, the questions were modified by removing the least chosen of the five multiple choice options, as indicated by the item response curves of Morris et al. (2012). In the current neuroimaging task implementation, each question was parsed into three self-paced presentation phases; participants were allowed to control the timing of these phases. The first phase of the question involved presentation of the text describing the phenomena and an accompanying diagram. The second phase posed the question, and the third phase presented the multi-choice answer options. FCI responses were assessed for overall and item-specific accuracy.

In addition to FCI questions, participants answered a series of "control questions" (**Figure 2B**), each of which had similar characteristics to the FCI questions in terms of reading requirements, visual complexity, and overall design. However, control questions did not inquire about physics-related content, instead these questions focused on reading comprehension and shape discrimination. Control questions allowed us to isolate cognitive processes presumably related to physics reasoning when contrasting FCI ("active task") vs. control questions ("control task").

FCI and control questions were presented in pseudo-random orders within three task runs. Each question was followed by 20 s of "rest," during which participants maintained their gaze on a fixation cross centrally projected on the screen. These three runs lasted approximately 6 min each. Participants received instruction and practice on the task in a carefully managed mock scanner training session to ensure correct performance during the MRI session. In addition to acquainting participants to the task, the mock scanner also allows participants to experience what the actual MRI scan will be like.

Data Analysis

Details on fMRI data acquisition parameters can be found in the Supplementary Materials. Prior to analysis, the data were preprocessed using commonly used neuroimaging analysis software packages: FSL (FMRIB Software Library, www.fmrib.ox.ac.uk/fsl) and AFNI (Analysis of Functional NeuroImages, http://afni.nimh.nih.gov/afni). Standard fMRI preprocessing procedures involved motion correction to remove signal artifacts associated with head motion, high-pass filtering to remove low frequency trends in the signal associated with non-brain noise sources (i.e., cardiac or respiratory), and spatial smoothing to increase signal to noise ratio during analysis. The data were then mapped to a standardized brain atlas (MNI152) to allow for group-level assessments.

We conducted two primary analyses to identify: (1) brain regions linked with physics reasoning (task effect) and (2) changes in brain activity associated with physics instruction (instruction effect). To delineate brain regions linked with

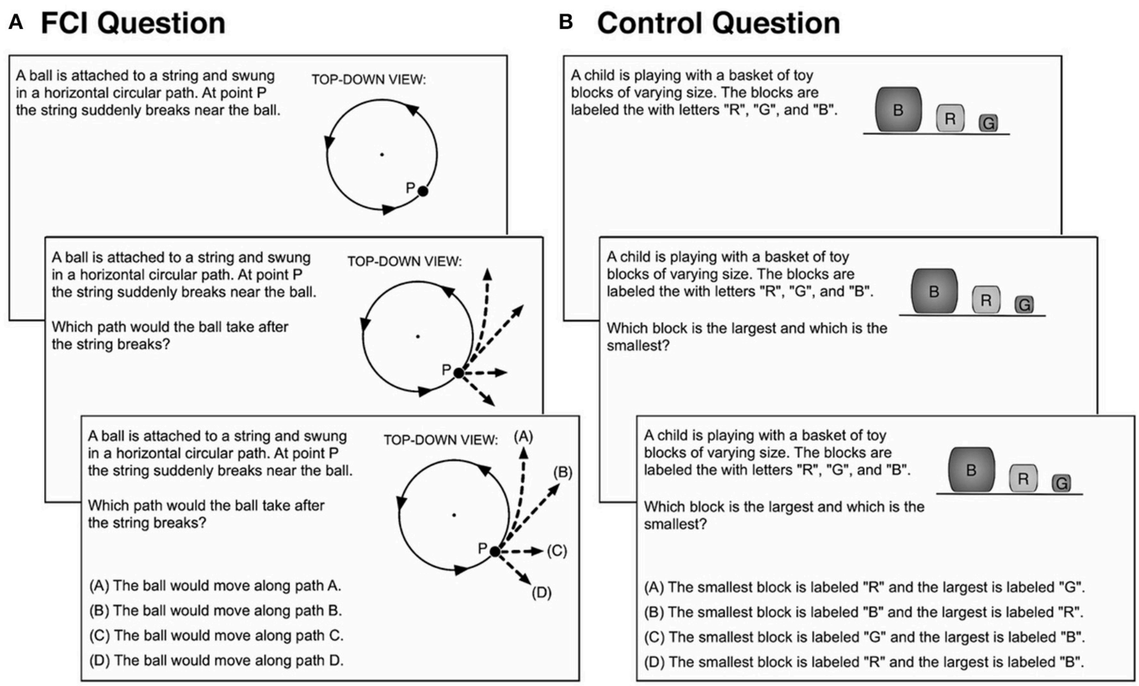

FIGURE 2 | Example items from the physics reasoning fMRI task. **(A)** FCI questions described a physical scenario using pictures and words and then asked a physics question followed by four potential answers. **(B)** Control question shared basic visual and linguistic features with FCI questions, however control questions did not ask students to engage in physics reasoning.

TABLE 2 | Overall and individual item accuracy for pre and post instruction FCI questions in the scanner.

FCI Question	Pre %	Post %	Change (Post–Pre) (%)	Item Difficulty Morris et al., 2006 (%)
2	29.5	39.3	+9.8	34.6
3	42.6	58.9	+16.3	51.5
6	78.7	78.6	−0.1	73.6
7	54.1	71.4	+17.3	66.4
8	39.3	46.4	+7.1	50.4
12	45.9	69.6	+23.7	65.2
14	24.6	41.1	+16.4	39.5
27	44.3	46.4	+2.1	59.4
29	42.6	85.7	+43.1	50.8
Total	44.6	59.7	15.1	

Item difficulty measures from Morris et al. (2006) are included for comparison.

physics reasoning at the pre-instruction time point, each preprocessed fMRI data set was input into a voxel-level General Linear Model (GLM) including regressors for the FCI and control task conditions (and various nuisance signals). Contrast images were created for each participant by subtracting the beta weights associated with the control questions from those for the FCI questions representing the degree to which each voxel responded more during physics reasoning as compared to the control condition (FCI > Control). These participant-level contrast images were then input into a group-level, one-sample t-test and significant physics reasoning-related brain activations were defined using a threshold of $P_{corrected} < 0.05$ ($P_{voxel-level} < 0.001$, family-wise error [FWE] cluster correction). To delineate brain regions showing physics reasoning-related activation changes following a MI course, the participant-level FCI > Control task contrast images (described above) from the pre- and post-instruction data collection sessions were input into a group-level, paired samples t-test. Both Pre > Post and Post > Pre contrasts were computed and significant instruction-related brain activity changes were defined using a $P_{corrected} < 0.05$ threshold ($P_{voxel-level} < 0.001$, FWE cluster correction). Follow up correlational analyses were also conducted between the BOLD signal change across instruction (Post > Pre) in the four largest significant clusters ($\geq$1,000 voxels) identified in the instruction effect analysis described above and accuracy post-instruction on the FCI using $P < 0.0125$, Bonferroni corrected,. Because the clusters probed showed significant extent across multiple brain areas, BOLD signal was extracted from spherical seeds centered at the peaks z-score of each cluster.

RESULTS

Accuracy

Table 2 includes the accuracy results of student responses for the nine questions in the pre and post-instruction scans along with item difficulties based in classical test theory, Morris et al.

(2006). A paired-samples t-test was conducted to compare post- vs. pre-instruction means. Cohen's d, was calculated to identify the magnitude of the effect, and 95% confidence intervals on the effect. The results of the t-test [$t_{(55)} = 6.31$, $p < 0.001$] and Cohen's d ($d = 0.84$) with a 95% confidence interval of 0.45–1.23 indicated with a high degree of confidence that response accuracy increased after instruction. These results are consistent with prior results examining increased FCI accuracy after course instruction (Brewe et al., 2010b). Furthermore, these accuracy results from participants in the scanner are in line with the classical test theory item difficulty (outside the scanner performance), where difficulty is calculated as the average score on a particular item.

Task Effect

MI students exhibited physics reasoning-related brain activity (FCI > Control) at the pre-instruction time point in four general brain areas, the prefrontal cortex, the parietal cortex, the temporal lobes, and the right cerebellum (**Figure 3**, red; Supplemental Table 1). More specifically, in the prefrontal cortex (PFC), activation peaks were observed in the left superior frontal gyrus (SFG), dorsomedial PFC (dmPFC), bilateral dorsolateral PFC (dlPFC), inferior frontal gyri (IFG), and orbitofrontal cortex (OFC). Within the posterior parietal cortex, brain activity was observed bilaterally in the supramarginal gyri, intraparietal sulcus (IPS), and angular gryi (AG). Large bilateral clusters of activation during physics reasoning were also observed in middle temporal (MT) and medial superior temporal (MST) areas. These same patterns of task-related brain activity from the pre-instruction stage were also observed when performing a similar assessment at the post-instruction stage (data not shown).

Instruction Effect

Significant increases in brain activity following instruction (Post > Pre) were observed within prefrontal and parietal cortices (**Figure 3**, blue; Supplemental Table 2). In particular, three clusters of increased PFC activity were identified in the left dlPFC along the inferior precentral sulcus, and bilaterally in the frontal poles. Parietal areas demonstrating increased activation after instruction were located in the posterior cingulate cortex (PCC) extending into retrosplenial cortex and the precuneus and in the left angular gyrus. No brain regions showed significantly more task-related activity at the pre-instruction stage as compared to post-instruction (Pre > Post). Follow up correlation analysis between the left PCC, left angular gyrus, left orbital frontal pole, and left DLPFC and accuracy on the FCI yielded no significant correlation ($r_{pcc} = -0.12$, $p_{corrected} = 1$; rag $= -0.07$, $p_{corrected} = 1$; $r_{ofc} = -0.01$, $p_{corrected} = 1$; $r_{dlpfc} = 0.02$, $p_{corrected} = 1$).

DISCUSSION

This neuroeducational study represents an initial effort to understand how physics reasoning may translate to the level of brain function assessed by fMRI and how instruction brings about changes in brain activity. To this end, we have provided fMRI results of brain activation from two main assessments. First, we observed that the physics reasoning task (FCI > Control questions) was associated with increased brain activity notably in lateral prefrontal and parietal regions. Second, we observed that students who completed the MI course showed increased activation during the physics reasoning task after the course in the posterior cingulate cortex and frontal pole regions.

Accuracy and Physics Reasoning

Participant responses to the FCI questions in the scanner show accuracy that is in line with published item difficulties and post course improvement in accuracy are consistent with Brewe et al. (2010b). This suggests that the MRI version of the task we developed is prompting physics reasoning that is consistent with that observed out of scanner environment. Effect sizes from pre- to post-instruction indicate similar performance on this task with modified FCI questions as on the full FCI. This improvement is indicative of a shift in physics reasoning as a result of instruction. We do not interpret these changes as recall effects for two reasons, the results of the FCI were not discussed with students, and the task itself was not identified as being derived from the FCI. Further, Henderson (2002) has shown that recall effects over the duration of a full semester are minimal. While accuracy is important for characterizing and to some degree validating the task that was developed for the fMRI environment, we did not expect accuracy to correlate with brain activity. Instead, physics reasoning, regardless of accuracy, is linked to brain activity.

Task Effect: Brain Activity Linked With Physics Reasoning

Our initial analysis identified brain activity among college students associated with physics reasoning (FCI > Control) in lateral prefrontal and parietal regions. One interpretation is that activity in these regions supports cognitive processes critical for answering physics reasoning problems such as attention, working memory, spatial reasoning, and mathematical cognition. More specifically, the lateral PFC's role in executive functions such as working memory and planning are well-characterized (Bressler and Menon, 2010) and these areas are important in manipulating representations in working memory and reasoning (Andrews-Hanna, 2012; Barbey et al., 2013). Lateral parietal regions are involved in motor functioning as well as spatial reasoning, mathematical cognition, and attention (Wendelken, 2015). Such an interpretation is reasonable in the context of the current task which likely involves generating mental simulations and representations in the service of identifying the correct answer choice. From a large-scale brain network perspective, the brain regions showing physics reasoning-related activation resemble one commonly observed functional brain network known as the central executive network (CEN). The CEN, consisting of lateral prefrontal and parietal regions (Bressler and Menon, 2010), is generally associated with externally oriented attentional and executive processes (e.g., working memory, response selection, and inhibition; Cole and Schneider, 2007; Seeley et al., 2007).

The task-related brain regions we observed were generally similar when separately considering data collected during the pre- and post-instruction scans. While speaking to the consistency of such brain activity, this analysis is not intended to determine which brain regions differ as a function of completing

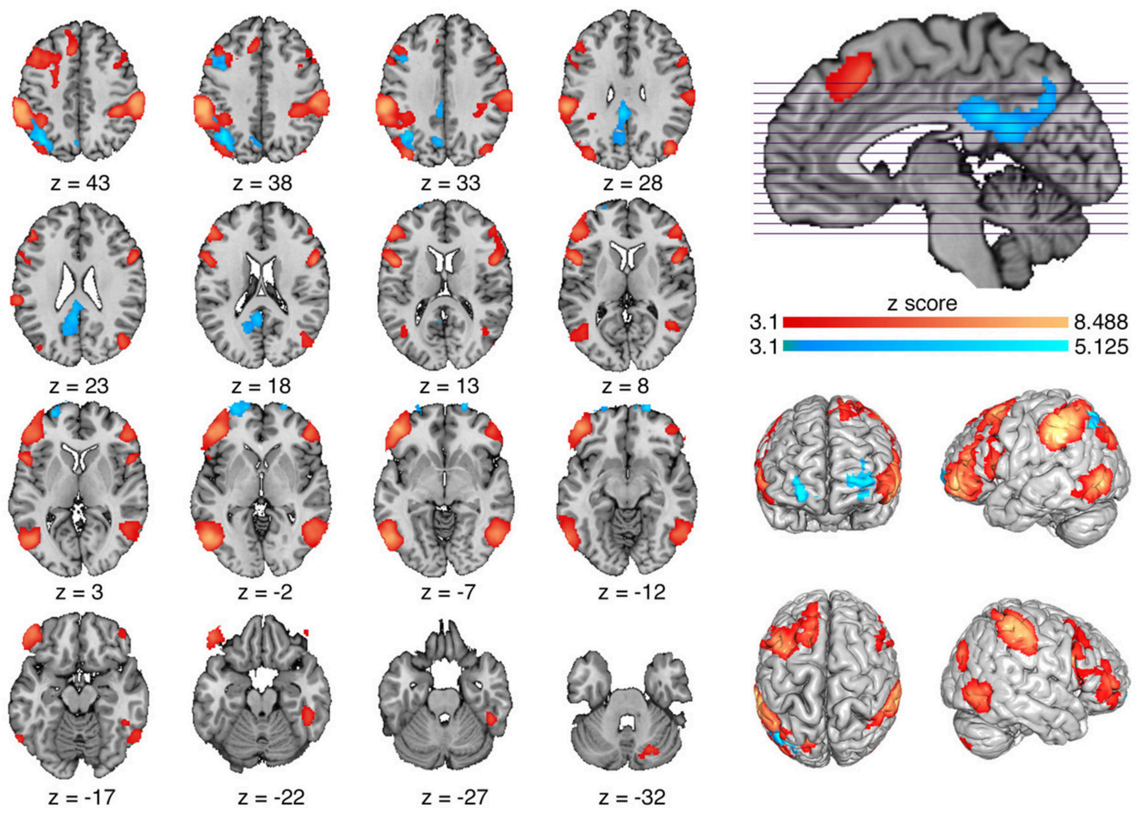

FIGURE 3 | Group-level fMRI results. (Red) Task effect: Brain regions showing increased activity during the physics reasoning task (FCI > Control) at the pre-instruction stage. (Blue) Instruction effect: Brain regions showing increased activity at the post- relative to pre-instruction (Post > Pre) scan during the physics reasoning task.

a MI course (see below). We suspect that such task-related brain activity would be similar among students in other instructional environments.

Instruction Effect: Changes in Brain Activity Post-instruction vs. Pre-instruction

Our second analysis identified increased brain activity among students completing the physics reasoning task after taking a MI course (Post > Pre) in the posterior cingulate cortex, frontal poles, dlPFC, and angular gyrus. These brain regions (PCC, angular gyrus) overlap with regions of another commonly observed large-scale functional brain network known as the default-mode network (DMN). The DMN, consisting of posterior cingulate cortex (PCC), angular gyri, medial PFC, and middle temporal gyri (Raichle et al., 2001; Laird et al., 2009), is generally associated with internally oriented cognitive processes (i.e., self-reflection, mind wandering, autobiographical memory, planning; Buckner et al., 2008). However, other lines of evidence also implicate DMN involvement in complex tasks such as narrative comprehension (Simony et al., 2016), semantic processing (Binder et al., 2009; Binder and Desai, 2011) or the generation and manipulation of mental images (Andrews-Hanna, 2012). In the context of the current task, one interpretation is that students may generate mental images to simulate events and formulate predictions. Additionally, post-instruction increase in DMN activity was observed during physics reasoning (which we show is supported by the CEN), and such coupling between the DMN and CEN during cognition has been hypothesized to arise during controlling attentional focus, thereby aiding in efficient cognitive function (Leech and Sharp, 2014).

Other brain regions showing greater activation during physics reasoning after the MI course included the dlPFC and the frontopolar cortex. The frontopolar cortex is a component of a decision-making network often involved with learning (Koechlin and Hyafil, 2007). The dlPFC is critically linked with the manipulation of verbal and spatial information in working memory (Barbey et al., 2013). Given previous links with, for example, mental simulation, working memory, mathematical calculations, and attention, we speculate that post-instruction increased activity in the PCC, angular gyrus, dlFPC and frontal pole may reflect enhanced mental operations and/or models involved with physics reasoning and/or generation of predictions about physical outcomes.

The PCC, left angular gyrus, left frontal pole, and left DLPFC were the four regions of greatest extent to show increased activity (Post > Pre), however, we did not see correlation between change in activity within these areas and accuracy on the FCI after instruction. The FCI is a cognitively demanding task which includes intuitive but wrong answers. Thus, it may simply be that even wrong answers on the FCI require significant mental effort. Inaccurate physics reasoning likely still involves many of the same mental operations successful physics reasoning does (i.e., mental imagery, visualization, prediction generation, and decision making, to name a few). Measures of accuracy in and of themselves may not display a simple one-to-one relationship with changes in brain activity across instruction. Rather, these changes in brain activity may be related to more complex behavioral changes in how student's reason through physics questions post-relative to pre-instruction. These might include shifts in strategy or an increased access to physics knowledge and problem solving resources.

We posit that the observed pre to post-instruction changes in brain activation during physics reasoning are consistent with what one may expect to observe as students develop refined mental models during classroom learning. Physics reasoning, regardless of an individual's familiarity with the material, is a process continually scaffolded by mental model use (Nersessian, 1995, 1999, 2002a,b; Giere, 2005; Koponen, 2006), and effective physics learning is engendered by building and deploying strategies to appropriately implement mental models during reasoning (Hestenes, 1987). In this study, we framed our exploration of learning-induced changes in brain activity in the context of the MI classroom because this pedagogical approach has been shown to effectively encourage the development and flexible implementation of models during physics reasoning (Brewe, 2008; Brewe et al., 2010b). Our experimental results do not go as far as to implicate MI as any more or less effective than other instructional strategies at supporting instructional-related changes in student's brain networks. However, if we accept that physics reasoning inherently relies on mental model use, we can begin to consider a more truly neuroeducational interpretation of physics learning in which shifts in network engagement across instruction bring about student conceptual change. Characterizing these neurobiological changes may ultimately help researchers and educators understand which instructional strategies may best support successful model development. We hold that the mental models student's deployed at the beginning of the semester during reasoning, upheld by a variety of CEN-supported attentional and executive processes, shifted after instruction, as evidenced by student's overall increased accuracy during reasoning. This instruction-induced shift in model use promoted increased involvement from key DMN and CEN regions within reasoning. This study represents an initial step in neuroeducational research demonstrating that such shifts, indicative of learning, are measurable and detectable using non-invasive brain imaging techniques. Additional work is needed to understand the relationship between external conceptual models as studied in science education, with mental models and related cognitive constructs as studied in neuroimaging literature.

This project has several limitations. First, we focused on the MI class and did not assess the brain activity of students from traditional lecture course sections or other active learning environments. Based on the data presented, we do not make claims that MI is a better or the only instructional tool capable of inducing brain network alterations. Rather, in the current study, we used MI as an exemplar case. It remains to be determined if different pedagogies differentially influence how physics reasoning-related brain networks develop. As noted above and consistent with recommendations (Freeman et al., 2014), we will explore this in the future and a future direction could investigate differences among active learning formats. Second, these analyses addressed brain activation and did not consider correlation with other behavioral measures, such as mental rotations, science anxiety, or academic performance measures which could further aid in the interpretation of these fMRI outcomes. Third, consideration of potential differences between female and male students remains for future investigations.

Notwithstanding these limitations and future direction, these preliminary outcomes implicate brain regions linked with physics reasoning and, critically, suggest that brain activity during physics reasoning is modifiable over the course of a semester of physics instruciton. Further work should investigate differences between MI and lecture instruction, as well as addressing differences among different active learning strategies across disciplines. Studying active learning broadly has the potential to more clearly elaborate how these pedagogies impact student learning and brain function.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of Florida International University's Institutional Review Board with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the FIU IRB.

AUTHOR CONTRIBUTIONS

EB, JB, and AL had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: EB, JB, MR, RL, MS, SP, and AL. Acquisition, analysis, or interpretation of data: EB, JB, MR, VS, TS, ERB, EIB, RO, AN, KB, RL, MS, SP, and AL. Drafting of the manuscript: EB and JB. Critical revision of the manuscript for important intellectual content: EB, JB, MR, VS, TS, ERB, EIB, RO, AN, KB, RL, MS, SP, and AL. Obtained funding: AL, EB, and SP. Administrative, technical, or material support: EB, RL, MS, SP, and AL. Study supervision: AL, MS, SP, MR, and EB.

ACKNOWLEDGMENTS

Primary funding for this project was provided by NSF REAL DRL-1420627 (AL, EB, SP, and JB). Contributions from co-authors were partially provided by NSF 1631325 (AL, MR, and TS), NIH R01 DA041353 (AL, MS, and MR), NIH U01 DA041156 (AL, MS, MR, KB, and ERB), NSF CNS 1532061 (AL), NIH K01DA037819 (MS), NIH U54MD012393 (MS), and FIU Graduate School Dissertation Year Fellowships (AN and RO). Additional thanks to the FIU Instructional and Research Computing Center (IRCC, http://ircc.fiu.edu) for providing HPC and computing resources that contributed to the research results reported within this paper, and to the Department of Psychology of the University of Miami for providing access to their MRI scanner. Lastly, special thanks to the FIU undergraduate students who volunteered, participated, and contributed to this project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fict.2018.00010/full#supplementary-material

REFERENCES

Andrews-Hanna, J. R. (2012). The brain's default network and its adaptive role in internal mentation. *Neuroscientist* 18, 646–656. doi: 10.1177/1073858411403316

Arsalidou, M., and Taylor, M. J. (2011). Is 2+2=4? Meta-analyses of brain areas needed for numbers and calculations. *NeuroImage* 54, 2382–2393. doi: 10.1016/j.neuroimage.2010.10.009

Barbey, A. K., Koenigs, M., and Grafman, J. (2013). Dorsolateral prefrontal contributions to human working memory. *Cortex* 49, 1195–1205. doi: 10.1016/j.cortex.2012.05.022

Bartley, J. E., Boeving, E. R., Riedel, M. C., Bottenhorn, K. L., Salo, T., Eickhoff, S. B., et al. (in press). Meta-analytic evidence for a core problem solving network across multiple representational domains. *Neurosci. Biobehav. Rev.*

Beichner, R. J., and Saul, J. M. (2003). "Introduction to the SCALE-UP (student-centered activities for large enrollment undergraduate programs) project." in *Proceedings of the International School of Physics, (July), 1–17*. Available online at: http://www.ncsu.edu/PER/Articles/Varenna_SCALEUP_Paper.pdf

Binder, J. R., and Desai, R. H. (2011). The neurobiology of semantic memory. *Trends Cogn. Sci.* 15, 527–536. doi: 10.1016/j.tics.2011.10.001

Binder, J. R., Desai, R. H., Graves, W. W., and Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* 19, 2767–2796. doi: 10.1093/cercor/bhp055

Bressler, S. L., and Menon, V. (2010). Large-scale brain networks in cognition: emerging methods and principles. *Trends Cogn. Sci.* 14, 277–290. doi: 10.1016/j.tics.2010.04.004

Brewe, E. (2008). Modeling theory applied: modeling instruction in introductory physics. *Am. J. Phys.* 76, 1155–1160. doi: 10.1119/1.2983148

Brewe, E. (2011). Energy as a substancelike quantity that flows: theoretical considerations and pedagogical consequences. *Phys. Rev. Spec. Top. Phys. Edu. Res.* 7:020106 doi: 10.1103/PhysRevSTPER.7.020106

Brewe, E., Kramer, L. H., and O'Brien, G. E. (2010a). "Changing Participation Through Formation of Student Learning Communities." in *AIP Conference Proceedings*. doi: 10.1063/1.3515255

Brewe, E., Kramer, L., and O'Brien, G. (2009). Modeling instruction: Positive attitudinal shifts in introductory physics measured with CLASS. *Phys. Rev. Spec. Top. Phys. Edu. Res.* 5:013102. doi: 10.1103/PhysRevSTPER.5.013102

Brewe, E., and Sawtelle, V. (2018). Modeling instruction for university physics: examining the theory in practice. *Eur. J. Phys.* doi: 10.1088/1361-6404/aac236

Brewe, E., Sawtelle, V., Kramer, L. H. O'Brien, G. E. Rodriguez, I., and Pamelá, P. (2010b). Toward equity through participation in Modeling Instruction in introductory university physics. *Phys. Rev. Spec. Top. Phys. Edu. Res.* 6:010106. doi: 10.1103/PhysRevSTPER.6.010106

Brewe, E., Traxler, A., de La Garza, J., and Kramer, L. H. (2013). Extending positive CLASS results across multiple instructors and multiple classes of modeling instruction. *Phys. Rev. Spec. Top. Phys. Edu. Res*. 9:20116. doi: 10.1103/PhysRevSTPER.9.020116

Buckner, R. L., Andrews-Hanna, J. R., and Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Ann. N. Y. Acad. Sci.* 1124, 1–38. doi: 10.1196/annals.1440.011

Cole, M. W., and Schneider, W. (2007). The cognitive control network: integrated cortical regions with dissociable functions. *Neuroimage* 37, 343–360. doi: 10.1016/j.neuroimage.2007.03.071

Craik, K. (1943). *The Nature of Explanation*. London: Cambridge University Press.

Crouch, C. H., and Mazur, E. (2001). Peer Instruction: ten years of experience and results. *Am. J. Phys.*, 69, 970–977. doi: 10.1119/1.1374249

Dou, R., Brewe, E., Zwolak, J. P., Potvin, G., Williams, E. A., and Kramer, L. H. (2016). Beyond performance metrics : examining a decrease in students' physics self-efficacy through a social networks lens. *Phys. Rev. Phys. Edu. Res.* 12:20124. doi: 10.1103/PhysRevPhysEducRes.12.020124

Etkina, E. E., Murthy, S., and Zou, X. (2006). Using introductory labs to engage students in experimental design. *Am. J. Phys.* 74:979. doi: 10.1119/1.2238885

Etkina, E., and Van Heuvelen, A. (2007). *Investigative Science Learning Environment – A Science Process Approach to Learning Physics. PER-Based Reforms in Calculus-Based Physics.* Available online at: http://www.compadre.org/PER/per_reviews/media/volume1/ISLE-2007.pdf

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., et al. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proc. Natl. Acad. Sci. U.S.A.* 111, 8410–8415. doi: 10.1073/pnas.1319030111

Frigg, R., and Hartmann, S. (2006). "Models in science," in *Stanford Encyclopedia of Philosophy*, Vol. 1, ed E. N. Zalta (Stanford, CA: Metaphysics Research Lab, Stanford University).

Giere, R. N. (2005). How models are used to represent reality. *Philos. Sci.* 71, 742–752. doi: 10.1086/425063

Gouvea, J., and Passmore, C. (2017). "Models of" versus "Models for"toward an agent-based conception of modeling in the science classroom. science and education, 26, 49–63. doi: 10.1007/s11191-017-9884-4

Greca, I. M. A., and Moreira, M. A. (2001). Mental, physical, and mathematical models in the teaching and learning of physics. *Sci. Edu.* 86, 106–121. doi: 10.1002/sce.10013

Greca, I. M., and Moreira, M. A. (2000). Mental models, conceptual models, and modelling. *Int. J. Sci. Educ.* 22, 1–11. doi: 10.1080/095006900289976

Halloun, I. A. (1996a). Schematic modeling for meaningful learning of physics. *J. Res. Sci. Teach.* 33, 1019–1041. doi: 10.1002/(SICI)1098-2736(199611)33:9<1019::AID-TEA4>3.0.CO;2-I

Halloun, I. A. (1996b). "Views About Science and physics achievement: The VASS story," in *AIP Conference Proceedings-Physics Education Research Conference,* Vol. 399, (College Park, MD), 605–614.

Halloun, I. A. (2004). *Modeling Theory in Science Education*. Dordrecht: Springer.

Hegarty, M. (2004). Mechanical reasoning by mental simulation. *Trends Cogn. Sci.* 8, 280–285. doi: 10.1016/j.tics.2004.04.001

Henderson, C. (2002). Common concerns about the force concept inventory. *Phys. Teacher* 40, 542–547. doi: 10.1119/1.1534822

Hestenes, D. (1987). Toward a modeling theory of physics instruction. *Am. J. Phys.* 55, 440–454. doi: 10.1119/1.15129

Hestenes, D. (1992). Modeling games in the Newtonian world. *Am. J. Phys.* 60:732. doi: 10.1119/1.17080

Hestenes, D., Wells, M., and Swackhamer, G. (1992). Force concept inventory. *Phys. Teacher* 30, 141–158. doi: 10.1119/1.2343497

Johnson-Laird, P. N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge, MA: Harvard University Press.

Justi, R., and Gilbert, J. (2000). History and philosophy of science through models: some challenges in the case of'the atom'. *Int. J. Sci. Educ.* 22, 993–1009. doi: 10.1080/095006900416875

Kober, N. (2014). *Reaching Students What Research Says about Effective Instruction in Undergraduate Science and Engineering*. Washington, DC: National Academies Press.

Koechlin, E., and Hyafil, A. (2007). Anterior prefrontal function and the limits of human decision-making. *Science* 318, 594–598. doi: 10.1126/science.1142995

Koponen, I. T. (2006). Models and modelling in physics education: a critical re-analysis of philosophical underpinnings and suggestions for revisions. *Sci. Educ.* 16, 751–773. doi: 10.1007/s11191-006-9000-7

Krieger, M. H. (1987). The physicist's toolkit. *Am. J. Phys.* 55, 1033–1038. doi: 10.1119/1.14929

Kuhn, T. S. (1970). *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.

Laird, A. R., Eickhoff, S. B., Li, K., Robin, D. A., Glahn, D. C., and Fox, P. T. (2009). Investigating the functional heterogeneity of the default mode network using coordinate-based meta-analytic modeling. *J. Neurosci.* 29, 14496–14505.doi: 10.1523/JNEUROSCI.4004-09.2009

Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind, Vol. 64*. Chicago, IL: University of Chicago Press.

Lasry, N., Rosenfield, S., Dedic, H., Dahan, A., and Reshef, O. (2011). The puzzling reliability of the force concept inventory. *Am. J. Phys.* 79:909. doi: 10.1119/1.3602073

Lasry, N., Watkins, J., Mazur, E., and Ibrahim, A. (2013). Response times to conceptual questions. *Am. J. Phys.* 81:703. doi: 10.1119/1.4812583

Leech, R., and Sharp, D. J. (2014). The role of the posterior cingulate cortex in cognition and disease. *Brain* 137, 12–32. doi: 10.1093/brain/awt162

Lehrer, R., and Schauble, L. (2006). "Cultivating model-based reasoning in science education," in *The Cambridge handbook of: The learning sciences,* ed R. K. Sawyer (New York, NY,: Cambridge University Press), 371–387.

Mackey, A. P., Miller Singley, A. T., and Bunge, S. A. (2013). Intensive reasoning training alters patterns of brain connectivity at rest. *J. Neurosci.* 33, 4796–4803. doi: 10.1523/JNEUROSCI.4141-12.2013

Mason, R. A., and Just, M. A. (2015). Physics instruction induces changes in neural knowledge representation during successive stages of learning. *Neuroimage* 111, 36–48. doi: 10.1016/j.neuroimage.2014.12.086

Mason, R. A., and Just, M. A. (2016). Neural representations of physics concepts. *Psychol. Sci.* 27, 904–913. doi: 10.1177/0956797616641941

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., and Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *J. Appl. Psychol.* 85, 273–283. doi: 10.1037/0021-9010.85.2.273

Matthews, M. R. (2007). Models in science and in science education: an introduction. *Sci. Educ.* 16, 647–652. doi: 10.1007/s11191-007-9089-3

McDermott, L. C., and Shaffer, P. S. (2001). *Tutorials in Introductory Physics*. New York, NY: Pearson.

McPadden, D., and Brewe, E. (2017). Impact of the second semester university modeling instruction course on students' representation choices. *Phys. Rev. Phys. Educ. Res.* 13:020129 doi: 10.1103/PhysRevPhysEducRes.13.020129

Morgan, M. S., and Morrison, M. (1999). *Models as Mediators: Perspectives on Natural and Social Science*. New York, NY: Cambridge University Press.

Morris, G. A., Branum-Martin, L., Harshman, N., Baker, S. D., Mazur, E., Dutta, S., et al. (2006). Testing the test: Item response curves and test quality. *Am. J. Phys.* 74, 449–453. doi: 10.1119/1.2174053

Morris, G. A., Harshman, N., Branum-Martin, L., Mazur, E., Mzoughi, T., and Baker, S. D. (2012). An item response curves analysis of the force concept inventory. *Am. J. Phys.* 80:825. doi: 10.1119/1.4731618

Nersessian, N. J. (1995). Should physicists preach what they practice? *Sci. Educ.* 4, 203–226. doi: 10.1007/BF00486621

Nersessian, N. J. (1999). "Model-based reasoning in conceptual change," in *Model-Based Reasoning in Scientific Discovery*, eds L. Magnani, N. J. Nersessian, and P. Thagard, (New York, NY: Kluwer Academic/Plenum Publishers), 5–22.

Nersessian, N. J. (2002a). Abstraction via generic modeling in concept formation in science. *Mind Soc.* 3, 129–154. doi: 10.1007/BF02511871

Nersessian, N. J. (2002b). *The Cognitive Basis of Model-Based Reasoning in Science.* Cambridge: Cambridge University Press.

Odenbaugh, J. (2005). Idealized, inaccurate but successful: a pragmatic approach to evaluating models in theoretical ecology. *Biol. Philos.* 20, 231–255. doi: 10.1007/s10539-004-0478-6

Passmore, C. M., and Svoboda, J. (2012). Exploring opportunities for argumentation in modelling classrooms. *Int. J. Sci. Educ.* 34, 1535–1554. doi: 10.1080/09500693.2011.577842

Prabhakaran, V., Smith, J. A., Desmond, J. E., Glover, G. H., and Gabrieli, J. D. (1997). Neural substrates of fluid reasoning: an fMRI study of neocortical activation during performance of the Raven's Progressive Matrices Test. *Cogn. Psychol.* 33, 43–63. doi: 10.1006/cogp.1997.0659

Prado, J., Chadha, A., and Booth, J. R. (2011). The brain network for deductive reasoning: a quantitative meta-analysis of 28 neuroimaging studies. *J. Cogn. Neurosci.* 23, 3483–3497. doi: 10.1162/jocn_a_00063

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. (2001). A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 676–682. doi: 10.1073/pnas.98.2.676

National Research Council (2012). "Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering," in *Committee on the Status, Contributions, and Future Directions of Discipline-Based Education Research; Board on Science Education; Division of Behavioral and Social Sciences and Education*, eds S. R. Singer, N. R. Nielsen, and H. A. Schweingruber (National Research Council; National Academies Press). Available online at: http://download.nap.edu/cart/download.cgi?andrecord_id=13362andfree=1

Rodriguez, I., Potvin, G., and Kramer, L. H. (2016). How gender and reformed introductory physics impacts student success in advanced physics courses and continuation in the physics major. *Phys. Rev. Phys. Educ. Res.* 12:9. doi: 10.1103/PhysRevPhysEducRes.12.020118

Rogoff, B. (1990)."Shared thinking and guided participation: conclusions and speculation BT - Apprenticeship in thinking: cognitive development in social context," in *Apprenticeship in Thinking: Cognitive Development in Social Context* (New York, NY: Oxford University Press), 189–210.

Sawtelle, V., Brewe, E., Goertzen, R. M., and Kramer, L. H. (2012). Identifying events that impact self-efficacy in physics learning. *Phys. Rev. Spec. Top. Phys. Educ. Res.* 8:20111. doi: 10.1103/PhysRevSTPER.8.020111

Sawtelle, V., Brewe, E., and Kramer, L. H. (2010). "Positive impacts of modeling instruction on self-efficacy," in *PERC Conference Proceedings,* Vol. 1289, eds C. Singh, N. S. Rebello, and M. Sabella (Portland, OR: American Institute of Physics), 289.

Schwartz, D. L., and Bransford, J. D. (1998). A time for telling. *Cogn. Instrum.* 16, 475–5223. doi: 10.1207/s1532690xci1604_4

Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., et al. Krajcik, J. (2009). Developing a learning progression for scientific modeling: making scientific modeling accessible and meaningful for learners. *J. Res. Sci. Teach.* 46, 632–654. doi: 10.1002/tea.20311

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., et al. Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci.* 27, 2349–2356. doi: 10.1523/JNEUROSCI.5587-06.2007

Simony, E., Honey, C. J., Chen, J., Lositsky, O., Yeshurun, Y., Wiesel, A., et al. (2016). Dynamic reconfiguration of the default mode network during narrative comprehension. *Nat. Commun.* 7:12141. doi: 10.1038/ncomms12141

Svoboda, J., and Passmore, C. (2011). *The Strategies of Modeling in Biology Education.* Science and Education. Available online at: http://www.springerlink.com/index/10.1007/s11191-011-9425-5

Traxler, A., and Brewe, E. (2015). Equity investigation of attitudinal shifts in introductory physics. *Phys. Rev. Spec. Top. Phys. Educ. Res.* 11, 1–7. doi: 10.1103/PhysRevSTPER.11.020132

Wells, M., Hestenes, D., and Swackhamer, G. (1995). A modeling method for high school physics instruction. *Am. J. Phys.* 63:606. doi: 10.1119/1.17849

Wendelken, C. (2015). Meta-analysis: how does posterior parietal cortex contribute to reasoning? *Front. Hum. Neurosci.* 8:1042. doi: 10.3389/fnhum.2014.01042

Windschitl, M., Thompson, J., and Braaten, M. (2008). Beyond the scientific method: model-based inquiry as a new paradigm of preference for school science investigations. *Sci. Educ.* 92, 941–967. doi: 10.1002/sce.20259

Zwolak, J. P., Dou, R., Williams, E. A., and Brewe, E. (2017). Students' network integration as a predictor of persistence in introductory physics courses. *Phys. Rev. Phys. Educ. Res.* 13, 1–14. doi: 10.1103/PhysRevPhysEducRes.13.010113

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in ICT

REVIEW
published: 29 August 2018
doi: 10.3389/fict.2018.00019

Three Active Learning Strategies to Address Mixed Student Epistemologies and Promote Conceptual Change

Calvin S. Kalman[1]* *and Mark J. Lattery*[2]

[1] Physics, Concordia University, Montreal, QC, Canada, [2] Department of Physics and Astronomy, University of Wisconsin Oshkosh, Oshkosh, WI, United States

Novice science learners or introductory science students vary greatly in their understanding of the nature of science. For example, many students do not conceive of scientific knowledge as a highly ordered, coherent, knowledge structure that contains a set of interrelated ideas. Such a framework enables the learner to relate new material to prior knowledge and, if warranted, assimilate the new material within the framework. Many students have strong beliefs that knowledge is conveyed by authorities, such as the instructor and the textbook. Also many student's own knowledge structure is fragmented or "in pieces," as described by diSessa. Fortunately, this portrayal is not valid for all students. Many other students enter the classroom with productive intellectual values and possess, or can quickly develop with little prompting, alternative, and coherent conceptions that conflict with target ideas. These students are able to relate new material to prior knowledge and, if warranted, assimilate new material into pre-existing conceptions. The challenge of contemporary science education reform is therefore to address the diverse needs of a "mixed student epistemology" classroom. In this paper we review three instructional strategies that show promise to address this challenge in the context of an introductory physics classroom: (1) the Reflective Writing and Labatorial interventions of Kalman et al. (2) the Conceptual Conflict Collaborative Group and Critique approaches of Kalman and Rohar, and (3) the integrated Elicit-and-Challenge and Bridging Technique strategies of Lattery. Each approach stresses the need for students to critically examine their own ideas in relation to target course ideas and discuss their ideas with peers. The second and third approaches emphasize the important role of the history and philosophy of science in science teaching. The aim of such efforts is not only to convey subject-matter content knowledge, but also to shape the student mindset, metacognitive practice, and understanding of the nature of science.

Keywords: reflective writing, critical thinking, knowledge in pieces, coherent theory, student's epistemological beliefs, cognitive dissonance theory, principle of counter induction, model-centered instruction

Edited by:
Elizabeth S. Charles, Dawson College, Canada

Reviewed by:
Phoebe A. Jackson, John Abbott College, Canada
Steven Paul Ehrlick, Ryerson University, Canada

***Correspondence:**
Calvin S. Kalman
calvin.kalman@concordia.ca

Specialty section:
This article was submitted to Digital Education, a section of the journal Frontiers in ICT

Received: *24 October 2017*
Accepted: *20 July 2018*
Published: *29 August 2018*

Citation:
Kalman CS and Lattery MJ (2018) Three Active Learning Strategies to Address Mixed Student Epistemologies and Promote Conceptual Change. Front. ICT 5:19. doi: 10.3389/fict.2018.00019

INTRODUCTION

An ongoing debate in the science-education community exists between those who believe that students enter the classroom with stable and coherent ideas about the natural world that differ from those presented in science textbooks and by their science teachers, and others who claim that student knowledge consists of isolated structures called phenomenological primitives (p-prims). The former view is referred to as the "Theory Theory" (TT; e.g., Posner et al., 1982; Vosniadou et al., 2008). An exemplar was the idea that students begin their studies with stable and coherent conceptions or theories about force and motion similar to the theories that were held by ancient philosophers and scientists (Wandersee et al., 1994). The second view is referred to as "Knowledge in Pieces" (KiP) and emerged in the 1990's as a dominant alternative to TT. Each view has important and different implications for the goals and methods of instruction that lead to productive changes in the student's knowledge structure (conceptual change).

Naturally, a student's scientific knowledge structure reflects their view of science learning. Students who see science learning as a passive activity believe that scientific knowledge is received knowledge, while students whose knowledge structure is highly fragmented see scientific knowledge as an enormous body of unrelated "bits" or trivia to be memorized. This private and usually implicit student epistemology can be characterized as "problem driven"; i.e., scientific knowledge consists of isolated structures, such as the equations, "cherry picked" from a chapter summary to solve problems related to specific situations. In this view, equations are not abstract representations of general ideas or principles, but merely instruments to calculate things. Conventional science instruction often exacerbates the problem. For example, inattention to the interplay of theory and experiment, or competing theories, leaves the student with the false impression that scientific progress is due entirely to experimentation. The modern scientific classroom must therefore be designed to address the diverse needs of a mixed student epistemology classroom.

Reflecting on his theory of student intuitive knowledge, diSessa (1996) writes: Do ordinary people have anything like theories of the physical world? It seems the most plausible a priori position is "no." Theories are things that belong to formal science (p. 711). Recently, however, Lattery (2017) presented detailed counter evidence for this claim. His research shows that many introductory physics students can and do think theoretically and even generate their own theories that differ from those found in their textbooks. The context of these observations is a university introductory physics classroom. Lattery's work provides evidence to support the claim "that students are authentic and creative scientific modelers" (p. 109), and asserts that, the "student view of force and motion does not appear to be incoherent or fragmented, but driven by [a] rule..." (p. 142). This observation suggests not only that many students enter the classroom with semi-coherent and stable conceptions about how the world works, but also an instinct for the nature of science or scientific knowledge (e.g., scientific knowledge must be coherent). Lattery's findings are consistent with the "theory theory" view of student knowledge that was previously dominant in the field.

The question then is, what might we make of these different perspectives? Given the possible epistemological diversity of our introductory science classrooms, we propose instructional strategies that help students build a formal understanding of science from productive "pieces" of knowledge *and* through a carefully guided analysis of their alternative conceptions. Key to combining these strategies is a learning environment that allows students to discuss their ideas with their peers and the instructor. Before describing these strategies in detail, we discuss (1) research that begins to evaluate the scope of the challenge presented by epistemological diversity in science education, (2) a concept critical to understanding learning environments that seek to address this challenge (*incommensurability*), and (3) a set of five research questions to guide the improvement of these learning environments.

OVERVIEW OF STUDENTS' VIEWS ON THE NATURE OF SCIENCE

Kalman (2002, 2010) demonstrates that student understanding of the nature of science can be advanced through direct study of the philosophy of science. In these research studies, university students are placed in small groups and assigned a philosopher of science to study throughout the course (e.g., Kuhn, 1962/1970; Popper, 1963; Lakatos, 1970; Feyerabend, 1993). Significant time (60 min a week or 60 min every second week or five 75 min sessions) is devoted to group presentations. The groups report periodically to the entire class. In addition to sharing these ideas in class the students hand in a written work. Only the latter is marked.

With respect to the theory-pieces debate, about half of the students in the experiment described by Kalman (2002, 2010), who held views of science consistent with Popper on the first day, view scientific knowledge as coherent and interrelated. At the end of the course, only three students (all categorized as "other") had a view of science that could be categorized as "knowledge in pieces" (KiP). All students who initially identified as Baconian and KiP had sharpened their viewpoint and showed evidence of a coherent view of the nature of science. It is now generally understood that the nature of a naïve student intuitive knowledge (whether coherent or fragmented) depends on the specific experiences and cognitive development of the student (Kalman, 2010). This knowledge is almost certainly correlated to the student's understanding of the scientific-thinking process itself.

INCOMMENSURABILITY AND SCIENCE LEARNING

Student conceptions in science are found to be *incommensurate* with accepted scientific theories (Chi, 2013), much as ancient

Greek ideas of force and motion are incommensurate to those of Newtonian mechanics. An idea is said to be incommensurate with another idea when there is no, or a very limited, basis of comparison between the two ideas. Kuhn (1962/1970) and Feyerabend (1962) independently introduced the idea of incommensurability to the philosophy of science. Kuhn used *incommensurability* to illustrate the holistic nature of the changes that take place in a scientific revolution; for example, Kuhn claims that many scholars initially rejected Newton's theory because it did not explain the attractive forces between matter, something required of any mechanics according to Aristotle, Descartes, and their followers (Kuhn, 1962/1970). In later publications such as Kuhn (1981/2000) he continued to emphasize the difference between normal, cumulative growth that is in accord with existing concepts and revolutionary discoveries that could not have been made wholly on the basis of previously known concepts such as the discovery of Newton's theory. Such new discoveries require replacing known concepts with new concepts that are incommensurable with antecedent ideas. The implications of these observations for classroom learning are just beginning to be explored.

In *Against Method*, Feyerabend states (p. 212): "In 1962, I called theories such as those containing 'impetus' and 'momentum' incommensurable theories." Feyerabend's incommensurability corresponds to questions that have meaning only in a particular theory. As Kalman (2009) notes, if an overlap between successive theories exists (i.e., shared ideas and/or concepts exist), then there can be interesting questions that are meaningful in the context of both theories. Thus within the context of both the wave and particle theories of light we can ask whether or not diffraction takes place. However, if there is no overlap, questions exist that have meaning only in the context of one theory, but not the other. As such, a question on the nature of the Ether makes sense in the context of the wave-ether theory, but has no meaning in the context of Einstein's special theory of relativity.

FIVE QUESTIONS OF SCIENCE INSTRUCTION

Addressing student ideas that differ in key respects to accepted (target) scientific knowledge is a significant challenge for science education. Kalman (2017) presents research questions in science educational research in that are examined in depth throughout his book. In the context of the design of science instruction, we consider five of these basic and local questions:

1. What is the nature of student knowledge: knowledge-in-pieces or a coherent theory.

The question of the nature of student knowledge has important implications for instruction. If student knowledge is fragmented and disorganized, instruction should build scientific concepts from the relevant "pieces" as done in the Bridging Technique (Clement and Rea-Ramirez, 2008). However, if student knowledge exhibits coherence (e.g., the student places increasing value on explanatory consistency in their modeling activities), instruction should confront student ideas, as done in the Elicit-and Challenge approach. Lattery (2017) presents evidence to justify an integration of these two approaches.

2. What is the stage of the students' intellectual development?

McKinnon and Renner (1971) state the hypothesis: "The majority of entering college freshmen do not come to college with adequate skills to argue logically about the importance of a given principle when the context in which it is used is slightly altered." In this context, Renner and Paske (1977) found that "approximately 50% of entering college freshmen are concrete operational." This itself is a gross simplification. Vygotsky (1978) critiqued the assumption that a student's developmental level is entirely given by a battery of tests of varying difficulties. In his opinion, what the student can do "with the assistance of others might be in some sense even more indicative of their mental development than what they can do alone" (p. 85). The role of the instructor is therefore to provide the necessary scaffolding (Wood et al., 1976) for students to progress through stages of development.

3. What are the student's epistemological beliefs or approaches to learning?

Elby (2001) notes "students' epistemological beliefs—their view about the nature of knowledge and learning affect their mindset, metacognitive practice, and study habits in a physics course." Another issue to consider is the views students have about the Nature of Science (NOS). Students generally start out with a Baconian perspective that scientific ideas develop by induction from experiment (Kalman, 2010). Clough (2006) points out that students' conceptions about the NOS are based upon "Teachers' language, cookbook activities, textbooks that report the end products of science without addressing how the knowledge was developed" (p. 467). Instruction should then focus also on students' views of the NOS, emphasizing the interaction of theory and experimentation as a method for adjudicating and refining theories.

4. What instructional supports are necessary for students to examine and develop their own ideas and compare them to the ideas presented by peers, the textbook, and the instructor??

Feyerabend (1993, p. 33) points out that critical evaluation of one's own ideas requires the consideration of an alternative, competing idea. This is the principle of counter induction, stated alternatively as changes in theories occur only when one theory is compared with another. In order to maximize empirical content, a scientist will compare theories with other theories rather than with experience, data, or facts. This pluralistic approach has often been used in the past. For example Newton did not try to prove, in advance of experimental evidence, that the assumptions he made in his theory of gravity were axiomatic or valid. His approach was to use them as working assumptions which would be accepted hypothetically only as long as their consequences threw light, in exact detail, on hitherto-unexplained phenomena. Thus, he made a practice of critiquing *theory qua theory*. Students, whose goal is solving problems alone, will have difficulty seeing the value of this approach. What would a student make of Newton's discovery that inertial mass and gravitational mass are the same? "It [mass] can also be known from a body's weight, for—by making very accurate experiments with pendulums—I have found it to be proportional to the weight..." (Newton, 1686, Opening paragraph of the *Principia*)." Instruction should then be designed to support such methods of comparing and contrasting theories.

5. How does the student deal with cognitive conflict or cognitive dissonance?

Cognitive dissonance theory (Festinger, 1957) continues to be the subject of new research; for reviews see Cooper (2007) and Harmon-Jones and Harmon-Jones (2007). Psychological discomfort, or dissonance is produced when relevant and inconsistent cognitions occur. Linenberger and Bretz (2012) discovered that cognitive dissonance occurring in interviews provides important information about how students understand enzyme substrate interactions. The student is convinced by experience with everyday phenomenon that their intuitive ideas about the natural world are correct (e.g., "motion implies force"). The natural student response to cognitive dissonance is to assimilate (in a Piagetian sense) scientific knowledge from the textbook or teacher into a pre-existing knowledge framework. Put in a different way, cognitive dissonance leads students to misread the textbook and mishear the teacher.

Our own individual and joint work has three instructional strategies that address the teaching and learning issues raised by these questions. We present these below. In each case, a combination of approaches is employed to meet the needs of a mixed student epistemology classroom. These active learning strategies engage the learner in scaffolded tasks and take them through one of the identified processes described above.

STRATEGY 1. REFLECTIVE WRITING AND LABATORIALS

Madsen et al. (2015) performed a meta-analysis synthesizing 24 studies and found that in typical physics classes students' beliefs are less expert-like at the end of the course than they were at the beginning. Kalman et al. (2015) considered the hypothesis that students' epistemological beliefs could become more expert-like with a combination of appropriate instructional activities: (a) pre-class reading with metacognitive reflection (Reflective Writing), and (b) in-class active learning (Labatorials) that produce cognitive conflict/dissonance, and (ideally) a transition to more productive ideas. Below we describe both Reflective Writing and Labatorials, as well as briefly report on the impact of a combined approach.

Reflective Writing

For many years, Kalman et al. designed new and innovative pedagogical tools to meet these instructional challenges: The Reflective Writing (RW) tool (Kalman and Rohar, 2010; Kalman, 2011; Huang and Kalman, 2012; El-Helou and Kalman, 2018) is a metacognitive activity, that prompts students to examine textual material, before coming to the classroom or laboratory in the manner of a hermeneutic circle (Gadamer, 1975/1960). The student begins an examination of a textual extract with preconceptions (misconceptions). The key quintessential experience occurs when the student is pulled up short by the textual extract. "Either it does not yield any meaning or its meaning is not compatible with what we had expected" (Gadamer, 1975/1960, p. 23). The metacognitive reflection begins when they question their understanding of the text within the entire "horizon" (Kalman, 2011, p. 163). Gadamer (1975/1960) expanded the notion of horizon that originated with "Heidegger (1962)". Gadamer takes the term "horizon" to be all that one can see defined by your pre-understandings. In reading a text, one encounters the horizon of the text and self-examination will produce a new horizon.

While reflective writing can be used as a tool for self-awareness it has also been employed as an assignment assessment with the aid of rubrics (Khanam and Kalman, 2017). This instructional strategy has been used in Grade 11 and across many post-secondary subject areas. One drawback to this approach is that it doesn't work well with younger students (El-Helou and Kalman, 2018); the authors speculate that this result is due to the stage of the students' intellectual development.

Labatorials (Laboratory + Tutorials)

The development of Labatorials at the University of Calgary (Ahrensmeier, 2013) was motivated by the introductory physics tutorial system used at the University of Washington (McDermott and Shaffer, 2001). Students are given a worksheet that contains instructions for experiments, calculation problems, computer simulations, and conceptual questions. At the onset, students are assigned to groups of 3 or 4 members and provided with conceptual questions and asked to make predictions. Each lab section has one lab instructor assigned to a maximum of 16 students. Each group completes a Labatorial worksheet that usually contains 3–6 checkpoints. On completing each checkpoint, the group reviews the answers with the lab instructor. If the answer to a question is incorrect the lab instructor will help the students to find the correct answer through exploration and discussion of alternative ideas. The worksheets are developed in such a manner that students who arrive on time and concentrate on the material can finish all checkpoints in the time allotted. Evidence suggest that Labatorials are useful to both instructors and students (Sobhanzadeh et al., 2017). Students identify their strengths and weaknesses and identify areas of their understanding that need to be strengthened. At the same time, instructors can recognize and address problems immediately.

A Combined Approach

Kalman et al. (2015) show that students' epistemological beliefs become more expert-like with a combination of (a) pre-class metacognitive reflection (Reflective Writing), and (b) in-class active learning (Labatorials) that produce cognitive dissonance. This research examined: an experimental group of 8 sections (110 students) and a control group of 7 sections (102 students) of an introductory physics course. Both groups performed Labatorials, however, the experimental group performed Reflective Writing while the control group performed summary writing.

To assess changes in students' epistemological beliefs, this study used Hofer's discipline-focused epistemological beliefs questionnaire (DFEBQ) (Hofer, 2000). A pre-test was administered in the Fall, and a post-test was administered after two semesters. All students performed Labatorials. Students who had taken Reflective Writing in the first semester continued with Reflective Writing in the subsequent semester; students who had taken summary writing in the first semester similarly continued in this mode. Results showed that the students' epistemological beliefs in the experimental group (Reflective Writing) become more expert-like in their thinking compared to the control group.

The strength of this combined approach is its emphasis on developing conceptual knowledge through writing and metacognition. Students/peers are also challenged to compare their ideas with accepted scientific views. Reflective Writing is done by students at home and does not require additional class time. Indeed, since students have read the textual material before coming to class, the instructor should cut back on the material that is presented and use the saved class time for "flipped classroom" activities such as those described in Strategy 2 below.

Summary of Strategy 1

Sobhanzadeh et al. (2017) found that Strategy 1 is useful to help students to explore the relationship among various physics concepts. Students improved their understanding of concepts, problem solving skills, engagement, and performance in the lab. However, the approach was not designed to produce profound and sustained learning, which (as we explain below) requires a thorough peer-centered discussion of both prior and target ideas.

We are currently testing Strategy 1 in high school physics classrooms. Preliminary results indicate that it works well in such a setting. The main challenge of implementing this approach is that Strategy 1 requires a complete redesign of conventional laboratories as described by Sobhanzadeh et al. (2017).

STRATEGY 2. CONCEPTUAL CONFLICT COLLABORATIVE GROUP AND CRITIQUE EXERCISE

Kalman and Lattery have each argued that deep science learning in the science classroom is not generally possible unless students have an opportunity to sort out their ideas with peers and consider alternative or competing ideas. Below we describe three approaches that show promise for generating these deeper levels of reflection, comparison and confrontation of opposing theories. These consist of the Conceptual Conflict Collaborative Group, the Critique Exercise, and lastly, the Combined Approach; each has been the subject of research by one of the current authors.

Conceptual Conflict Collaborative Group

In a university course (Kalman and Aulls, 2003) students considered two alternative frameworks: pre-Galilean Physics and Newtonian Physics. The idea of the course design is for students at first to view the frameworks almost in a theatrical sense involving a conflict of actors (Aristotle, Galileo, Newton, and others) in the history of science. The study showed some students gradually identify with the conceptual positions taken by the proponents of the alternative frameworks and become themselves a part of the action.

During the course, students gradually realize that the positions defended by the actors are connected to concepts from different parts of the course. Armed with this knowledge, students evaluate the two competing, alternative frameworks through the Conceptual Conflict Collaborate Group (CG) exercise (Kalman et al., 1999) and through an argumentative essay (*critique*) (Kalman et al., 2004). Three to four students are assigned to a collaborative group. Within each group students take on roles such as scribe, reporter, critic, or timekeeper. Although students remain in the same group throughout the semester, students are given the option to change roles in each activity.

For each exercise, students are asked to discuss for a fixed time limit a demonstration or qualitative problem. Time limits are set so that none of the groups need to wait for other groups to complete the task. The lesson impresses on the student that there are at least two ways of looking at the problem. Having two groups with different concepts report to the class produces the desired conceptual conflict. Then, representatives of each group debate the issue between themselves. Afterwards, the rest of the class is invited to present questions to these representatives. (The use of personal scientific conceptions by an "expert" did not appear to have negative connotations, an issue examined by presenting qualitative essay questions on the final exam).

To underscore that two conflicting concepts have been presented, the class is asked to vote on which concept resolves the demonstration or qualitative problem. Voting is essential because students due to cognitive dissonance students often misinterpret what they hear or read. Due to the vote, students are anxious to find out which point of view is correct. The professor resolves the conflict by using demonstrations.

To evaluate this approach, Kalman et al. (1999) studied two sections of the same calculus-based mechanics course taught by the same instructor. Four concepts were examined. In one section concepts A and C were examined using the collaborative group method while concepts B and D were taught by conventional methods. In the other section, the procedure was reversed: concepts B and D were examined using the collaborative group method while concepts A and C were taught by conventional methods. Pre- to post-test gains for question sets based on an enhanced version of the force concept inventory (FCI; Hestenes et al., 1992), showed that the group that used the collaborative group method was more successful in making a conceptual change than the group taught by conventional methods (Kalman et al., 1999).

Kalman et al. (2010) also compared the above approach with Peer Instruction (PI) (Mazur, 1997; Crouch and Mazur, 2001; Lasry et al., 2008). This experiment made use of two equally experienced instructors teaching an introductory first year physics course for science majors at a large public university. Students were randomly allotted to the two sections of the course. Both teachers had often used PI with clickers to cover other concepts than those covered in this paper. The Force Concept Inventory (FCI) was used as a pre- and post-test to compare the two classes.

A comparison of CG and PI for two classes and three tasks are as follows: For the first task, the collaborative group (CG) method produced a statistically significant higher score ($p = 0.017$). There was no statistically significant difference between the methods for the second task, even though the CG method produced a slightly higher result. For the third task, the class using CG produced a higher score with virtually no overlap within the statistical error. Overall, the CG method seems to be more effective than the PI method.

Critique Exercise

The critique activity was introduced to promote critical examination of the alternatives produced in the collaborative group exercise. Essentially the critique activity is an argumentative essay. Students have to produce as many possible arguments that favor all of the conceptual ideas raised in class and then indicate which viewpoint is in accord with the experimental evidence. The critiques are designed to encourage the students to undergo a "critical discussion to decide which natural interpretations can be kept and which must be replaced" (Nelson, 1994). To write critiques, students had to clearly contrast two perceptions of physics principles, specifically students must provide convincing arguments both for an explanation arising from the Newtonian viewpoint and an alternative explanation. Furthermore, they must clearly state which viewpoint is "correct" based on experimental results.

A Combined Approach

The Conceptual Conflict Collaborate Group exercise was used in conjunction with the critique exercise by Kalman et al. (2004). Students were presented with two scenarios drawn from an earlier conceptual-conflict collaborative-group activity. One scenario corresponded to an explanation that does not have experimental validity and the other to the Galileo-Newtonian framework. Both scenarios were generated by students in the classroom. All in all, three conflict exercises were used in conjunction with the critique activity. Comparison was made with students who had in a previous year used the CG exercise alone. Analysis was done using only those students in the second year who took both the pre- and post-tests, who were present at all three conflict exercises and additionally who wrote all three critiques. The addition of the critique produced a statistically significant improvement for those students exposed to both the Conceptual Conflict Collaborate Group exercise and the Critique exercise compared with those students exposed to collaborative groups alone.

The strength of this combined use of peer conflict and writing (argumentative essay) is the depth of critical analysis it produces. Students quickly become invested in their positions through peer interactions (oral and written); and they see a stake in defending their ideas and evaluating others. The effects of this immersive experience of scientific communication are stable learning outcomes. The approach is most helpful to students who entered the classroom with a view of scientific knowledge as coherent and highly ordered—a necessity for instructional strategies that place importance on experiment and logical consistency to induce conceptual change. A limitation of this approach is that the competing/ideas (i.e., those of Aristotle, Galileo, Newton, and others) considered by students are presented from the beginning, rather than uncovered through their own experimental work and logical reasoning.

As discussed at the end of Strategy 1, RW is done by students at home and does not require additional class time. Indeed, since students have read the textual material before coming to class, the instructor should cut back on the material that is presented and use the saved class time for "flipped classroom" activities such as those described in Strategy 2. Typically to implement Strategy 2, we have students do Reflective Writing so that time is available for the Conceptual Conflict Collaborative Group activity, as well as other activities.

STRATEGY 3. THE ELICIT-AND-CHALLENGE APPROACH AND THE BRIDGING TECHNIQUE

The third and final strategy involves the use of an Elicit-and-Challenge approach along with a Bridging Technique. This method starts with the assumption of a mixed epistemology and is implemented within a "model-centered" physics classroom. Instead of conflicting explanations, students build models that need to be explained. As with the Reflective Writing (RW) and conceptual conflict techniques and the joint Labatorial-RW interventions, these strategies draw on the history and philosophy of science for both inspiration and implementation.

Elicit-and-Challenge Approach

The Elicit-and-Challenge approach begins with a set of hands-on activities to expose students to common set of concepts, ideas, and skills for the lesson unit. Then, students are placed in small groups to complete a modeling task. A consensus model is developed and shared with the entire class. Students articulate and defend their models before peers and the instructor.

For example, in Lattery (2017) students are asked to develop and present dynamical models of fan-cart phenomena based on their understanding of statics, the concept of net force, and numerous motion detector activities. Two primary cases are considered: the one-way trip of the fan cart (the mechanical analog of the vertical ball drop) and the two-way trip of the fan cart (the mechanical analog of the vertical ball toss). The ultimate goal of such activities is for students to acquire a classical (Newtonian) force concept through an extended and carefully guided process of model building.

As students struggle to respond to new information (contrary experimental evidence, logic arguments, and resources from related physical system), new models are generated. After multiple competing models are thoroughly considered, students write a paper on their ideas, receive a peer review, and provide a rebuttal. In the above example, students are observed to generate only four basic pre-Newtonian models of the one- and two-way trips of the fan cart; these models map onto those generated by ancient philosophers and scientists for the analogous cases of the vertical ball drop and toss.

As a cumulative activity, students "teach and defend" their model to others using a whiteboard presentation—a key technique used in the Modeling Method of Instruction (Jackson et al., 2008). The goal of this Elicit-and-Challenge approach is not necessarily to alter the student's alternative conceptions or epistemologies, but to give students space to explore the limits and exhaust the defense of their models. This often-regressive process highlights the distinct weakness of student's prior ideas, and challenges students to evaluate their assumptions. In short, the process primes them for target ideas.

In the above implementation, the history of science is incorporated in two ways. First, the instructor employs detailed arguments for/against competing models of the vertical ball drop and toss to probe student's views of the associated fan-cart phenomena. This can be done either in large-group class discussions or through anonymous peer reviews. Second, the instructor revisits the historical connections at the end of the unit to validate the students modeling efforts ("great minds think alike") and highlight the nature of the scientific modeling process (e.g., models are tentative, models to specific phenomena provide the key means to evaluate ideas, and multiple competing models are the norm in frontline science).

The use of a peer-review process to explicate and evaluate student competing models is very similar to the Critique Exercise previously described. In either case, the teaching principle is the same: for deep and sustained learning to occur, students must be given the opportunity to consider multiple competing models—whether generated through the student modeling process (Lattery, 2017) or through comparison of theories proposed by different groups (Kalman et al., 2004; Kalman and Rohar, 2010).

The primary challenge of the Elicit-and-Challenge approach is teacher training; in order to negotiate/challenge the various student models presented, teachers must have a sound understanding of the subject-matter content *and* a strong technical knowledge of how students think and learn in the domain. Additionally, this approach does not always end in a "tidy" resolution—the attainment of the intended learning objective. After "exploring the limits and exhausting the defenses," a student may be unable to make the intellectual leap to target ideas. This can be frustrating to the student. The purpose of the Bridging Technique in the next section is to marshal a set of resources (prior to or in parallel with the Elicit-and-Challenge process) so that the students can, with teacher guidance, discover target ideas.

Bridging Technique

In the Bridging Technique "students are guided by the instructor through a chain or network of related modeling tasks intended to bridge the student's prior knowledge with target knowledge." (Lattery, 2017, p. 254). For example, in the above activity, a network of cases involving a double fan cart is used to bridge the student's intuitive understanding of the one-way trip to a classical understanding of the two-way trip. The Bridging Technique is implemented either after or in parallel with the Elicit-and-Challenge tasks, although students do not generally recognize the bridging tasks as relevant to those tasks.

In contrast to Elicit-and-Challenge activities, bridging activities:

> do not challenge student alternative conceptions, but develop the capacity of students to recognize knowledge drawn from one physical case as relevant (literally similar or analogous) to another, and extend the range of applicability of a single unifying idea ("things go back to their original shape" or "net force steps produce velocity kinks") over a range of related physical cases (p. 255).

The Bridging Technique relies on the ability of students to extend prior knowledge (commonsense intuitions, folk science, anchoring intuitions, p-prims, etc.) to new domains through formal literal similarity and analogical comparisons. It should be noted that the Bridging Technique used in isolation, does not lead to sustainable learning outcomes because "it circumvents the student's initial high-priority ideas" (p. 255). In other words, in the above example, the student may be able to follow the "Newtonian agenda" of the bridging sequence, but not acquire the specific tools to understand why previous commonsense ideas fail. As a result, these less-productive ways of thinking remain central to the student's thinking and tend to resurface in new contexts.

A Combined Approach

Science learning gains achieved by combining the above approaches are documented through several detailed case studies in an introductory physics classroom (Lattery, 2017). A combined approach reflects openness to the question of student knowledge. In other words, it accepts the possibility that student knowledge consists of either "knowledge-in-pieces," for which the Bridging Technique is appropriate; or, more coherent structures, for which the elicit-and-challenge approach is suitable. A strength of this approach is that the competing ideas of the students flow naturally from their own experimental work and scientific reasoning; resolution of conflicts is guided by peer and instructor questioning—the latter inspired by the detailed analysis of model justifications in the history of science.

Strategy 3 requires a classroom learning environment that immerses students in the scientific modeling process, such as the Modeling Method of Instruction (Jackson et al., 2008). This strategy has been used successfully in a university-level physics course for non-science majors seeking general education credit. Future studies are being planned to evaluate this approach in middle and high school classrooms (grades 6–12).

SUMMARY

Contemporary science education reform must address the diverse needs of a "mixed epistemology" classroom. In this article, we presented three strategies that show promise to address these complex issues: Labatorials and Reflective Writing of Kalman et al. (2015); Conceptual Concept Collaborative Group and Critique essays of Kalman and Rohar (2010); and the Elicit-and-Challenge and Bridging Technique described by Lattery (2017).

Common to these approaches is an emphasis on the nature of science, subject matter content, and the role of the history of science in science education. Also common are activities that enable students to think through multiple competing ideas of the same phenomena with peers. Sorting through the strengths and weaknesses of multiple competing ideas enables students to understand not only why target conceptions succeed, but also why initial conceptions fail; both types of understanding are essential for deep learning in science (Lattery, 2017). Note that while the above three dual or combined instructional approaches target the same level of students, these interventions are sophisticated and not designed to be implemented all in one course. For a complete discussion of issues that teachers should take into consideration in employing these strategies, please consult the sources and references therein.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

Ahrensmeier, D. (2013). A practical application of Physics Education Research-informed teaching interventions in a first-year physics service course. *J. Tech. Educ.* 1, S166–S178.

Chi, M. T. H. (2013). "Two kinds and four sub-types of misconceived knowledge, ways to change it, and the learning outcomes," in *International Handbook of Research on Conceptual Change, 2nd Edn,* ed S. Vosniadou (New York, NY: Routledge Press), 49–70.

Clement, J., and Rea-Ramirez, M. (eds.). (2008). *Model Based Learning and Instruction in Science.* Dordrecht: Springer.

Clough, M. P. (2006). Learners' responses to the demands of conceptual change: considerations for effective nature of science instruction. *Sci. Educ.* 15, 463–494. doi: 10.1007/s11191-005-4846-7

Cooper, J. (2007). *Cognitive Dissonance: 50 Years of a Classic Theory*. Thousand Oaks, CA: Sage Publications.

Crouch, C. H., and Mazur, E. (2001). Peer instruction; ten years of experience and results. *Am. J. Phys.* 69, 970–977. doi: 10.1119/1.1374249

diSessa, A. (1996). "What do just plain folk know about physics?," in *The Handbook of Education and Human Development: New Models of Learning, Teaching, and Schooling,* eds D. Olson and N. Torrance (Cambridge: Blackwell), 709–730.

Elby, A. (2001). Helping students learn how to learn. *Am. J. Phys. Educ. Res. Suppl.* 69, S454–S464. doi: 10.1119/1.1377283

El-Helou, J., and Kalman, C. S. (2018). Reflective writing for a better understanding of scientific concepts in high school. *Phys. Teach*. 56, 88–91. doi: 10.1119/1.5021434

Festinger, L. A. (1957). Cognitive dissonance. *Sci. Am.* 207:93.

Feyerabend, P. (1993). *Against Method, 3rd Edn.* Brooklyn, NY: Verso.

Feyerabend, P. K. (1962). "Explanation, reduction and empiricism," in *Minnesota Studies in the Philosophy of Science*, *Vol. III,* eds H. Feigl and G. Maxwell (Minneapolis, MN: University of Minneapolis Press), 28–59.

Gadamer, H. G. (1975/1960). *Truth and Method*, eds G. Barden and J. Cumming (New York, NY: Crossroads).

Harmon-Jones, E., and Harmon-Jones, C. (2007). Cognitive dissonance theory after 50 years of development. *Z. Sozialpsychol.* 38, 7–16. doi: 10.1024/0044-3514.38.1.7

Heidegger, M. (1962). *Being and Time (John Macquarrie & Edward Robinson, Trans.)*. New York, NY: Harper and Row.

Hestenes, D., Wells, M., and Swackhamer, G. (1992). Force concept inventory. *Phys. Teach*. 30, 141–158. doi: 10.1119/1.2343497

Hofer, B. K. (2000). Dimensionality and disciplinary differences in personal epistemology. *Contemp. Educ. Psychol.* 25, 378–405. doi: 10.1006/ceps.1999.1026

Huang, X., and Kalman, C. S. (2012). A case study on reflective writing. *J. Coll. Sci. Teach.* 42, 92–99.

Jackson, J., Dukerich, L., and Hestenes, D. (2008). Modeling instruction: an effective model for science education. *Sci. Educ.* 17, 10–17.

Kalman, C. S. (2002). Developing critical thinking in undergraduate courses: a philosophical approach. *Sci. Educ*. 11, 83–94. doi: 10.1023/A:1013071130538

Kalman, C. S. (2009). The need to emphasize epistemology in teaching and research. *Sci. Educ*. 18, 325–348. doi: 10.1007/s11191-007-9135-1

Kalman, C. S. (2010). Enabling students to develop a scientific mindset. *Sci. Educ*. 19, 147–163. doi: 10.1007/s11191-009-9186-6

Kalman, C. S. (2011). Enhancing students' understanding of concepts by getting students to approach text in the manner of a hermeneutical circle. *Sci. Educ*. 20, 159–172. doi: 10.1007/s11191-010-9298-z

Kalman, C. S. (2017). *Successful Science and Engineering Teaching: Theoretical and Learning Perspectives, 2nd Edn*. New York, NY: Springer.

Kalman, C. S., and Aulls, M. W. (2003). Can an analysis of the contrast between pre-Galilean and Newtonian theoretical frameworks help students develop a scientific mindset? *Sci. Educ.* 12, 761–772. doi: 10.1023/B:SCED.0000004554.35266.7a

Kalman, C. S., Milner-Bolotin, M., and Antimirova, T. (2010). Comparison of the effectiveness of collaborative groups and peer instruction in a large introductory physics course for science majors. *Can. J. Phys.* 88, 325–332. doi: 10.1139/P10-024

Kalman, C. S., Morris, S., Cottin, C., and Gordon, R. (1999). Promoting conceptual change using collaborative groups in quantitative gateway courses. *Am. J. Phys.* 67, S45–S51. doi: 10.1119/1.19080

Kalman, C. S., and Rohar, S. (2010). Toolbox of activities to support students in a physics gateway course. *Phys. Rev. Spec. Top. Phys. Educ. Res*. 6:020111. doi: 10.1103/PhysRevSTPER.6.020111

Kalman, C. S., Rohar, S., and Wells, D. (2004). Enhancing conceptual change using argumentative essays. *Am. J. Phys.* 72, 715–717. doi: 10.1119/1.1645285

Kalman, C. S., Sobhanzadeh, M., Thompson, R., Ibrahim, A., and Wang, X. (2015). Combination of interventions can change students' epistemological beliefs. *Phys. Rev. Spec. Top. Phys. Educ. Res.* 11:020136. doi: 10.1103/PhysRevSTPER.11.020136

Khanam, N., and Kalman, S. (2017). Implementation and evaluation of the course dossier methodology. *Can. J. Scholarsh. Teach. Learn.* 8, 1–20. doi: 10.5206/cjsotl-rcacea.2017.1.7

Kuhn, T. S. (1962/1970). *The Structure of Scientific Revolutions.* Chicago, IL: University of Chicago Press.

Kuhn, T. S. (1981/2000). "What are scientific revolutions?," in *The Road Since Structure*, eds Conant and J. Haugeland (Chicago, IL: University of Chicago Press), 13–32.

Lakatos, I. (1970). "Falsification and the methodology of scientific research programs," in *Criticism and the Growth of Knowledge*, eds I. Lakatos and A. Musgrove (New York, NY: Cambridge University Press), 91–196.

Lasry, N., Mazur, E., and Watkins, J. (2008). Peer instruction: from harvard to the two-year college. *Am. J. Phys.* 76, 1066–1069. doi: 10.1119/1.2978182

Lattery, M. J. (2017). *Deep Learning in Introductory Physics: Exploratory Studies of Modeling-Based Reasoning* (Charlotte, NC: Information Age Publishing), 279.

Linenberger, K. J., and Bretz, S. L. (2012). Generating cognitive dissonance in student interviews through multiple representations. *Chem. Educ. Res. Pract.* 13, 172–178. doi: 10.1039/C1RP90064A

Madsen, A., McKagan, S. B., and Sayre, E. C. (2015). How physics instruction impacts students' beliefs about learning physics. *Phys. Rev. Spec. Top. Phys. Educ. Res.* 11:010115. doi: 10.1103/PhysRevSTPER.11.010115

Mazur, E. (1997). *Peer Instruction; A User's Manual*. Upper Saddle Hill, NJ: Prentice Hall.

McDermott, L. C., and Shaffer, P. S. (2001). *Tutorials in Introductory Physics and Homework Package*. Englewood Cliffs, NJ: Prentice Hall

McKinnon, J. W., and Renner, J. W. (1971). Are colleges concerned with intellectual development. *Am. J. Phys.* 39, 1047–1052. doi: 10.1119/1.1986367

Nelson, C. (1994). "Critical thinking and collaborative learning," in *Collaborative Learning: Underlying Processes and Effective Techniques*, eds K. Bosworth and S. J. Hamilton (San Francisco, CA: Jossey-Bass Higher and Adult Education Series), 45–48.

Newton, I. (1686). *Philosophiæ Naturalis Principia Mathematica*. London: Royal Society.

Popper, K. R. (1963). *Conjectures and Refutations, 2nd Rev. Edn.* London: Routledge and Kegan Paul.

Posner, G., Strike, K., Hewson, P., and Gertzog, W. (1982). Accommodation of a scientific conception: toward a theory of conceptual change. *Sci. Educ.* 66, 211–227. doi: 10.1002/sce.3730660207

Renner, J. W., and Paske, W. C. (1977). Comparing two forms of instruction in college physics. *Am. J. Phys.* 45, 851–859. doi: 10.1119/1.11061

Sobhanzadeh, M., Kalman, C. S., and Thompson, R. I. (2017). Labatorials in introductory physics courses. *Eur. J. Phys.* 38, 1–18. doi: 10.1088/1361-6404/aa8757

Vosniadou, S., Vamvakoussi, X., and Skopeliti, I. (2008). "The framework theory approach to the problem of conceptual change," in *International Handbook of Research on Conceptual Change,* ed S. Vosniadou (New York, NY: Routledge).

Vygotsky, L. S. (1978). *Mind Insociety: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.

Wandersee, J., Mintzes, J., and Novak, J. (1994). "Research on alternative student conceptions in science," in *Handbook of Research on Science teaching and Learning,* ed D. Gabel (New York, NY: Simon & Schuster and Prentice Hall International), 177–210.

Wood, D. J., Bruner, J. S., and Ross, G. (1976). The role of tutoring in problem solving. *J. Psychol. Psychiatry* 17, 89–100. doi: 10.1111/j.1469-7610.1976.tb00381.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Education

ORIGINAL RESEARCH
published: 31 August 2018
doi: 10.3389/feduc.2018.00073

A Learning Community Approach for Post-Secondary Large Lecture Courses

Steven P. Ehrlick[1]* and James Slotta[2]

[1] RTA School of Media, Ryerson University, Toronto, ON, Canada, [2] Lynch School of Education, Boston University, Boston, MA, United States

This design-based research (DBR) study examined the ways in which a learning community approach can be enacted in large undergraduate lecture courses through a scaffolded, complex curricular design that utilizes active and inquiry-based learning. By combining a traditional lecture with breakout tutorials, the study involved two iterations, firstly by adopting the Fostering Communities of Learners (FCL) pedagogical model, then by augmenting the model by blending its methodology with elements from a more recent model called Knowledge Community and Inquiry (KCI). Both iterations were evaluated for adherence to, and enactment of, the FCL principles. The second iteration was further evaluated to determine the impact of adding a KCI collective knowledge base. Measures included the enactment of the curricular design, achievement of course learning outcomes, the group inquiry project, tutorial activities, and focus groups for teaching assistants and students. Findings provided evidence of the viability and effectiveness of a learning community approach in large lecture courses at the undergraduate level when combining the learning principles of the FCL model with the student-populated dynamic knowledge base. Students achieved both individual and group success in meeting learning outcomes through individual inquiry and collaborative, active learning, with the knowledge base providing a forum for students to share their research and access ideas for their inquiry.

Keywords: learning community, active learning, collaborative learning, inquiry learning, reciprocal teaching, lectures, knowledge community

OPEN ACCESS

Edited by:
Rob Cassidy,
Concordia University, Canada

Reviewed by:
Andrew Joyce-Gibbons,
Durham University, United Kingdom
Krista Mincey,
Xavier University of Louisiana,
United States

***Correspondence:**
Steven P. Ehrlick
sehrlick@ryerson.ca

Specialty section:
This article was submitted to Digital Education, a section of the journal Frontiers in Education

Received: *18 April 2018*
Accepted: *13 August 2018*
Published: *31 August 2018*

Citation:
Ehrlick SP and Slotta J (2018) A Learning Community Approach for Post-Secondary Large Lecture Courses. Front. Educ. 3:73. doi: 10.3389/feduc.2018.00073

INTRODUCTION

Few institutional practices have survived the centuries as intact as the university lecture. "The one teaching the many" is the bedrock upon which most professors rely for the transfer of information, with the goal of inculcating knowledge through repeated oral presentations. This instructivist, didactic method of instruction is viewed as a practical means by which learned scholars can transmit their knowledge to the many with the greatest level of efficiency. However, in recent decades this didactic approach has been assailed as ineffective (Bloom and Shuell, 1981; Bransford et al., 2000).

This study presents an account of two iterations of the design and enactment of a Fostering Communities of Learners (FCL) pedagogical model, first conceived and applied by Brown and Campione (1996), which is itself part of the larger domain of study, called learning communities (Bielaczyc and Collins, 2006). FCL has influenced other learning community theoretical models, but has not seen much advancement theoretically or even applications in research studies over the past two decades. The present study afforded the opportunity to develop the FCL model for an undergraduate environment, augment it with theoretical elements from the Knowledge Community and Inquiry (KCI) model (Slotta and Najafi, 2013), and apply it to the design of a large lecture university course—a domain where it has not been previously attempted.

Recognition of the need for students to acquire Twenty-first century skills is widespread, but pedagogical practices remain largely entrenched in a behaviorist mode of top-down instruction and quantifiable testing. In contrast (e.g., Brown, 1994) viewed students as "active constructors" of their knowledge within a community of learners. Brown and Campione (1996) developed the FCL model to address her learning science conceptual underpinnings.

However, the adoption of such alternative pedagogical approaches in higher education requires the willingness of professors to innovate, spending significant time on their course designs, which would also entail epistemic challenges to students who are not accustomed to such forms of learning. But if the university experience is to be a vital factor in students' future success, then it is important for students to see their university education as providing them with the academic skills they will use throughout their lives. Active learning is a pedagogical approach well-suited to the development of Twenty-first century skills by engaging students in activities designed to promote collaboration, reflection, and problem solving, with the goal of achieving learning outcomes, developing critical thinking and providing applied course content (Bonwell and Eison, 1991; Prince, 2004; Felder et al., 2009). Informed by a constructivist perspective, active learning almost always includes collaborative or co-operative activities (Prince, 2004) where students participate in hands-on, real-life activities which help them connect their experience in school with later experiences after graduation and reinforce a positive attitude toward the institution where they learned these skills.

Any study that includes an active learning component should attempt as far as possible to make a comprehensive assessment of learning outcomes as opinions may vary and data deemed unreliable without it. This study has been purposeful in its curricular design to not only input active learning components but to create criteria to measure the epistemological impact of those activities.

Interpretive Frameworks

The FCL set of principles is both a set of learning science principles and a pedagogical model designed to help students develop expertise in service to their peers, collaborate, and advance their collective understanding through active learning with a community ethos. The structure of FCL consists of individual and group research on core topics followed by the sharing of research by way of several active learning activities, including cross-talk, jigsaw, and reciprocal teaching. The model culminates with the creation and presentation of a consequential task (Brown and Campione, 1996). The model is designed to work with content that requires deep understanding and this works best with "big ideas," transforming the classroom into a learning community. Students begin to specialize, expanding their own potential (the more adept described by Brown as "majoring"), as the group proceeds toward consensus. Students who understand the topic become advisors to those who are less adept (Vygotsky, 1978). FCL then, at its simplest, is a three-step process—research, sharing, and a consequential task (Brown and Campione, 1996) (**Figure 1**).

Slotta and his colleagues (Slotta and Peters, 2008; Slotta and Najafi, 2013) created KCI with FCL as a foundation, to guide the design of learning community curricula that scaffolds students and teachers in carefully designed inquiry scripts. A main feature of KCI is the creation of a collective knowledge base that is indexed to the specific learning goals of the curriculum. Students provide content for the knowledge base during individual and small group inquiry, argumentation and discussion. This knowledge base becomes as a persistent resource for all inquiry, as students refine their understanding through scaffolded activities.

KCI is based upon three guiding principles: (1) that students work collectively and collaboratively to build their knowledge base, which is both a product of, and resource for inquiry activities; (2) that inquiry activities are connected to themes emerging from the community's collective interests, and (3) that inquiry activities provide assessable outcomes that are linked to the required learning goals (Slotta and Najafi, 2013).

Another contribution of KCI is its inclusion of metacognitive orientation for students and teachers (an "icebreaker") that explains the learning process inherent in the KCI model. This creates for the students an awareness of, and strategies for the mechanics of learning, the execution of the curricular design and an initial understanding of the dynamics at play in the building of a learning community (Slotta and Najafi, 2010).

With the addition of a collective knowledge base to the FCL framework, the design under study approached that of KCI (differing in the retention of the core FCL structures). FCL and KCI share the goal of making the learning community approach more accessible, permanent, and practicable for instructors.

To date, almost all research and experimentation in inquiry-based learning communities has been situated in K-12 classrooms, with much smaller cohorts of students, and relatively lighter content requirements. And as Scardamalia and Bereiter (2010) point out, knowledge building cultures do not pop up spontaneously. They require the diligence and creativity of the teacher in order to maintain a community where ideas are constantly being generated and approved upon.

This study sought to migrate the FCL model from its study and application in middle school to a university setting, swapping children for young adults and the typical K-12 classroom with a large lecture hall. Brown and Campione's (1996) agenda for their research was to "contribute to a theory of learning that can capture and convey the essential features of the learning environments that we design" (p. 290). This too was the intention of the present study—to capture the essential features of the FCL design, particularly when augmented by the artifacts from other learning community theoretical approaches, especially from KCI. Fostering a learning community was the paramount goal, but the researchers recognized that pragmatism also has high currency. Enactment of this complex curricular design first relied on strict adherence to the learning science principles that infused it, and

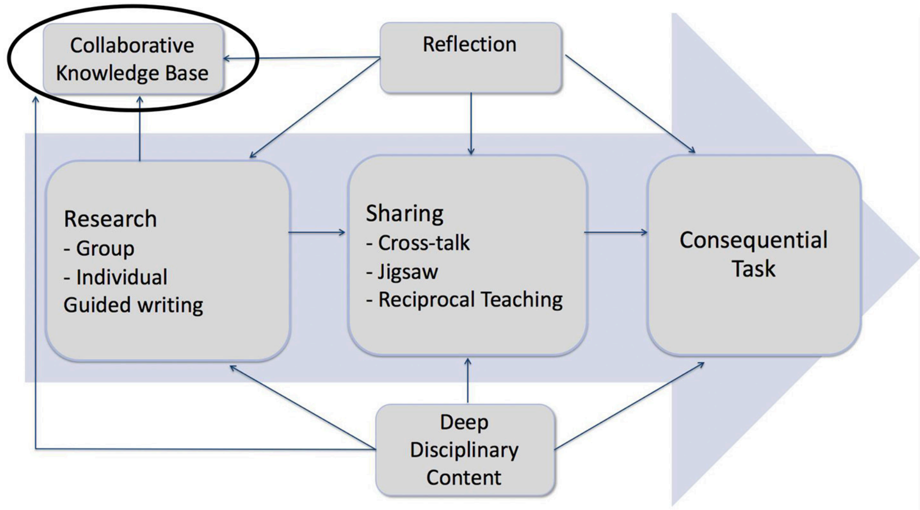

FIGURE 1 | FCL model with the addition of KCI-inspired Collaborative Knowledge Base.

once satisfied, the enactment was analyzed not only for adherence to the design, but whether the model was replicable and useful; a model that is pragmatic was an important ancillary goal.

The purpose of this study is to be the first scientific investigation of the FCL curricular model enacted in a large university course setting, and to investigate the impact of the introduction of a KCI-inspired collaborative knowledge base. This design-based research (DBR) studied the efficacy of a complex pedagogical model in two iterations, situated in a lecture hall setting, populated in each iteration by ∼235 undergraduate students in order to investigate the model's viability and to make recommendations as to improving its affordances for future research and application. Specifically, this study sought to address the following two research questions:

1. How can the FCL model be applied as a learning community approach within a large undergraduate course?
2. What are the limitations of the model, and what adaptations can help respond to those limitations?

METHODOLOGY

Design-Based Research

Applying FCL principles to a university-level course required a methodology that would facilitate the study of multiple iterations of a curricular design and the ability to modify, change, and augment that curriculum. DBR provides certain affordances well-suited for a study of a theoretical curricular design, with successive iterations that follow design, analysis, and re-design cycles (Shavelson et al., 2003), thereby allowing for any modifications and augmentations in the design to reveal themselves in the data. DBR is commonly used in the learning sciences to study "complex educational systems" where theories of learning are given practical application through the construction of an effectively designed learning framework (Bannan-Ritland, 2003). This integrative design is characterized by iteration of design, enactment and evaluation. This methodology is therefore a logical choice for the design and enactment of a complex curricular design guided by the FCL principles laid out by Brown and Campione (1996) (**Figure 2**).

DBR may draw upon a mix of qualitative and quantitative measures. DBR sets out to provide solutions to perceived pedagogical obstacles. Past educational research tended to concentrate on individual differences and the causal effect of interventions. DBR is characterized by a more grounded theory approach, where goals manifest during the running of a course design which spawns more design ideas (Bereiter, 2002). McKenney and Reeves (2013) lament however, that DBR literature tends to focus on design interventions without enough emphasis on new understandings of educational phenomena. Bielaczyc and Collins (2007) identify three tensions in DBR, one between improving practice and refining theory, the second between the individual components and the integrity of the whole design, and finally, the challenges inherent in multidisciplinary research. To be effective then, DBR must attend to both innovative curricular design and a rigorous assessment of tangible results that educators may then be able to implement, mindful of not allowing the improvement of practice to forsake theory development, and vice versa, and with care not to favor particular components at the expense of the overall design.

Participants

This section provides a description of the teaching assistants and students who participated in the study.

Teaching Assistants (TAs)

Four TAs were provided for the course under study. TAs led 11 tutorials and one of the researchers led the 12th. The TAs were fourth year undergraduate students from the same program in which the course was offered, three of whom had previously taken this course.

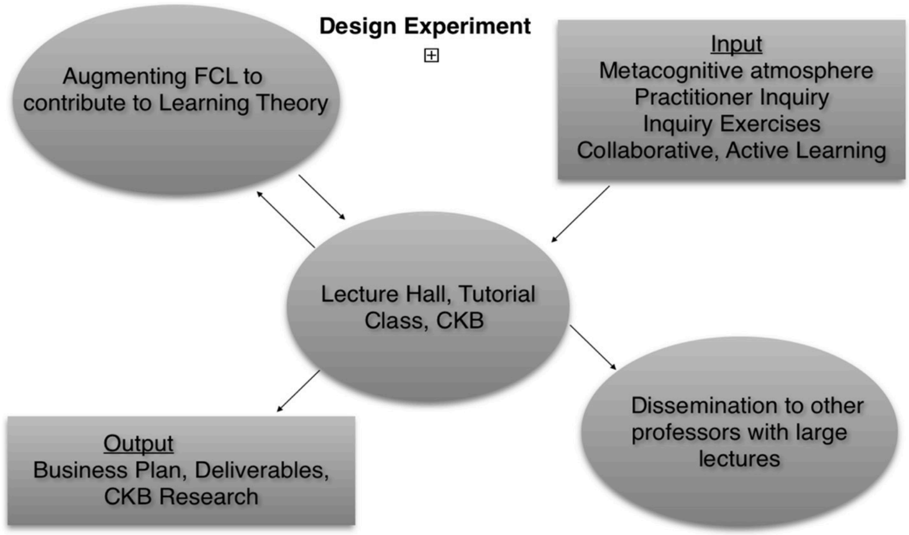

FIGURE 2 | Modified Brown (1992) graphic showing features of design experiments supplemented with particulars from this study.

Students

The course under study, *Business of Creative Media*, is an introductory general business course designed to introduce undergraduate media students to basic business, legal, and financial aspects of the media industry. The course is a required/elective hybrid. The first iteration included 179 Media Production B.A. program students. The second iteration had 145 Media Production students enrolled in what is for them, a required course. This course was also an elective for the first year Sport Media B.A. program students. There were 54 of these students in the first and 24 in the second iteration of the study. The reasons for the drop off was due to a change in their elective choices. The two iterations also had students from the Creative Industries B.A. program. In the first iteration, there were only two such students. But again, due to changes in electives, there were 46 Creative Industries students in the second iteration. A third program, a New Media B.F.A. program, brought one student to the first and 15 students in the second iteration of the study. In total, there were 236 and 231 students in the two iterations, respectively.

Study Context—Needs Assessment

The first iteration was implemented in the Fall of 2015, followed a year later by the second iteration in the Fall of 2016. The setting was a unique kind of lecture hall: a movie theater, still active in the evenings, rented by the university during daytime hours for large lecture courses. In general, students in media courses are provided intensive labs and lectures in all forms of media production, from television and radio broadcast, to transmedia digital platform story construction.

In the years preceding the study, the course had run as a large lecture, 3 h in duration. There were no breakout tutorials and two TAs were assigned as graders. Workshops were run within the lecture hall, but these were limiting in a number of ways. Moving students into their groups in an amphitheater was chaotic. There were only 30–40 min of the lecture time devoted to active learning activities and monitoring and proffering advice was limited to those groups closest to the aisles or in the front. Clearly, there was a need for breakout tutorials, and the researcher/practitioner of this study campaigned to have them added, which request was acceded to in time for this study.

The course as of Fall 2015 was structured as a traditional 3-h lecture during odd-numbered weeks while even-numbered weeks consist of a 1-h lecture plus a 2-h tutorial (20 students per tutorial).

Materials

Students have ubiquitous access to the Internet while on campus. The World Wide Web was accessed continually during tutorials and for active learning exercises in lectures (planned and *ad hoc*). Students also had access via the Internet to a learning management program called D2L, utilized by professors for course shells. Access to materials such as lecture slides, course documents, readings and tutorial instructions were housed in the D2L shell.

Methods

Various qualitative methods were used in the collection of data for this study (**Figure 3**). Built into the curricular design were student-generated artifacts that are natural data sources. Others were conceived as additional data sets. This abundance of data has the potential to create confusion, however, methodological choices can be seen to be valid if they provide the researcher with the tools with which to solve a problem (Trow, 1957).

The following forms of data collection were used to collect student experiences: (1) student and TA focus groups; (2) student generated artifacts (tutorial directed writing exercises, a reflective test and a group business plan); and, instructor field notes. By

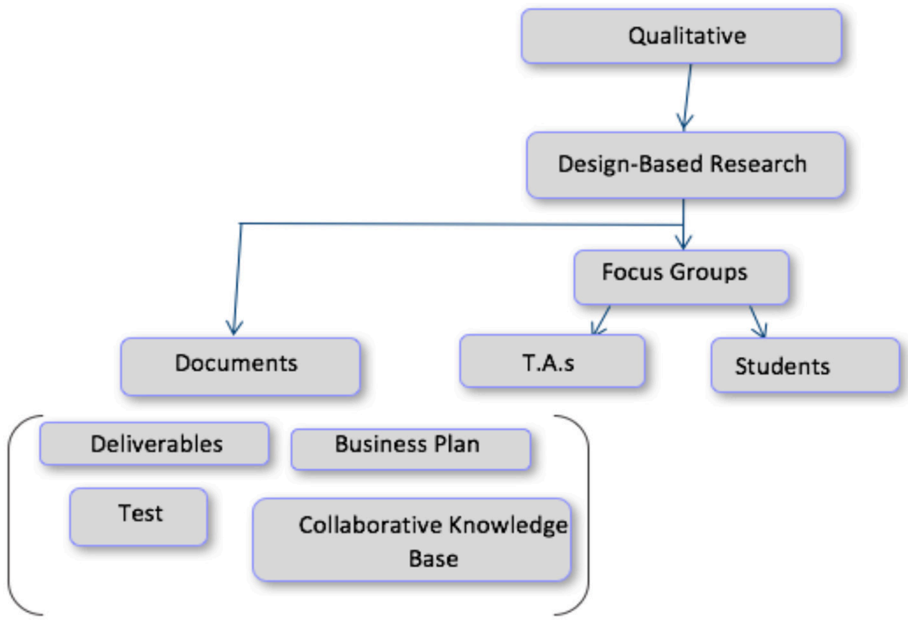

FIGURE 3 | Flow chart of methodology and data collection.

collecting and utilizing different qualitative methods, the plan was to prevent important research phenomena from escaping notice, rather, allowing such phenomena to surface (Erzberger and Kelle, 2003).

The test for the second iteration (which was reflective in nature for both iterations), was uploaded in digital form to D2L. This allowed for content analysis of each student test answer using the software NVivo, as well as for both focus groups (student and TA). Content analysis is a versatile method for analyzing text data (Cavanagh, 1997; Chi, 1997). Grades provided the data for some basic statistical comparisons.

COURSE DESIGN—ITERATION ONE

In this section we address our first research question: How can the FCL model be applied as a learning community approach within a large undergraduate course? The first iteration was designed to study both adherence and enactment of the FCL pedagogical model first developed by Ann Brown and her colleagues. The first iteration was intended as a baseline test of an FCL-designed curriculum enacted in a large lecture course with tutorials. Once satisfied with the adherence to the model, the curriculum was enacted to determine whether the design successfully fostered community amongst students. Specifically, did design elements such as cross-talk and jigsaw enhance student learning; were the lectures effective "benchmark lessons" and did they provide linkage to tutorial activities? The researchers analyzed the model in order to highlight deficiencies and learn how modifications and augmentations might be added to the design to create and sustain a more robust and effective learning community in a university setting with a large number of subjects.

The course was 12 weeks long. On odd numbered weeks a 3 h "benchmark" lecture was delivered. On even weeks a 1 h lecture was delivered then students spent 2 h in a tutorial (a total of six throughout the term). There were 12 tutorial sections, 11 run by the four TAs and one run by the researcher/practitioner. In essence, the 12 tutorials provided the researchers with 12 discreet classes in which to run an FCL curriculum. Two sessions were spent with the TAs instructing them on the framework of constructivism and the principles of FCL. The tutorials were scripted so that as much as possible, TAs served as time keepers and advisors, moving activities along on a schedule scripted by the FCL design. TA Instructor Notes were prepared, offering guidelines and potential issues as well as providing an explanation of the learning outcomes for each tutorial.

The purpose of the tutorials was to effectuate learning into action by having each group "incorporate" a company with the group members as its shareholders, directors and officers. Each tutorial was scripted to firstly reinforce the lecture topics and then to provide hands-on active learning exercises to increase individual and group knowledge funds in order to enable them to draft a comprehensive business plan. Each tutorial workshop culminated with a directed writing task, a short summary of each group's research activities ("Deliverables") sent to their TA for assessment. The tutorials were designed to reinforce the lecture material, provide hands-on individual and collective research and communal sharing of that research through cross-talk and jigsaw. In cross-talk, the TA would designate an officer, for example, the VP Legal, to explain and summarize to the tutorial class what their group had learned after researching a sub-topic. In other tutorials, an emissary from each group was sent into the other three groups to explain a sub-topic which had been researched by their group; at other times, officers with the same office convened to discuss their group's sub-topics. The idea behind jigsaw is to disseminate subtopics in order for all groups to comprehend the entire topic. These topics must have rich content, able to be subdivided so that by being exposed to all the sub-topics students can then understand the entire topic. A topic such as the law of copyright is one such example, where the sub-topics of copyright term, copyright ownership, infringement, and what constitutes a copyrighted work would be assigned to different groups for individual and group research prior to dissemination to the entire class by way of cross-talk or jigsaw. In this way the entire class is exposed to the totality of a given topic, gaining access to research applicable to their company. The tutorials were thusly scripted in such a way as to increase each student group's knowledge fund in order to complete the culminating consequential task, the drafting of a business plan for their business.

The tutorials contained scripted components that helped guide the "officers" (CEO, CFO, etc.) of the company to contribute their knowledge fund to the business plan. An emphasis on business innovation was highlighted in lectures and tutorials. These students live in a start-up world and as such, were encouraged to come up with innovative businesses and new ways of conducting their businesses. This coincides, for example, with Scardamalia and Bereiter's (2010) knowledge community principles of applying real ideas to authentic problems, of improving on existing ideas, exploring idea diversity, and building knowledge that has value to others.

Each tutorial ended with a Deliverable, a guided writing exercise that answered a research question, summarized the group's individual and collective research, clarified questions that related to their roles and factors particular to their companies and made predictions on the impact of their research on future issues related to their business plan. Deliverables were emailed to their TA for assessment. Each Deliverable was worth a maximum of 5% of a student's total grade, graded collectively. The scaffolded tutorial design offered a real-life scenario for students. For example, in the first tutorial, groups were instructed to incorporate a company[1]. Each group was required to decide the type of media company they would become and nature of their product or service. Would it be a production company making documentaries or a talent agency? The choice was left to the students, the only criteria being that the company be media-related. This lead to an interesting variable in the scripting of the design as there were instances where companies would be required to find ways to complement each other's companies (e.g., one company providing a service or product to another) and even pitch their companies collectively to a virtual investor. In some tutorial sections the companies complimented each other as if planned; in others the connections were tenuous. But all groups started with the same corporate structure.

The first tutorial established the corporations, as outlined in **Box 1**. The second tutorial reinforced the benchmark lecture on the law of copyright, each officer conducting individual research on a sub-topic followed by a roundtable discussion of what each officer discovered regarding their assigned copyright issue and how it might affect their company. Each group's VP Legal officer then described to the entire tutorial class the nature of their company's business and assessed if any of the other three companies might be ones they could do business with. The third tutorial dealt with finance, again reinforcing a lecture by the professor and a guest media industry CFO. Groups were encouraged to work with a dynamic budget spreadsheet, adding revenue, and costs in preparation of the budget they would produce for their business plans. Four financial documents were then assigned, one to each group (balance sheet, income statement, etc.) followed by a jigsaw activity where the CFO from each company made the rounds to the other three companies explaining the nature and purpose of the document that their group had researched.

During the fourth tutorial, a screenplay sample with Errors & Omissions issues was analyzed by each group. Groups were then asked to discuss the intellectual property issues they may have with their own company model and these were shared across companies. In the fifth tutorial, students took an abridged Meyers Briggs Personality Test. After discovering their personality type, a group discussion ensued, intended to shed light on each company's group dynamic. The companies were then instructed to downsize and reduce salaries. This had the dual intention of bringing to the surface any discord amongst group members and to give students a platform for discussing who was not pulling their weight in the company. There was also a jigsaw activity where officers with the same title could swap companies but no students in any of the tutorial sections took the opportunity to do so.

In the sixth and final tutorial, a surprise presentation was sprung on the groups. Each company had an hour to pull a pitch together for investors. CEOs were instructed that these investors were looking for four companies to invest in, so the four pitches needed to have a common theme and the complimentary aspects of the four companies addressed.

Concurrently, lectures covered "benchmark" topics, including law of copyright, law of contract, corporate structure, leadership, corporate culture, emotional intelligence, legal issues in media and the art of negotiation.

Assessments

The business plan was worth 50% of each student's total grade, with 20% of the scoring rubric devoted to individual assessment. The tutorial workshops Deliverables were worth 5% each. The test was worth 20%.

Enactment of Iteration One

To the extent that adherence to, and enactment of the FCL curricular design was the primary goal of the first iteration, this phase was a success. The previous year's running of the course had led the researchers to believe that introducing FCL elements to the newly added breakout tutorials would provide the environment for achievable enactment of the FCL model.

[1]Groups filled in an actual government-issued Articles of Incorporation form, setting out shareholders, first directors, and issuance of shares, but only sent it to their TA.

BOX 1 | **Excerpt from Deliverable 1—first tutorial.**

Deliverable 1 - Company Creation/Board of Directors and Officers Election. Corporation research.

1. Your GA will put you in your groups of 5. Each group has a number 1 through 4. This number will be used throughout the term.

2. Download articlesform.pdf at: **D2L>Content>Deliverables>Deliverable 1**. This is a fillable pdf file. Fill out sections 1, 2, 3, 4, 6 and 10 in the Articles of Incorporation Form document (name of corporation, address, names and addresses of first directors). For item 6 write in "Five Common Shares". You are issuing one common share to each group member. Come up with a name for your company or create a numbered company like 123456 Ontario Inc.

3. You are all now shareholders and board members with equal equity shares in the company. You each own one Common Share. Your task is to hold your first Shareholders Meeting to formally elect the Board of Directors. The Board will consist of:

DIRECTORS

Chairman of the Board (1)

Directors (4)

The Chairman of the Board chairs director meetings and casts a vote to break a tie. Directors advise as to overall corporate activities and direct its course strategically. Directors do not run the company day-to-day. They appoint Officers to run the company. Often, especially in small private companies, directors are also officers. This will be the case in your company.

Hold your first Shareholders Meeting and elect the Board of Directors. If you wish to be Chairman you may address the Board as to your qualifications (which may just be the desire to be the Chairman!). Appoint one member of the board as the Chairman of the Board. There are no significant extra responsibilities for being Chairman other than being the final arbiter when there is not consensus on a particular issue.

4. Once the Board is elected, the Shareholders Meeting ends and the Board will meet for the first time to elect officers. The Chairman will call a meeting of the Board of Directors.

5. The following offices must be filled and salaries allotted. **The budget for salaries is 50 coins. You do not have to spend all of the coins.**

President and CEO

Secretary and COO (production)

Treasurer and CFO

Vice President Legal

Vice President Marketing and Sales

The Officers are responsible for the day-to-day running of the company. They report to the Board. In this case, you are reporting to yourselves.

You must appoint and fill all of the positions above and assign salaries. This will impact responsibility over different aspects of the business plan. YOU CAN CHANGE THESE POSITIONS AT ANY TIME BY CALLING A BOARD OF DIRECTORS MEETING.

6. End the board meeting. **Create a resolution for the Board meeting**. A template can be found in:

D2L>Content>Deliverables>Deliverable 1.

Make sure to include all of the offices in paragraph 5 above. You will also assign Secretary and Treasurer offices to the COO and CFO respectively.

7. Next, the Officers will meet once they are appointed to discuss what their production mandate will be. Is this a film documentary company? A web design company? An app development company? Create a short list of ideas. Now do some preliminary research on your own (10 min). Go around the group and pitch ideas. Come to a consensus. **Write an elevator pitch on the type of production or service you are forming the company around**. You are not stuck with this business idea. You may well find that you need to pivot later in the term. Keep in mind that your main assignment is a business plan, not a production bible. The production or service will dictate the financial, personnel and infrastructure needs of your business. Make sure you can handle the production from a business point of view. There are only five of you. Imagine you are really going to start this company. How will you execute your vision realistically?

But the lectures remained largely instructivist, and only partially fulfilled the FCL requirement of "benchmark lessons." Some were merely instructional and used to clarify and highlight the previous 3 h lecture. To whatever extent the lectures could be considered benchmark lessons, they still did little to reinforce the learning community ethos within the class as a whole.

In essence, the first iteration design created 12 distinct communities of learners - without any student perception of belonging to a community of learners outside of their tutorials and small groups. Students within their tutorial sections were unaware of the research being conducted in other tutorial sections. There was no opportunity, no means for students to interact with other student's research activities by way of active learning activities outside of their own tutorial class. It became clear that a more global (i.e., whole class level) repository for research activities could help promote better awareness and exchange across the tutorial sections.

SECOND ITERATION

In this section the researchers addressed the second research question: What are the limitations of the model, and what adaptations can help respond to those limitations? The second iteration of the curriculum design involved the introduction of a Collaborative Knowledge Base (CKB) and the student research activities that populated it (see **Figure 4** for an example). Thus, a digital repository for individual student inquiry and "knowledge" was conceived, to provide a means by which students could share work, learn from each other, and create a sense that the entire class was working together as a learning community. The CKB was added to the curricular design in order to connect off-campus research activity, active learning in lectures, and group activity in the tutorials. Hence, the first iteration laid the groundwork for the second, in part by exposing the limitations in the curriculum.

As in the first iteration, the analysis focused on (1) adherence of the design to the pedagogical model, and (2) faithful enactment

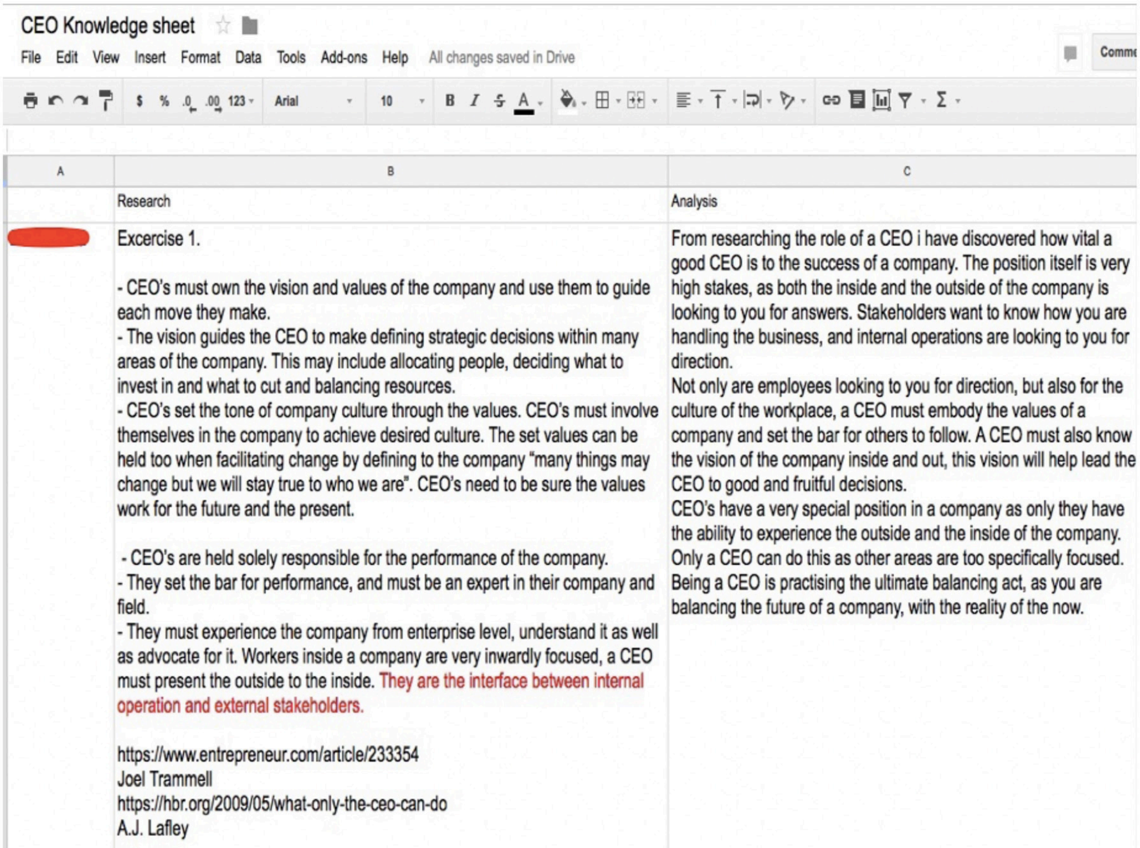

FIGURE 4 | Sample CKB entry for CEOs - Exercise 1. Students research a topic, provide analysis, then link to a personal page where they leave questions and comments for their tutorial group members.

TABLE 1 | NVivo Main and Child Codes used to analyze test question and focus groups.

Main codes	Child codes		
CKB Use	CKB as future reference	Referring to research of others	CKB and lecture linkage
Community of Learners	Collaborative learning	CKB and lecture linkage	Referring to research of others
Inquiry Learning	Research above and beyond	Compelled (pushed, forced) to do research	Gained expertise
Future Application	Real life experiences	CKB as future reference	Gained expertise

of the design during the course itself (i.e., the instructor did what was designed), as well as (3) student learning outcomes. Student assessment would remain virtually unchanged except for 10% of their grade which was allotted to the research exercises that populated the CKB. Unlike the first iteration of this study, particular attention was paid to the test, which was written on digital devices and uploaded to D2L for marking. One question was formulated to require students to reflect on their use of the CKB. Coding of this question was developed to identify the following major themes: CKB use, community of learners, inquiry learning, and future application. **Table 1** outlines the codes used to analyze the test and the student and TA focus group transcripts.

Analysis of focus groups included the adoption of a micro-interlocutor analysis based upon the work of Onwuegbuzie et al. (2009), whereby the focus group is assessed both as a group and as a series of individuals, which enables the researcher to record the responses of members who may not be contributors of a particular theme, but whom are nonetheless recorded and acknowledged in the overall analysis. For instance, such analysis might reveal a student who is silent, contrary or who tends to go along with the majority view. Eight randomly selected students partook in the focus group. All officer positions were represented.

Design Changes—Curriculum

Collaborative Knowledge Base (CKB)

The decision to add a Collaborative Knowledge Base to the curriculum was influenced by two theoretical models. Early on in their research, Scardamalia and Bereiter (2010) were interested in examining whether students reading other student's work would improve their ideas. Slotta and Najafi (2013) point out that Web 2.0 technologies have provided new tools for innovative pedagogical approaches. The researchers were faced with two significant issues before turning theory into practice: (1) what technology to use and (2) how to design it to allow a learning community to manifest? It was decided to experiment with a Google Doc spreadsheet. The following criteria were needed for a CKB to succeed as a user-friendly, collaborative research repository: ease of use, coherent structure, and usefulness. As each tutorial was composed of four groups containing five corporate officers, a research page for each officer was created. This meant 48 students assigned to each page—a more manageable number technically as well as from a student interaction point of view. Students were given research exercises catered to their particular office. The CKB exercises were assigned on a Thursday and were due to be uploaded by the following Monday's lecture.

During the lecture, the Google Doc was brought onscreen and students were called upon to explain their research and analysis. If, for example, a CFO student was selected for commentary, other CFOs were asked to join the discussion. The instructor would then switch to another student with a different role in his or her company, and the discussion would begin anew. In this manner, individual research was brought into the large lecture. During tutorials, groups were instructed to go back into the CKB and, using reciprocal teaching, explain to the other officers their research and its impact on the company.

This discussion of CKB research occurred during the beginning of five of the six tutorials (the sixth was the presentation day). Students gained knowledge related to their roles, then brought that knowledge back to their group for dissemination. In other words, they participated in a "conference" with their fellow officers, allowing them to instruct each other and better understand each officer's role. This would hopefully improve their business and be reflected in their business plan. As one student put it, "There were many times our group referred to a CKB exercise (not just the previous weeks, but all CKB exercises) to help build proper financial decks, or create a proper business plan."

By creating a permanent repository for individual research, students were able to share their findings and analysis both with the wider audience of the entire class as well as in their tutorial groups, thus strengthening the distributed expertise of each individual who had the opportunity to share that expertise with peers, sometimes in the large lecture, but always with their fellow officers. The CKB thus served as a permanent collection of research by individuals, which benefited all groups in all tutorials as a resource that aided them in the formation of their business plans. Student participation in the CKB added a research element missing from the first iteration and provided students with an opportunity to conduct deep inquiry related to their particular office. The curriculum became more ambitious with the addition of the CKB; it provided a valuable addition to the overall design. The CKB reinforced, in the minds of the student subjects, research, helping, understanding, learning, and knowledge, among other concepts.

Enactment of Iteration Two

The CKB was analyzed for evidence of its effectiveness as a medium for individual inquiry and as a viable research repository where students would populate the CKB with their own research and possibly benefit from the research of others. It was analyzed to see if there was any discernable increase in the achievement of learning outcomes across the student population.

Grades were first analyzed to detect any statistical differences in the performance of students from both iterations. The test conducted during the first iteration produced a median score of 81.7%. The median test score for the second iteration was 82.4%. The difference between the two medians is negligible (SD of 0.35), however, students in second iteration produced a 9.4% increase in the number of scores over 90% with a standard deviation of 10.5. Notable as well, is that while the frequency of students with the highest marks (A+) was significantly higher in the second iteration, overall, students in the first iteration fared better in the test (**Figure 5**). This may be as a result of the increased individual workload for students in the second iteration phase of the study which provided the opportunity for fatigue or apathy.

Grades for the business plan show the most significant disparity between the two iterations. The student frequency for the grade range of A− to A+ (between 80 and 100%) was 51.2% for the first and 61.5% for the second iteration (SD of 5.15). The grading rubric for the two iterations was identical which was designed to mitigate discrepancies in marking by the TAs during both iterations. In the first iteration it was necessary to have one of the TAs more normally distribute her grades as they initially fell significantly lower than the assessments of the researcher and the other TAs. This adjustment was unnecessary in the second iteration.

An examination of the content of student focus group transcripts and the answers to Question 2 of the test revealed several themes, including the impact of the research contributions of others. With regards to the test, 50% percent of the student population who wrote the test digitally ($N = 223$) made favorable comments concerning the benefit they gained from reading the research of other officers from other companies in the CKB, many of the students offering more than one example. As one student noted:

> *"Everyone in this class had a different view on each exercise and everyone's company is different. Reading through all their answers gave me so many different perspectives and helped me to grasp some concepts more easily when put into different words. For example, while I had only listed three aspects of being a CEO which I had deemed most important, others had listed different aspects that I realized were also important points and which gave me a much better understanding of my position."*
>
> *-Student test response*

The ability of the CKB to help students understand the parameters of their role by reading the work of fellow officers was a constant theme in answers to Question 2. The CKB provided some students with a leg up when they experienced frustration or anxiety due to being thrust into corporate positions they knew little about. As another student comments in the test:

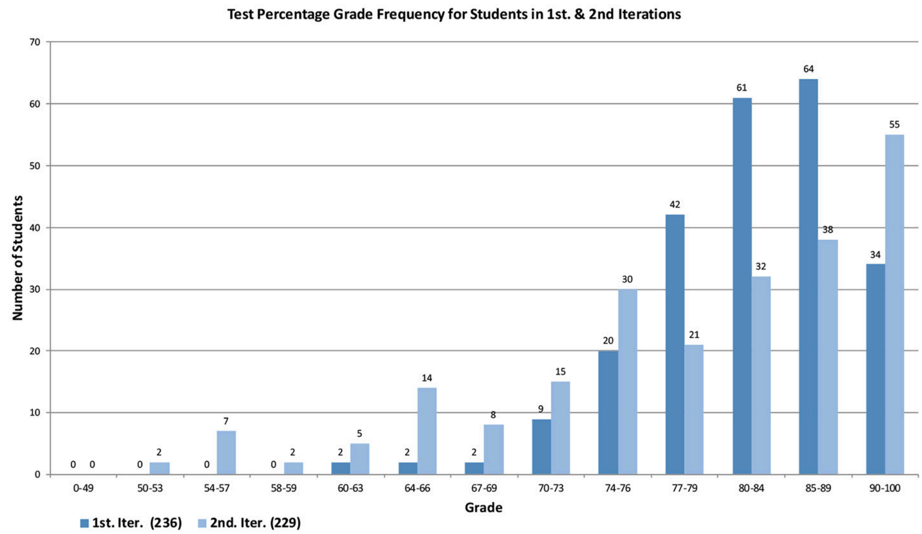

FIGURE 5 | Grade frequency for test. 236 and 229 students took the test during the 1st and 2nd iterations, respectively.

> *"Another CKB exercise I remember specifically is #3, which required the examination of different types of business models. I remember being unsure which business model would work best for the company we were designing, so I looked at other COO's businesses that were similar to ours, and the models they decided to use. This helped me vastly in deciding what would be best for our company, and the CKB exercises in general deepened my overall understanding of the entire class."*
>
> *-Student test response*

Thus, a recurring theme found in student answers to Question 2 related to them reviewing other student's work in order to better understand their own roles and the tasks which lay before them when it came time to draft their company's business plan. Another theme that arose from analysis of the focus group transcript was the notion that the CKB removed the perceived glamor of such positions and replaced idealized perceptions with realistic expectations and foundational knowledge of the actual job description.

In the student focus group ($N = 8$), participants were asked a similar question, regarding the extent to which the CKB had augmented their understanding of their roles with regard to the business plan. This question evoked strong positive responses, with focus group members stating without any exception that reading the work of others had enhanced their understanding of their roles and reinforced that they were on the right track with their own research. They also noted that reading articles other students had posted enhanced their understanding of their role, that different perspectives added to their own, and that conducting research with real- life examples had enabled connections between theory and more practical applications.

Inquiry-based learning is a bedrock principle in the learning community approach, whether FCL or KCI. The first iteration provided instances of individual inquiry but this was greatly expanded with the addition of the CKB research exercises. The hope was that university students would recognize the benefit of researching deeply into topics, gaining valuable opportunities to critically in the general sense, and more specifically, to hone professionally relevant Twenty-first century skills. After coding Question 2 of the test, certain words appeared and re-appeared that were associated students' research exercises for the CKB. These words were *pushed, forced* and *helped*. Twenty (20%) percent of respondents ($N = 233$) made reference to how the CKB research exercises forced/pushed/helped them do research they otherwise would not have participated in. There were 36 individual student mentions of being "forced" to go beyond their academic comfort level.

Another recurring theme was acknowledgement that the parts helped create the whole, that individual research when shared with the group, created a collaborative environment where information was shared, aiding in the completion of the business plan. Students became aware of their own learning, describing the metacognitive nature of the curriculum. In the student focus group, a participant made this statement, with which all others in the group readily agreed:

> *I just thought that the tutorials served as like a perfect bridge between the CKB and the business plan, because you researched what you had to work on for your CKB and then you would have to translate that into your collaboration with your group members within your own role, in a very explicit way, which we would end up using for your business plan. So - the CKB was a perfect bridge that tied the two elements of the course together.*
>
> *-Student in a Focus Group*

As to be expected with 48 groups, not all groups achieved the same high level of collaboration. Some groups exhibited the all-too-familiar characteristic where some members were dedicated, conscientious and willing to carry more of the load than other members. Some groups thus fell short of achieving the goal of collective cognitive responsibility. One student talked about hoping each week that her group members would begin to work collaboratively, but was disappointed each time. Another described the common situation where two or three members picked up the slack, students who saw the value in the course and were high achievers. The researcher/practioner estimates that approximately one group in five were underperforming in this fashion.

DISCUSSION

The above sections describe an unprecedented opportunity to enact a complex learning community curricular design in a large undergraduate class and to run two iterations of the course to allow for re-design and reenactment of the model. Despite the impossibility of having the same cohort as subjects for both iterations, the specialized nature of the undergraduate programs from which students were enrolled ensured a certain amount of student coherence between iterations.

In response to the first research question (How can the FCL model be applied as a learning community approach within a large undergraduate course?), the goal of the first iteration was to successfully design and enact an FCL modeled curriculum, modified to enable the curriculum to be delivered at a university level course with a large body of students. The design was guided by five assumptions, namely, that (i) the middle-school design of previous studies would have to be modified to facilitate the learning capabilities and expectations of university students, (ii) the activities, particularly the consequential task, would need to be grounded in real-life activities in order to be perceived as relevant and maintain student interest throughout the term, (iii) the consequential task should be directly connected to the learning domain in order to provide a basis for analyzing the effectiveness of the learning undertaken by students, (iv) consideration would have to be given to the limitations imposed by the physical context of a large movie theater converted into a lecture hall room, and, (v) the model would have to be adapted to run in 12 different tutorial sections with the researcher present in only one, relying on TAs to lead students through FCL-devised scripting.

This course covers a spectrum of business concepts as previously enumerated. The design thus had to be flexible enough to incorporate these topics and still remain true to the FCL model. This was accomplished by making all topics tethered to the consequential task (i.e., the business plan), a culminating inquiry project that was indexed to the full space of content. The topics were the tributaries and the business plan was the river. By the time students reached the mouth, to carry the metaphor forward, the river of knowledge was at its deepest and widest.

In response to our second research question, (What are the limitations of the model, and what adaptations can help respond to those limitations?), the addition of the Collaborative Knowledge Base provided the linkage between individual inquiry, lectures, and the collaborative work being conducted in tutorials. Instead of 12 discreet learning community pockets, the CKB provided an umbrella of shared individual research that at times the entire population viewed in lectures, and in others where such research was shared amongst group members and the other groups in the respective tutorials.

The introduction of the CKB into the curriculum involved a major overhaul of the curricular design, allowing for more individual inquiry, more sharing of research in lectures and providing a permanent repository of student research that all students and groups had access to. With the introduction of the CKB, lectures shed much of their previous instructivist flavor by replacing instruction and guest time with the display of CKB research onscreen, providing the opportunity for class-wide discussion and analysis on a myriad of topics. This promoted a learning community ethos in the lectures, an element missing in the first iteration.

Student participation in the CKB added a research element missing from the first iteration and provided students with an opportunity to conduct deep inquiry related to their particular office. This had a cumulative effect as research traveled from the CKB to lecture to the tutorials where groups reviewed each other's work, conducted reciprocal teaching, and developed skills necessary to collectively build a business plan.

The researchers made other significant findings related to active learning at this academic level and with young adults as subjects. They observed that the more the course subject matter and activities resemble real-life experiences, the more likely university students will perceive the course, the activities, and the professor as being credible. And if a learning community model adopts an approach of presenting students with real-life questions and provides exercises that produce tangible, authentic artifacts by way of active learning (i.e., if students detect a direct link from the learning community activity and getting a job), the course will more likely be accepted as having intrinsic value. Students must see concurrent value in their research and the artifacts they create or these activities will be relegated to an exercise that must be completed for a grade and nothing more. This real-life aspect of the curriculum should be further infused into future iterations.

The fundamental ambition of developing this augmented FCL model was—revisiting the acronym—to "foster a community of learners" in a large undergraduate lecture class. Students, by way of individual inquiry and collaborative knowledge building, worked together to create a real-life artifact, the business plan. Students also demonstrated their acquired knowledge in a curriculum that spanned a wide range of topics. Learning outcomes were achieved and overall the course was well-received. But how well was a learning community really established, beyond the tutorial sections, and what impact did this have on learning? The addition of the CKB and the ensuing discussions generated by the students' research within the lectures is a good start. But despite some evidence that a sense of community was established class-wide, it is impossible to state what effect this may have had on learning. Thus, it is important to reserve any claims about the effect of a learning community approach on actual learning by these subjects—if only to spur future research in this area and guide future iterations that aim to further reinforce this sense of community.

Finally, this study greatly benefited from having one of the researchers in the classroom, which allowed for a first-hand, unmediated experience regarding the execution of the design in both iterations. We therefore conclude with these final observations from one of the authors.

A professor conducting practitioner inquiry has a different role than in a co-design with a teacher. There is less contact with the students (once a week) and given the number of students in a large lecture, less ability to script one's role. Lectures are fluid and dynamic, therefore the researcher must balance his or her lecture between engagement and instruction, allowing for unintended or unscripted variances to occur.

Secondly, it is vitally important that the instructor thoroughly communicate to students the metacognitive aspects of learning communities and their responsibilities within it in order to achieve "buy-in" from the students. This can be achieved by informing student not only of how the course will run, but how this approach will provide students with critical thinking skills, collaborative learning, and learning community.

Thirdly, the instructor of a large lecture with breakout tutorials must accept that delegation is part of the design. It is therefore vitally important to select TAs who understand the intent of the design, can enact it as the instructor's proxy in their tutorial sections, and can observe the enactment critically so as to provide relevant input during the TA focus group. This is a different scenario than that of the researcher who creates a co-design with a teacher then stands back to let the enactment occur. The temptation of the practitioner-researcher who does not have complete supervision of the enactment is to hover over the TAs, visit their tutorials, and affect the scripting merely by their presence. This impulse must be resisted and satisfied by trust in the design. In other words, the instructor must be mindful of the Hawthorne effect.

Finally, it is important to leave the "researcher" outside of the classroom. In the role of instructor, it is important to be guided solely by the lesson plan, by the curricular design, and not allow one's researcher mind to influence what is happening in the classroom.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the Research Ethics Boards of the University of Toronto and Ryerson University with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

Bannan-Ritland, B. (2003). The role of design in research: the integrative learning design framework. *Educ. Res.* 32, 21–24. doi: 10.3102/0013189X032001021

Bereiter, C. (2002). Design research for sustained innovation. *Cognit. Stud. Bull. Jap. Cognit. Sci. Soc.* 3, 321–327. doi: 10.11225/jcss.9.321

Bielaczyc, K., and Collins, A. (2006). "Fostering knowledge-creating communities," in *Collaborative Learning, Reasoning, and Technology,* eds A. M. O'Donnell, C. E. Hmelo-Silver, and G. Erkens (Mahwah, NJ: Erlbaum), 37–60.

Bielaczyc, K., and Collins, A. (2007). "Design research: foundational perspectives, critical tensions, and arenas for action," in *Children's Learning in Laboratory and Classroom Contexts: Essays in Honor of Ann Brown,* eds J. Campione, K. Metz, and A. S. Palincsar (New York, NY: Routledge), 89–111.

Bloom, K. C., and Shuell, T. J. (1981). Effects of massed and distributed practice on the learning and retention of second-language vocabulary. *J. Educ. Res.* 74, 245–248. doi: 10.1080/00220671.1981.10885317

Bonwell, C., and Eison, J. (1991). *Active Learning: Creating Excitement in the Classroom*. ASHE-ERIC Higher Education Report No. 1, Washington, DC: Jossey-Bass.

Bransford, J. D., Brown, A. L., and Cocking, R. R. (2000). *How People Learn: Brain, Mind, Experience, and School.* Washington, DC: National Academy Press.

Brown, A. L. (1992). Design experiments: theoretical and methodological challenges in creating complex interventions in classroom settings. *J. Learn. Sci.* 2, 141–178. doi: 10.1207/s15327809jls0202_2

Brown, A. L. (1994). The advancement of learning. *Educ. Res.* 23, 4–12. doi: 10.3102/0013189X023008004

Brown, A. L., and Campione, J. C. (1996). "Psychological theory and the design of innovative learning environments: on procedures, principles, and systems," in *Innovations in Learning: New Environments for Education,* eds L. Schauble, and R. Glaser (Mahway, NJ: Erlbaum), 289–325.

Cavanagh, S. (1997). Content analysis: concepts, methods and applications. *Nurse Res.* 4, 5–16. doi: 10.7748/nr1997.04.4.3.5.c5869

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: a practical guide. *J. Learn. Sci.* 6, 271–315. doi: 10.1207/s15327809jls0603_1

Erzberger, C., and Kelle, U. (2003). "Making inferences in mixed methods: the rules of integration," in *Handbook of Mixed Methods in Social & Behavioral Research,* eds A. Tashakkori, and C. Teddlie (Thousand Oaks, CA: Sage), 457–488.

Felder, R. M., Celanese, H., and Brent, R. (2009). Active learning: an introduction. *ASQ High. Educ. Brief* 2, 1–5.

McKenney, S., and Reeves, T. C. (2013). Systematic review of design-based research progress: is a little knowledge a dangerous thing? *Educ. Res.* 42, 97–100. doi: 10.3102/0013189X12463781

Onwuegbuzie, A. J., Dickinson, W. B., Leech, N. L., and Zoran, A. G. (2009). A qualitative framework for collecting and analyzing data in focus group research. *Int. J. Qual. Methods* 8, 1–21. doi: 10.1177/160940690900800301

Prince, M. (2004). Does active learning work? A review of the research. *J. Eng. Educ.* 93, 223–232. doi: 10.1002/j.2168-9830.2004.tb00809.x

Scardamalia, M., and Bereiter, C. (2010). A brief history of knowledge building. *Can. J. Learn. Techn.* 36, 1–16. doi: 10.21432/T2859M

Shavelson, R. J., Philips, D. C., Towne, L., and Feuer, M. J. (2003). On the science of education design studies. *Educ. Res.* 32, 25–28. doi: 10.3102/0013189X032001025

Slotta, J. D., and Najafi, H. (2010). "Knowledge communities in the classroom," in *International Encyclopedia of Education Vol. 8*, eds P. Peterson, E. Baker, and B. McGaw (Oxford: Elsevier), 189–196.

Slotta, J. D., and Najafi, H. (2013). "Supporting collaborative knowledge construction with Web 2.0 technologies," in *Emerging Technologies for the Classroom,* (New York, NY: Springer), 93–112.

Slotta, J. D., and Peters, V. L. (2008). "A blended model for knowledge communities: embedding scaffolded inquiry. International perspectives in the learning sciences: crea8ing a learning world,"in *Proceedings of the Eighth International Conference for the Learning Sciences - ICLS 2008* (Utrecht: International Society of the Learning Sciences, Inc), 343–350.

Trow, M. (1957). Comment on participant observation and interviewing: a comparison. *Hum. Organ.* 16, 33–35. doi: 10.17730/humo.16.3.cx277m417x00w647

Vygotsky, L. S. (1978). *Mind in society: The Development of Higher Psychological Processes*. eds and Transl. by M. Cole, V. John-Steiner, S. Scribner, and E. Souberman (Cambridge, MA: Harvard University Press).

Conflict of Interest Statement: The handling editor is currently co-organizing a Research Topic with one of the authors, JS, and confirms the absence of any other collaboration.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Education

ORIGINAL RESEARCH
published: 05 July 2018
doi: 10.3389/feduc.2018.00052

CKBiology: An Active Learning Curriculum Design for Secondary Biology

Alisa Acosta[1]* *and James D. Slotta*[2]

[1] Department of Curriculum Teaching and Learning, Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, Canada, [2] Lynch School of Education, Boston College, Chestnut Hill, MA, United States

This research paper presents the design of an active learning curriculum and corresponding software environment called *CKBiology*, reporting on its implementation in two sections of a Grade 12 Biology course across three design cycles. Guided by a theoretical framework called Knowledge Community and Inquiry (KCI), we employed a design-based research methodology in which we worked closely with a high school biology teacher and team of technology developers to co-design, build, test, implement, and revise this curriculum within a blended learning context. We first present the results of a needs assessment and baseline analysis in which we identify the design constraints and challenges associated with infusing a "traditional" Grade 12 Biology course with a KCI curriculum. Next, we present the design narrative for CKBiology in which we respond to these constraints and challenges, detailing the activity sequences, pedagogical aspects, and technology elements used across three design iterations. Finally, we provide a qualitative analysis of student and teacher perspectives on aspects of the design, including activity elements as well as the CKBiology interface. Findings from this analysis are synthesized into design principles which may serve the wider community of active learning researchers and practitioners.

Keywords: computer-supported collaborative learning, active learning, inquiry-based learning, learning communities, science education, K-12 education

Edited by:
Tom Crick,
Swansea University, United Kingdom

Reviewed by:
Emilia Matallana,
Universitat de València, Spain
Maria Antonietta Impedovo,
Aix-Marseille Université, France

****Correspondence:***
Alisa Acosta
alisa.acosta@utoronto.ca

Specialty section:
This article was submitted to Digital Education, a section of the journal Frontiers in Education

Received: *28 April 2018*
Accepted: *15 June 2018*
Published: *05 July 2018*

Citation:
Acosta A and Slotta JD (2018) CKBiology: An Active Learning Curriculum Design for Secondary Biology. Front. Educ. 3:52. doi: 10.3389/feduc.2018.00052

INTRODUCTION

In today's era of "alternative facts," the importance of a scientifically literate citizenry cannot be overstated. Combatting complex global problems such as climate change, new viral epidemics, economic disparity and nuclear threats will require a sustained collaborative effort among knowledgeable scientists, engineers, politicians, and a scientifically literate public. Thus, producing graduates who are prepared for occupations in science, technology, engineering, and mathematics (STEM) has become a global priority (OECD., 2012). However, dropout rates for STEM programs at the post-secondary level remain high. For example, in the United States 48% of bachelor's degree students and 69% of associate's degree students who enter STEM programs never complete them, with approximately half of these students switching to a non-STEM major, and the other half dropping out before earning a degree or certificate (U.S. Department of Education, 2013). In Ontario, Canada (the context of the present study), Computer Science and Physical Sciences are among the top three undergraduate programs with the lowest graduation rates, with 38.3 and 33.9% of students failing to complete these degrees, respectively (MAESD., 2016).

One factor influencing students' performance in STEM courses is related to the instructional strategies that are employed. As Kober (2015) describes, "A single course with poorly designed instruction or curriculum can stop a student who was considering a science or engineering major in her tracks" (p. xi). Bloom (1984) has observed that nearly any means of instruction is superior to lecture, and yet this is the approach that many undergraduate STEM courses maintain. In order to learn science and engineering well, students need to be able to understand and apply the practices of the discipline, develop skills in problem solving, communication, and collaboration, and critically evaluate new information in the field (Kober, 2015). Scholars and educational leaders have called for new pedagogical approaches that better prepare students to face the complex challenges of an increasingly globalized, technology-driven, knowledge economy (Tapscott and Williams, 2012; Pellegrino and Hilton, 2013; OECD., 2016).

In response to such calls for change, science educators have explored new modes of learning and instruction such as "flipped classrooms," wherein students spend their homework time watching video lectures and reading texts so that classroom time can be devoted to more active forms of collaborative group work, inquiry and problem solving (Bens, 2005; DeLozier and Rhodes, 2017). Referred to broadly as "Active Learning" (AL), these approaches have become increasingly prominent, resulting in professional societies (e.g., SALTISE.ca) and university-based centers to support the design of AL courses (e.g., Charles et al., 2011),. Several studies have measured the benefits of AL (Dori and Belcher, 2005; Code et al., 2014), and evidence has begun to accumulate that AL methods achieve better educational outcomes than lecture-based approaches (Freeman et al., 2014; Waldrop, 2015).

However, despite these indications of its efficacy, AL remains largely ill-specified in its formulation (Ruiz-Primo et al., 2011; Brownell et al., 2013). For example, while particular group strategies are often invoked (e.g., cooperative learning, collaborative projects, or jigsaw groups) very little is known about the learning processes that occur within such methods, the materials or assessments they require, nor the role of the instructor (Henderson and Dancy, 2007). What makes a collaborative group activity effective? When should it be used within the curriculum? How will students collaborate, and to what end? How should their progress, process, or products be assessed? Simply naming or broadly describing an AL approach does not provide sufficient information about the content, structure or sequencing of activities or interactions (amongst students, materials, instructors, and the classroom environment) that it entails. Additionally, most forms of AL employ some form of technology, leveraging the valuable resources of student laptops, mobile phones, Smart boards, and a wide range of software applications and classroom management tools. These technologies can offer new opportunities for teachers to increase the sophistication of interactions and ideas in their courses, however their integration within the classroom adds a layer of complexity to the curriculum, making it challenging for teachers to enact or "orchestrate" any given design. Thus, technology can offer both a means of achieving active learning as well as a barrier to implementing it.

To advance the study of AL, this paper offers a detailed account of a full-course AL curriculum, and a custom-designed software environment called *CKBiology*. We describe a design-based research project implemented in two sections of a Grade 12 Biology course that comprised three iterative design cycles over the course of one academic year. We worked closely with a high school biology teacher and team of technology developers to co-design, build, test, and enact this curriculum to address the following two research questions:

1. What are the design opportunities and constraints associated with infusing a traditional Grade 12 Biology course with active learning designs?
2. What forms of active learning can address those constraints and challenges, and what technology elements are needed to support them?

THEORETICAL FOUNDATIONS

Active Learning (AL) is rooted in the theoretical perspective of social constructivism, which emphasizes the importance of social interactions, cultural tools and activities in shaping the learning and development of an individual (Woolfolk et al., 2009). Here, the learner is seen as playing an active role in constructing her own knowledge, building understandings, and making sense of information. This contrasts with instructionist theories of learning in which the learner is seen as a recipient of knowledge transmitted from an external authoritative source. Examples of AL include solving ill-structured problems, negotiating diverse ideas and perspectives, engaging in inquiry and critical thinking, and developing a sense of responsibility for one's learning. Ruiz-Primo et al. (2011) characterize AL using the following four attributes: (1) conceptually-oriented tasks, (2) collaborative learning activities, (3) technology elements, and (4) inquiry-based projects.

One topic of great relevance to AL, particularly in regard to the role of technology and classroom learning environments, is that of scripting and orchestration (Dillenbourg and Jermann, 2007; Kollar et al., 2007). Similar to a theatrical script, which specifies all aspects of a play (i.e., stage, props, lines, actions, and behaviors), a *pedagogical script* explicates a learning design in terms of the participants, roles, goals, groups, activities, materials, and logical conditions or determinants of activity boundaries (Fischer et al., 2013). Like its theatrical counterpart, a pedagogical script is only an abstract or idealized description until it is actually performed. *Orchestration* refers to the enactment of the script, binding it to the local context of learners, classrooms, curriculum and instructor, and giving it concrete form in terms of materials, activities and interactions amongst participants (Tchounikine, 2013). Pedagogical scripts are orchestrated in the classroom, online or across contexts (i.e., home, school, or mobile), with the "orchestrational load" shared by (1) the instructor, who can tell students what to do, pause activities to hold short discussions, or advance the lesson from one point in the script to another; (2) the materials, including text or other media, instructions, or interactive Web sites; (3) the technology environment, such as online portals, discussion forums, note sharing, wikis or Google Docs; and (4) the physical learning environment such as the classroom configuration, furniture, walls, or lighting (Slotta, 2010).

Studies have shown that AL can have a variety of positive effects on teaching and learning, including improvements in student affect and motivation (Dori and Belcher, 2005), student engagement (Fisher, 2010), group interactions (Mercier et al., 2016), shared responsibility for learning (Baepler and Walker, 2014), and student learning outcomes (Brooks, 2011). In the largest and most comprehensive meta-analysis to date, Freeman et al. (2014) analyzed 225 studies that reported data on student performance in undergraduate science, technology, engineering, and mathematics (STEM) courses under traditional lecturing vs. active learning approaches. Taking into account factors such as class size, discipline, student/instructor quality, and methodological rigor within the included studies, their findings indicated that average student performance on examinations and concept inventories increased by 0.47 SDs (i.e., around 6%) in AL sections, and that students in classes with traditional lectures were 1.5 times more likely to fail than were students in classes with AL (Freeman et al., 2014).

Classrooms as Learning Communities

One promising approach to the design of AL curricula is to consider the classroom as a learning community. For many years, theories on collaborative learning tended to focus on how participating in a group would affect an individual's performance (Stahl, 2015), however in the late 1980s two programs of research emerged that gave focus to groups of learners at the community level: Fostering Communities of Learners (FCL; Brown and Campione, 1994) and Knowledge Building (KB; Scardamalia et al., 1989). Both of these research programs upheld the notion that the activities occurring in school classrooms should mirror those of authentic research communities, incorporating aspects of collective epistemology and community-level knowledge advancement (Brown, 1994; Scardamalia and Bereiter, 2006). The theoretical perspectives of FCL and KB are distinct with respect to the objectives of the community, the centrality of student-generated ideas, and the level of emphasis placed on prescribed learning goals and activity structures (Scardamalia and Bereiter, 2007; Carvalho, 2017). However, they share a commitment to helping students and teachers identify as a coherent learning community, the sharing of information and dissemination of knowledge and practices (Slotta and Najafi, 2010).

The learning community approach has been defined as "a culture of learning in which everyone is involved in a collective effort of understanding" (Bielaczyc and Collins, 1999, p. 2). Students bring diverse interests and expertise to the classroom and the teacher helps them to work collectively to advance knowledge, with all individual members benefiting along the way. However, scholars have noted that it is challenging for teachers or researchers to coordinate such an approach (Kling and Courtright, 2003; van Aalst and Chan, 2007). Slotta (2014) articulated four key challenges to this approach: (1) to establish an epistemological context such that each student understands the collective nature of the curriculum; (2) to ensure that community knowledge is accessible as a resource during student activities; (3) to ensure that scaffolded inquiry activities advance the community's progress as well as that of all individual learners; and (4) to foster productive discourse that helps individual students and the community to progress.

In response to the challenges of constructing effective learning communities, *the Knowledge Community and Inquiry (KCI)* model was developed to guide the design of collective inquiry curricula that integrate whole class, small group and individual activities (Slotta and Peters, 2008; Slotta and Najafi, 2013). KCI provides structural requirements and design principles that allow (1) an epistemological orientation to help students understand the nature of science and learning communities, (2) a knowledge base that is indexed to the targeted science domain, (3) an inquiry script that specifies collective, collaborative and individual activities in which students construct the knowledge base and then use it as a resource for subsequent inquiry, and (4) student outcomes that allow assessment of progress on targeted learning goals (see **Figure 1**). The model guides the design of activity sequences including individual, group (e.g., jigsaw) and whole-class activities (e.g., brainstorm, resource collecting), ensuring that all students progress on the learning goals.

To date, KCI curriculum designs have been enacted in elementary school, secondary school, and higher educational contexts. In elementary schools, work has included units in astronomy (Cober et al., 2013; Fong and Slotta, 2015), and ecology (Cober et al., 2013, 2015a). In secondary schools, KCI units have been designed on the topics of human disease (Peters and Slotta, 2010), climate change (Slotta and Najafi, 2013), evolution (Lui and Slotta, 2014), forces and motion (Tissenbaum et al., 2012), and literary studies (Carvalho and Hall, 2016). Recent work in secondary school contexts has extended beyond single curricular units to entail full course designs, including courses in Grade 12 Health Science (Serevetas, 2017) as well as the current work in Grade 12 Biology (Slotta and Acosta, 2017). Similarly, KCI research in higher educational contexts has included full course designs in pre-service teacher education (Slotta and Najafi, 2013), business and media (Ehrlick and Slotta, 2017), as well as a Massive Open Online Course for in-service teachers (Håklev and Slotta, 2017).

METHODOLOGY

Design-Based Research

This project employed a design-based research (DBR) methodology to support the creation and development of innovative learning environments through the parallel processes of design, evaluation, and theory-building (Brown, 1992; Collins, 1992; Edelson, 2002). DBR emerged in the early 1990s in response to the experienced limitations of traditional psychological research methods, which required controlled experimentation and regarded cognition as something that "takes place inside the learner and only inside the learner"

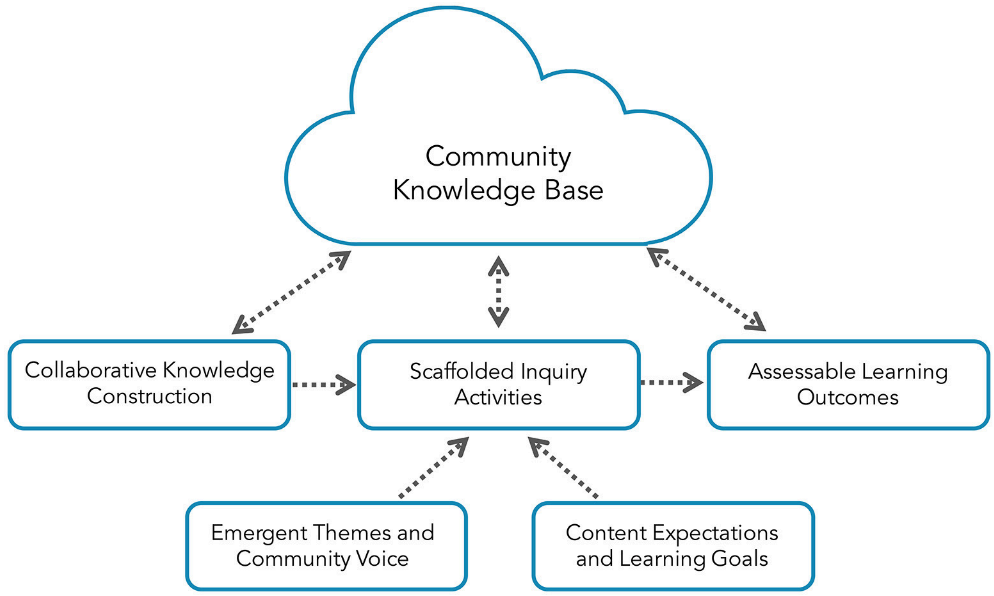

FIGURE 1 | The Knowledge Community and Inquiry (KCI) model.

(Simon, 2001, p. 210). In contrast, DBR activities are situated in *naturalistic* contexts and focus on understanding the messiness of real-world practice (Barab and Squire, 2004; Bell, 2004). Within such complex environments, it would be difficult—if not impossible—to test the causal impact of specific independent variables on specific dependent variables using experimental designs (Barab, 2014). Consequently, DBR is not concerned with so-called "learning outcomes," but rather with the design of innovations to transform "existing situations into preferred ones" (Stahl, 2015, p. 15). In this sense, DBR draws from an engineering ethos, wherein success is seldom defined by the ability to provide theoretical accounts of how the world operates, but rather by the development of solutions to problems that satisfy existing conditions and meet the stated design goals within prevailing constraints (Nathan and Sawyer, 2014).

DBR activities are inherently *iterative*, involving cycles of design, enactment, detailed study, and revision (Bell et al., 2004). What sets DBR apart from other forms of educational research is its commitment to the development of sustained *innovations* in education (Bereiter, 2002). Beyond merely understanding the usability or feasibility of new educational technologies, DBR researchers seek to understand *how* these technologies can be productively embedded into educational systems (e.g., curriculum designs, activity structures, pedagogical practices; Bell et al., 2004) as well as the relative *improvability* of these designs within such systems (Bereiter, 2002).

Co-design

The effectiveness of any research that is situated within a real classroom context is critically dependent upon the classroom teacher's understanding and enactment of the designed approaches and materials (Slotta and Peters, 2008). Studies on the adoption of educational innovations have shown that the level and nature of adoption is strongly influenced by teachers' interpretations of their classroom ecologies, including how they perceive the designs to align with their goals, teaching strategies, and learning expectations (Blumenfeld et al., 2000; Means et al., 2001; Roschelle et al., 2006). Furthermore, practitioners who adopt research-based approaches must be receptive to innovations and willing to experiment with unproven methods (Bereiter, 2002).

As such, researchers in the learning sciences have developed a collaborative approach to the design of educational innovations that are deeply situated within the context of real-world classrooms. In contrast to top-down approaches to educational reform, in which teachers are simply provided with an approach that they are expected to adopt, the co-design method engages teachers as active participants in the design process, positioning them as professional contributors to an interdisciplinary co-design team (Collins, 1992). Roschelle et al. (2006) define co-design as "a highly-facilitated, team-based process in which teachers, researchers, and developers work together in defined roles to design an educational innovation, realize the design in one or more prototypes, and evaluate each prototype's significance for addressing a concrete educational need" (p. 606).

The co-design approach offers several benefits, including providing teachers with a high level of ownership and agency over the designed innovation (Roschelle et al., 2006). Because teachers remain actively involved throughout the entire design process, they not only develop a strong understanding of the underlying research but also firmly believe in the curricular materials that are produced (Cober et al., 2015b). Consequently, co-design has the potential to transform teachers into advocates for innovation within their school districts (Penuel et al., 2007).

Participants and Sampling

The co-design team in this project included five members: One Grade 12 Biology teacher, two technology developers, and two researchers. A purposeful sampling approach was used to select the teacher participant, based upon her prior experience in KCI research as well as her availability to design and implement a KCI curriculum during the 2016–2017 academic year. This teacher held a PhD in biological sciences and has been teaching at our study school since 2010. Student participants consisted of two sections of a Grade 12 Biology course ($n = 29$), both taught by the same co-design teacher. The student participants were an incidental sample, in that they happened to be those who were assigned to the classes of our co-design teacher. Student participants were high-achieving and culturally diverse, reflecting the overall population of the school.

Ethics Protocol

This study was carried out in accordance with the Canadian Tri-Council Policy Statement on Ethical Conduct for Research Involving Humans. The student participants in this study were between 16 and 18 years of age, however the risk to the participants was low as they were simply participating in classroom activities that were co-designed and led by their teacher. The ethics protocol for this study was approved by the Social Sciences, Humanities, and Education Research Ethics Board at the University of Toronto. Before the research began, both classes were given an orientation session in which the general purpose of the study was explained and a letter of information was provided to all participants and their parents/guardians. Additionally, a consent form was provided to students and their parents/guardians requesting permission for them to participate in video recorded and/or photographed classroom sessions. For collaborative activities, only groups for which *all* members returned signed consent forms were recorded and/or photographed. All subjects, as well as their parents/guardians, gave written informed consent to participate in the study in accordance with the Declaration of Helsinki.

Research Setting

This research was conducted at a university laboratory school in a large urban area. Activities took place within three settings: (1) At home (online) using the CKBiology platform, (2) in a traditional science classroom with a "bring your own device" (BYOD) policy, and (3) in a specially-designed AL Classroom, constructed by the school with the explicit aim of fostering productive collaborations between students. The AL Classroom featured six large multi-touch displays positioned around the perimeter of the room, racks of portable white-boards and markers, and flexible furniture (i.e., on casters) that enabled students to be grouped according to a variety of configurations.

CKBiology Technology Environment

In order to support a KCI approach throughout this course, we developed a custom technology environment called *CKBiology*, adapting the more general "Common Knowledge" (CK) platform that was designed to support KCI in previous studies (Fong et al., 2015). CKBiology was designed in close collaboration with our co-design teacher and reflects the unique design constraints of her course structure, her students, and her school context. Accordingly, CKBiology is a bespoke technology that was custom tailored to support our KCI script, enabling the teacher to orchestrate our various curricular activities and configurations (e.g., grouping students, distributing materials and activities), providing information at-a-glance to students and teachers about progress within the community, and scaffolding students in specific activities within the various learning contexts.

One important feature of this environment was a layer of intelligence, implemented on the Web server—invisible to any user interface, but supporting the scripting and orchestration conditions of our design. We sought to track the progress of individual students and groups, as well as the community as a whole, providing valuable information that could serve as input into teacher decisions or be automatically processed on our server. For example, the tracking of student activities could be used to provide real-time feedback or displays of progress, which could inform students and teachers alike in their timing, assessment and orchestration of the activities (e.g., by showing progress bars of students, groups, and community). For each iteration of our curriculum, *CKBiology* was adapted to support our specific scripting and orchestration conditions. The software thus, served to implement our designs, as well as to capture the data that could be analyzed, and can be seen as a product or outcome of this design-based research. While this software was developed for research purposes and was not intended to serve as a standalone product, the software repository has been made freely available on GitHub under an open-source MIT license to anyone who wishes to use, copy, expand, or adapt this software for their own purposes.

Sources of Data and Approach to Analysis

In order to enhance the validity of findings throughout this project, data was triangulated from the following sources:

1. *Design documents*, including co-design meeting minutes, lesson planning documents, and software mockups;
2. *Audio and video recordings*, used to document small groups during in-class review sessions;
3. *Researcher field notes*, which provided a thick description of the research context/setting and curriculum enactment, including details surrounding the collaborative processes and interactions that occurred among individual students, groups, and the teacher;
4. *Learning artifacts and data logs*, including text-based notes, images, relationships between terms, review reports, and metadata captured by the CKBiology platform; and
5. *Teacher interviews* conducted at the end of each design cycle.

For each design cycle, findings from each of these data sources were synthesized into design recommendations to be incorporated into subsequent iterations of *CKBiology*. Specifically, we organized all of our enactment data according to the following three categories: (1) Pedagogical challenges, (2) technological challenges, and (3) epistemological challenges. These categories were chosen because they mirrored the overarching design principles of the KCI model (Slotta, 2014).

We then prioritized our findings from each of these categories using an informal scale ranging from "urgent" to "nice to have." Working in consultation with the teacher, we negotiated which of these items we would address in the next design iteration and which items would/could be saved for future iterations.

Limitations

Overall, this research project was fairly context-specific, which makes it difficult to generalize findings to the broader population. In general, DBR addresses issues associated with replicability through the provision of detailed descriptions of the research context as well as an ongoing record of the design's history in the form of a "design narrative" (Cobb et al., 2003; Bell et al., 2004). A good design narrative provides an account of which design elements were intentional or accidental, successful or unsuccessful, explains why certain trade-offs were made, and provides justification as to why particular changes to the design over time were warranted (Bell et al., 2004). A strong design narrative allows others to judge the value of the design contribution and to connect its underlying ideas and findings to new contexts of innovation (Barab, 2014). Additionally, active involvement by the classroom teacher throughout the design process also means that the designs are likely to be enacted faithfully, giving researchers confidence that any measures collected throughout the intervention will truly reflect the underlying theory (Slotta and Peters, 2008).

NEEDS ASSESSMENT

This section addresses our first research question: What are the design opportunities and constraints associated with infusing a traditional Grade 12 Biology course with AL designs?

Co-design Meetings

During the year leading up to our CKBiology implementation, we held a series of co-design meetings with our teacher participant to discuss the opportunities and constraints that existed at the course-level and the school-level which would guide our designs of an AL component for the following year's course. The school in which this work was situated offered full-year courses (as opposed to a semester system), which ran from September to mid-June. Our teacher had two sections of a Grade 12 Biology course for the 2016–2017 academic year, and wanted to separate the theoretical and practical (i.e., lab) portions of the course, such that September to April would be devoted to theory and April to June would be reserved for labs and experiments. In co-design, it is essential that designs accommodate all interests, so we agreed to this approach and suggested that our designs could fit within the earlier (theoretical) portions, readying students for the later lab-based activities. The teacher indicated that the school as a whole was seeking to promote inquiry-based approaches in many of their courses, but that such approaches were particularly challenging to implement in Grade 12 Biology, since it was notoriously content-heavy. To address the heavy content needs, we decided on an approach of developing a KCI component for homework and review activities that would complement the traditional instructional approaches used in class (e.g., lectures and worksheet activities). Our designs would also include a series of end-of-unit "review challenge" activities that would provide students with an opportunity for more creative, inquiry-oriented collaborations in a face-to-face context.

The Grade 12 Biology course was divided into five curricular units, as mandated by the Ontario Ministry of Education: (1) Biochemistry, (2) Metabolic Processes, (3) Molecular Genetics, (4) Homeostasis, and (5) Population Dynamics. Each unit spanned a period of ∼6 weeks, with the exception of Unit 5 (Population Dynamics) which was only 2 weeks in duration. As shown in **Figure 2**, each of these units was treated as one design cycle, allowing us to evaluate and improve our designs from one unit to the next. In this paper, we report on results from the first four units only.

Baseline Observations: Biochemistry Unit

We collected baseline data in the form of lesson plans, researcher field notes, and teacher interviews, during the Biochemistry Unit?the first of five units in the course. We have labeled this "Unit 0" so that the numbering of subsequent units would align with our numbered design iterations (e.g., Unit 1 for Iteration 1). Lessons were taught using a lecture-based format, where PowerPoints were made available to students in advance of each lesson using the Moodle learning management system. Students were also given a paper booklet of handouts to help guide their note-taking for each lesson topic. These booklets were created by two teachers in the biology department, including our co-design teacher. Prior to each lecture, students were instructed to review the PowerPoints and arrive to class with the relevant pages of notes completed. The lectures served to reinforce concepts and provided an opportunity for students to ask questions to the teacher if there was something they did not understand.

In considering this unit as a baseline to inform the designs, our co-design teacher sought to try out the kinds of review activities that students would be performing in our subsequent KCI designs (i.e., a collective, learning community approach), but without the technology supports or structured materials. The purpose of this pilot effort was to inform our subsequent designs. The Unit 0 review activity included two parts, which took place over two class sessions. On the first day, students were given a printed copy of a research article on one of four topics related to biochemistry. Articles were distributed to students by the teacher based on physical proximity, such that students sitting close together received the same article. Students were free to choose their own seats upon entering the classroom, with most choosing to sit next to their friends. Each student was also given a paper handout containing a list of key terms and concepts they had learned throughout the unit. After reading their article independently, students were asked to highlight any terms or concepts from the list that applied to their article. Working in their same-"article groups," consisting of 3–4 members, students negotiated the relevance of the terms and concepts each had selected, and generated a master list of terms with explanations justifying how each was applicable to the article. The master list generated by each group was collected by the teacher at the end of class. Prior to the second review period, the teacher made

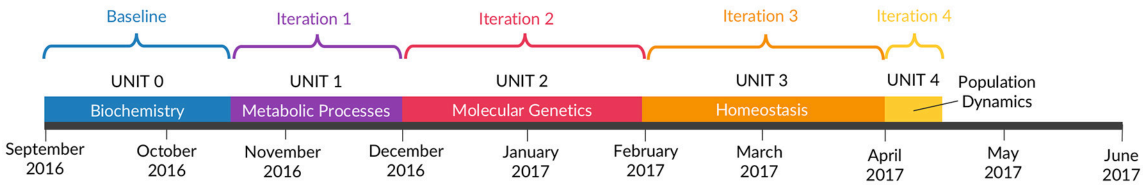

FIGURE 2 | Course timeline showing curricular units and corresponding design iterations.

photocopies of each group's master list such that every group member received his/her own copy.

On the second day of review, students worked in jigsaw groups (i.e., with one representative from each of the previous article groups). Each of these groups was assigned an overarching theme by the teacher (e.g., "matter and molecular interactions," "form and function") and were asked to identify cross-cutting "big ideas" that emerged across all four of the articles with respect to these themes. Groups were also given a paper handout with a series of questions/prompts for each article, a space on which to record their big ideas, as well as their master list of terms from the previous day. The "big ideas" handout was collected at the end of the period and assessed by the teacher.

Findings

The students in both class sections were high-achieving and performance-driven, reflecting the overall population of this school. Throughout the unit, the teacher reported that class time was mostly spent with her talking through the PowerPoints. As she was lecturing, she would assess students' understanding based on factors such as facial expressions as well as the questions that students asked aloud in class. For the review activities, the teacher indicated that there was a good mapping of the terms/vocabulary that students had learned throughout the unit and the terms that were included in the activity handouts. However, our field notes as well as the teacher interview revealed that students were unclear on the purpose of the review activities—and, in particular, how the "big ideas" they were describing would help them perform better on their unit test.

Throughout the review activities, the teacher walked around the room fairly randomly to check up on how students/groups were doing. According to the teacher, *"I was just, like, walking around and checking on people like, 'What are you doing?' 'Show me what you have done.' 'Please do your work.' Stuff like that. And sometimes it worked, and sometimes it didn't work."* Researcher field notes indicated that, while students seemed engaged and on-task in their group discussions, they didn't write very much down on their handouts for submission. According to the teacher, *"they did some work, but it wasn't magnificent work."*

As an outcome of our consultations and baseline observations, we identified the following opportunities and constraints to implementing our AL curriculum design within this Grade 12 Biology course:

Design Opportunities:

1. *Adding a learning community "layer" onto the existing course structure*—As part of their homework activities, there was an opportunity for students to work together to co-create a persistent, shared, community knowledge base which would later serve as a resource for their review activities. Engaging students as a learning community would require an explicit epistemic treatment such that they would view each other as collaborators rather than as independent learners working in parallel and competing for grades.
2. *Supporting real-time formative feedback*—There was an opportunity to support students and the teacher in tracking their progress at various levels of granularity (i.e., as individuals, small groups, and as a whole class community). Providing the teacher with an overview of the progress of the learning community would enable her to make more informed decisions concerning when and where to intervene or provide assistance.
3. *Designing conceptually rich and meaningful "review challenge" activities*—There was an opportunity to design a "consequential task" (Brown and Campione, 1996) that would require students to draw from their community knowledge base in order to perform an engaging inquiry activity.
4. *Active Learning Classroom & BYOD support*—The school had recently completed construction on their own AL Classroom, which was available to be booked for our review challenge activities. Additionally, the school provided IT support for students to bring their own devices to class.

Design Constraints:

1. *Course structure*—Our designs were constrained to fit within the "theoretical" portion of the course only. With the exception of the review challenge activities, our designs would mostly be enacted by students outside of class time (i.e., for homework).
2. *Curriculum expectations*—Our designs had to conform to the content expectations of the Ontario Ministry of Education Grade 12 Biology (University Preparation) course.
3. *Review challenge activities*—Our review challenge designs were constrained to the (theoretical) material that students had already learned; there were limited opportunities to engage students in projects or labs in which they would learn or research a new topic.
4. *CKBiology activities could not be for marks*—To comply with our ethics protocol, students could not be directly evaluated

for the work they completed as part of our design intervention. While this was not seen as a major issue at the outset (given the high performance of these students), the fact that it was a senior year course, together with the extraordinary level of student activities and commitments (e.g., visiting universities) did make this a factor.

DESIGN ITERATIONS: CKBIOLOGY AND ACTIVE LEARNING

In this section, we respond to our second research question: What forms of active learning can address our constraints and challenges, and what technology elements are needed to support them? Given our co-design approach of adding AL designs as a culminating activity for each unit of the course, we report each iteration as a separate sub-section, summarizing what was learned and how it informed subsequent designs. In this way, we describe the complete arc of our design-based research, organized according to the temporal sequence in which it occurred. We close with a summary of the limitations of our study and potential applications of our findings to future work.

Iteration 1: Metabolic Processes Unit

In the Metabolic Processes Unit—hereafter referred to as "Unit 1" —we introduced KCI and the CKBiology platform for the "lessons" portion of the unit only, in part because we required this iteration to inform the full features of CKBiology. Thus, dedicating our efforts toward the "lessons" portion of this unit enabled us to carry forward our design and programming into Unit 2, where we added the review activities.

Design of Unit 1

At the beginning of the unit, we visited both class sections to provide students with an orientation to KCI and CKBiology. After making introductions, we began by discussing the idea of "Science 2.0," explaining how the nature of science is changing and how large, collaborative research projects—facilitated by the social web—are becoming increasingly prevalent. Students were asked to imagine scientists working as collaborators across large distances and scales, rather than as independently isolated individuals working alone in a lab. Next, we introduced our research project and explained to students that they would have an opportunity to experience Science 2.0 as part of their school science activities. Throughout these activities, they would be asked to think of each other as collaborators rather than as independent, parallel learners competing for grades. At this time, students were informed that their participation in this research project would have no direct bearing on their grades, and that in choosing to participate they would be making a valuable contribution to CSCL research. Students were then introduced to the CKBiology platform. We performed a demonstration of the lesson activities and other functionality of CKBiology, which we projected on a display at the front of the room. The orientation session concluded with a question and answer period, at which time students asked questions and offered comments related to CKBiology and the overall research project.

CKBiology activities were completed as part of students' homework and served as a complement to classroom lectures. There were two lesson topics in Unit 1—photosynthesis and cellular respiration—which were taught over five class sessions. As in the previous unit, students were asked to view PowerPoints and complete the appropriate pages of notes/handouts before arriving to class. In class, lectures were held which served to reinforce these concepts and provided students with an opportunity to ask questions and clarify their understandings. Following each lecture, students logged on to CKBiology for homework where they were assigned three different types of tasks. The first type of task was to explain a term or concept related to that day's lesson (see **Figure 3A**). The list of terms associated with a given lesson was established in advance by the co-design team, and the terms were divvied up evenly among students in the class. Students' explanations for these terms were contributed to the community knowledge base in the form of text-based notes with optional images (see **Figure 3B**). On average, students were assigned to explain three terms for each of the two lessons.

The second type of task was to identify relationships between terms or concepts in the knowledge base. Within the CKBiology interface, students were presented with two terms separated by a drop-down list of relationship types (see **Figure 4**). In this case, there was actually a "correct relationship" between each pair of terms, established in advance by the co-design team and programmed into the software. If a student chose the correct relationship, they were free to advance to the next task and a line would appear connecting the two terms in the knowledge base. The relationship would also appear as a sentence within each note involved in the relationship. For example, the sentence "chloroplast *contains* lumen" would appear in both the "chloroplast" note and the "lumen" note. If a student specified an incorrect relationship, a numeric counter would appear above their response indicating the number of attempts they had made at selecting the correct relationship. Since students would not be able to advance until they had chosen the correct relationship, the purpose of the counter was to discourage students from "gaming the system" by clicking through all possible answers without giving thoughtful consideration to each one. On average, students were assigned three or four relationships for each of the two lessons in Unit 1.

The third and final task was to peer review, or "vet," the explanations submitted by other students. Students were presented with an anonymized note followed by the prompt: "Is this explanation complete and correct?" If the student responded "yes," that student's name would be appended to the note along with the statement "This explanation is complete and correct." If the student responded "no," a text box and image uploader would appear beneath the original note, and the student would be asked to add any new ideas and/or corrected information (see **Figure 5**). Any additional information entered by the student would be appended to the original note along with the student's name. Subsequent vetting decisions performed on that note would be appended in the same fashion. On

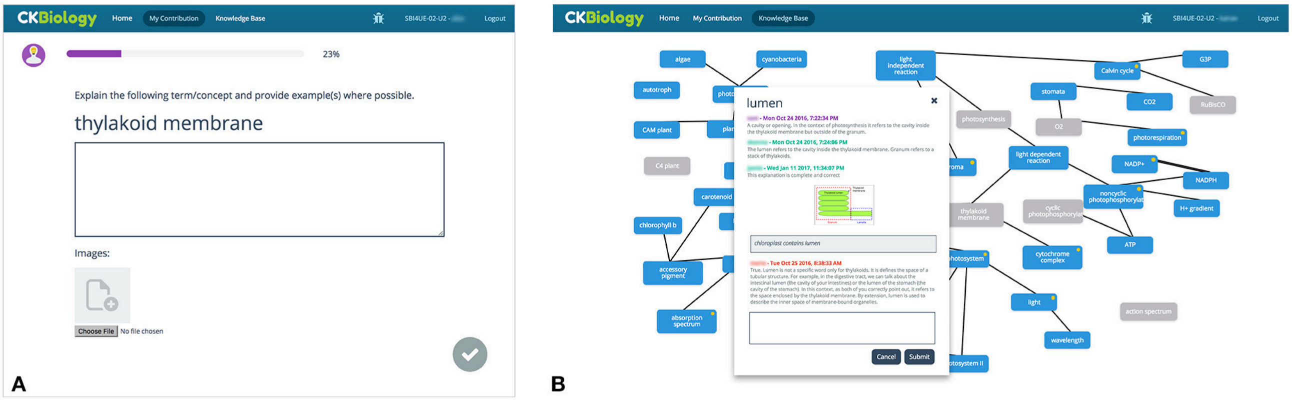

FIGURE 3 | (A) CKBiology term explanation screen; **(B)** CKBiology Knowledge Base view, showing an open note for "lumen." Within each note, the original explanation is shown with a purple heading, vetting is shown with a green heading, and additional comments are shown with a red heading.

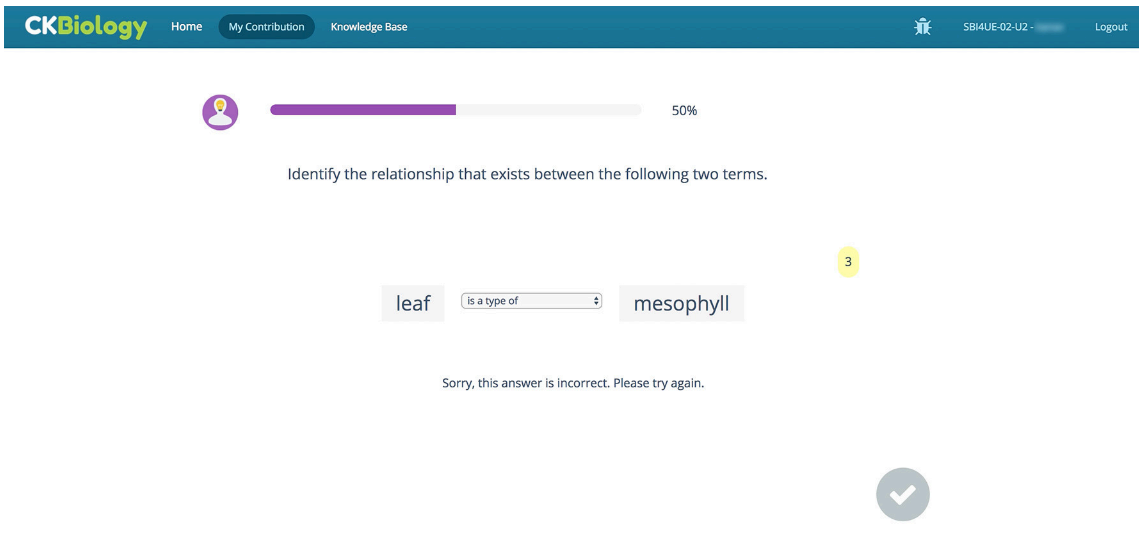

FIGURE 4 | CKBiology relationships screen. A numerical counter (shown in yellow) indicates the number of attempts made at establishing a correct relationship.

average, students were assigned four or five vets per lesson in Unit 1.

Within the knowledge base, a yellow dot was used to identify notes that had been deemed incomplete or incorrect as a result of student vetting. This yellow dot served as a cue to the teacher to take a closer look at these notes and potentially initiate a follow-up discussion to negotiate or improve upon these ideas as a class. As well, students and the teacher had the ability to comment upon any note within the knowledge base. Comments were appended to the note along with the commenter's name, and appeared below the rest of the note content.

As students progressed through each of their assigned tasks, a progress bar at the top of their screen would indicate the percentage of work they had completed and the percentage of work that remained. Additionally, on their home screen (which showed information about all lessons and units) students could see their individual progress bar for each lesson as well as an overall progress bar for the whole learning community (see **Figure 6**). If a student saw that the progress level of the community was below 100%, they could choose to go "above-and-beyond" their own assigned tasks and make additional contributions to the knowledge base to boost community-level progress. Anyone going beyond their assigned tasks earned a gold star icon and additional progress points for that lesson. These additional contributions typically took the form of extra vetting tasks and did not detract from the assigned work of other students. Thus, no single student could dominate the knowledge base by populating an inordinate number of terms and relationships, and every student was still accountable for making their fair share of contributions.

The product of these homework activities was a shared community knowledge base that aggregated students'

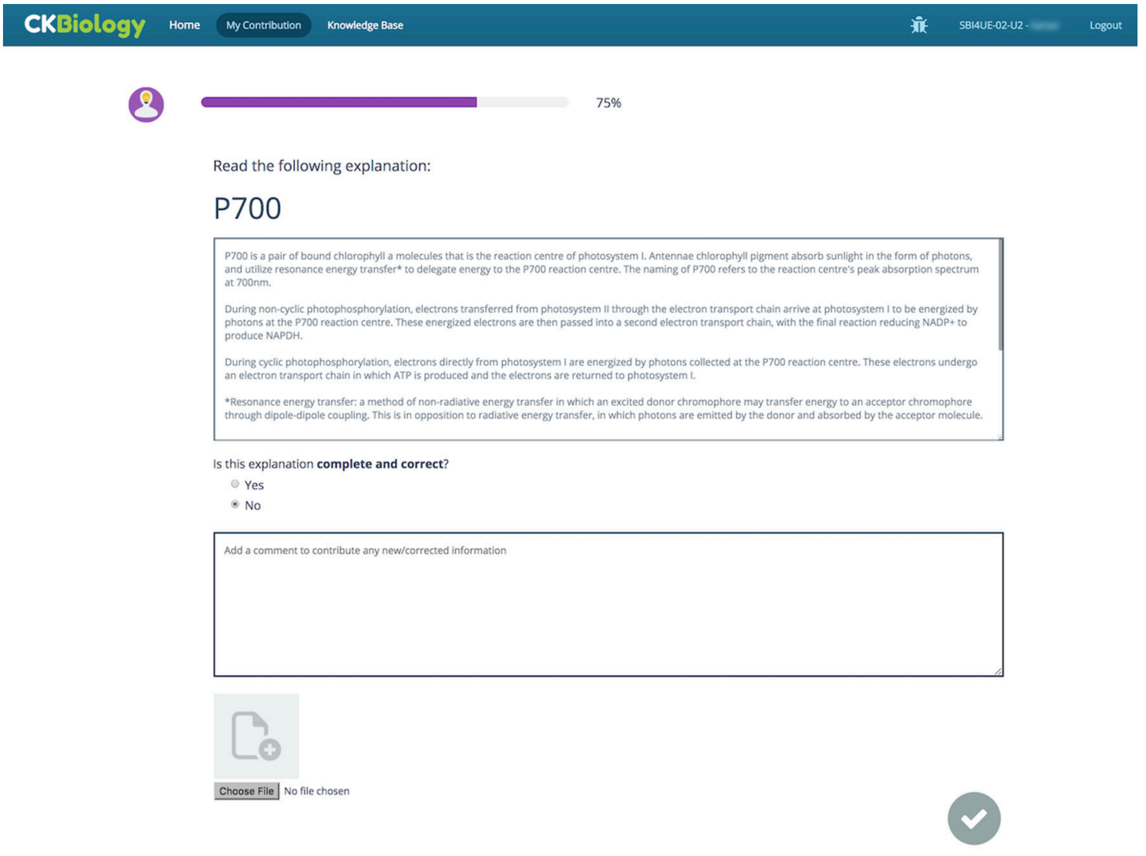

FIGURE 5 | CKBiology vetting screen.

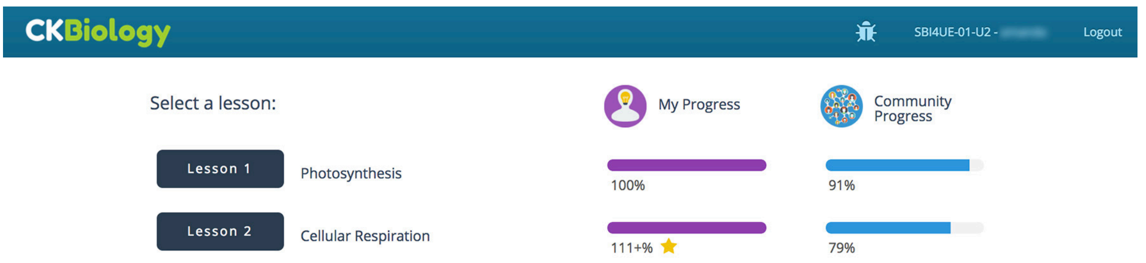

FIGURE 6 | CKBiology student home screen, showing individual progress bars (purple) and community-level progress bars (blue). Students who contribute more than the minimum requirement earn a gold star icon, and additional progress points above 100%.

contributions in the form of a concept map for each lesson. Following the homework activities in CKBiology, the teacher could look at the knowledge base to assess students' understanding, and initiate a follow-up discussion in class if warranted. The teacher was also provided with a dashboard that provided an overview of students' progress for each lesson. In cases where a student was not contributing their fair share to the knowledge base, the teacher would consult with the student and try to remedy the situation.

Enactment of Unit 1

Students completed their CKBiology homework on a regular basis throughout Unit 1. The average student progress across all lessons was 93% for both course sections, with many students choosing to go above-and-beyond their own assigned work. At the same time, an average of three students per class section did not make any contributions to the CKBiology knowledge base (i.e., their progress was at 0%) throughout Unit 1. For this design cycle, these

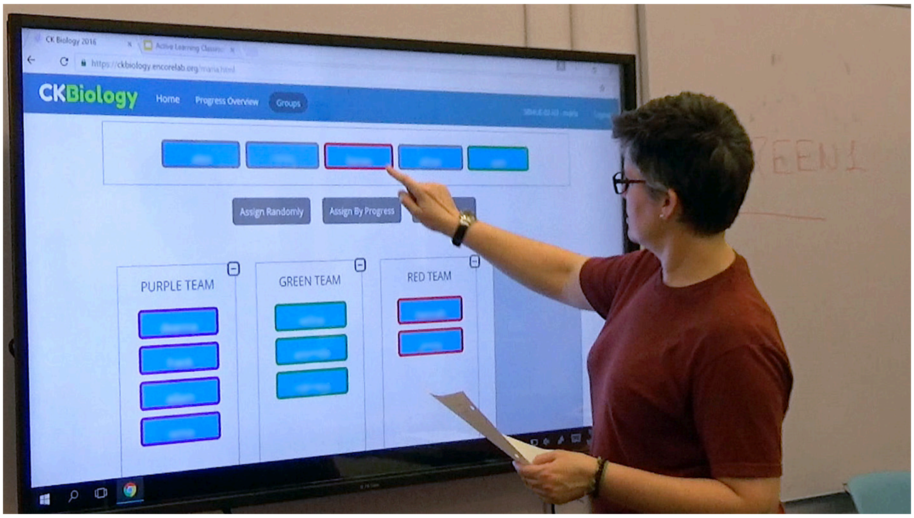

FIGURE 7 | Teacher using the CKBiology group formation tool on a multi-touch display within the Active Learning Classroom. Students' names have been blurred to preserve anonymity. Written informed consent was obtained from the depicted individual for the publication of this image.

missing contributions were left as gaps in the knowledge base.

Regarding the teacher's use of the knowledge base, she explained that she did not have time to refer to the knowledge base either in class or at home, but acknowledged that she wished she had: *"I should have been using this more. I think it's really helpful when we're looking at different concepts to go in and check how much they're doing….and just point out to them, like, 'there is issues with this and this and this' and give them feedback. I did not have the time to actually go in and do that, which I think is a shame because I believe this is a great way to show them how things relate to each other and to also check their knowledge."*

The teacher also commented that since we did not design any review activities for this unit (i.e., a task in which students were asked to apply their knowledge base), students may have had difficulty seeing the relevance of their CKBiology work: "*Just because [the knowledge base] exists that doesn't make it relevant to them, right?*" When considering what form of review activities we should add in the next unit, the teacher commented that students would benefit from more structured review sessions rather than periods of free study: *"If you give [students] review time in class they don't review. They just do their other homework and then they go home and then they stay up late at night the day before the test… and they pretend that that's enough to do well."* More generally, the teacher noted that these students tend to be resistant to pedagogical change: *"When you ask them to do something different, they're very resistant. But I think they're coming around, or I feel that there has been a change or a turn on their perception and I think they're starting to see the value of what we're doing."*

Iteration 2: Molecular Genetics Unit

In the Molecular Genetics Unit—hereafter referred to as "Unit 2" —we maintained the same format and structure for the "lessons" portion of the unit, and introduced several new review activities where students made constructive use of their knowledge base. Additionally, we introduced a group-formation tool as a component of the teacher dashboard to facilitate transitions between the "individual" and "small group" social planes. These improvements are described below.

Group Formation Tool

The group formation tool enabled the teacher to form groups of students "in the moment," according to the following protocols:

- *Group by progress*—matches students with similar mean progress scores. Mean progress scores are calculated based on all lessons within a given unit.
- *Jigsaw*—shuffles previously existing groups such that each new group contains at least one representative from each of the previous groups.
- *Random*—distributes students into groups randomly.
- *Manual mode*—allows the teacher to modify any of the above groups, or to form groups by manually dragging-and-dropping student names into teams.

Although other grouping protocols could have been included, this initial set of protocols was chosen based on the teacher's input as to the kinds of groups she wished to form during the Unit 2 review activities. In subsequent iterations, we created additional grouping protocols based on the teacher's input for those activity designs (e.g., a group recommender, and a group-by-specialization protocol).

The interface for the group formation tool is presented in **Figure 7**. To use the tool, the teacher began by adding the desired number of teams or groups, which appeared as a series of empty boxes. After moving any absent students to the "absent" box, the teacher then selected the desired grouping protocol using one of the buttons on the screen, or by manually adding members to each group by dragging and dropping student names. At her option, the teacher could also modify group membership manually if adjustments were required.

Lessons

Once again, CKBiology activities were completed as part of students' homework and served as a complement to the traditional classroom lectures. There were five lesson topics in Unit 2, which were taught over nine class sessions. Before arriving to class, students were asked to view the PowerPoints and complete the appropriate pages of notes/handouts. In class, lectures served to reinforce these concepts and provided students with an opportunity to ask questions and clarify their understandings. Following each lecture, students logged on to CKBiology where they completed their explanation notes, relationships, and vetting. On average, students were assigned three explanations, three relationships, and seven vets for each lesson in Unit 2.

Review Activities

We designed four review activities for Unit 2. The goal of these review activities was for students to draw upon the knowledge base they had co-constructed throughout the unit, and to apply this knowledge to a new context of inquiry.

Review 1

The first review activity was completed individually. Within the CKBiology interface, students were asked to select a field of research from among four choices: (1) Cell biology, (2) Food science, (3) Pathology, and (4) Pharmacology. There was a maximum of four students per topic, with students receiving a notification if their chosen specialization was full. Students were then presented with a short article related to their chosen research field, and were instructed to "tag" any terms/concepts from the knowledge base that were relevant to the article. Next, students had to explain *how* each term/concept they had tagged was applicable within the context of their article. There was no minimum or maximum number of tags required for this activity, which was considered completed as long as students had provided explanations for all of the tags they had applied. The teacher's dashboard showed which students had completed the activity, were still in progress, or hadn't yet started.

Review 2

For the second review activity, students were assigned to jigsaw groups containing one representative from each of the four fields of research. The CKBiology interface contained each of the four articles, as well as an aggregation of all of the tags that each student had applied. The color intensity of each tag varied from pale blue to dark blue, depending on how many of the four articles contained that tag. Clicking on each tag brought up a "cross-cutting ideas" screen that prompted students to "explain how this term/concept is *common* across all of these articles." Beneath the text box appeared each of the explanations that individual group members had submitted in Review 1 (i.e., of how the tag was related to one specific article). Students were also given the option to remove a term/concept if no cross-cutting ideas could be identified.

Review 3

The third review activity was a group challenge completed in the AL classroom. Students worked in groups of five, with all groups performing the same activity in parallel. The teacher created groups with the group formation tool using the "assign randomly" protocol. The progress of each group was visible to students and the teacher on a "progress overview" screen located at the front of the room. Tapping on any of the group names allowed the teacher to see the responses they had submitted so far, thereby informing her of which groups, if any, required her attention at a given moment. The premise of Review Activity 3 was that each group had been hired by a research funding agency to evaluate a research proposal in order to decide if the proposed project was both possible and scientifically sound. As part of their evaluation, groups had to prepare a report in which they explained key elements of the research and commented on its plausibility. Students' creation of this report was scaffolded by CKBiology, wherein students responded to a series of questions and virtual analyses (e.g., gene sequencing, protein synthesis, PCR, plasmid cloning). Ten question were displayed, in turn, on a large multi-touch screen, with responses entered using a shared wireless keyboard. Group members also used their own personal devices to consult the knowledge base and other online resources throughout this activity.

Review 4

In the final review activity, students were assigned to jigsaw groups consisting of at least one representative from each of the Review 3 groups. To begin, each group was given one of the 10 questions from the Review 3 activity along with the three versions of responses submitted by each of the Review 3 groups. Their task was to discuss the three responses and improve upon the ideas therein, arriving at a "best version" of the response to submit to the funding agency. Groups were also asked to tag concepts from the knowledge base that reviewers would need to understand in order to be able to respond to that question. Once a group had submitted a best response with tags, it received another question to work on until all 10 questions had been reviewed by at least one group. The output of the Review Activity 4 was a whole-class version of the review report, which served to consolidate students' ideas and informed a final discussion about whether the proposed research project should be funded.

Enactment of Unit 2

Several pedagogical challenges arose during the enactment of Unit 2. Firstly, it seemed that the novelty of the CKBiology lesson activities had started to fade, and students simply weren't keeping up with their CKBiology homework. Second, while the teacher continued to activate lessons in CKBiology as the unit progressed, due to time constraints she did not engage students in follow-up discussions wherein gaps in the knowledge base would have been revealed and discussed. This removed the social pressures that would have served to motivate students to do their homework. Before the final lesson, the teacher explored the knowledge base on her own and noticed that students' progress was low. However, this observation occurred right before the winter break, and at that point little could be done to catch up.

The "review challenge" activities were scheduled to occur on return from winter break. However, since these activities relied upon completed explanations from the knowledge base, they could not proceed as planned. Instead, students spent the first review day catching up on outstanding CKBiology homework. We had booked a total of 3 days in the ALC for the purposes of the review activities, and for various reasons it was not possible to postpone or reschedule any of these sessions. Thus, we simplified Review 1, and decided to skip Review 2 altogether.

Students were quite engaged in the Review 3 challenge activity. In one of the sessions the teacher commented, *"This is the most lively I've seen this class all year!"* However, the pace at which students progressed through the activity was much slower than anticipated, partly because of the impact of the winter break (i.e., on their memories), and partly because of how meticulous they were in their responses. The teacher stated that she didn't want to hurry students along just for the sake of reaching the end of the activity, seeing as how they were so deeply engaged and having such rich discussions. Consequently, by the end of the second review day most groups had only completed two or three out of the 10 questions. On the third and final review day, we decided to continue with the Review 3 activity, having no choice but to forgo Review 4. Despite the extra time allotted for Review 3, none of the groups were able to finish, with most groups ending on question six (of 10) before the period had ended.

In a debrief interview following this unit, the teacher commented that she wanted her students to be more motivated to complete their CKBiology activities, acknowledging the negative of not being able to assign grades to students' work. Therefore, on the unit test, whose design was solely under the teacher's control?she decided to include several questions that were modeled after the CKBiology review activities. For this unit, students were provided with a research proposal related to gene expression and alternative splicing in aging, and were asked to evaluate the research proposal as well as "tag" (using pencil and paper) and explain any concepts necessary for evaluating the proposal. In this manner, the inclusion of similar types of questions on the unit test meant that completing the CKBiology work would be beneficial for their performance.

Iteration 3: Homeostasis Unit

In response to some of the pedagogical challenges that arose during Unit 2—most notably students not keeping up with their CKBiology homework—one of the changes that was implemented for the Homeostasis Unit (i.e., Unit 3) was to provide class time for students to complete their CKBiology work. While the activity structure for the "lessons" portion of the script remained the same, the context in which the CKBiology work took place was now in the science classroom rather than at home. Within the classroom, the knowledge base was projected at the front of the room while students were working. This meant that students' contributions were physically prominent within the space, making knowledge gaps more public, an also allowing for more frequent discussions about aspects of the knowledge base (e.g., when there were evident vetting disagreements).

With respect to the review activities, we wanted to establish a more meaningful connection between the articles (i.e., Review 1) and the subsequent review activities. We therefore changed Review 2, introducing a "specialist certification" activity. We also exchanged our use of research articles in Review 1 with medical case studies, which students would apply toward solving a series of medical problems. In Review 3, students worked in jigsaw groups containing one representative from each specialization, with each group acting as a medical clinic. These changes are elaborated below. Based on the timing issues we had experienced in the previous unit, we shortened Review 3 considerably, from 10 questions to five, and eliminated the fourth Review activity altogether.

Several new technological features were added to Unit 3. First, we added a "specialization recommender" to Review 1, which made a recommendation to each student about which specialization they might choose (i.e., Immunology, Endocrinology, Nephrology, and Neurology), based on their contributions to the CKBiology knowledge base. We also enhanced the information provided on the teacher dashboard for Review 1, including the number of terms each student had tagged in addition to their level of completion. As well, complementing the student-facing recommender, we also added a teacher-facing recommender to the group formation tool for Review 2. Here, each specialization was assigned a color, and the names of students who had not chosen a specialization would appear with a colored outline corresponding to their recommended group. For example, **Figure 8** shows that the students "gaoxia" and "rokham" are recommended for the "Immunology" group, and that "stian" is recommended for the "Neurology" group.

A final technological design revision that was made in Unit 3 was the addition of a "call a conference" function. When students were working in their medical clinics (i.e., jigsaw groups) and a situation arose in which a particular specialist needed to consult with his/her fellow specialists, the "call a conference" button sent out a bat-signal-like alert to the other clinics, requesting the relevant specialists to convene in the designated conference area within the room. We did not put any restrictions on the number or frequency of conferences that could be called throughout Review 3.

Lessons

The activity structure for the "lessons" portion of Unit 4 was the same as in the previous two units—the only difference being that students now completed their CKBiology work in their classroom rather than at home. There were eight lesson topics in Unit 4, which were taught over 14 class sessions. In CKBiology, students were assigned an average of four to five explanations, five relationships, and 30 vets per lesson throughout Unit 4. (The high number of vetting tasks was attributed to a bug in the code).

Review Activities

There were three review activities for Unit 3:

Review 1—Upon logging into CKBiology, students were asked to select an area of specialization from among four

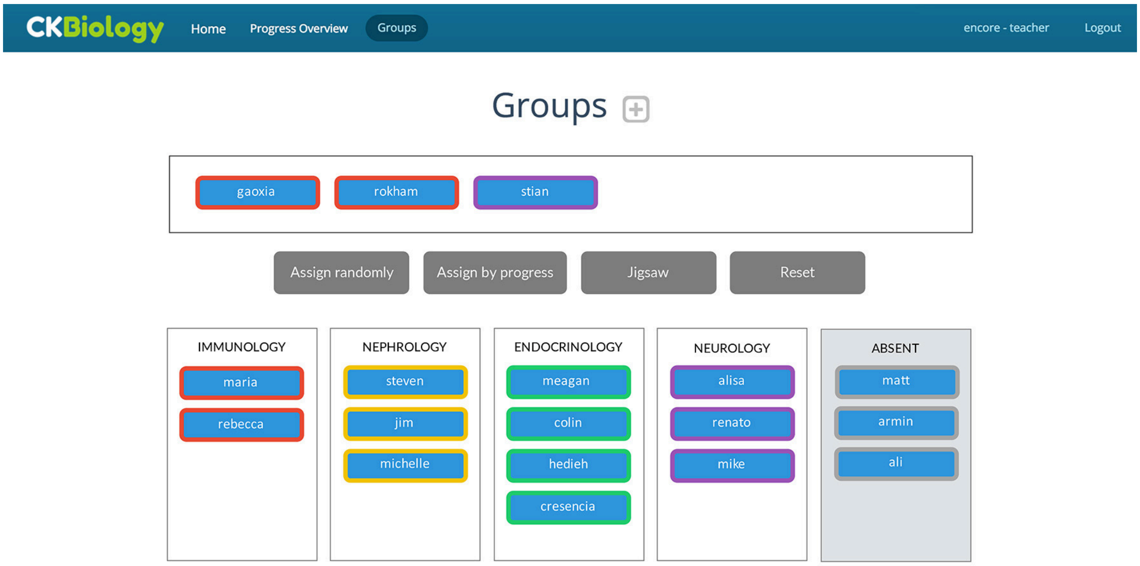

FIGURE 8 | Teacher-facing specialization recommender. The colored ring around each students' name indicates their recommended group. Student names are pseudonyms.

choices: (1) Immunology, (2) Endocrinology, (3) Nephrology, and (4) Neurology. As mentioned above, students were given a recommendation about which specialization would be well-suited to them. To do so, we calculated a score for each specialization based on the student's contributions to the knowledge base. Accounting for a maximum of four students per specialization, we generated a recommendation for each student based on their highest score for a non-full group. This recommendation (shown in **Figure 9**) was presented to students as entirely optional, with students free to choose whichever specialization they wished. Once students had chosen a specialization, they were presented with a medical case study whose purpose was to introduce students to various symptoms, lab analyses, test results, and treatment options related to a disorder within their area of specialization. For example, students who had selected "endocrinology" were given a case study about Graves' Disease, and students who had selected "nephrology" were given a case study about Glomerulonephritis. Students were then instructed to tag their medical case study with terms/concepts from the knowledge base, and to provide explanations as to how these terms were applicable within the context of their case study.

Review 2—The second review activity was performed in the AL Classroom. Students worked within their specialist groups to solve a series of challenge questions related to their area of specialization. Questions were presented in CKBiology using a shared group display, and responses were entered by different group members using a wireless keyboard. Specialist groups also received a selection of paper handouts, which contained information on how to interpret various lab test results. For example, the Nephrology group was given handouts to assist them in interpreting urinalysis and urine microscopy test results. Likewise, the Neurology group was given handouts on how to interpret an EEG, the Endocrinology group received handouts on various blood tests, and the Immunology group was given handouts on autoantibodies. Specialist groups who successfully completed all of their challenge questions received "certification" in their area of specialization, which included a personalized paper certificate signed by their teacher.

Review 3—For the third review activity, students worked in jigsaw groups (i.e., "medical clinics") containing one representative from each specialization. Playing the role of medical practitioners, students had to bring together their diverse expertise in order to diagnose a virtual patient with ambiguous symptoms. This included ordering the appropriate tests, explaining the reasoning behind their diagnosis, and identifying possible treatment options—thereby consolidating the knowledge they had acquired over the course of the unit. Students were guided through this activity via a series of five scaffolded questions in the CKBiology platform. Within the interface, the "call a conference" button was displayed next to each question. As in the previous unit, the progress of each group was visible on a public display at the front of the ALC. The teacher could also view each group's responses in real-time to get a sense of when and where students would most benefit from her assistance.

Enactment of Unit 3

While completing the CKBiology work during class time reduced the amount of time available for lecture, it had several benefits to the learning community. First, because the knowledge base

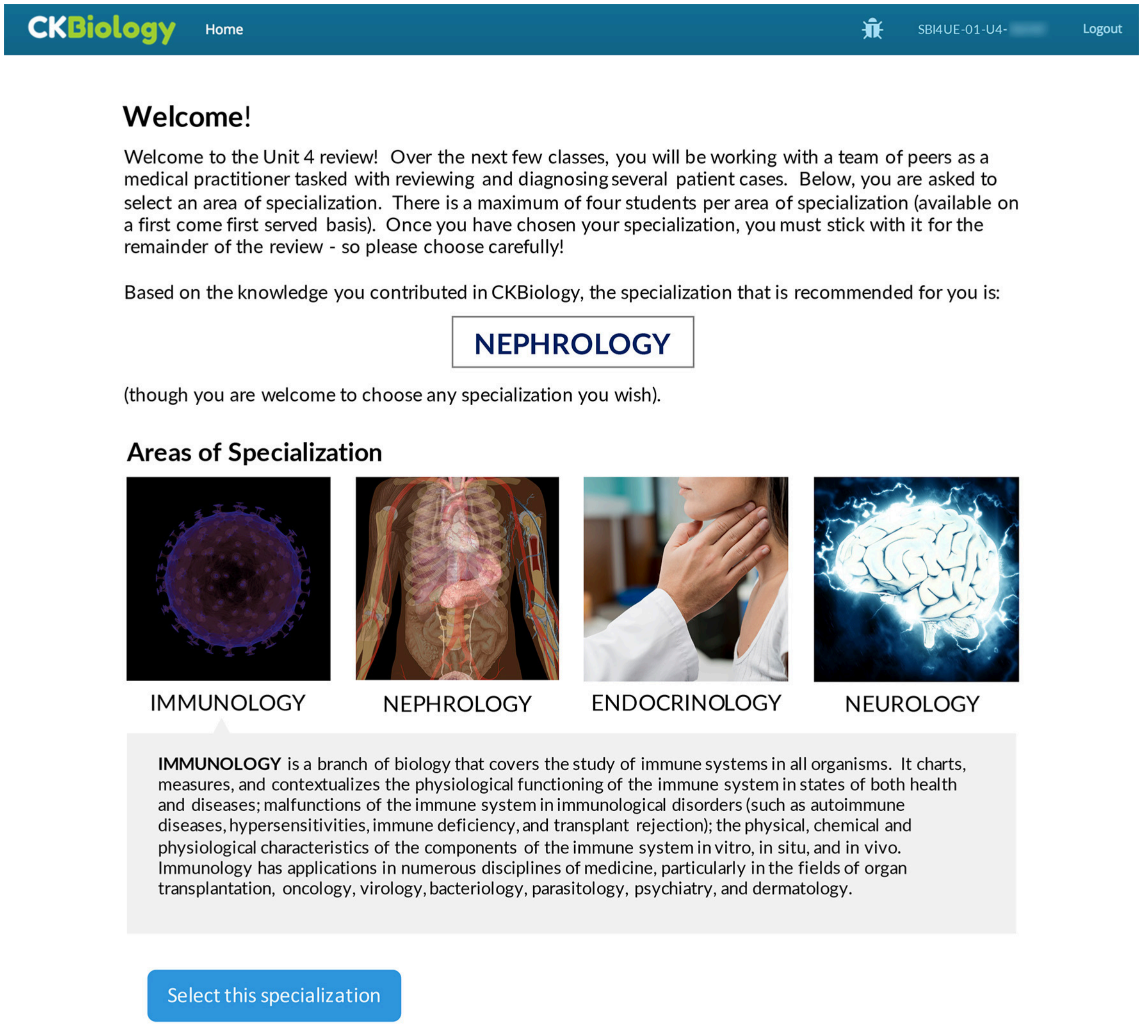

FIGURE 9 | Student-facing specialization recommender showing a recommendation for "Nephrology." Beneath this recommendation, students could explore all four areas of specialization before making their final selection.

was projected at the front of the classroom while students were working, any gaps or conflicts that existed in the knowledge base were made visible and salient. Consequently, discussions around the knowledge base occurred with greater frequency—whether they were initiated formally by the teacher, or informally among peers while they were working. Additionally, student progress for each of the CKBiology lessons frequently exceeded 100%, with many students performing two or three times the amount of work that had been assigned to them (i.e., earning progress scores of 200–300%). The average student progress across all eight lessons in Unit 4 was 109.7%. This figure is particularly impressive given a "vetting bug" where students were accidentally assigned more items to vet due to a software coding error.

The teacher commented that the Unit 3 review activities seemed more cohesive than in previous units, and that the articles/case studies were more meaningfully connected. Regarding students' use of the specialization recommender, only 26.3% of students ended up choosing the specialization that was recommended to them. An additional 5.3% of students indicated that they *would have* chosen their recommended specialization, except it had already filled up. The low uptake of recommendations may have partly been related to the fact that students completed the Review 1 activity synchronously in class as opposed to asynchronously for homework, as planned. With all students working simultaneously, the system was generating recommendations at the same time as they were being filled. Consequently, a student may have been presented with a recommendation that, moments later, was no longer available.

Attendance for the review activities remained a challenge, and became particularly problematic when trying to form jigsaw groups of specialists. In some cases, there were specialists present for Review 3 who had been absent for Review 2 and had not yet earned their certification. In other cases, specialists who *had* earned their certification in Review 2 were absent for Review 3, leaving some groups without expertise in these specializations. These absences were handled in two ways. First, each medical clinic was provided with a folder containing all of the specialist resources that had been generated during Review 2, including the paper handouts for each specialization as well as access to the Review 2 reports in CKBiology. In this sense, the "knowledge" of that specialist was still present at the table, even if the person wasn't. Second, students could use the "call a conference" button if they needed further information related to a particular specialization. This "call a conference" functionality was used a total of six times across both class sections, with all specialist groups conferring at least once.

An additional design challenge that arose during the enactment of the Unit 3 review activities was related to the way group progress was measured and displayed. Technically, students could enter a single character as a response to a challenge question and then proceed to the next as if that response was complete (Students could later go back and revise their responses). It was thus up to the teacher to identify such cases (e.g., using the "group report" function on her dashboard) and intervene when a particular answer wasn't up to par. However, for the purposes of the progress bar calculation, this single-character response was considered "complete," and groups would earn progress points for submitting such a placeholder response. Students quickly caught on to this, and began entering single-character responses to the challenge questions—however their reason for doing this wasn't because they wanted to earn 100% progress for doing little/no work. Instead, they did this so that they could read all of the challenge questions ahead of time (i.e., to see where this activity was going) before going back and carefully considering each response. Consequently, several groups *appeared* to have earned 100% progress at the beginning of the activity, even though their responses were virtually empty.

This "false progress" made it challenging for the teacher to decide when and where to intervene. The teacher used the Reports screen on her teacher dashboard to look at the responses for each group, however she generally waited until a group *claimed* to be finished before reviewing their responses. According to the teacher: "*What I did was like... when I would see that they were done... I would go and check [their answers]. 'Ok... this is not great,' 'Mmm, this needs to be looked after... ' So then I would go back to them and say, 'Listen people. Yes, you are on the right track, but you need to look at this and this and this,' and 'What about blablabla' and 'Did you consider blablabla.' And that's how I used it.*"

Overall, the enactment of Unit 3 was successful in that students co-constructed a quality knowledge base with many exceeding what was required of them, and then applied the knowledge base to a new context of inquiry (i.e., a medical case study). They were engaged in their review challenge activities, and completed everything within the time available. The teacher also noted that "*The certificates were a big hit. Who knew? [laughs] If I had known this I would be giving them certificates every single class!*" She also responded positively to the group formation tool:

"It is so useful. I LOVED it. I thought that was fantastic... Because it makes it really easy to see what you're doing with your groups. It makes it really easy to see, for example, when we have the jigsaw, that you were actually jigsawing people properly... it's not something that I have to, you know, look at people or change them afterwards or whatever. Like I can really quickly do that and do it right. I thought it was great."

DISCUSSION

The sections above describe an uncommon opportunity to iteratively develop an active learning design over four distinct cycles during a single course offering. This opportunity arose because of the cyclical nature of our course context, with active learning elements occurring at the end of each curricular unit in the form of review activities. Because there was a month or so between iterations, we were able to examine the previous enactment, revise our designs, and develop the corresponding materials and technology environments (i.e., CKBiology). While this approach introduces the confound of having a single cohort of students engage with each successive iteration, by the same token it allowed us to develop our designs in a single coherent context, building upon the knowledge and experience of community members. Our plans for future work will extend this research to four new school contexts with a comparative study of all participants—including the ways that teachers adopt and adapt our designs for their particular curricula, students, and schedules.

In response to our first research question (i.e., What are the design opportunities and constraints associated with infusing a traditional Grade 12 Biology course with active learning designs?), this work advanced a general active learning progression, as epitomized by the Unit 3 designs, wherein students worked as a community to explain, connect, and review all the salient concepts from the unit, and then use the resulting "knowledge base" as a resource for inquiry-oriented challenge activities. We employed a jigsaw group strategy for the review activities, first creating a set of expert groups, with an activity designed to enhance group members' knowledge of their respective specializations, then regrouping such that one member from each expert group was present in a more general team. These groups were charged with creating reports and summaries, and applying their knowledge to contextually relevant challenges (e.g., reviewing grant proposals or addressing a medical diagnosis). Through three successive units (and one baseline unit), we progressively refined and adapted the review activities, including new supports for student groups, for teacher and community awareness (i.e., of community progress), and for teacher orchestration.

In response to our second research question (i.e., What forms of active learning can address those constraints and challenges, and what technology elements are needed to support them?), this iterative design study allowed us to progress in our understanding of the role of technologies for supporting students and teachers in AL. For students, we investigated and iteratively refined the role of progress bars for their individual, group and community efforts (Acosta and Slotta, 2018). We also examined group process supports during review activities, including grouping strategies and a specialization recommender. We also emphasized two forms of ambient technologies for our AL classroom: First was the inclusion of the concept network as a central display, showing terms that had or had not yet been defined, whether and to what extent they had been vetted, and relationships amongst them. This omnipresent display allowed the teacher to occasionally find certain concepts or terminologies within the display, touch them to reveal their definition, comment on relationships or conflicts in vetting, etc. She could also spot gaps in the network, and encourage greater progress. Another ambient technology was the teacher dashboard, which was visible only to the teacher and was always available for reference as a source of information about specific group products and productivity.

Throughout this effort, we were cognizant of several ongoing tensions, which challenged our successful enactment. The first was concerned with the culture of assessment in the school, and the need felt by students for grading and recognition of their contributions. Because our research ethics protocol disallowed assigning grades for participation, we were forced into a position of focusing on review activities that were perceived by students as supplementary. This perception was addressed by the teachers' decision to use our designs as a basis for part of her unit tests. However, we recognize the general need for epistemological coherence within a learning community approach. Students who are situated within an otherwise lecture and test-based course will have a difficult time identifying with and participating in any collective elements. Another challenge was concerned with the fact that this course was taught in the senior year of a university-preparatory program, where the students have substantial extracurricular activities and commitments during their final year.

Future research will more closely examine the group formation processes, as well as specific supports for group processes, representations of community knowledge, and orchestration supports for the teacher. We emphasize the need for co-design in such approaches as the only viable means of ensuring that partner teachers are fully aware of all designs, feel a sense of ownership, and succeed in orchestrating them during the time of enactment. We will also study the epistemology of our designs, with an effort to shift this and other course designs into more fully community-oriented curricula. KCI provides an excellent context for active learning, as it emphasizes collective products, and their application as resources in community-based inquiries. We will continue developing Common Knowledge (CK) in various forms, further examining its role in supporting a KCI community during active learning designs.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

This work was supported by the Social Sciences and Humanities Research Council of Canada.

REFERENCES

Acosta, A., and Slotta, J. D. (2018). "Representations of progress in a learning community curriculum for grade 12 biology," in *Proceedings of the 13th International Conference on the Learning Sciences* (London: ISLS Inc,).

Baepler, P., and Walker, J. D. (2014). Active learning classrooms and educational alliances: changing relationships to improve learning. *New Direct. Teach. Learn.* 2014, 27–40. doi: 10.1002/tl.20083

Barab, S. (2014). "Design-based research: a methodological toolkit for engineering change," in *Cambridge Handbook of the Learning Sciences,* ed R. K. Sawyer (Cambridge, MA: Cambridge University Press), 151–170.

Barab, S., and Squire, K. (2004). Design-based research: putting a stake in the ground. *J. Learn. Sci.* 13, 1–14. doi: 10.1207/s15327809jls1301_1

Bell, P. (2004). On the theoretical breadth of design-based research in education. *Educ. Psychol.* 39, 243–253. doi: 10.1207/s15326985ep3904_6

Bell, P., Hoadley, C. M., and Linn, M. C. (2004). "Rationale for design-based research in education," in *Internet Environments for Science Education,* eds M. C. Linn, E. A. Davis, and P. Bell (Mahwah, NJ: Lawrence Erlbaum Associates Inc,). 73–85.

Bens, I. (2005). *Facilitating with Ease! Core Skills for Facilitators, Team Leaders and Members, Managers, Consultants, and Trainers, 2nd Edn*. San Francisco, CA: Jossey Bass.

Bereiter, C. (2002). Design research for sustained innovation. *Jap. Cogn. Sci. Soc.* 9, 321–327. doi: 10.11225/jcss.9.321

Bielaczyc, K., and Collins, A. (1999). "Learning communities in classrooms: a reconceptualization of educational practice," in *Instructional-Design Theories and Models: A New Paradigm of Instructional Theory*, Vol. 2, ed C. M. Reigeluth (Mahwah, NJ: Lawrence Erlbaum Associates Inc.), 269–292.

Bloom, B. S. (1984). The 2 sigma problem: the search for methods of group instruction as effective as one-to-one tutoring. *Educ. Res.* 13, 4–16. doi: 10.3102/0013189X013006004

Blumenfeld, P., Fishman, B. J., Krajcik, J., Marx, R., and Soloway, E. (2000). Creating usable innovations in systemic reform: scaling up technology-embedded project-based science in urban schools. *Educ. Psychol.* 35, 149–164. doi: 10.1207/S15326985EP3503_2

Brooks, D. C. (2011). Space matters: the impact of formal learning environments on student learning. *Br. J. Educ. Technol.* 42, 719–726. doi: 10.1111/j.1467-8535.2010.01098.x

Brown, A. L. (1992). Design experiments: theoretical and methodological challenges in creating complex interventions in classroom settings. *J. Learn. Sci.* 2, 141–178. doi: 10.1207/s15327809jls0202_2

Brown, A. L. (1994). The advancement of learning. *Educ. Res.* 23, 4–12. doi: 10.3102/0013189X023008004

Brown, A. L., and Campione, J. C. (1994). "Guided discovery in a community of learners," in *Classroom Lessons: Integrating Cognitive Theory and Classroom Practice*, ed K. McGilly (Cambridge, MA: MIT Press), 229–270.

Brown, A. L., and Campione, J. C. (1996). *Psychological Theory and the Design of Innovative Learning Environments: On Procedures, Principles, and Systems*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Brownell, S. E., Kloser, M. J., Fukami, T., and Shavelson, R. J. (2013). Context matters: volunteer bias, small sample size, and the value of comparison groups in the assessment of research-based undergraduate introductory biology lab courses. *J. Microbiol. Biol. Educ.* 14, 176–182. doi: 10.1128/jmbe.v14i2.609

Carvalho, R. (2017). *A Design Framework for Collective Inquiry Learning Environments.* M.A. Thesis, Ontario Institute for Studies in Education, University of Toronto, Toronto, ON.

Carvalho, R., and Hall, A. H. (2016). "Literary inquiry: exploring literature collaboratively in a digital learning space," in *Proceedings of the 1st Learning Sciences Graduate Student Conference* (Chicago, IL: Learning Sciences Research Institute), 58–59.

Charles, E. S., Tissenbaum, M., Whittaker, C., Lui, M., Dugdale, M., and Slotta, J. D. (2011). "Co-design of collaborative collective knowledge environment," in *Proceedings of the 9th International Conference on Computer-Supported Collaborative Learning Vol. 1* (Hong Kong: International Society of the Learning Sciences Inc.,), 641–645.

Cobb, P., Confrey, J., Lehrer, R., and Schauble, L. (2003). Design experiments in educational research. *Educ. Res.* 32, 9–13. doi: 10.3102/0013189X032001009

Cober, R. M., Acosta, A., Lui, M., Moher, T., Kuhn, A., Quintana, C., et al. (2015a). "The use of visual evidence for planning and argumentation," in *Proceedings of the 11th International Conference on Computer Supported Collaborative Learning Vol. 2* (Gothenburg: International Society of the Learning Sciences Inc), 548–551.

Cober, R. M., McCann, C., Moher, T., and Slotta, J. D. (2013). "Aggregating students' observations in support of community knowledge and discourse," in *Proceedings of the 10th International Conference on Computer-Supported Collaborative Learning (CSCL) Vol. 1* (Madison, WI: International Society of the Learning Sciences Inc), 121–128.

Cober, R. M., Tan, E., Slotta, J. D., So, H.-J., and Könings, K. D. (2015b). Teachers as participatory designers: two case studies with technology-enhanced learning environments. *Instruct. Sci.* 43, 203–228. doi: 10.1007/s11251-014-9339-0

Code, W., Piccolo, C., Kohler, D., and MacLean, M. (2014). Teaching methods comparison in a large calculus class. *ZDM* 46, 589–601. doi: 10.1007/s11858-014-0582-2

Collins, A. (1992). "Toward a design science of education, in *New Directions in Educational Technology,* eds E. Scanlon and T. O'Shea (New York, NY: Springer-Verlag), 15–22.

DeLozier, S. J., and Rhodes, M. G. (2017). Flipped classrooms: a review of key ideas and recommendations for practice. *Educ. Psychol. Rev*. 29, 141–151. doi: 10.1007/s10648-015-9356-9

Dillenbourg, P., and Jermann, P. (2007). "Designing integrative scripts," in *Scripting Computer-Supported Collaborative Learning,* eds F. Fischer, I. Kollar, H. Mandl, and J. M. Haake (New York, NY: Springer), 275–301.

Dori, Y. J., and Belcher, J. (2005). How does technology-enabled active learning affect undergraduate students' understanding of electromagnetism concepts? *J. Learn. Sci.* 14, 243–279. doi: 10.1207/s15327809jls1402_3

Edelson, D. C. (2002). Design research: what we learn when we engage in design. *J. Learn. Sci.* 11, 105–121. doi: 10.1207/S15327809JLS1101_4

Ehrlick, S., and Slotta, J. D. (2017). The impact of a collaborative knowledge base on a Fostering Communities of Learners model in post-secondary large lectures. *Int. J. Infonomics* 10, 1333–1340. doi: 10.20533/iji.1742.4712.2017.0163

Fischer, F., Slotta, J. D., Dillenbourg, P., Tchounikine, P., Kollar, I., and Wecker, C. (2013). "Scripting and orchestration: recent theoretical advances," in *Proceedings of the Tenth International Conference of Computer-Supported Collaborative Learning. Vol. 1,* eds N. Rummel, M. Kapur, M. Nathan, and S. Puntambekar (Madison, WI: International Society of the Learning Sciences Inc), 564–571.

Fisher, K. (2010). *Technology-enabled Active Learning Environments: An Appraisal* (CELE Exchange). Paris: Centre for Effective Learning Environments.

Fong, C., Cober, R. M., Messina, R., Moher, T., Murray, J., Peebles, B., et al. (2015). "Common knowledge: iterative design, scripting, and orchestration of inquiry discourse in elementary science," in *Presented at the American Educational Research Association Annual Meeting* (Chicago, IL).

Fong, C., and Slotta, J. D. (2015). "Common knowledge: supporting the orchestration of multi-modal note-sharing in a heterogeneous classroom ecology," in *Presented at the Workshop conducted at the 11th International Conference on Computer Supported Collaborative Learning (CSCL)* (Gothenburg), 30–33.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., et al. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proc. Natl. Acad. Sci. U.S.A.* 111, 8410–8415. doi: 10.1073/pnas.1319030111

Håklev, S., and Slotta, J. D. (2017). "A principled approach to the design of collaborative MOOC curricula," in *Digital Education: Out to the World and Back to the Campus*, eds C. Delgado Kloos, P. Jermann, M. Pérez-Sanagustín, D. Seaton, and S. White (Madrid; Cham: Springer), 58–67.

Henderson, C., and Dancy, M. H. (2007). Barriers to the use of research-based instructional strategies: the influence of both individual and situational characteristics. *Phys. Rev. Special Topics Phys. Educ. Res.* 3:020102. doi: 10.1103/PhysRevSTPER.3.020102

Kling, R., and Courtright, C. (2003). Group behavior and learning in electronic forums: a sociotechnical approach. *Inform. Soc*. 19, 221–235. doi: 10.1080/01972240309465

Kober, N. (2015). *Reaching Students: What Research Says About Effective Instruction in Undergraduate Science and Engineering*. Washington, DC: National Academies Press.

Kollar, I., Fischer, F., and Slotta, J. D. (2007). Internal and external scripts in computer-supported collaborative inquiry learning. *Learn. Instruct.* 17, 708–721. doi: 10.1016/j.learninstruc.2007.09.021

Lui, M., and Slotta, J. D. (2014). Immersive simulations for smart classrooms: exploring evolutionary concepts in secondary science. *Technol. Pedagogy Educ.* 23, 57–80. doi: 10.1080/1475939X.2013.838452

MAESD. (2016). *Graduation, Employment and Government Loan Default Rates 2016*. Available online at: http://www.provost.utoronto.ca/link/students/osap/Graduation__Employment_and_Government_Loan_Default_Rates_2016.htm (Accessed September 30, 2017).

Means, B., Penuel, W. R., and Padilla, C. (2001). *The Connected School: Technology and Learning in High School*. San Francisco, CA: Jossey-Bass.

Mercier, E. M., Higgins, S. E., and Joyce-Gibbons, A. (2016). The effects of room design on computer-supported collaborative learning in a multi-touch classroom. *Interact. Learn. Environ.* 24, 504–522. doi: 10.1080/10494820.2014.881392

Nathan, M., and Sawyer, R. K. (2014). "Foundations of the learning sciences," in *Cambridge Handbook of the Learning Sciences,* ed R. K. Sawyer (Cambridge, MA: Cambridge University Press), 21–43.

OECD. (2012). *Strengthening Education for Innovation* (Science, Technology and Industry e-Outlook). Available online at: https://www.oecd.org/sti/outlook/e-outlook/

OECD. (2016). *Innovating Education and Educating for Innovation: The Power of Digital Technologies and Skills*. Paris: OECD Publishing.

Pellegrino, J. W., and Hilton, M. L. (2013). *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century*. Washington, DC: National Academies Press.

Penuel, W. R., Roschelle, J., and Shechtman, N. (2007). Designing formative assessment software with teachers: an analysis of the co-design process. *Res. Pract. Technol. Enhanc. Learn*. 2, 51–74.

Peters, V. L., and Slotta, J. D. (2010). "Content analysis of collaboratively constructed knowledge artifacts: issues and opportunities for research," in *Learning in the Disciplines: Proceedings of the 9th International Conference of the Learning Sciences*, Vol. 2, eds K. Gomez, L. Lyons, and J. Radinsky (Chicago, IL: International Society of the Learning Sciences Inc.), 117–124.

Roschelle, J., Penuel, W. R., and Shechtman, N. (2006). "Co-design of innovations with teachers: definition and dynamics," in *Proceedings of the 7th International Conference on Learning Sciences* (Bloomington, CA: International Society of the Learning Sciences), 606–612.

Ruiz-Primo, M. A., Briggs, D., Iverson, H., Talbot, R., and Shepard, L. A. (2011). Impact of undergraduate science course innovations on learning. *Science* 331, 1269–1270. doi: 10.1126/science.1198976

Scardamalia, M., and Bereiter, C. (2006). "Knowledge building: theory, pedagogy and technology," in *Cambridge Handbook of the Learning Sciences,* ed R. K. Sawyer (Cambridge, UK: Cambridge University Press), 97–118.

Scardamalia, M., and Bereiter, C. (2007). "Fostering communities of learners and knowledge building: an interrupted dialogue," in *Children's Learning in Laboratory and Classroom Contexts: Essays in Honor of Ann Brown,* eds J. C. Campione, K. Metz, and A. S. Palincsar (New York, NY: Routledge), 197–212.

Scardamalia, M., Bereiter, C., McLean, R. S., Swallow, J., and Woodruff, E. (1989). Computer-supported intentional learning environments. *J. Educ. Comput. Res.* 5, 51–68. doi: 10.2190/CYXD-6XG4-UFN5-YFB0

Serevetas, M. (2017). "The teacher's role in orchestrating a Knowledge Community and Inquiry classroom," in *Presented at the Paper Presented at* *the 45th Canadian Society for the Study of Education (CSSE) Conference* (Toronto, ON).

Simon, H. (2001). "Learning to research about learning," in *Cognition and Instruction: Twenty-five Years of Progress,* eds S. Carver and D. Klahr (Mahwah, NJ: Lawrence Erlbaum), 205–226.

Slotta, J. D. (2010). "Evolving the classrooms of the future: the interplay of pedagogy, technology and community," in *Classroom of the Future: Orchestrating Collaborative Spaces,* eds K. Mäkitalo-Siegl, J. Zottmann, F. Kaplan, and F. Fischer (Rotterdam: Sense Publishers), 215–242.

Slotta, J. D. (2014). "Knowledge community and inquiry: new opportunities for scripting and orchestration," in *Paper presented at the Network of Associated Programs in the Learning Sciences (NAPLES)* (Munich).

Slotta, J. D., and Acosta, A. (2017). "Scripting and orchestrating learning communities: a role for learning analytics," in *Proceedings of the 12th International Conference on Computer Supported Collaborative Learning Vol. 1* (Philadelphia, PA), 343–350.

Slotta, J. D., and Najafi, H. (2010). "Knowledge communities in the classroom," in *International Encyclopedia of Education. Vol. 8*, eds P. Peterson, E. Baker, and B. McGaw (Oxford: Elsevier), 189–196.

Slotta, J. D., and Najafi, H. (2013). "Supporting collaborative knowledge construction with Web 2.0 technologies," in *Emerging Technologies for the Classroom: A Learning Sciences Perspective,* eds C. Mouza and N. Lavigne (New York, NY: Springer), 93–112.

Slotta, J. D., and Peters, V. (2008). "A blended model for knowledge communities: embedding scaffolded inquiry," in *Proceedings of the 8th International Conference on International Conference for the Learning Sciences, Vol. 2* (Utrecht: International Society of the Learning Sciences), 343–350.

Stahl, G. (2015). "The group as paradigmatic unit of analysis: The contested relationship of CSCL to the learning sciences," in *Reflections on the Learning Sciences*, eds M. A. Evans, M. J. Jacker, and R. K. Sawyer (Cambridge, UK: Cambridge University Press), 76–102.

Tapscott, D., and Williams, A. D. (2012). *Macrowikinomics: New Solutions For A Connected Planet*. Toronto, ON: Portfolio Penguin Canada.

Tchounikine, P. (2013). Clarifying design for orchestration: orchestration and orchestrable technology, scripting and conducting. *Comput. Educ.* 69, 500–503. doi: 10.1016/j.compedu.2013.04.006

Tissenbaum, M., Lui, M., and Slotta, J. D. (2012). Co-designing collaborative smart classroom curriculum for secondary school science. *J. Univ. Comput. Sci.* 18, 327–352. doi: 10.3217/jucs-018-03-0327

U.S. Department of Education (2013). *STEM Attrition: College Students' Paths Into and Out of STEM Fields (No. NCES 2014-001)*. National Center for Education Statistics. Available online at: https://nces.ed.gov/pubs2014/2014001rev.pdf

van Aalst, J., and Chan, C. K. (2007). Student-directed assessment of knowledge building using electronic portfolios. *J. Learn. Sci*. 16, 175–220. doi: 10.1080/10508400701193697

Waldrop, M. M. (2015). Why we are teaching science wrong, and how to make it right. *Nat. News* 523, 272–274. doi: 10.1038/523272a

Woolfolk, A., Winne, P. H., and Perry, N. E. (2009). *Educational Psychology (Fourth Canadian Edition)*. Toronto, ON: Allyn and Bacon.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in ICT

ORIGINAL RESEARCH
published: 07 March 2018
doi: 10.3389/fict.2018.00003

Challenge of Helping Introductory Physics Students Transfer Their Learning by Engaging with a Self-Paced Learning Tutorial

Emily Megan Marshman[1], Seth DeVore[2] and Chandralekha Singh[1]*

[1] Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA, United States, [2] Department of Physics, West Virginia University, Morgantown, WV, United States

OPEN ACCESS

Edited by:
James Davis Slotta,
Boston College, United States

Reviewed by:
S. Raj Chaudhury,
University of South Alabama,
United States
Licia Maria Rita Montagna,
Humanitas Research Hospital, Italy

***Correspondence:**
Emily Megan Marshman
emm101@pitt.edu

Specialty section:
This article was submitted to Digital Education, a section of the journal Frontiers in ICT

Received: 24 October 2017
Accepted: 12 February 2018
Published: 07 March 2018

Citation:
Marshman EM, DeVore S and Singh C (2018) Challenge of Helping Introductory Physics Students Transfer Their Learning by Engaging with a Self-Paced Learning Tutorial. Front. ICT 5:3. doi: 10.3389/fict.2018.00003

With advances in digital technology, research-validated self-paced learning tools can play an increasingly important role in helping students with diverse backgrounds become good problem solvers and independent learners. Thus, it is important to ensure that all students engage with self-paced learning tools effectively in order to learn the content deeply, develop good problem-solving skills, and transfer their learning from one context to another. Here, we first provide an overview of a holistic framework for engaging students with self-paced learning tools so that they can transfer their learning to solve novel problems. The framework not only takes into account the features of the self-paced learning tools but also how those tools are implemented, the extent to which the tools take into account student characteristics, and whether factors related to students' social environments are accounted for appropriately in the implementation of those tools. We then describe an investigation in which we interpret the findings using the framework. In this study, a research-validated self-paced physics tutorial was implemented in both controlled one-on-one interviews and in large enrollment, introductory calculus-based physics courses as a self-paced learning tool. We find that students who used the tutorial in a controlled one-on-one interview situation performed significantly better on transfer problems than those who used it as a self-paced learning tool in the large-scale implementation. The findings suggest that critically examining and taking into account how the self-paced tools are implemented and incentivized, student characteristics including their self-regulation and time-management skills, and social and environmental factors can greatly impact the extent and manner in which students engage with these learning tools. Getting buy in from students about the value of these tools and providing appropriate support while implementing them is critical for ensuring that students, who otherwise may be constrained by motivational, social, and environmental factors, engage effectively with the tools in order to learn deeply and transfer their learning.

Keywords: self-paced learning, adaptive learning, personalized learning, transfer of learning, physics education research

INTRODUCTION

Background: Self-Paced Learning Tools

Research-validated self-paced learning tools provide a valuable opportunity for personalized learning and can supplement learning even in brick and mortar classrooms (Kulik and Kulik, 1991; Azevedo et al., 2004, 2005; Azevedo, 2005; Allen and Seaman, 2013; Breslow et al., 2013; Colvin et al., 2014; Seaton et al., 2014; Alraimi et al., 2015; Bower et al., 2015; Margaryan et al., 2015). Adaptive self-paced learning tools can allow students with diverse prior preparations to obtain feedback and support based upon their needs, and students can work through them at their own pace and receive help as needed (Reif, 1987; Lenaerts et al., 2002; Chen et al., 2010; Chandra and Watters, 2012; Debowska et al., 2012; Chen and Gladding, 2014). Appropriate use of research-validated self-paced adaptive learning tools can be particularly beneficial for under-prepared students and provide a variety of students an opportunity to learn. These tools can play a central role in scaffolding student learning, helping them gain a deep understanding of the content (Yalcinalp et al., 1995; McDermott, 1996; Korkmaz and Harwood, 2004; Singh, 2008a; Kohnle et al., 2010; Marshman and Singh, 2015; Sayer et al., 2017), develop their problem-solving, reasoning, and meta-cognitive skills (Reif and Scott, 1999; Singh, 2004; Demetriadis et al., 2008; Singh and Haileselassie, 2010), and facilitate transfer of learning from one context to another (Chi et al., 1994).

However, even in a brick-and-mortar class, ensuring that students engage effectively with available self-paced learning tools to learn is challenging, especially among students who are struggling with the course material and are in need of out of class help to learn. For example, students may lack the motivation, self-regulation, and time-management skills necessary for effective engagement with self-paced learning tools (Bandura, 1997; Zimmerman and Schunk, 2001; Wigfield et al., 2008; Moos and Azevedo, 2008, 2009; Marshman et al., 2017), and their environments may not be conducive to effective engagement with these tools without explicit additional support. Thus, without a critical focus on effective implementation of these tools and sufficient help and incentives to ensure effective engagement with the tools, students may not follow the guidelines for using these tools even if they are research-validated and continuously available *via* the internet. The ineffective use of these self-paced learning tools can significantly reduce their efficacy and impede transfer of learning to new situations.

It is therefore important to investigate how students engage with self-paced learning tools (e.g., in a controlled environment where their interactions with the tool are monitored vs. when they are not monitored) and contemplate strategies that can provide additional support and incentives to students who otherwise may not engage with them as intended. We have been investigating how students engage with optional, web-based tutorials outside of class in introductory physics courses when told that engaging with them would help them with their homework and quizzes (DeVore et al., 2017). These investigations suggest that students often do not engage with the research-validated self-paced tutorials in a manner in which the researchers intended. In particular, many students skimmed through the tutorials or tried to memorize procedures from them without developing a functional understanding of the concepts and a large fraction of students did not engage with them at all. The findings of these investigations motivated us to develop a theoretical framework that holistically takes into account the characteristics of the students and the self-paced learning tools, as well as the environments in which the tools are implemented. The framework can be used as a guide in the development and implementation of self-paced learning tools that encourage all students to engage with them effectively in order to learn.

Goal

The goal of this paper is to first provide an overview of the theoretical framework that focuses on the factors that can impact effective student engagement with the self-paced learning tools and the extent to which students learn from those tools and are able to transfer their learning to new contexts. Then, we report on an investigation in an introductory physics course involving a self-paced learning tutorial on angular momentum and how the findings were interpreted using the framework. In particular, we discuss how introductory physics students who were asked to engage with a research-validated tutorial on the conservation of angular momentum as a self-paced learning tool did not benefit as much as those who used the same tool in a controlled environment and they especially struggled to transfer their learning to a new situation. Then, we summarize the findings of the investigation vis-à-vis the framework that shed light on the aspects of the implementation of research-validated self-paced learning tools that should be critically considered in order to improve their effectiveness.

Overview of the Strategies for Engaged Learning Framework (SELF)

The SELF (see **Figure 1**), is a holistic framework which suggests that, for effective learning from self-paced learning tools, the instructional design and learning tools, their implementation, student characteristics, and social and environmental factors collectively play a role and determine how effectively a majority of students will engage with them (DeVore et al., 2017). The framework consists of four quadrants and posits that all of them must be considered holistically to help students learn effectively. The horizontal dimension involves the characteristics of learning tools and students, both of which should be taken into account when developing and implementing learning tools effectively. The vertical dimension involves internal and external characteristics of the learning tools and the learners. This dimension focuses on how the characteristics of the learning tools and students as well as the environments in which the tools are implemented are important to consider ensuring that students will engage with them effectively and learn from them.

The internal characteristics of the tool pertain to the tool itself (e.g., whether it includes formative assessment). The external characteristics of the tool pertain to how the tool is implemented and incentivized (e.g., whether the tool is framed appropriately to ensure student buy in). The internal characteristics of the students pertain to, e.g., students' prior preparation, motivation, goals, and epistemological beliefs about learning in a particular discipline that can impact their level of engagement with the tools. The external characteristics of the students pertain to social and environmental factors such as support from mentors and balance of coursework. These four factors should be taken into account holistically to develop and implement self-paced learning tools effectively in order to help students learn content,

	Learning Tool Characteristics	Student Characteristics

Internal Characteristics	Factor 1. Learning tool characteristics (internal) – pertaining to features embedded in the learning tools that help students learn
	• Based on “cognitive apprenticeship model” to promote mastery of material for a variety of students
	• Include material providing scaffolding support
	• Involve efficiency and innovation in learning
	• Incorporate elements of productive engagement and productive struggle
	• Involve formative assessment
		Factor 2. Student characteristics (internal)
		• Prior knowledge and skills
		  ○ Prior preparation
		  ○ Cognitive / metacognitive skills
		• Motivational and affective factors
		  ○ Goals
		  ○ Interest and value
		  ○ Self-efficacy
		  ○ Epistemology beliefs
		  ○ Intelligence mindset
		  ○ Grit
		• Self-regulation

External Characteristics	Factor 3. Learning tool characteristics (external) – pertaining to how the tool is implemented in a particular course
	• Embed features to frame the importance of learning from tools and to get student buy in
	• Include motivational features conducive to effective learning during implementation of self-study tools
	• Reinforce learning by coupling learning of different students via creation of learning communities
	• Make explicit connection between in-class lessons and out of class assignments and assessments
	• Incentivize students to engage with tools via grades and other motivational factors
	• Support to help students manage their time better
	• Support to improve students’ self-efficacy and epistemological beliefs
		Factor 4. Student characteristics (external) - pertaining to the student-environment interaction
		• Collaboration skills
		• Balance of coursework and/or work
		• Family encouragement and support
		• Support and mentoring from advisors and counselors
		• Time-management
		  ○ Minimizing unimportant activities that appear urgent (e.g., socializing)
		  ○ Maximizing important activities that may not appear to be urgent (e.g., learning physics)

FIGURE 1 | Strategies for Engaged Learning Framework (SELF). This figure was first presented in DeVore et al. (2017).

develop problem-solving skills, and transfer their learning to new contexts.

Factors 1 and 2: Internal Characteristics of Learning Tools and Students

Factors 1 and 2 of the framework (the internal characteristics of the learning tool and students) are informed by several cognitive theories that point to the importance of knowing students' prior knowledge and difficulties in order to develop effective instructional tools. For example, Hammer's "resource" model suggests that students' prior knowledge and learning difficulties can be used as a resource to help students learn better (Hammer, 1994a,b). Similarly, the Piagetian model of learning emphasizes an "optimal mismatch" between what the student knows and is able to do and the instructional design (Piaget, 1978). In particular, this model focuses on the importance of knowing students' prior knowledge, skills, and difficulties and using this knowledge to design instruction to help them assimilate and accommodate new ideas and build a good knowledge structure. Vygotsky's "zone of proximal development" (ZPD) refers to the zone defined by the difference between what a student can do on his/her own and what a student can do with the help of an instructor who is familiar with his/her prior knowledge and skills (Posner et al., 1982). Scaffolding is a crucial component of this learning model and can be used to stretch students' learning beyond their current knowledge by carefully crafted instruction. Bransford and Schwartz's "preparation for future learning" (PFL) framework suggests that instructional design should include elements of both innovation and efficiency to help students transfer their learning from one context to another (Schwartz et al., 2005). Transfer of learning involves applying knowledge flexibly to new situations other than those in which the knowledge was initially learned and is a hallmark of expertise (Gick and Holyoak, 1983, 1987; Singh, 2008c,d; Nokes-Malach and Mestre, 2013). One interpretation of the PFL model posits that efficiency and innovation can be considered to be two orthogonal dimensions in the instructional design. If instruction only focuses on efficiently transmitting information, cognitive engagement will be diminished and learning will not be effective. On the other hand, if the instruction is solely focused on innovation, students will struggle to connect what they are learning with their prior knowledge and learning and transfer will be inhibited. An appropriate balance of efficiency and innovation builds on students' prior knowledge and difficulties appropriately and helps them decontextualize their learning (i.e., apply their learning in many different contexts), which can facilitate transfer of learning. All of these cognitive theories ("resources," "optimal mismatch," "ZPD," and "PFL" learning models) point to the fact that one must determine the prior knowledge, motivation, and self-regulation of students (Kulik, 1994; Mangels et al., 2006; Hsieh et al., 2007; Sungur, 2007; Fryer and Elliott, 2008; Hulleman et al., 2008; Greene et al., 2010; Song et al., 2016) in order to design effective instruction commensurate with students' current knowledge and skills.

Moreover, instructional design that conforms to the field tested cognitive apprenticeship framework (Collins et al., 1989) can help students learn effectively (see Factor 1). The cognitive apprenticeship model involves three major components: modeling, coaching and scaffolding, and weaning. In this approach, "modeling" means that the instructor demonstrates and exemplifies the skills that students should learn. "Coaching and scaffolding" refer to providing students suitable practice, guidance, and feedback so that they learn the skills necessary for good performance. "Weaning" means gradually fading the support and feedback with a focus on helping students develop self-reliance. Much research in physics education has focused on the cognitive factors in developing effective pedagogical tools and assessment. For example, in physics and other related disciplines, tutorials (Chang, 2001; Singh et al., 2006; Wagner et al., 2006; Singh, 2008b; Zhu and Singh, 2011, 2012a,b, 2013; Brown and Singh, 2015; DeVore and Singh, 2015; Sayer et al., 2015; Singh and Marshman, 2015; DeVore et al., 2016a,b; Marshman and Singh, 2016, 2017a,b,c), peer instruction (clicker questions with peer discussion) (Mazur, 1997), collaborative group problem solving with context-rich physics problems (Heller and Hollabaugh, 1992), POGIL (process-oriented guided-inquiry learning) activities (Farrell et al., 1999), etc. have been found effective in helping students learn (Shaffer and McDermott, 1992; Singh, 2009; Yerushalmi et al., 2012a,b; Stewart et al., 2016; Wood et al., 2016).

Factors 3 and 4: External Characteristics of Learning Tools and Students

We note that instructional tools and student characteristics (Factors 1 and 2) do not exist in a "vacuum." One must also take into account the environment, which includes the setting in which the learning tools are implemented and students' social environments. Factors 3 and 4 of the SELF framework focus on how learning tools are implemented in a particular course and students' environments, respectively. In particular, Factor 4 (the student-environment interaction) can either encourage or discourage effective engagement with learning tools. For example, having supportive parents, teachers, and mentors can be beneficial in fostering students' motivation and engagement with learning tools (Grolnick et al., 2002). Students' time management skills have also been shown to correlate with performance in college (Britton and Tesser, 1991). Students' self-regulation can also either hinder or enhance engagement with self-paced learning tools. In addition, Factor 3 (how learning tools are implemented and incentivized in a course) can affect the extent and manner in which students engage with the learning tools. For example, "framing" instruction to achieve student "buy-in" (e.g., why students should deliberately engage with self-paced learning tools) can help in motivating students to engage with them. Studies have shown that providing to students rationales for why a particular learning activity is worth the effort and why it is useful for them both in the short and long run can help them engage with it more constructively (Deci et al., 1994; Jang, 2008). Motivational researchers also posit that providing stimulating and interesting tasks that are personally meaningful, interesting, relevant, and/or useful to students can increase their interest and value in a subject, increasing their motivation and engagement in learning (Pintrich, 2003). For example, physics education researchers have developed "context-rich" physics problems, i.e., problems that involve "real-world" applications of physics principles and are complex and ill-defined (Heller and Hollabaugh, 1992). These types of problems can often increase students' interest and value associated with physics. Furthermore, instruction that fosters a community of learners can also encourage productive engagement in learning. Within this community of learners, students are encouraged to construct their own knowledge while being held accountable to others, which can, in part, encourage them to engage deeply with the content (Brown, 1997; Engle and Conant, 2002). These types of peer collaborations can be exploited and incentivized when students are learning by engaging with self-paced learning tools.

In sum, the four factors of the SELF Framework can be considered holistically when designing instruction to help students engage effectively with learning tools, including in a self-paced learning environment. We note that each of the factors can interact with other factors. For example, the way in which instructional tools are implemented (Factor 3) is impacted by the characteristics of the learning tool (Factor 1), the characteristics of the students (Factor 2), and the way in which students interact with their social environments (Factor 4). Furthermore, the student characteristics in Factor 2 can inform the learning tool characteristics (Factor 1), the implementation of learning tools (Factor 3), as well as how the student interacts with the environment (Factor 4). Below, we describe a study in which we investigated how students engaged with a self-paced (web-based) tutorial in an introductory physics course and describe the findings of the study in light of the SELF framework.

Research Objectives and Questions

In this study, we investigated how students engage with a research-validated, self-paced introductory physics tutorial on angular momentum conservation in (1) one-on-one interview settings and (2) a large-scale implementation as a self-paced learning tool in a calculus-based introductory physics course at a large research university in the US and interpreted the findings using the framework described in the preceding section. The self-paced tool used in this study was a research-validated web-based tutorial that focused on quantitative problem solving involving angular momentum conservation principle and was designed to aid students *via* a guided inquiry-based approach to learning. In the interview setting, the researchers assured student engagement by requiring them to work through the tutorial in a deliberate manner as prescribed. Students in the large-scale implementation of the tutorial were encouraged to use the self-paced tutorial as preparation for homework and quizzes and had the option to use it as a self-study tool outside of class, so that the researchers did not have control over how the students engaged with it. Student learning was evaluated by their performance on a pre-quiz problem that was identical to the tutorial problem (that either provided scaffolding support or did not—the scaffolding support provided will be detailed later) and a transfer quiz problem (called paired problem) that was comparable to the tutorial problem in that it was an angular momentum conservation problem for introductory physics but was posed in a different context. We compared the performance of students in the large-scale implementation and one-on-one interview settings on the paired quiz problem that required transfer of learning. In particular, the researchers focused on four research questions:

1. In the large scale-implementation of the self-paced tutorial, how does the performance of students who worked through the tutorial compare to the performance of students who did not work through the tutorial on a "pre-quiz" problem that is identical to the tutorial problem?
2. In the large scale-implementation of the self-paced tutorial, how does the performance of students who only worked through an "unscaffolded pre-quiz" problem compare with the performance of the students who worked through the tutorial on a "paired" quiz problem (i.e., is there any difference in the performance of students who worked on the tutorial vs. those who only worked through an "unscaffolded pre-quiz" problem on a follow up transfer problem)?
3. In the large-scale implementation of the self-paced tutorial, how does the performance of students who only worked through a "scaffolded pre-quiz" problem compare with the performance of the students who worked through the tutorial on a "paired" quiz problem (i.e., is there any difference in the performance of students who worked on the tutorial vs. those who only worked through a scaffolded "pre-quiz" problem on a follow up transfer problem)?
4. How does the performance of students who worked through the tutorial in a monitored, one-on-one interview setting compare to the performance of students who worked through the tutorial as a self-paced learning tool in the large-scale implementation on a "paired quiz" problem that involved transfer of learning?

We discuss the findings of these research questions vis-à-vis the holistic framework. The answers to these research questions and their interpretation using the holistic framework can shed light on the characteristics and implementation of the self-paced learning tutorial that resulted in effective or ineffective student engagement and transfer of learning.

METHODOLOGY

Overview

In this investigation involving introductory student engagement with a self-paced tutorial on angular momentum conservation, students were asked to work through a research-validated tutorial in a one-on-one interview situation in which the researchers monitored them and required that they work through the tutorial deliberately while thinking aloud. The same tutorial was also implemented in large, introductory calculus-based physics courses in which students were given the option to use it as a self-paced learning tool outside of class in order to prepare for their homework and quiz on the same content. The tutorial focused on a quantitative problem involving angular momentum conservation and was designed to aid students *via* a guided approach to learning. Student learning was evaluated by their performance on in-class "scaffolded or unscaffolded pre-quiz" problems. The "scaffolded pre-quiz" consisted of the same problem as in the tutorial and broke the problem down into sub-problems in a multiple-choice format. The "unscaffolded pre-quiz" consisted of an open-ended problem identical to the tutorial problem and did not break the problem into sub-problems. In addition, students were assessed on their ability to transfer their learning from the tutorial problem to a "paired" quiz problem, which was similar to the tutorial problem in the underlying physics principles but had different "surface" features (the paired quiz problem was given immediately after the pre-quiz problem).

This study was carried out in accordance with the recommendations of the Human Research Protection Office. The research study and protocols used in the study were reviewed and approved by the University of Pittsburgh Institutional Review Board committee. All interviewed subjects gave written informed consent.

Learning Tools and Assessments Used

The details of the development of the web-based tutorial were reported in a prior study (DeVore et al., 2017). It was guided by the cognitive apprenticeship learning framework (Collins et al., 1989). In this approach, "modeling" implies that the instructor demonstrates and exemplifies the skills that students should learn (e.g., how to solve physics problems systematically). "Coaching" involves providing students opportunities for practice and guidance so that they are actively engaged in learning the skills necessary for good performance. "Weaning" consists of reducing the support and feedback gradually so as to help students develop self-reliance. The web-based tutorial includes modeling *via* breaking the tutorial problem down into sub-problems and including a systematic approach to problem solving. It also involves coaching by providing immediate feedback and support based on students' difficulties. The coaching and scaffolding are adaptive in that the help and guidance provided to students after they answer each multiple-choice sub-problem are tailored to the student's specific difficulty. The adaptive web-based tutorial also involves weaning by gradually providing less scaffolding as student understanding improves and they become more confident in solving the problem on their own. In addition, the tutorial also includes reflection sub-problems that require students to transfer their learning to different contexts and develop self-reliance.

As described in DeVore et al. (2017), similar to other self-paced introductory physics tutorials, the angular momentum conservation tutorial starts with an overarching problem, which is quantitative in nature. **Figure 2** shows the overarching problem for this tutorial. Before working through the tutorial, students are asked to attempt the problem to the best of their ability. The tutorial then divides this overarching problem into a series of sub-problems, which take the form of research-guided, conceptual, multiple-choice questions. These sub-problems help students learn effective approaches for successfully solving a physics problem, e.g., analyzing the problem conceptually, planning and implementing the solution, and reflecting on the final answer and the entire problem-solving process. The alternative

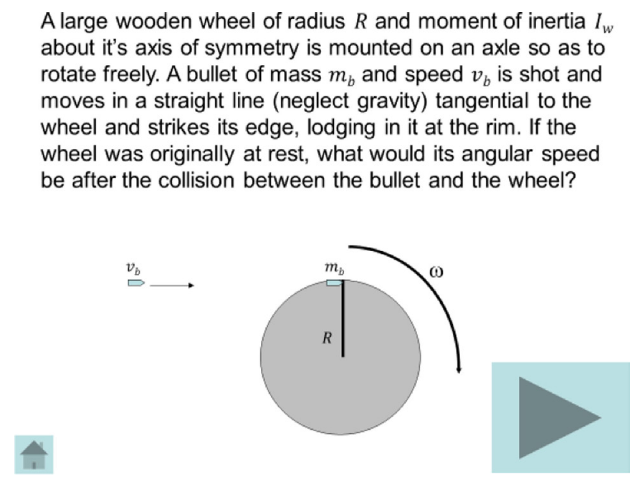

FIGURE 2 | The overarching problem in the Conservation of Angular Momentum web-based tutorial.

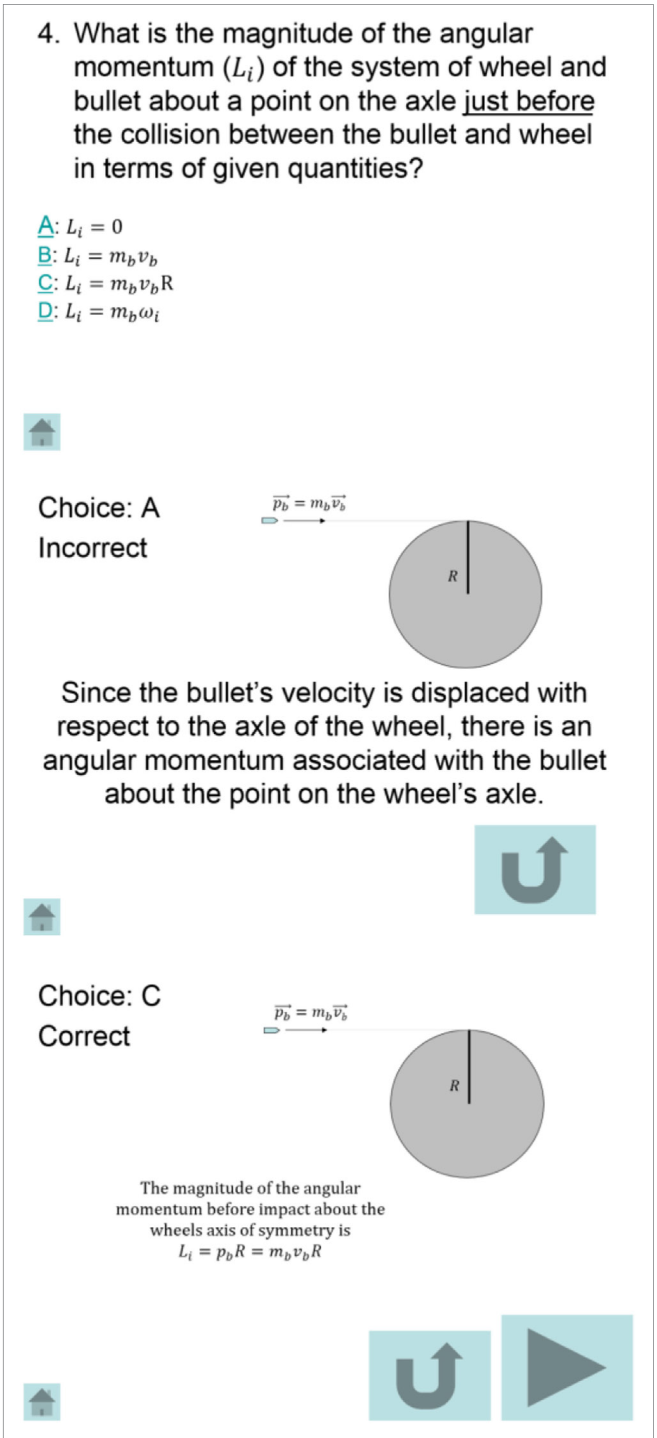

FIGURE 3 | Example of a sub-problem from the tutorial focusing on conservation of angular momentum. Appropriate adaptive support is provided, depending on whether the student selects a certain incorrect or a correct response.

choices in these multiple-choice questions bring out common difficulties students have with the concepts. Incorrect responses direct students to appropriate help sessions in which students are provided suitable feedback and conceptual explanations with diagrams and/or appropriate equations to learn relevant physics concepts. The correct responses to the multiple-choice questions advance students to a brief statement affirming their selection followed by the next sub-problem.

Figure 3 shows an example of a sub-problem in the conservation of angular momentum web-based tutorial and the adaptive feedback provided to students. The top image in **Figure 3** shows the sub-problem in which students are provided an opportunity to determine the magnitude of the initial angular momentum of a particular system. If the students select answer option A (which is incorrect), the adaptive web-based tutorial provides feedback that helps students think about the angular momentum associated with the bullet (middle image in **Figure 3**). If the student selects answer option C (which is correct), the adaptive feedback confirms that the students' answer is correct and gives a reason for why it is correct (bottom image in **Figure 3**).

After students work on other sub-problems, they answer several reflection sub-problems. These reflection sub-problems focus on helping students reflect upon what they have learned and apply the concepts learned in different contexts (to help them decontextualize their learning and promote transfer). If students have difficulty answering the reflection sub-problems, the tutorial again provides adaptive feedback that caters to student difficulties.

The development of the tutorials, of which the tutorial on angular momentum conservation described here is a subset, went through a cyclic, iterative process detailed elsewhere (DeVore et al., 2017). For the angular momentum conservation tutorial problem, three graduate student researchers and one professor (all physics education researchers) performed a cognitive task analysis (Wieman, 2015) to decompose it into a series of sub-problems dealing with different stages of problem solving. Each sub-problem was posed as a multiple-choice question. The incorrect options for each multiple-choice question included common difficulties that were discovered by having introductory physics students solve similar problems in an open-ended format in think-aloud interviews. Explanations for each multiple choice option were written and refined based on one-on-one student interviews to aid students in repairing and extending their knowledge structure when they select an incorrect option. Using this approach, the initial draft of the web-based tutorial was created. The initial draft of the tutorial was revised many times based on interviews with introductory physics students and feedback from graduate students and several professors who worked through it to ensure that they agreed with the wording of the sub-problems and progression of the tutorial. During this revision process, the fine-tuned version of the tutorial was implemented in one-on-one think aloud (Ericsson and Simon, 1993; Chi, 1994) interviews with introductory physics students and were shown to improve student performance on the paired problem that was developed in parallel with the tutorial to assess transfer of learning to a new context (see **Box 1** for the paired problem for the tutorial focusing on conservation of angular momentum discussed here).

The paired quiz problem associated with the tutorial requires the same underlying physics concepts to solve it but is posed in a different context, i.e., it focuses on assessing transfer of learning (Gick and Holyoak, 1983, 1987; Singh, 2002, 2008c,d; Nokes-Malach and Mestre, 2013) from the tutorial problem to a new context (application of the conservation of angular momentum principle is required in order to solve the paired problem-see **Box 1**). The paired problem assesses whether students have learned to de-contextualize the problem solving approach and concepts learned *via* the tutorial. The paired quiz problem is an open-ended problem that is not broken up into sub-problems. This type of a problem can play an important role in the weaning part of the learning model and can assess whether students have developed self-reliance and are able to solve other problems based upon the same underlying concepts as the tutorial problem without any guidance or support.

After students had the opportunity to use the tutorial as a self-paced learning tool as a part of the introductory physics class, a pre-quiz problem was administered immediately followed by a paired quiz problem. While the paired quiz problem was the same for all students, those in some recitation classes were randomly administered the scaffolded version of the pre-quiz problem while those in the other recitation classes were administered unscaffolded (US) version of the pre-quiz problem. The scaffolded pre-quiz consists of multiple-choice questions, structured in the same way as the associated tutorial. In other words, the multiple-choice questions that students answer as part of the scaffolded pre-quiz involve the same questions as the tutorial sub-problems (in the same order as in the tutorial, but students are not provided feedback on whether their choices are correct, unlike the immediate feedback that is available for the sub-problems in the tutorial). Thus, the difference between the tutorial and the scaffolded pre-quiz is that the tutorial provides feedback to students after they choose an answer. On the other hand, the scaffolded pre-quiz offers no such feedback or reinforcement when an answer is selected for each multiple-choice question. The US pre-quiz is identical to the tutorial problem except that it is open-ended—students are provided no additional scaffolding (the problem is not broken into sub-problems).

BOX 1 | Paired problem for the conservation of angular momentum tutorial.

Suppose that a merry-go-round, which can be approximated as a disk, has no one on it, but it is rotating about a central vertical axis at 0.2 revolutions per second. If a 100kg man quickly sits down on the edge of it, what will be its new speed? (A disk of mass m and radius R has a moment of inertia I = (1/2)mR2, mass of merry-go-round = 200kg, radius of merry-go-round = 6m).

Student Demographics and Implementation Approach

Below, we describe the student demographics and methodology for the implementation of the angular momentum tutorial in one-on-one implementation to student volunteers and as a self-paced learning tool in a calculus-based introductory physics course (taken primarily by freshman undergraduate students interested in pursuing engineering or physical science majors) at the University of Pittsburgh, which is a large, typical state-affiliated university in the US.

We first determined whether the tutorial was effective in one-on-one interview settings before implementing and assessing its impact in large introductory physics course as a self-paced learning tool. Twenty 2–3 h long, one-on-one, think-aloud interviews were conducted with students who were enrolled in either an algebra or calculus-based introductory physics course. Approximately half of the students were enrolled in an algebra-based physics course and the other half were enrolled in a calculus-based physics course. These students were paid volunteers who responded to a flyer distributed in the introductory physics classes. They had traditional, lecture-based classroom instruction related to physics concepts covered in the tutorial. The interview data were de-identified so it is not possible to match students' interview data with whether they were enrolled in an algebra-based or calculus-based course. In this deliberate one-on-one interview implementation, students were observed and audio-recorded by a researcher as they worked on the tutorial. The researcher required that the students follow the instructions for working through the tutorial. For example, students were first asked to outline the solution to the tutorial problem to the best of their ability before they started the tutorial. They were required to answer each sub-problem in the appropriate order. Throughout this one-on-one implementation process, each student was asked to think aloud so that the researcher could understand their thought processes and the researcher recorded observations of each student's interaction with the web-based tool. The researcher remained silent while the students worked and only prompted them to keep talking if they remained silent for a long time. After working through the tutorial, the students worked on the paired problem. Students in the one-on-one interview situation spent between 15 and 30 min on the tutorial. All students had enough time to finish working through the tutorial and the paired problem.

After we found that the tutorial was effective in one-on-one interview settings, it was then implemented as a self-paced learning tool as part of a large, calculus-based introductory physics course. The course was a first semester physics course with 220 students (split into two sections). These students came from varied backgrounds with a majority of them pursuing engineering or physical science degrees. Approximately 60% of the students had taken a high school calculus course and were concurrently enrolled in a college level calculus course. On average, the students were between 18 and 19 years of age. The course was comprised of ~30% females. We note that none of the students in the course had visual disabilities. This implementation allowed the researchers to determine the effectiveness of the tutorial for students in a typical introductory physics course at a research university in the US in which researchers had no control over how the tutorial was used by the students as a self-paced learning tool. **Table 1** shows the sequence of the self-study tool activities and recitation quizzes in the introductory physics course.

The tutorial was posted on the course website after students had received classroom instruction in relevant concepts. It could be used at a time convenient to each student, but the amount of time each student spent working through it could not be tracked. The tutorial and associated homework problems were assigned in the same week. Instructors incentivized the self-study tutorial by telling students that the tutorial would be helpful for solving assigned homework problems and in-class quiz problems (scaffolded and US pre-quiz problems and paired quiz problems) for that week. Although students were made aware that no points would be awarded simply for completing the web-based tutorial, announcements were made in class, posted on the course website, and sent via email informing students that the tutorial was available after relevant concepts were covered in class.

The scaffolded or US pre-quiz and paired problem were administered in a recitation class in the following week after students had access to the associated web-based tutorial to use as a self-paced learning tool for an entire week. Students in different recitation classes were randomly assigned to either an US or scaffolded pre-quiz condition. Immediately after students submitted the solution to the pre-quiz problem (that was identical to the tutorial problem), they were given the corresponding paired quiz problem (a problem that involves the same physics principles as the tutorial problem shown in **Box 1**). All students had sufficient time to complete the tutorial and quizzes. Students were given a grade based on their performance on the pre-quiz and paired quiz problems as their weekly quiz grade. On top of each sheet with the paired problem quizzes that were administered in the recitation classes, students were asked questions such as whether they had worked through the online tutorial, whether they thought the tutorial was effective at helping them solve the problem, and how much time they spent on the tutorial (they were told that their answers to these questions would not influence their score on the quiz). These questions allowed us to separate students into "tutorial" or "non-tutorial" groups and determine the performance of students who engaged with the tutorial on the pre-quizzes and paired quizzes.

TABLE 1 | Sequence of activities involving the self-study tool in a calculus-based introductory physics course and number of students (*N*) in each group.

In-class	Outside of class	In-class recitation	In-class recitation
Traditional instruction in relevant topics	Worked on tutorial (*N* = 128)	Scaffolded pre-quiz problem (multiple-choice tutorial sub-problems) (*N* = 61)	Paired quiz problem (open-ended transfer problem) (*N* = 200)
		Unscaffoled pre-quiz problem (open-ended tutorial problem) (*N* = 67)	
	Did not work on tutorial (*N* = 74)	Scaffolded pre-quiz problem (multiple-choice tutorial sub-problems) (*N* = 31)	
		Unscaffolded pre-quiz problem (open-ended tutorial problem) (*N* = 43)	

The number of students who worked on the pre-quiz problem and paired quiz problem does not match because not all students took the paired quiz problem.

The purpose of administering the pre-quiz was twofold. First, we wanted to examine whether students who worked through the tutorial as a self-paced learning tool were able to solve the same tutorial problem successfully in a quiz setting without the adaptive support of the tutorial (catering to specific student difficulties). We note that the scaffolded pre-quiz involved breaking the tutorial problem down into sub-problems, but there was no adaptive feature in the pre-quiz and students were not given any feedback or support if they selected an incorrect response to any of the sub-problems on the pre-quiz. The US pre-quiz consisted of an open-ended problem identical to the tutorial and did not break the problem into sub-problems and gave no adaptive support to students. Thus, the pre-quizzes allowed us to examine, in part, how effectively the students engaged with the self-study tutorial (but not necessarily the extent to which they could transfer their learning) by evaluating their performance on the pre-quiz problems that were identical to the tutorial problem with and without scaffolding (and without adaptive support that the tutorial provided). The second purpose of giving the pre-quiz was to compare the performance of students who worked through the tutorial with those who only worked on the corresponding scaffolded pre-quiz (but did not work through the tutorial) on the paired quiz problems. The pre-quizzes enabled us to evaluate whether students who worked on only the pre-quizzes performed better or worse than those who engaged with the tutorial as a self-study tool on paired problems (which were transfer problems involving the same underlying concepts). In this way, we were able to investigate, in part, whether students who worked through the self-paced tutorial engaged with it effectively and were able to transfer their learning to a new context (as opposed to students who only worked through a scaffolded pre-quiz that did not include adaptive learning support).

To compare the performance of students who worked on the tutorial in a one-on-one interview setting with those who used it as a self-study tool, we examined student performance on the paired problem in these two settings. Three graduate students and a professor who do research in physics education iteratively developed a rubric for the paired problem. Once the final version of the rubric was agreed upon, 10% of the paired problem quizzes were graded independently. When the scores were compared, the inter-rater agreement was better than 90% across the three graduate students and professor. In this way, we were able to examine, in part, the level of student engagement with the self-study tutorial in a one-on-one implementation and a large-scale, self-study implementation.

RESULTS

In this section, we refer to students who worked through the tutorial as group "T" and students who did not work through the tutorial as group "NT." For the different scaffolding pre-quiz conditions, students who worked through a scaffolded pre-quiz are referred to as the "S" group and students who worked through an US pre-quiz are referred to as the "NS" group. Students who were in the one-on-one interview condition are referred to as the "INT" group.

In regards to research question 1 (In the large scale-implementation of the self-paced tutorial, how does the performance of students who worked through the tutorial compare to the performance of students who did not work through the tutorial on "pre-quiz" problems that are identical to the tutorial problem?), we focus on students' performance on the pre-quiz problems shown in **Table 2**. For the scaffolded and US pre-quiz in the large-scale implementation of the tutorial in calculus-based physics courses, **Table 2** shows that the T group performed better than the NT group on the pre-quiz that involved the same problem as the tutorial problem. A *t*-test indicated that the T group performed significantly better than the NT group on the scaffolded pre-quiz ($p < 0.001$) and the US pre-quiz ($p < 0.001$). This finding suggests that the tutorial was beneficial for helping students be able to at least reproduce the solution to the same problem as the tutorial problem, whether the problem was broken into multiple-choice sub-problems (scaffolded pre-quiz) or an open-ended format (US pre-quiz).

Regarding research question 2 (In the large-scale implementation of the self-paced tutorial, how does the performance of students who only worked through an unscaffolded "pre-quiz" problem compare with the performance of the students who worked through the tutorial on a "paired" quiz problem?), we discuss the students' performance on the paired problems that required transfer of the learning from the tutorial. **Table 2** shows that in the large-scale implementation of the tutorial in calculus-based physics courses, students in the T + US group had an average score of 67.7% on the paired problem that required transfer of learning. On the other hand, students in the NT + US group had an average score of 50.1% on the paired problem. A *t*-test revealed that students in the T + US group performed significantly better than students in the NT + US group on the paired problem ($p = 0.002$). This finding suggests that students who worked through the tutorial that included scaffolding support performed better on the paired quiz transfer problem than students who had not been given any scaffolding support *via* the tutorial or pre-quiz problem.

In regards to research question 3 (In the large-scale implementation of the self-paced tutorial, how does the performance of students who only worked through a scaffolded "pre-quiz" problem compare with the performance of the students who worked through the tutorial on a "paired" quiz problem?), **Table 2** shows that students in the NT + S group had an average of 56.3% on the paired quiz problem involving transfer of learning. Students in the T + S group had an average of 66.5% on the paired quiz problem involving transfer of learning. Students in the T + US group had an average of 67.7% on the paired quiz involving transfer of learning, which was not statistically significantly different from the average score of the students in the T + S group on the paired problem (66.5%). A *t*-test revealed that the average score of the students in the NT + S group on the paired problem (56.3%) is not statistically significantly different from the average score of the students in the T + S group (66.5%) ($p = 0.124$), nor is it statistically significantly different from the average score of the students in the T + US group (67.7%) ($p = 0.798$). It appears that students who worked through the tutorial that included both scaffolding support and adaptive features did not perform significantly better on the paired problem that required transfer of learning than students who had only worked through a pre-quiz that included scaffolding support. This finding suggests that students who stated that they worked through the tutorial may not have taken advantage of the adaptive features of the tutorial to help them transfer their learning to new contexts.

We now discuss findings related to research question 4 (How does the performance of students who worked through the tutorial in a monitored, one-on-one interview setting compare to the performance of students who worked through the tutorial as a self-paced learning tool in the large-scale implementation

TABLE 2 | Student performance on scaffolded (S) and unscaffolded (US) pre-quizzes and paired quizzes for the conservation of angular momentum tutorial with SDs for the tutorial group (T), non-tutorial group (NT), and one-on-one interview group (INT).

	Prequiz average score (SD)		Paired problem average score (SD)		
	Scaffolded prequiz (S)	**US prequiz**	**Scaffolded prequiz (S)**	**US prequiz**	**No prequiz**
T	95.8% (14.9%) $N = 61$	82.0% (32.7%) $N = 67$	66.5% (25.6%) $N = 60$	67.7% (26.8%) $N = 66$	X
NT	75.6% (23.8%) $N = 31$	34.4% (40.4%) $N = 43$	56.3% (31.3%) $N = 31$	50.1% (29.5%) $N = 43$	X
INT	X	X	X	X	83.3% (16.0%) $N = 22$

The number of students in each group is denoted with N.

on "paired quiz" problems that involved transfer of learning?). We found that students in the INT group performed significantly better on the paired problem than the students in the T group. **Table 2** shows that the students in the INT group had an average of 83% on the paired problem (compared to ~67% for the students in the T group). We note that the INT group was comprised of students in algebra-based and calculus-based physics courses, but we were unable to separate out the scores of students in algebra-based and calculus-based physics courses because the data were de-identified without separating them. However, the SDs of the scores of the INT group on the paired problems are small compared to the SDs of the scores of the T group (the SD on the paired problem for the INT group is 16%, compared to 25.6% in the T group). Thus, it appears that students in the INT group (i.e., students in both algebra-based and calculus-based physics courses) had comparable scores on the paired problems, and so a comparison between the INT group and the calculus-based T group is appropriate. Moreover, prior research suggests that students in the algebra-based introductory physics courses on average perform worse than those in the calculus-based introductory physics (DeVore et al., 2017). Therefore, if we had not de-identified the data and could separate the interview group into two sub-groups (algebra-based vs. calculus-based), the average scores of the calculus-based group would most likely be higher than 83%. In sum, students who worked through the tutorial as a self-study tool in the large-scale implementation of the tutorial performed significantly worse compared to the students who engaged with the tutorial in one-on-one interview settings on the paired quiz problem that required transfer of learning.

DISCUSSION AND INTERPRETATION OF FINDINGS IN TERMS OF THE SELF FRAMEWORK

Our findings suggest that introductory physics students who reported that they worked through the self-paced tutorial on angular momentum conservation performed better than those who did not on pre-quiz problems (even in the unscaffolded version when no support was provided) that were identical to the tutorial problem. This implies that students who reported working through the tutorial were better at reproducing the solution of the tutorial problem than those who did not work through the tutorial. Furthermore, students who worked through the tutorial performed better than those who only worked through an unscaffolded pre-quiz but did not work through the tutorial on the paired problem involving transfer of learning. However, overall, students who worked through the tutorial struggled on the transfer problem and did not perform significantly better than those who only worked through a scaffolded pre-quiz on the paired problem. We also found that students who worked on the tutorial in a one-on-one setting performed significantly better on the paired quiz problem that required transfer of learning than those who worked on the tutorial in the large-scale implementation as a self-paced learning tool.

Our findings suggest that many students who worked through the tutorial without supervision in the large-scale implementation may not have engaged with it in an effective manner. While the students who took advantage of the tutorial performed better on the paired problem than those who worked only through an unscaffolded pre-quiz, the students in the tutorial group did not perform significantly better on the paired problem than those who worked only through a scaffolded pre-quiz. This finding indicates that students may have benefited somewhat from the scaffolding support from the tutorial (i.e., when the tutorial problem was broken down into sub-problems) on the paired problems, but they may not have taken full advantage of the adaptive features of the tutorial that were meant to help them repair their knowledge structure and transfer their learning to new contexts. Furthermore, students who engaged with the tutorial as a self-study tool in the large-scale implementation of the tutorial performed significantly worse than those who worked through the tutorial in the one-on-one interview setting on the paired problem. The students in the one-on-one interview setting were required to work through the tutorial in a deliberate and engaged manner, and they performed well on the paired quiz that involved transfer of learning from the tutorial. This finding indicates that the tutorial was effective in helping students learn physics concepts and transfer their learning to new situations when students engaged with it in a deliberate manner. However, students who used the tutorial as a self-study tool in a large-scale implementation without supervision did not perform as well as those in the one-on-one interview settings on the paired transfer problems, indicating that they may not have taken advantage of the adaptive features of the tutorial in a deliberate and engaged manner. This dichotomy between the performance of the self-study group and one-on-one implementation group on the paired problem suggests that a carefully designed tutorial, when used as intended, can be a powerful learning tool for introductory physics students across diverse levels of prior preparation. However, ensuring that students engage with it effectively as a self-paced learning tool can be challenging.

The significantly worse performance of the tutorial group on the paired quiz problem in the large-scale, self-study implementation (compared to the one-on-one implementation group) may be due to the fact that students engaged superficially with the tutorial. Although these students were given explicit instructions on how to work through the tutorial effectively, they could have taken short cuts and skipped sub-problems if they decided not to adopt a deliberate learning approach while using the web-based tool. Indeed, based upon student comments and other data gathered with their responses to the paired problem in the self-study group, some students explicitly commented that they "skimmed" or "looked over" the tutorial but that type of engagement with the adaptive web-based tool may not help them learn deeply and transfer their learning in order to apply the concepts learned to new situations. Additionally, they may not have attempted to first solve the tutorial problem on their own without the scaffolding provided by the web-based tutorial (even though explicitly told to do so), even though this step would have allowed them to productively struggle with the problem and prime them to learn from the tutorial (Kapur, 2008; Clark and Bjork, 2014). We note that even some of the students in one-on-one interviews needed to be prompted several times to make a prediction for each sub-problem rather than randomly guessing an answer. Furthermore, the written responses of students who used the tutorial as a self-study tool on the paired problem suggest that many of them may have memorized a few equations by skimming through the tutorial. These students may have expected that those equations would help them in solving the paired quiz problem instead of engaging with the self-paced tool in a systematic manner. Interestingly, in a survey given at the end of the course, a majority of students who claimed that they had used the self-paced tutorial stated that they thought that it was helpful. However, their performance on the paired problem reflected that they had not learned effectively from it.

The findings of this study can be interpreted in terms of the SELF framework. We note that the development of the self-paced tutorial discussed here was based upon a cognitive task analysis of the underlying concepts, built on students' prior knowledge and difficulties, and drew upon cognitive learning theories (i.e., Factors 1 and 2 of the SELF framework). However, we found that the tutorial was not as effective in helping students transfer their learning when implemented as a self-paced learning tool in a large physics course (in which students' engagement with the tutorial was unsupervised) compared to supervised, one-on-one interview situations, as measured by their performance on a transfer problem. This dichotomy in students' performance on the problem involving transfer of learning supports the notion in the SELF framework that a major challenge in effectively implementing a research-validated interactive tutorial as a self-study tool is likely to be related to how it was implemented and incentivized (whether students had sufficient incentives to effectively engage with it in a self-paced learning environment), whether students had the motivation, self-regulation, and time-management skills to engage with it and how the constraints of the social environments impacted their engagement (i.e., Factor 4 of the SELF framework) (Ericsson et al., 1993; Winne, 1996; Pintrich, 2003; Narciss et al., 2007; Mason and Singh, 2010; Brown et al., 2016). It appears that without sufficient support to help students develop self-management and time-management skills and providing incentives to motivate students to engage with the self-paced tutorial, many students may not have effectively engaged with it. In particular, the SELF framework supports that haphazard use of a research-validated self-paced tool can reduce its effectiveness significantly and inhibit transfer of learning to new contexts as we found in this investigation. Therefore, it is important for educators and education researchers to contemplate how to provide appropriate incentives and support in order for students to engage effectively and benefit from self-paced learning tools.

While students' environments are challenging to account for when implementing self-paced learning tools, Factor 3 of the SELF framework focuses on incentives and support students can be provided during the implementation of the self-paced learning tools to improve their level of engagement. Factor 3 posits that the external characteristics of the tools (i.e., how the tools are implemented in classes) may improve student engagement with the tools by taking into account students' characteristics and environments. In our study, students may have engaged more effectively with the self-paced tutorial if elements from Factor 3 were included in the implementation of the self-study tools. For example, it may be helpful to get student buy-in by having students think carefully about why they should engage deliberately with a self-paced learning tool. Students who struggle with managing their time can be provided guidance in making a daily schedule that includes enough time for learning from self-paced tools. Additionally, the instructor can strive to make connections between self-paced learning assignments and other in-class lessons or out of class activities and assessments to help students engage with the self-paced learning tools more effectively.

Moreover, some students may not have engaged effectively with the self-paced tutorial in our study due to self-efficacy issues or unproductive beliefs about learning as suggested in the SELF framework. In particular, students who have low self-efficacy (Bandura, 1997; Moos and Azevedo, 2009) and/or unproductive epistemological beliefs about learning in a particular discipline such as physics (e.g., physics is just a collection of facts and formulas, only a few smart people can do physics, and that learning physics involves memorizing physics formulas and reiterating them on exams) (Hammer, 1994a,b; Redish et al., 1998; Maries et al., 2016) are unlikely to productively engage with the self-paced learning tools. Therefore, students who have difficulty engaging with the self-study tools due to lack of self-efficacy or unproductive epistemological beliefs can be helped to improve their self-efficacy and develop productive epistemological beliefs. For example, a short online intervention has been shown to improve student self-efficacy (Mangels et al., 2006). These issues are important to address in order to ensure that students who are most in need of learning from self-paced learning tools benefit from them and can flexibly transfer their learning to new situations (Tinto, 1993, 1997; Braxton, 2000; Braxton et al., 2004; Herzog, 2005; Diaz and Cartnal, 2006; Anderson, 2011; Boston and Ice, 2011; Boston et al., 2011; DeAngelo et al., 2011; Campbell and Mislevy, 2013).

Moreover, the SELF framework proposes that another factor that may help students effectively engage with self-paced learning tools is encouraging them to engage in learning communities. In these learning communities, all students would be expected to learn from the self-study tools and then engage in some follow up activities in a group environment (either electronically or physically, depending on the class). Thus, individual students are accountable to their group members and are encouraged to engage with self-study assignments and activities deliberately to prepare for the group activities. For example, in the study discussed here, if students were assigned to work in a learning community on a complex physics problem after engaging with the self-paced learning tool, they may have had more incentive to engage deeply with the tutorial individually in order to prepare for the group work.

Furthermore, incorporating grade incentives (Morrison et al., 1995; Li et al., 2015) to engage with the self-study tool is another factor that can increase student engagement (see Factor 3 of the framework). For example, instructors can give course credit to students based on their answers to each sub-problem with decreasing scores if they answer the same sub-problem multiple times. This strategy might encourage students to answer each sub-problem carefully instead of guessing the answers. In addition, if students in the study described here were asked to submit a copy of the correct answer to each sub-problem and explain their reasoning, this practice may have increased their motivation to deliberately engage with the self-study tool and help them transfer their learning to new contexts.

It is also important to note that we cannot disentangle any of the factors in the SELF framework and how they impact student engagement with a self-paced learning tool. For example, students who are lacking prior preparation (factor 2 of the framework) may also have difficulty with time-management (Factor 4 of the framework). Furthermore, when students work in learning communities that keep each student accountable while providing mutual support (see Factor 3 of the framework), they may manage their time better (see Factor 4 of the framework). To help students learn effectively from the self-study tools and transfer their learning to new contexts, the characteristics of the learning tools (Factors 1 and 3) should take into account the characteristics of the students and their environments (Factors 2 and 4). In particular, Factor 3, which is often ignored by educators who develop and/or implement self-study tools, is a critical aspect of ensuring that students effectively engage with self-paced learning tools and learn to transfer their learning to new situations.

Our investigation suggests that, despite the ease and convenience of accessing adaptive, self-paced tutorials, there are challenges in ensuring that students, especially those who in need of out-of-class scaffolding support, engage with them effectively. Even well-designed self-study tools that take into account students' prior knowledge and difficulties and are based on cognitive learning theories may not necessarily help students transfer their learning if students do not engage with them effectively due to motivational or environmental factors. Thus, if a learning tool is aimed at improving students' transfer of learning, deep engagement is necessary and it is crucial that the developers and implementers of the self-paced tool appropriately take into account factors such as students' motivation, self-regulation and time-management skills, and social environments (i.e., Factors 2 and 4 of the SELF framework). It is possible that if students are provided supports and incentives such as those in Factor 3 of the SELF framework, they may engage with research-validated self-study tool such as the tutorial described here more effectively and their transfer of learning to new contexts will be improved.

ETHICS STATEMENT

We obtained IRB approval to work with human subjects.

AUTHOR CONTRIBUTIONS

Each other contributed equally to the article.

ACKNOWLEDGMENTS

We thank the US National Science Foundation for financial support (grant numbers PHY-1505460 and DUE-1524575) and members of the University of Pittsburgh physics education research group for their help with the rubric development and grading to establish inter-rater reliability.

FUNDING

US National Science Foundation for financial support (grant numbers PHY-1505460 and DUE-1524575).

REFERENCES

Allen, I., and Seaman, J. (2013). *Changing Course: Ten Years of Tracking Online Education in the U.S. Sloan Consortium*. Available at: http://www.onlinelearningsurvey.com/reports/changingcourse.pdf

Alraimi, K., Zo, H., and Ciganek, A. (2015). Understanding the MOOCs continuance: the role of openness and reputation. *Comput. Educ.* 80, 28. doi:10.1016/j.compedu.2014.08.006

Anderson, K. (2011). *ProQuest (UMI No. 3458241, 2011)*. Ph.D. thesis, Iowa State University.

Azevedo, R. (2005). Special issue on computers as metacognitive tools for enhancing student learning, computer environments as metacognitive tools for enhancing learning. *Educ. Psychol.* 40, 193. doi:10.1207/s15326985ep4004_1

Azevedo, R., Cromley, J., Winters, F., Moos, D., and Greene, J. (2005). Special issue on scaffolding self-regulated learning and metacognition: implications for the design of computer-based scaffolds, adaptive human scaffolding facilitates adolescents' self-regulated learning with hypermedia. *Instr. Sci.* 33, 381. doi:10.1007/s11251-005-1273-8

Azevedo, R., Guthrie, J., and Seibert, D. (2004). The role of self-regulated learning in fostering students' conceptual understanding of complex systems with hypermedia. *J. Educ. Comput. Res.* 30, 87. doi:10.2190/DVWX-GM1T-6THQ-5WC7

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York, NY: W.H. Freeman.

Boston, W., and Ice, P. (2011). Assessing retention in online learning: an administrative perspective. *Online J. Dist. Learn. Admin.* 14, 2.

Boston, W., Ice, P., and Gibson, A. (2011). Comprehensive assessment of student retention in online learning environments. *Online J. Dist. Learn. Admin.* 14, 1.

Bower, M., Dalgarno, B., Kennedy, G., Lee, M., and Kenney, J. (2015). Design and implementation factors in blended synchronous learning environments: outcomes from a cross-case analysis. *Comput. Educ.* 86, 1. doi:10.1016/j.compedu.2015.03.006

Braxton, J. (2000). *Reworking the Student Departure Puzzle*. Nashville, TN: Vanderbilt University Press.

Braxton, J., Hirschy, A., and McClendon, S. (2004). *Understanding and Reducing College Student Departure, ASHE-ERIC Higher Education Report*, Vol. 30. San Francisco, CA: Jossey-Bass.

Breslow, L., Pritchard, D., DeBoer, J., Stump, G., Ho, A., and Seaton, D. (2013). Studying learning in the worldwide classroom research into edX's first MOOC. *Res. Pract. Assess.* 8, 13.

Britton, B., and Tesser, A. (1991). Effects of time-management practices on college grades. *J. Educ. Psychol.* 83, 405–410. doi:10.1037/0022-0663.83.3.405

Brown, A. (1997). Transforming schools into communities of thinking and learning about serious matters. *Am. Psychol.* 52, 399–413. doi:10.1037/0003-066X.52.4.399

Brown, B., Mason, A., and Singh, C. (2016). Improving performance in quantum mechanics with explicit incentives to correct mistakes. *Phys. Rev. Phys. Educ. Res.* 12, 010121. doi:10.1103/PhysRevPhysEducRes.12.010121

Brown, B., and Singh, C. (2015). "Development and evaluation of a quantum interactive learning tutorial on Larmor precession of spin," in *Proceedings of the Physics Education Research Conference 2014*, eds P. Engelhardt, A. Churukian, and D. Jones (Minneapolis, MN). doi:10.1119/perc.2014.pr.008

Campbell, C., and Mislevy, J. (2013). Students' perceptions matter: early signs of undergraduate student retention/attrition. *J. Coll. Stud. Ret.* 14, 467. doi:10.2190/CS.14.4.c

Chandra, V., and Watters, J. (2012). Re-thinking physics teaching with web-based learning. *Comput. Educ.* 58, 631. doi:10.1016/j.compedu.2011.09.010

Chang, C. (2001). A problem-solving based computer-assisted tutorial for the earth sciences. *J. Comput. Assist. Learn.* 17, 263–274. doi:10.1046/j.0266-4909.2001.00181.x

Chen, Z., and Gladding, G. (2014). How to make a good animation: a grounded cognition model of how visual representation design affects the construction of abstract physics knowledge. *Phys. Rev. ST Phys. Educ. Res.* 10, 010111. doi:10.1103/PhysRevSTPER.10.010111

Chen, Z., Stelzer, T., and Gladding, G. (2010). Using multi-media modules to better prepare students for introductory physics lecture. *Phys. Rev. ST Phys. Educ. Res.* 6, 010108. doi:10.1103/PhysRevSTPER.6.010108

Chi, M. (1994). "Thinking aloud," in *The Think Aloud Method*, eds M. W. van Someren, Y. F. Barnard, and J. A. C. Sandberg (London: Academic Press).

Chi, M., Slotta, J., and de Leeuw, N. (1994). From things to processes: a theory of conceptual change for learning science concepts. *Learn. Instruct.* 4, 27–43. doi:10.1016/0959-4752(94)90017-5

Clark, C. M., and Bjork, R. A. (2014). "When and why introducing difficulties and errors can enhance instruction," in *Applying the Science of Learning in Education: Infusing Psychological Science into the Curriculum*, eds V. A. Benassi, C. E. Overson, and C. M. Hakala. Available at: www.teachpsych.org/ebooks/asle2014/index.php

Collins, A., Brown, J., and Newman, S. (1989). "Cognitive apprenticeship: teaching the crafts of reading, writing and mathematics," in *Knowing, learning, and instruction: Essays in honor of Robert Glaser*, ed. L. B. Resnick (Hillsdale, NJ: Lawrence Erlbaum), 453–494.

Colvin, K., Champaign, J., Liu, A., Zhou, Q., Fredericks, C., and Pritchard, D. (2014). Learning in an introductory physics MOOC: all cohorts learn equally, including an on-campus class. *Int. Rev. Res. Open Dist. Learn.* 15, 263.

DeAngelo, L., Franke, R., Hurtado, S., Pryor, J., and Tran, S. (2011). *Completing College: Assessing Graduation Rates at Four-Year Institutions*. Los Angeles, CA: UCLA Higher Education Research Institute.

Debowska, E., Girwidz, R., Greczyło, T., Kohnle, A., Mason, B., Mathelitsch, L., et al. (2012). Report and recommendations on multi-media materials for teaching and learning electricity and magnetism. *Eur. J. Phys* 34, L47–L54. doi:10.1088/0143-0807/34/3/L47

Deci, E., Eghrari, H., Patrick, B., and Leone, D. (1994). Facilitating internalization: the self-determination theory perspective. *J. Pers.* 62, 119–142. doi:10.1111/j.1467-6494.1994.tb00797.x

Demetriadis, S., Papadopoulos, P., Stamelos, I., and Fischer, F. (2008). The effect of scaffolding students' context-generating cognitive activity in technology-enhanced case-based learning. *Comput. Educ.* 51, 939. doi:10.1016/j.compedu.2007.09.012

DeVore, S., Gauthier, A., Levy, J., and Singh, C. (2016a). Development and evaluation of a tutorial to improve students' understanding of a lock-in amplifier. *Phys. Rev. Phys. Educ. Res.* 12, 020127. doi:10.1103/PhysRevPhysEducRes.12.020127

DeVore, S., Gauthier, A., Levy, J., and Singh, C. (2016b). Improving student understanding of lock-in amplifiers. *Am. J. Phys.* 84, 52. doi:10.1119/1.4934957

DeVore, S., Marshman, E., and Singh, C. (2017). Challenge of engaging all students via self-paced interactive electronic learning tutorials for introductory physics. *Phys. Rev. PER* 13, 010127. doi:10.1103/PhysRevPhysEducRes.13.010127

DeVore, S., and Singh, C. (2015). "Development of an interactive tutorial on quantum key distribution," in *Proceedings of the Physics Education Research Conference 2014*, eds P. Engelhardt, A. Churukian, and D. Jones (Minneapolis, MN). doi:10.1119/perc.2014.pr.011

Diaz, D., and Cartnal, R. (2006). Term length as an indicator of attrition in online learning. *Innov. J. Online Educ.* 2, 5.

Engle, R., and Conant, F. (2002). Guiding principles for fostering productive disciplinary engagement: explaining an emergent argument in a community of learning classroom. *Cogn. Instruct.* 20, 399–483. doi:10.1207/S1532690XCI2004_1

Ericsson, A., Krampe, R., and Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychol. Rev.* 100, 363. doi:10.1037/0033-295X.100.3.363

Ericsson, A., and Simon, H. (1993). *Protocol Analysis: Verbal Reports as Data, revised ed*. Cambridge, MA: MIT Press.

Farrell, J., Moog, R., and Spencer, J. (1999). A guided inquiry chemistry course. *J. Chem. Educ.* 76, 570–574. doi:10.1021/ed076p570

Fryer, J., and Elliott, A. (2008). "Self-regulation of achievement goal pursuit," in *Motivation and Self-Regulated Learning: Theory, Research and Applications*, eds D. H. Schunk and B. J. Zimmerman (New York, NY: Lawrence Erlbaum), 53.

Gick, M., and Holyoak, K. (1983). Schema induction and analogical transfer. *Cog. Psy.* 15, 1–38. doi:10.1016/0010-0285(83)90002-6

Gick, M., and Holyoak, K. (1987). "The cognitive basis of knowledge transfer," in *Transfer of Learning: Contemporary Research and Applications*, eds S. Cormier and J. Hagman (San Diego, CA: Academic Press), 9–46.

Greene, J., Costa, I., Robertson, J., Pan, Y., and Deekens, V. (2010). Exploring relations among college students' prior knowledge, implicit theories of intelligence, and self-regulated learning in a hypermedia environment. *Comput. Educ.* 55, 1027. doi:10.1016/j.compedu.2010.04.013

Grolnick, W., Gurland, S., Jacob, K., and Decourcey, W. (2002). "The development of self-determination in middle childhood and adolescence," in *Development of Achievement Motivation*, eds A. Wigfield and J. Eccles (San Diego, CA: Academic Press), 147–171.

Hammer, D. (1994a). Epistemological beliefs in introductory physics. *Cogn. Instruct.* 12, 151. doi:10.1207/s1532690xci1202_4

Hammer, D. (1994b). Students' beliefs about conceptual knowledge in introductory physics. *Int. J. Sci. Educ.* 16, 385. doi:10.1080/0950069940160402

Heller, P., and Hollabaugh, M. (1992). Teaching problem solving through cooperative grouping. Part 2: designing problems and structuring groups. *Am. J. Phys.* 60, 637. doi:10.1119/1.17118

Herzog, S. (2005). Measuring determinants of student return vs. dropout/stopout vs. transfer: a first-to-second year analysis of new freshmen. *Res. High. Educ.* 46, 883. doi:10.1007/s11162-005-6933-7

Hsieh, P., Sullivan, J., and Guerra, N. (2007). A closer look at college students: self-efficacy and goal orientation. *J. Adv. Acad.* 18, 454. doi:10.4219/jaa-2007-500

Hulleman, C., Durik, A., Schweigert, S., and Harachiewicz, J. (2008). Task values, achievement goals, and interest: an integrative analysis. *J. Educ. Psychol.* 100, 398. doi:10.1037/0022-0663.100.2.398

Jang, H. (2008). Supporting students' motivation, engagement, and learning during an uninteresting activity. *J. Educ. Psychol.* 100, 798–811. doi:10.1037/a0012841

Kapur, M. (2008). Productive failure. *Cognit. Instruct.* 2, 379–424. doi:10.1080/07370000802212669

Kohnle, A., Douglass, M., Edwards, T. J., Gillies, A. D., Hooley, C. A., and Sinclair, S. D. (2010). Developing and evaluating animations for teaching quantum mechanics concepts. *Eur. J. Phys.* 31, 1441. doi:10.1088/0143-0807/31/6/010

Korkmaz, A., and Harwood, W. (2004). Web-supported chemistry education: design of an online tutorial for learning molecular symmetry. *J. Sci. Educ. Technol.* 13, 243–253. doi:10.1023/B:JOST.0000031263.82327.6e

Kulik, C., and Kulik, J. (1991). Effectiveness of computer-based instruction: an updated analysis. *Comput. Hum. Behav.* 7, 75. doi:10.1016/0747-5632(91)90030-5

Kulik, J. (1994). "Meta-analytic studies of findings on computer-based instruction," in *Technology Assessment in Education, and Training*, eds E. Baker and H. O'Neil Jr. (New York, NY: Routledge), 9.

Lenaerts, J., Wieme, W., Janssens, F., and Van Hoecke, T. (2002). Designing digital resources for a physics course. *Eur. J. Phys.* 23, 175–182. doi:10.1088/0143-0807/23/2/311

Li, L., Li, B., and Luo, Y. (2015). Using a dual safeguard web-based interactive teaching approach in an introductory physics class. *Phys. Rev. ST Phys. Educ. Res.* 11, 010106. doi:10.1103/PhysRevSTPER.11.010106

Mangels, J., Butterfield, B., Lamb, J., Good, C., and Dweck, C. (2006). Why do beliefs about intelligence influence learning success? A social cognitive neuroscience model. *Soc. Cogn. Affect. Neurosci.* 1, 75. doi:10.1093/scan/nsl013

Margaryan, A., Bianco, M., and Littlejohn, A. (2015). Instructional quality of massive open online courses. *Comput. Educ.* 80, 77. doi:10.1016/j.compedu.2014.08.005

Maries, A., Lin, S., and Singh, C. (2016). "The impact of students' epistemological framing on a task requiring representational consistency," in *Proceedings of the Physics Education Research Conference 2016* (Sacramento, CA). doi:10.1119/perc.2016.pr.048

Marshman, E., Kalender, Z., Schunn, C., Nokes-Malach, T., and Singh, C. (2017). A longitudinal analysis of students' motivational characteristics in introductory physics courses: gender differences. *Can. J. Phys.* doi:10.1139/cjp-2017-0185

Marshman, E., and Singh, C. (2016). Interactive tutorial to improve student understanding of single photon experiments involving a Mach-Zehnder Interferometer. *Eur. J. Phys.* 37, 024001. doi:10.1088/0143-0807/37/2/024001

Marshman, E., and Singh, C. (2017a). Investigating and improving student understanding of quantum mechanics in the context of single photon interference. *Phys. Rev. Phys. Educ. Res.* 13, 010117. doi:10.1103/PhysRevPhysEducRes.13.010117

Marshman, E., and Singh, C. (2017b). Investigating and improving student understanding of the probability distributions for measuring physical observables in quantum mechanics. *Eur. J. Phys.* 38, 025705. doi:10.1088/1361-6404/aa57d1

Marshman, E., and Singh, C. (2017c). Investigating and improving student understanding of the expectation values of observables in quantum mechanics. *Eur. J. Phys.* 38, 045701. doi:10.1088/1361-6404/aa6d34

Marshman, E., and Singh, D. (2015). "Developing an interactive tutorial on a quantum eraser," in *Proceedings of the Physics Education Research Conference 2014*, eds P. Engelhardt, A. Churukian, and D. Jones (Minneapolis, MN). doi:10.1119/perc.2014.pr.040

Mason, A., and Singh, C. (2010). Do advanced students learn from their mistakes without explicit intervention? *Am. J. Phys.* 78, 760. doi:10.1119/1.3318805

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

McDermott, L. (1996). Development of a computer-based tutorial on the photoelectric effect. *Am. J. Phys.* 64, 1370–1379. doi:10.1119/1.18360

Moos, D., and Azevedo, R. (2008). Exploring the fluctuation of motivation and use of self-regulatory processes during learning with hypermedia. *Instr. Sci.* 36, 203. doi:10.1007/s11251-007-9028-3

Moos, D., and Azevedo, R. (2009). Learning with computer-based learning environments: a literature review of computer self-efficacy. *Rev. Educ. Res.* 79, 576. doi:10.3102/0034654308326083

Morrison, G., Ross, S., Gopalakrishnan, M., and Casey, J. (1995). The effects of feedback and incentives on achievement in computer-based instruction. *Contemp. Educ. Psychol.* 20, 32. doi:10.1006/ceps.1995.1002

Narciss, S., Proske, A., and Koerndle, H. (2007). Promoting self-regulated learning in web-based learning environments. *Comput. Hum. Behav.* 23, 1126. doi:10.1016/j.chb.2006.10.006

Nokes-Malach, T., and Mestre, J. (2013). Toward a model of transfer as sense-making. *Educ. Psy.* 48, 184–207. doi:10.1080/00461520.2013.807556

Piaget, J. (1978). *Success and Understanding*. Cambridge, MA: Harvard University Press.

Pintrich, P. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *J. Educ. Psychol.* 95, 667–686. doi:10.1037/0022-0663.95.4.667

Posner, G., Strike, K., Hewson, W., and Gertzog, W. (1982). Accommodation of a scientific conception: toward a theory of conceptual change. *Sci. Educ.* 66, 211. doi:10.1002/sce.3730660207

Redish, E., Saul, J., and Steinberg, R. (1998). Student expectations in introductory physics. *Am. J. Phys.* 66, 212. doi:10.1119/1.18847

Reif, F. (1987). Instructional design, cognition, and technology: applications to the teaching of scientific concepts. *J. Res. Sci. Teach.* 24, 309. doi:10.1002/tea.3660240405

Reif, F., and Scott, L. (1999). Teaching scientific thinking skills: students and computers coaching each other. *Am. J. Phys.* 67, 819l–883l. doi:10.1119/1.19130

Sayer, R., Maries, A., and Singh, C. (2015). "Developing and evaluating a tutorial on the double-slit experiment," in *Proceedings of the Physics Education Research Conference 2015*, eds A. Churukian, D. Jones, and L. Ding (College Park, MD). doi:10.1119/perc.2015.pr.070

Sayer, R., Maries, A., and Singh, C. (2017). A quantum interactive learning tutorial on the double-slit experiment to improve student understanding of quantum mechanics. *Phys. Rev. PER* 13, 010123. doi:10.1103/PhysRevPhysEducRes.13.010123

Schwartz, D., Bransford, J., and Sears, D. (2005). "Efficiency and innovation in transfer," in *Transfer of Learning from a Modern Multidisciplinary Perspective*, ed. J. Mestre (Greenwich, CT: Information Age), 1.

Seaton, D., Bergner, Y., Chuang, I., Mitros, P., and Pritchard, D. (2014). Who does what in a massive open online course? *Commun. ACM* 57, 58. doi:10.1145/2500876

Shaffer, P., and McDermott, L. (1992). Research as a guide for curriculum development: an example from introductory electricity. Part II: design of instructional strategies. *Am. J. Phys* 60, 1003. doi:10.1119/1.16979

Singh, C. (2002). When physical intuition fails. *Am. J. Phys.* 70, 1103. doi:10.1119/1.1512659

Singh, C. (2004). Interactive video tutorials for enhancing problem-solving reasoning, and meta-cognitive skills of introductory physics students. *AIP Conf. Proc.* 720, 177. doi:10.1063/1.1807283

Singh, C. (2008a). *Computer-Based Tutorials to Develop Expertise in Introductory Physics Students, American Physical Society (APS) Forum on Education Newsletter, Summer, 13*. Available at: https://www.aps.org/units/fed/newsletters/summer2008/singh.cfm

Singh, C. (2008b). Interactive learning tutorials on quantum mechanics. *Am. J. Phys.* 76, 400. doi:10.1119/1.2837812

Singh, C. (2008c). Assessing student expertise in introductory physics with isomorphic problems. I. Performance on a non-intuitive problem pair from introductory physics. *Phys. Rev. ST Phys. Educ. Res.* 4, 010104. doi:10.1103/PhysRevSTPER.4.010104

Singh, C. (2008d). Assessing student expertise in introductory physics with isomorphic problems. II. Effect of some potential factors on problem solving and transfer. *Phys. Rev. ST Phys. Educ. Res.* 4, 010105. doi:10.1103/physrevstper.4.010105

Singh, C. (2009). Problem solving and learning. *AIP Conf. Proc.* 1140, 183. doi:10.1063/1.3183522

Singh, C., Belloni, M., and Christian, W. (2006). Improving students' understanding of quantum mechanics. *Phys. Today* 59, 43. doi:10.1063/1.2349732

Singh, C., and Haileselassie, D. (2010). Developing problem solving skills of students taking introductory physics via web-based tutorials. *J. Coll. Sci. Teach.* 39, 34. Available at: https://www.learntechlib.org/p/108402/

Singh, C., and Marshman, E. (2015). "Developing an interactive tutorial on a Mach-Zehnder interferometer with single photons," in *Proceedings of the Physics Education Research Conference 2014*, eds P. Engelhardt, A. Churukian, and D. Jones (Minneapolis, MN). doi:10.1119/perc.2014.pr.056

Song, H., Kalett, A., and Plass, J. (2016). Interplay of prior knowledge, self-regulation and motivation in complex multimedia learning environments. *J. Comput. Assist. Learn.* 32, 31. doi:10.1111/jcal.12117

Stewart, J., DeVore, S., Stewart, G., and Michaluk, L. (2016). Behavioral self-regulation in a physics class. *Phys. Rev. PER* 12, 010125.

Sungur, S. (2007). Contribution of motivational beliefs and metacognition to students' performance under consequential and nonconsequential test conditions. *Educ. Res. Eval.* 13, 127. doi:10.1080/13803610701234898

Tinto, V. (1993). *Leaving College: Rethinking the Causes and Cures of Student Attrition*. Chicago, IL: University of Chicago Press.

Tinto, V. (1997). Classrooms as communities: exploring the educational character of student persistence. *J. Higher Educ.* 68, 599. doi:10.2307/2959965

Wagner, A., Altherr, S., Eckert, B., and Jorg Jodl, H. (2006). Multimedia in physics education: two teaching videos on the absorption and emission spectrum of sodium. *Eur. J. Phys.* 27, L31–L35. doi:10.1088/0143-0807/27/6/L01

Wieman, C. (2015). Comparative cognitive task analyses of experimental science and instructional laboratory courses. *Phys. Teach.* 53, 349. doi:10.1119/1.4928349

Wigfield, A., Hoa, L., and Klauda, S. (2008). "The role of achievement values in the regulation of achievement behaviors," in *Motivation and Self-Regulated Learning: Theory, Research and Applications*, eds D. Schunk and B. Zimmerman (New York, NY: Lawrence Erlbaum), 169.

Winne, P. (1996). A metacognitive view of individual differences in self-regulated learning. *Learn. Individ. Differ.* 8, 327. doi:10.1016/S1041-6080(96)90022-9

Wood, A., Galloway, R., Donnelly, R., and Hardy, J. (2016). Characterizing interactive engagement activities in a flipped introductory physics class. *Phys. Rev. PER* 12, 010140.

Yalcinalp, S., Geban, O., and Ozkan, I. (1995). Effectiveness of using computer assisted supplementary instruction for teaching the mole concept. *J. Res. Sci. Teach.* 32, 1083–1095. doi:10.1002/tea.3660321007

Yerushalmi, E., Cohen, E., Mason, A., and Singh, C. (2012a). What do students do when asked to diagnose their mistakes? Does it help them? I. An atypical quiz context. *Phys. Rev. ST Phys. Educ. Res.* 8, 020109. doi:10.1103/PhysRevSTPER.8.020109

Yerushalmi, E., Cohen, E., Mason, A., and Singh, C. (2012b). What do students do when asked to diagnose their mistakes? Does it help them? II. A more typical quiz context. *Phys. Rev. ST PER* 8, 020110. doi:10.1103/physrevstper.8.020110

Zhu, G., and Singh, C. (2011). Improving students' understanding of quantum mechanics via the Stern-Gerlach experiment. *Am. J. Phys.* 79, 499. doi:10.1119/1.3546093

Zhu, G., and Singh, C. (2012a). Improving students' understanding of quantum measurement: I. Investigation of difficulties. *Phys. Rev. ST Phys. Educ. Res.* 8, 010117. doi:10.1103/PhysRevSTPER.8.010117

Zhu, G., and Singh, C. (2012b). Improving students' understanding of quantum measurement: II. Development of research-based learning tools. *Phys. Rev. ST Phys. Educ. Res.* 8, 010118. doi:10.1103/PhysRevSTPER.8.010118

Zhu, G., and Singh, C. (2013). Improving student understanding of addition of angular momentum in quantum mechanics. *Phys. Rev. ST Phys. Educ. Res.* 9, 010101. doi:10.1103/PhysRevSTPER.9.010101

Zimmerman, B., and Schunk, D. (2001). *Self-Regulated Learning and Academic Achievement: Theory, Research and Practice*. Mahwah, NJ: Lawrence Erlbaum.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Education

ORIGINAL RESEARCH
published: 22 January 2018
doi: 10.3389/feduc.2017.00068

Check for updates

What Types of Instructional Shifts Do Students Experience? Investigating Active Learning in Science, Technology, Engineering, and Math Classes across Key Transition Points from Middle School to the University Level

Kenneth Akiha[1,2], Emilie Brigham[1,2], Brian A. Couch[3], Justin Lewin[1,2], Marilyne Stains[4], MacKenzie R. Stetzer[2,5], Erin L. Vinson[1,2] and Michelle K. Smith[1,2]*

[1] School of Biology and Ecology, University of Maine, Orono, ME, United States, [2] Maine Center for Research in STEM Education, University of Maine, Orono, ME, United States, [3] School of Biological Sciences, University of Nebraska, Lincoln, NE, United States, [4] Department of Chemistry, University of Nebraska, Lincoln, NE, United States, [5] Department of Physics and Astronomy, University of Maine, Orono, ME, United States

OPEN ACCESS

Edited by:
Rob Cassidy,
Concordia University, Canada

Reviewed by:
Sue Wick,
University of Minnesota Twin Cities,
United States
Deborah Donovan,
Western Washington University,
United States

***Correspondence:**
Kenneth Akiha
kenneth.akiha@maine.edu

Specialty section:
This article was submitted to Digital Education, a section of the journal Frontiers in Education

Received: 27 October 2017
Accepted: 14 December 2017
Published: 22 January 2018

Citation:
Akiha K, Brigham E, Couch BA, Lewin J, Stains M, Stetzer MR, Vinson EL and Smith MK (2018) What Types of Instructional Shifts Do Students Experience? Investigating Active Learning in Science, Technology, Engineering, and Math Classes across Key Transition Points from Middle School to the University Level. Front. Educ. 2:68. doi: 10.3389/feduc.2017.00068

Despite the need for a strong Science, Technology, Engineering, and Math (STEM) workforce, there is a high attrition rate for students who intend to complete undergraduate majors in these disciplines. Students who leave STEM degree programs often cite uninspiring instruction in introductory courses, including traditional lecturing, as a reason. While undergraduate courses play a critical role in STEM retention, little is understood about the instructional transitions students encounter upon moving from secondary to post-secondary STEM courses. This study compares classroom observation data collected using the Classroom Observation Protocol for Undergraduate STEM from over 450 middle school, high school, introductory-level university, and advanced-level university classes across STEM disciplines. We find similarities between middle school and high school classroom instruction, which are characterized by a large proportion of time spent on active-learning instructional strategies, such as small-group activities and peer discussion. By contrast, introductory and advanced university instructors devote more time to instructor-centered teaching strategies, such as lecturing. These instructor-centered teaching strategies are present in classes regardless of class enrollment size, class period length, or whether or not the class includes a separate laboratory section. Middle school, high school, and university instructors were also surveyed about their views of what STEM instructional practices are most common at each educational level and asked to provide an explanation of those perceptions. Instructors from all levels struggled to predict the level of lecturing practices and often expressed uncertainty about what instruction looks like at levels other than their own. These findings suggest that more opportunities need to be created for instructors across multiple levels of the education system to share their active-learning teaching practices and discuss the transitions students are making between different educational levels.

Keywords: active-learning, classroom observation, secondary education, undergraduate education, educational transitions

INTRODUCTION

Science, Technology, Engineering, and Math (STEM) education plays an essential role in building the foundational knowledge needed to solve global problems. For decades, this importance has been highlighted by both researchers and policy, yet the United States continues to produce fewer STEM graduates than the economy demands (President's Council of Advisors on Science and Technology, 2012). Despite an increased interest in STEM degrees from well-prepared students, there is a dramatic attrition rate once students begin college-level programs (Pryor and Eagan, 2013; Eagan et al., 2014). Half of intended STEM bachelor's degree majors do not end up earning a STEM degree within 6 years of entering college (Eagan et al., 2014), and the majority of those leaving do so in the first 2 years of their degree (Watkins and Mazur, 2013). The attrition rates are even greater at the 2-year college level, where two-thirds of students intending to earn a STEM associates degree do not do so within 4 years (Van Noy and Zeidenberg, 2014). These attrition rates overwhelm any gains from increased interest in STEM degrees, leading to a shortfall in the number of students entering the STEM workforce.

One of the proposed solutions to meet the need for one million more STEM graduates by 2022 is to increase student retention rates in STEM majors by 33% (President's Council of Advisors on Science and Technology, 2012). In order to work toward this goal, it is important to examine why students, who were previously interested in STEM in high school, are leaving STEM degree programs at such a high rate. One main source of student attrition in STEM fields is the types of experiences students have upon arriving in their college classes. A seminal study conducted in the late 1990s found that students switch from STEM degrees for a variety of reasons related to their experiences as students (Seymour and Hewitt, 1997). Both switching and non-switching students said that one of the most common concerns was uninspiring teaching in STEM courses, with over 90% of switchers mentioning it as a part of their interviews and almost three-quarters of non-switchers mentioning it. An example of uninspiring teaching is a class solely dedicated to lecturing about information in the textbook. While this type of instruction has been the predominant method at the undergraduate level for centuries (Brockliss, 1996), alternative methods, such as active-learning strategies, have been shown to promote greater learning and better outcomes for students (e.g., Prince, 2004; Freeman et al., 2014).

Recently, the efficacy of active-learning methods was quantified. In a meta-analysis of 225 studies that reported on exam scores and/or failure rates comparing undergraduate STEM courses using lecture-based instruction with ones using active-learning, researchers found two significant trends (Freeman et al., 2014). Students in active-learning classrooms earned exam scores half a letter grade higher than students in lecture-based classrooms for the same course. In addition, students in active-learning courses are one and a half times more likely to pass the course compared to students in sections that predominately use traditional lecturing. Additional studies found that requiring participation in a number of active-learning interventions improved achievement for all students, especially traditionally underrepresented students, without requiring any additional staffing or financial resources (Haak et al., 2011; Eddy and Hogan, 2014).

While active learning provides an effective means to engage students and improve student outcomes, it remains unclear how the amount and type of active learning used in classes changes as students progress through different instructional levels, from middle school to advanced undergraduate courses. Understanding the instructional transitions students experience has the potential to help explain why students choose to leave STEM majors. However, there are a number of challenges when trying to meaningfully describe the amount and types of active learning taking place in classrooms across different instructional environments. Studies that characterize instructional practices are typically performed in either high school *or* undergraduate classrooms, often as part of the evaluation of professional development programs (Rockoff et al., 2008; Smith et al., 2014; Garrett and Steinberg, 2015; Campbell et al., 2016). The observation tools and research methods used in studies that examine instructional practices often differ, further complicating comparisons between them.

One exception is the Reformed Teaching Observation Protocol (RTOP) (Sawada et al., 2002), which has been used in both high school and undergraduate settings. The RTOP includes Likert-scale items that observers score to measure the instructional practices implemented in the classroom on a scale from lecture based and teacher centered (0) to inquiry based and student centered (100). The RTOP was originally developed as part of an evaluation system for a program designed for preparing K-12 teachers. Since its development, it has also been used to track changes in undergraduate faculty practices due to participation in different types of professional development (Ebert-May et al., 2015; Manduca et al., 2017). A survey of studies on high school STEM teachers indicates that average RTOP scores range from 37.3 to 53.5 (Roehrig and Kruse, 2005; Yezierski and Herrington, 2011). Similar studies at the undergraduate-level are limited; however, one study found that 20 different first-year college science instructors had an average RTOP score of 35.9 (Lund et al., 2015), and a study of biology instructors who participated in extensive professional development programs reported an average score of 37.1 (Ebert-May et al., 2011). Thus, there are likely meaningful differences in the instructional practices employed in these educational environments; however, the RTOP protocol does not offer the resolution required to understand these differences in a meaningful manner.

Observation protocols that have been developed more recently, including the Teaching Dimensions Observation Protocol (TDOP; Hora et al., 2013) and the Classroom Observation Protocol for Undergraduate STEM (COPUS; Smith et al., 2013), record instructor and student instructional behaviors in 2-min time intervals and provide additional tools that can be used to examine practices at different educational levels. The TDOP was designed as a supplement to survey data when characterizing classroom practice and involves observers marking codes, such as Interactive Lecture or Student Comprehension Question, from a set of over 40 observable classroom behaviors and actions every 2 min (Hora et al., 2013). The COPUS has 25 total codes and was adapted from the TDOP as a more basic instrument that requires less training time and could be used by a variety of individuals to provide feedback to instructors and identify professional development needs (Smith et al., 2013).

Since its development, COPUS has been used in a variety of studies at the undergraduate level to describe general campus-wide instructional practices as well as to examine more specific active-learning strategies, such as the use of clickers and worksheet-based activities (Smith et al., 2014; Lund et al., 2015; Lewin et al., 2016; Cleveland et al., 2017). On a campus-wide scale, COPUS has been useful in describing the variation in instructional practices present across STEM disciplines and in creating profiles of commonly observed types of classrooms. One study using COPUS data from 55 different courses across 13 STEM departments found a diverse range of teaching practices (Smith et al., 2014). Specifically, the study found a wide range in the frequency at which instructors used Lecturing, with some using it as little as 2% of their total instructional behaviors and others using it as 98% of their total instructional behavior. This observed continuum showed that the binary categorization of instructional practice as either lecture-based or active learning represents an oversimplification. Another study used COPUS data from 269 class periods taught by 73 different instructors across 28 universities to create 10 classroom profiles ranging from teacher centered to student centered (Lund et al., 2015). The creation and application of these profiles provides a finer resolution for describing the instructional practices utilized at research universities. Taken together, this work demonstrates that COPUS can be a meaningful tool in characterizing classroom experiences in undergraduates STEM courses.

To explore why students who were interested in STEM in high school leave during their undergraduate years, we need to understand the instructional transitions students encounter as they progress through the educational system. Using COPUS and instructor survey data from middle school, high school, and undergraduate STEM classes, this study sought to characterize how STEM classroom experiences compare across the transition from secondary to post-secondary educational institutions. Specifically, we asked: (1) How do instructional experiences in middle school and high school STEM classes compare with first-year and advanced-level undergraduate classes? (2) Do the instructional experiences at the undergraduate level depend on variables, such as class size, class length, or whether the class also includes a laboratory section? and (3) What perceptions do middle school, high school, and university instructors hold about instructional practice across all educational levels and how do instructors' perceptions compare with observed practices? The answers to these questions can help to clarify specific instructional transitions and explanations for the associated issues, which can serve as both areas for future research and targets for professional development aimed at increasing student retention in STEM fields.

MATERIALS AND METHODS

Observation Data Collection

This study includes classroom observation data from middle school, high school, and university level classrooms. To observe university classrooms, we emailed University of Maine STEM instructors asking them if they would allow secondary school (i.e., middle and high school) teachers to visit their classrooms and collect observation data; 74% of those emailed agreed. Middle and high school teachers performed the observations as part of their participation in the University Classroom Observation Program, which was designed to give faculty formative feedback on their teaching from external observers without conflating that feedback with review procedures for tenure and promotion (Smith et al., 2014). The program occurred over four semesters and each semester we had more applicants than slots available for middle and high school teacher observers (average acceptance rate = 34%), so we were able to select teachers with a range of experiences (e.g., numbers of years teaching, socioeconomic needs of the community) from a variety of school districts.

Altogether, the teachers conducted 364 class observations. These observations included 153 instructors who taught 128 courses in 21 different departments (anthropology; biology, and ecology; chemical and biological engineering; chemistry; civil and environmental engineering; computer sciences; earth sciences; ecology and environmental sciences; economics; electrical and computer engineering; electrical engineering technology; food and agriculture; forest resources; marine science; mathematics and statistics; mechanical engineering; molecular and biomedical science; nursing; physics and astronomy; plant, soil, and environmental science; psychology; and wildlife, fisheries, and conservation biology) as shown in **Table 1**. Observations from 270 classes taught in Spring 2014, Fall 2014, and Spring 2015 have been reported in earlier studies (Smith et al., 2014; Lewin et al., 2016).

We conducted middle (grades 6–8) and high school (grades 9–12) class observations in public secondary schools located within a 140-mile radius of the University of Maine (Orono, ME, USA). We asked secondary teachers who had participated as observers of university classes if they would allow their classes to also be observed. In addition, many of the secondary teachers identified other STEM teachers in their districts who were willing to have their classes observed. In total, investigators observed 118 secondary school class periods. These observations included 82 teachers from 37 schools (**Table 1**).

Observer Training

Secondary teachers who observed university classes received COPUS training and carried out observations in pairs as described in Smith et al. (2013). Briefly, the 2-h training introduced the teachers to the 25 COPUS codes shown in **Figure 1** and gave them a chance to practice coding using short video clips from real university classrooms. Sample observation sheets can be found in Smith et al. (2013) and at http://www.cwsei.ubc.ca/resources/COPUS.htm. After watching the videos, the middle and high school teachers listed the codes they selected and discussed any disagreements. When the training was complete, the teachers observed in pairs, but were instructed to record their COPUS results independently. Inter-rater reliability (IRR) was calculated using Cohen's kappa scores as described in Lewin et al. (2016). The mean Cohen's kappa score for all of the university observations was 0.91 (SE $\pm$ 0.01), indicating strong IRR

TABLE 1 | Demographic information about all the secondary and university courses observed.

	Courses	Instructors	Schools (HS/MS), departments (University)	Observations	STEM Breakdown	Class size range	Mean class size
Middle school	39	24	15	43	S—60% TE—0% M—40%	8–27	16.7
High school	68	58	22	75	S—75% TE—4% M—21%	2–24	13.1
University first-year	36	58	20	131	S—60% TE—15% M—25%	16–339	99.3
University advanced	92	95	21	233	S—65% TE—28% M—7%	11–322	68.8

Observations were categorized as S (Science), TE (Technology and Engineering), or M (Mathematics) based on course title at the middle school and high school level and by department at the university level.

	Collapsed Codes	Individual Codes
Instructor is:	Presenting (P)	**Lec**: Lecturing or presenting information **RtW**: Real-time writing **D/V**: Showing or conducting a demo, experiment, or simulation
	Guiding (G)	**FlUp**: Follow-up/feedback on clicker question of activity **PQ**: Posing non-clicker question to students (non-rhetorical) **CQ**: Asking clicker question (entire time, not just when first asked) **AnQ**: Listening to and answering student questions to entire class **MG**: Moving through class guiding ongoing student work **1o1**: One-on-one extended discussion with individual students
	Administration (A)	**Adm**: Administration (assign homework, return tests, etc.)
	Other (OI)	**W**: Waiting (instructor late, working on fixing technical problems) **O**: Other
Students are:	Receiving (R)	**L**: Listening to instructor
	Talking to class (STC)	**AnQ**: Student answering question posed by instructor **SQ**: Student asks question **WC**: Students engaging in whole-class discussion **SP**: Students presenting to entire class
	Working (W)	**Ind**: Individual thinking/problem solving **CG**: Discussing clicker question in groups of students **WG**: Working in groups on worksheet activity **OG**: Other assigned group activity **Prd**: Making a prediction about a demo or experiment **TQ**: Test or quiz
	Other (OS)	**W**: Waiting (instructor late, working on fixing technical issues) **O**: Other

FIGURE 1 | Classroom Observation Protocol for Undergraduate STEM instrument codes and abbreviated descriptions used to describe instructor and student behavior during in-class observations. The individual codes are further grouped into collapsed codes.

(Landis and Koch, 1977). Only codes marked by both observers in a given time interval were included in the data set for this study.

Three observers conducted observations in secondary school classes and included two Master of Science Teaching students who are now high school teachers (co-authors Kenneth Akiha and Justin Lewin) and one University of Maine professional development coordinator who is a former high school teacher (co-author Erin L. Vinson). These observers received similar training on conducting classroom observations using the COPUS protocol (e.g., discussion of codes and practice coding common videos). IRR was determined by observing a video of the same class period and observing at least three different live classes in pairs. Each of these comparisons yielded a Cohen's kappa score of greater than 0.9, demonstrating strong IRR (Landis and Koch, 1977). Given the dependable IRR and that traveling to observe in pairs would have greatly limited the number of observations, subsequent secondary classes were observed by only one individual.

Observation Data Sorting

We sorted secondary class observation data based on the level of the class: either middle school or high school. For the university classroom observation data, we sorted the data based on three categories of STEM courses: general education, first-year, and advanced (**Figure 2**). If data came from a course not required for a STEM major, we categorized the course as "general education" and subsequently excluded the associated data from our analysis because our focus is on how instructional practices can affect student retention in STEM majors. Of the courses required for STEM majors, we classified courses with greater than 33% first-year student enrollment and less than two pre-requisites in any given department as "First-Year" courses. Required courses with less than 33% first-year enrollment or more than two pre-requisites in a given department were sorted as "Advanced" courses.

Data Analysis

For this study, we analyzed the COPUS data using two different strategies described in Lewin et al. (2016): relative abundance, as described by percentage of collapsed codes, and relative frequency, as described by percentage of 2-min time intervals containing specific codes (**Figure 1**). For relative abundance, collapsed codes refer to categories that describe more general instructor and student behaviors, usually consisting of multiple individual codes. For example, the Instructor Presenting collapsed code category consists of three individual codes: Lecturing (Lec), Real-time Writing, and Demo/Video (**Figure 1** shows all collapsed code categories). To visualize and compare relative abundance of each COPUS code, we calculated the percentage of each collapsed code by totaling the number of codes in that category during a class and dividing by the total number of codes marked during the class. For example, if there were 20 codes marked under the Instructor Presenting collapsed code category and 50 codes marked in total, then 20/50 or 40% of the codes correspond to the Instructor Presenting collapsed code.

However, when trying to compare the frequency of a single code, such as Instructor Lec or Student Listening (L), percent code calculations can be misleading because multiple COPUS codes can be marked at the same time, which can impact the denominator of the calculation. Therefore, we also quantified relative frequency by calculating the percentage of 2-min time intervals in which a given code was marked. To do this, the number of 2-min time intervals marked for each code was divided by the total number of time intervals that were coded in that class session. For example, if instructor Lec was marked in 18 time intervals out of a possible 30 time intervals, then 18/30 or 60% of the possible 2-min time intervals contained lecture.

We were also interested in comparing the amount of time students worked in groups because it is one way to generally compare teacher-centered versus student-centered teaching practices. COPUS has multiple student codes involving group work: Clicker Group Work (CG), Worksheet Group Work (WG), and Other Group Work (OG). These three codes measure finer distinctions of what can be broadly classified as students working in groups (Lund et al., 2015), so if any of those codes were marked then we counted them in the general Group Work (GW) code.

COPUS Use in Middle and High School Classrooms

Because the COPUS instrument was developed and validated at the undergraduate level, we needed to determine if it adequately captures the classroom experiences in middle and high school in addition to those in undergraduate STEM classes. In particular, we were concerned that there might be certain activities or teaching modes that would go undetected. To address this possibility, we looked at the relative frequency and relative abundance of the Instructor Other (OI) and Student Other (OS) codes (**Figure 1**) documented in middle school and high school observations combined and all undergraduate observations (**Table 2**). We chose to compare Other codes to examine whether certain behaviors that were not observed in undergraduate classes, and therefore unable to be captured by the COPUS instrument, were present in middle school and high school classrooms. On average, Other codes made up less than 5% of the total codes marked in middle school and high school classes, while the same codes made up less than 3% of the total codes marked in undergraduate classes.

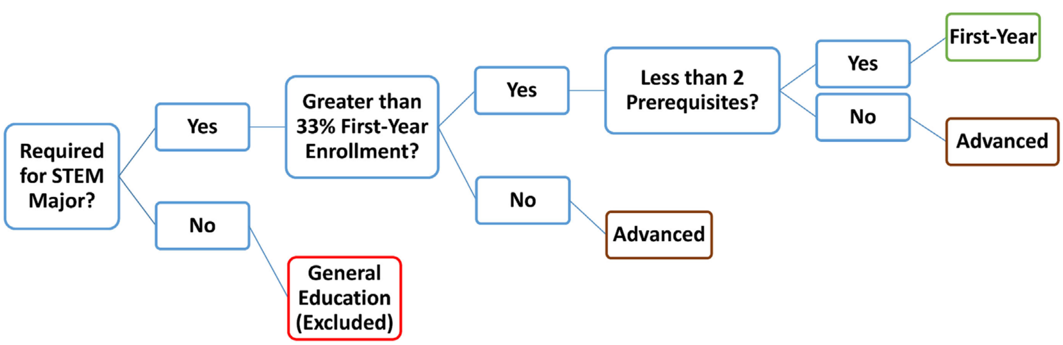

FIGURE 2 | Sorting scheme used to determine the university course level.

TABLE 2 | Relative abundance and relative frequency of Instructor Other (OI) and Student Other (OS) codes in middle and high school classes and university level classes.

	Relative abundance		Relative frequency	
	OI (%)	OS (%)	OI (%)	OS (%)
Middle and High School ($n = 118$)	4.3	4.0	8.6	6.7
University ($n = 364$)	2.2	0.8	3.2	1.5

Also, on average, Other codes were marked in less than 9% of the total number of 2-min time intervals in a middle school and high school class period, while the same codes were marked in less than 4% of the total number of 2-min time intervals in a university class period (**Table 2**). Based on observer comments, the most common OI code behaviors across all levels were listening to student presentations; setting up technology, materials, or equipment; and facilitating and guiding class discussions. The most common OS code behaviors across all levels were students writing on the board, forming groups, and students getting or putting away materials. At the middle school and high school levels, observers noted more time for students getting or putting away materials. Overall, the overlap in Other code behaviors for both instructors and students, combined with the relatively low and similar abundances and frequencies at both levels, provided evidence that the COPUS instrument was not systematically missing important activities that may be present in middle school and high school STEM classrooms.

Survey Responses

Because the results from our study may be used to design professional development for instructors at multiple education levels, we wanted to determine how our data matched the perceptions and expectations instructors have of the type of instruction their students are either coming from or heading to in the future. To learn more about instructors' perspectives on instructional behaviors at different educational levels, university faculty who were observed by middle and high school teachers and/or attended a variety of professional development opportunities at the University of Maine (e.g., workshops, speakers) were sent an email asking them to take a short survey. Similarly, middle and high school teachers who participated in University Classroom Observation Program or other professional development events at the University of Maine (e.g., workshops, summer teaching institutes) were sent the same email and asked to share it with their colleagues. The survey included a multiple-choice question in which respondents were asked to select one of four graphs that showed different result patterns describing average percent Instructor Lec code in classes at the middle school, high school, first-year college, and advanced college levels. The survey respondents also answered a follow-up open-response question in which they were asked to explain why they selected a specific multiple-choice answer.

To examine the range of answers chosen, the percent of each choice was calculated for the middle school, high school, and university educator groups. To analyze the open-response question answers, we used a content analysis process (Miles et al., 2013). Specifically, one co-author (Emilie Brigham) read the answers, created categories based on large themes, and scored the short-answer responses based on the presence or absence of each category in an individual's response. A second co-author (Michelle K. Smith) used the categories, independently scored the responses, and suggested new categories for the scheme. The coding between the two authors was compared and any coding differences were resolved through discussion.

IRB Information

All faculty members and secondary teachers who agreed to be observed were given a human subjects consent form. The Institutional Review Board at the University of Maine granted approval to evaluate observation data of classrooms and survey instructors about the observation results (exempt status, protocol no. 2010-04-3 and 2013-02-06). Because of the delicate nature of sharing observation data with other instructors and administrators, the consent form explained that the data would only be presented in aggregate and would not be subdivided according to variables such as department or school. We provided instructors access to observation data from their own course(s) upon request after we collected observation data for this study.

RESULTS

Instructional Practices across Education Levels

We used the COPUS to obtain a comprehensive view of classrooms at each educational level and started by comparing the relative abundance of all the instructor collapsed COPUS codes (**Figure 3**). In the middle school and high school classes, the Instructor Presenting collapsed code, which is more frequently seen in traditional lecture classes, comprised between 0 and 66% of instructor collapsed codes. In first-year and advanced university-level courses, the Instructor Presenting collapsed code represented between 0 and 100% of instructor collapsed codes at both levels.

Another way to compare data across multiple educational levels is to examine the frequency of particular COPUS codes across the 2-min time intervals. When examining the 2-min relative frequency of the Instructor Lec code, the interquartile ranges were lower for middle school and high school classrooms when compared to first-year and advanced university courses (**Figure 4**). Furthermore, a Kruskal–Wallis Test showed very strong evidence of a difference ($p < 0.001$) between the mean ranks of at least one pair of groups. A Dunn's pairwise test of all six pairs of levels showed instructors in first-year and advanced university classes spent significantly more time using the Instructor Lec code than instructors in middle school and high school classes ($p < 0.001$ adjusted using the Bonferroni correction). In particular, the difference between the median percentage of 2-min time intervals marked with the Instructor Lec code in high school and first-year university classes was 48% (32% in high school to 80% in first-year university classes), more than 10-fold greater than any other difference between chronologically adjacent levels.

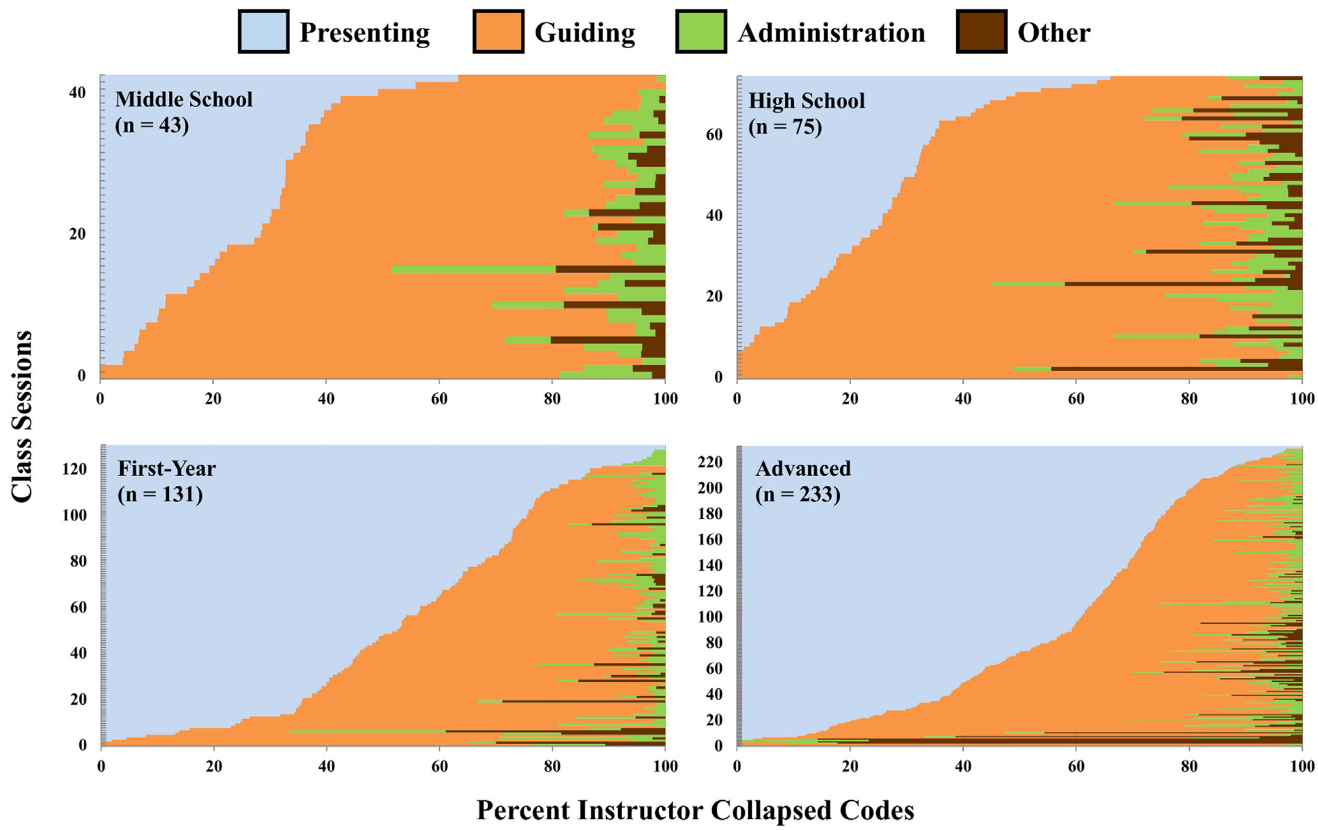

FIGURE 3 | Percentage of instructor collapsed Classroom Observation Protocol for Undergraduate STEM codes for middle school, high school, first-year, and advanced university classes. Each horizontal bar represents a different class session. The classes are ordered by the collapsed code Instructor Presenting. **Figure 1** describes the Instructor Collapsed Codes.

There was no significant difference in the median percentage of 2-min time intervals, including Instructor Lec between first-year and advanced university courses.

In addition to comparing traditional instructional codes, we compared relative frequency using instructional codes often associated with student-centered classrooms, such as Instructor Moving and Guiding (MG) throughout the classroom. Middle school and high school classes showed a greater range of percent 2-min time intervals containing the Instructor MG code (**Figure 5**). For both university levels, more than half of the observations captured no Instructor MG during the entire class. When comparing mean ranks, a Kruskal–Wallis Test showed very strong evidence of a difference ($p < 0.001$) between at least one pair of groups. A Dunn's pairwise test of all six pairs of levels showed instructors in middle school and high school classes spent significantly more time MG than university instructors ($p < 0.001$ adjusted using the Bonferroni correction).

Student Classroom Experiences across Educational Levels

To view the instructional experience from the student perspective, we analyzed the student collapsed COPUS codes and saw a difference in the ranges of classroom behaviors at different educational levels (**Figure 6**). Because students sitting quietly and taking notes is often associated with lecture-based classrooms, we also compared the Student Receiving collapsed code. Student Receiving made up a range of 0–60% of the student collapsed codes in middle school and high school classes compared to 0–100% of student collapsed codes in both levels of university classes.

We also looked at the relative frequency of individual student codes, beginning with codes for traditional instruction student behaviors such as Listening (L). Our data showed that middle school and high school classes exhibited a greater interquartile range of percent 2-min time interval values, while first-year and advanced university classes had higher median values (**Figure 7**). Moreover, a Kruskal–Wallis Test showed very strong evidence of a difference ($p < 0.001$) between the mean ranks of at least one pair of groups. A Dunn's pairwise test of all six pairs of levels showed students in middle school and high school classes spent significantly less time listening and taking notes than students in first-year and advanced university classes ($p < 0.001$ adjusted using the Bonferroni correction).

In addition, we examined the relative frequency of codes for student behaviors typical of student-centered classrooms, such as

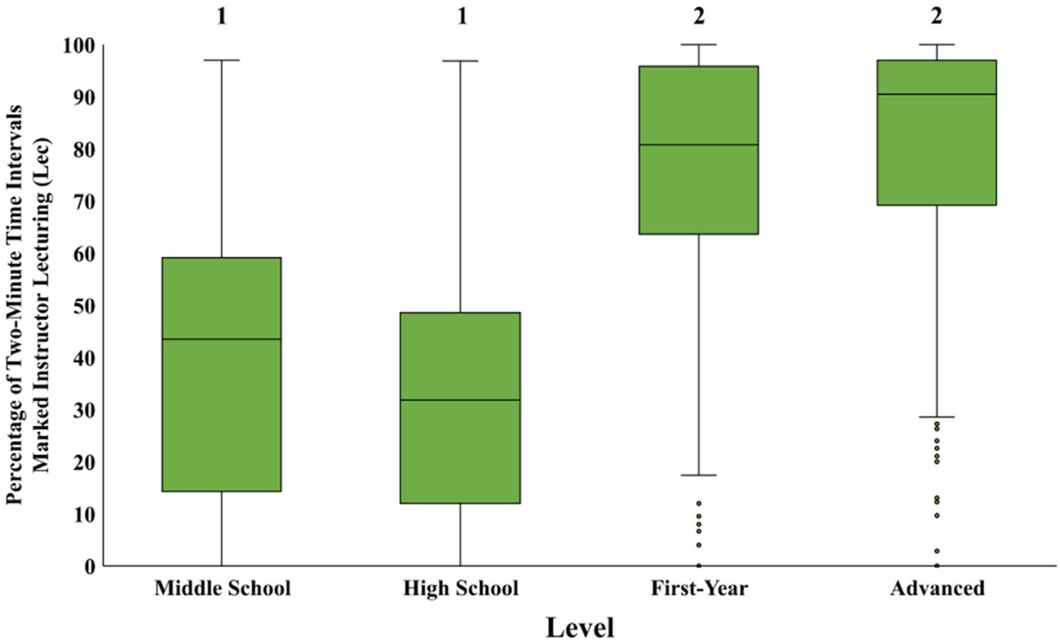

FIGURE 4 | Comparison of the relative frequency of the Classroom Observation Protocol for Undergraduate STEM code Instructor Lecturing (Lec) for middle school, high school, first-year university, and advanced university classes. Box-and-whisker plots showing the median and variation between the four instructional levels. The line in the middle of the box represents the median percentage of 2-min time intervals for the class sessions in each level. Boxes represent the interquartile range, whiskers represent 1.5 times the interquartile range, and data points not included in 1.5 times the interquartile range are shown with dots. Levels labeled with different numbers indicate a significant difference between mean ranks by a *post hoc* Dunn's pairwise comparison (Kruskal–Wallis Test, $\chi^2 = 153.03$, df = 3, $N = 482$, $p < 0.001$; Dunn's pairwise comparisons, $p < 0.001$ for levels labeled with different numbers).

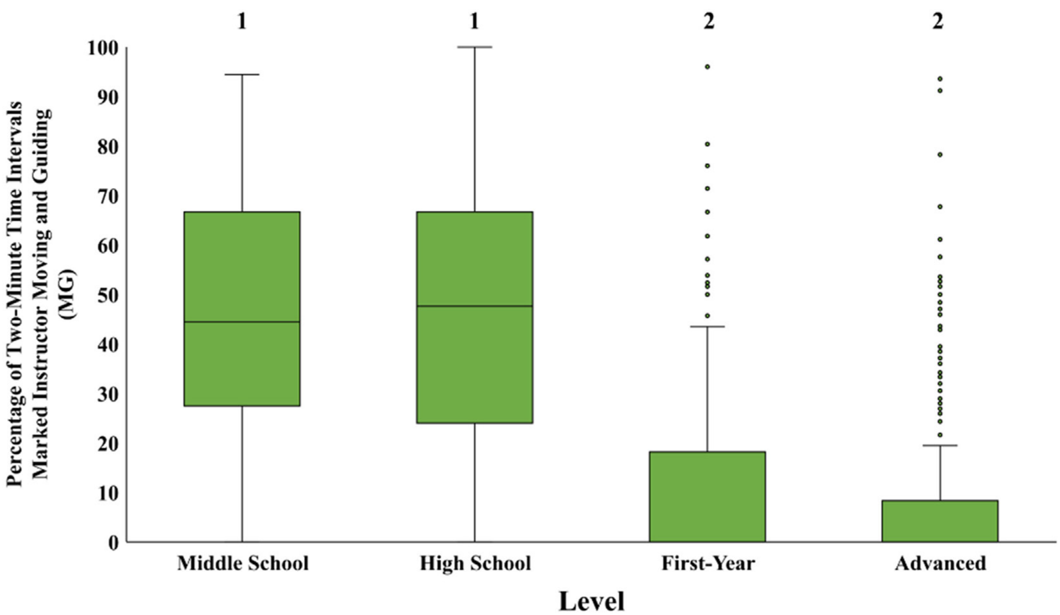

FIGURE 5 | Comparison of the relative frequency of the Classroom Observation Protocol for Undergraduate STEM code Instructor Moving and Guiding (MG) for middle school, high school, first-year university, and advanced university class sessions. Box-and-whisker plots showing the median and variation between the four instructional levels. The line in the middle of the box represents the median percentage of 2-min time intervals for the class sessions in each level. Boxes represent the interquartile range, whiskers represent 1.5 times the interquartile range, and data points not included in 1.5 times the interquartile range are shown with dots. Levels labeled with different numbers indicate a significant difference between mean ranks by a *post hoc* Dunn's pairwise comparison (Kruskal–Wallis Test, $\chi^2 = 169.56$, df = 3, $N = 482$, $p < 0.001$; Dunn's pairwise comparisons, $p < 0.001$ for levels labeled with different numbers).

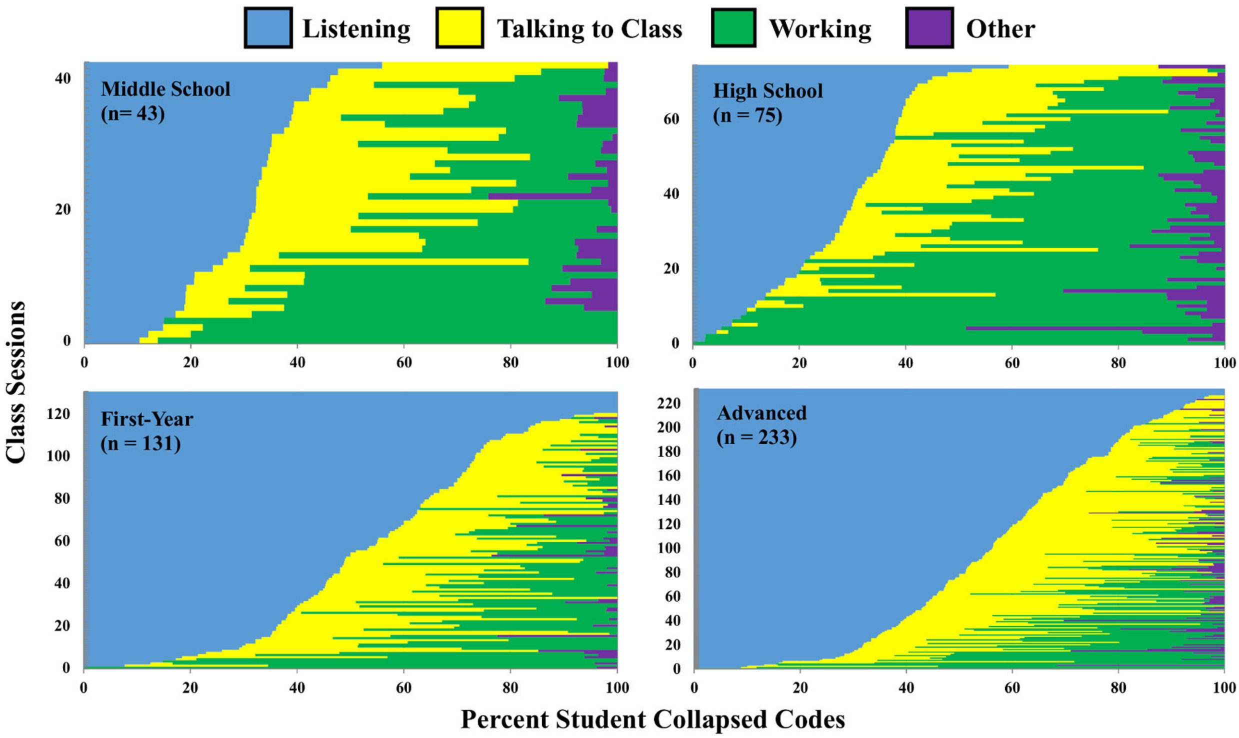

FIGURE 6 | Percentage of student collapsed Classroom Observation Protocol for Undergraduate STEM codes for middle school, high school, first-year, and advanced university classes. Each horizontal bar represents a different class session. The classes are ordered by the collapsed code Student Receiving. **Figure 1** describes the Student Collapsed Codes.

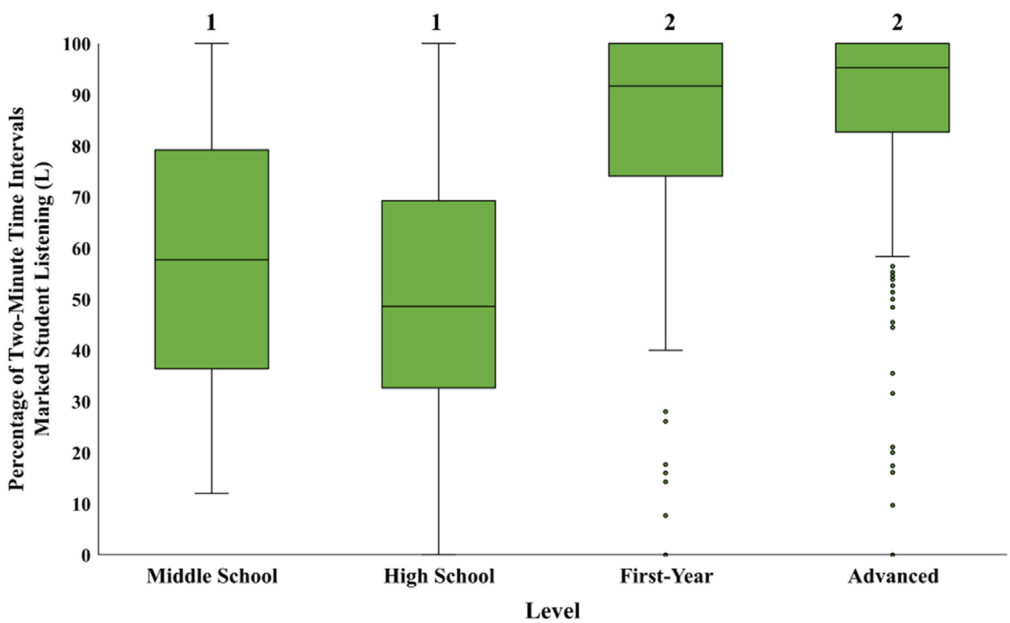

FIGURE 7 | Comparison of the relative frequency of Classroom Observation Protocol for Undergraduate STEM code Student Listening (L) for middle school, high school, first-year university, and advanced university classes. Box-and-whisker plots showing the median and variation between the four instructional levels. The line in the middle of the box represents the median percentage of 2-min time intervals for the class sessions in each level. Boxes represent the interquartile range, whiskers represent 1.5 times the interquartile range, and data points not included in 1.5 times the interquartile range are shown with dots. Levels labeled with different numbers indicate a significant difference between mean ranks by a *post hoc* Dunn's pairwise comparison (Kruskal–Wallis Test, $\chi^2 = 137.37$, df = 3, $N = 482$, $p < 0.001$; Dunn's pairwise comparisons, $p < 0.001$ for levels labeled with different numbers).

GW, a combination of three individual codes: CG, WG, and OG. Middle school and high school classes had a larger interquartile range of percent 2-min time intervals containing a student GW code (**Figure 8**). For both university levels, half of the observations documented no student GW during the entire class. In addition, a Kruskal–Wallis Test showed very strong evidence of a difference ($p < 0.001$) between mean ranks of at least one pair of levels. A Dunn's pairwise test of all six pairs of levels showed students in middle school and high school classes spent significantly more time working in groups than students in first-year and advanced university classes ($p < 0.001$ adjusted using the Bonferroni correction).

Class Size Effect

One common difference between middle school, high school, and university courses is class size. To investigate whether the instructional differences we observed between these educational levels were due to class size, we compared data from all classrooms with fewer than 30 students enrolled. We chose 30 students as the benchmark for small university classes because all of the middle school and high school classrooms we observed contained 30 or fewer students. In total, we had observation data for 74 small university class periods (8 First-Year and 66 Advanced). Even when focusing exclusively on small university classes, we observed that Instructor Presenting and Student Receiving collapsed codes were more common when compared to middle school and high school classrooms (**Figure 9**).

In addition, we compared the median percentages of 2-min time intervals for the same four codes as above: Instructor Lec, Instructor MG, Student Listening, and Student GW in classes with fewer than 30 students enrolled. A Kruskal–Wallis Test with a *post hoc* Dunn's pairwise comparison of all three pairs for all four codes showed instructors in middle school and high school classes spent significantly less time Lec (**Figure 10A**) and significantly more time MG (**Figure 10B**) compared with instructors in small enrollment university classes ($p < 0.001$). In addition, students in middle school and high school classes spent significantly less time Listening (**Figure 10C**) and more time Working in Groups (**Figure 10D**) compared with students in small enrollment university classes ($p < 0.001$).

Length of Class Effect

Another explanation for differences in instructional practices is length of class time. For example, longer class periods may provide more opportunities for active learning. To investigate, we examined the correlation between the total number of 2-min time intervals and percentage of 2-min time intervals with the Instructor Lec code for middle school, high school, first-year university, and advanced university classes. For middle school, high school, and first-year university classes, there is a non-significant correlation between length of time and percent time lecturing (middle school and high school: $r = 0.007$, $R^2 < 0.01$, $p > 0.05$, first-year university: $r = -0.09$, $R^2 < 0.01$, $p > 0.05$). For advanced university classes, there is a significant negative correlation

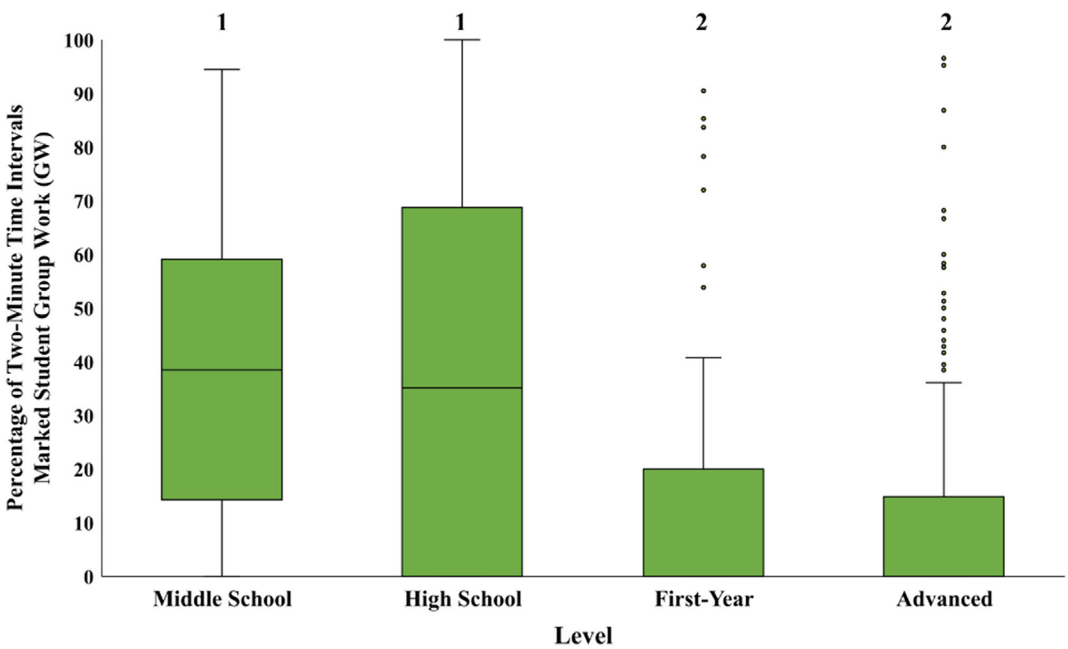

FIGURE 8 | Comparison of the relative frequency of Classroom Observation Protocol for Undergraduate STEM code Student Group Work (GW) for middle school, high school, first-year university, and advanced university classes. Box-and-whisker plots showing the median and variation between the four instructional levels. The line in the middle of the box represents the median percentage of 2-min time intervals for the class sessions in each level. Boxes represent the interquartile range, whiskers represent 1.5 times the interquartile range, and data points not included in 1.5 times the interquartile range are shown with dots. Levels labeled with different numbers indicate a significant difference in mean ranks by a *post hoc* Dunn's pairwise comparison (Kruskal–Wallis Test, $\chi^2 = 81.60$, df = 3, $N = 482$, $p < 0.001$; Dunn's pairwise comparisons, $p < 0.001$ for levels labeled with different numbers).

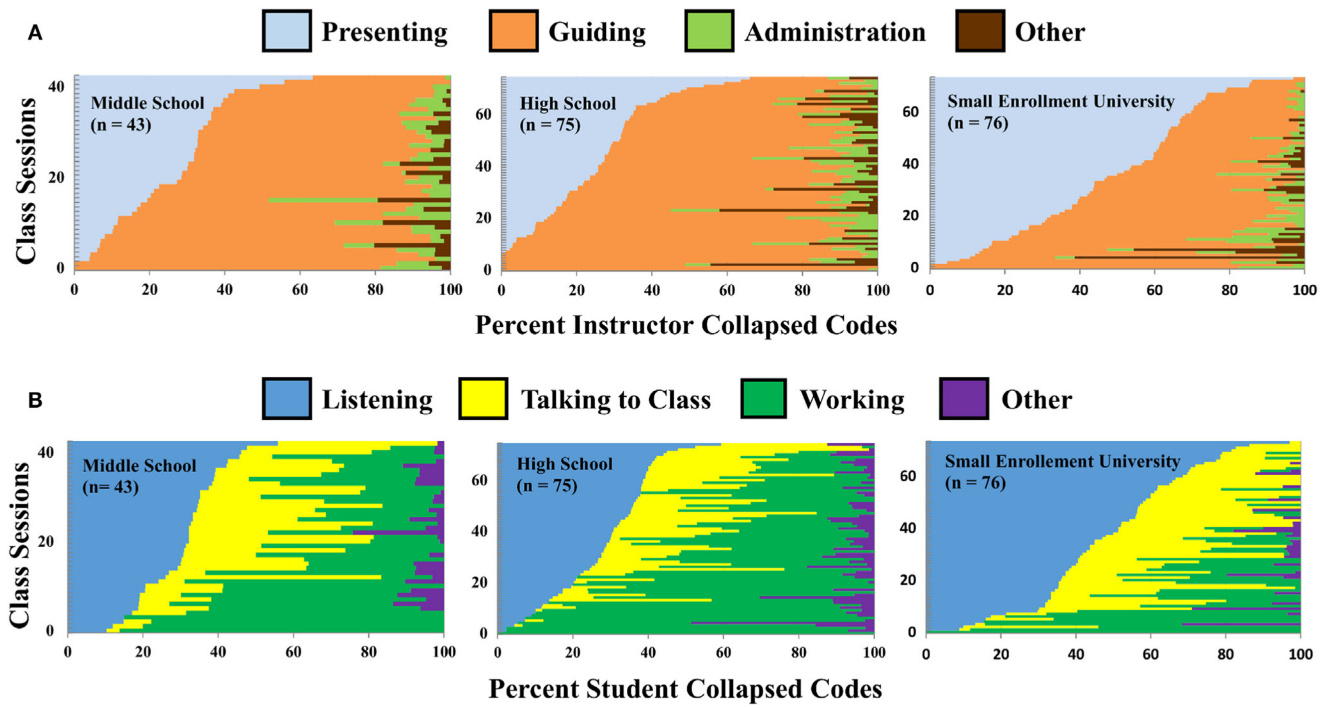

FIGURE 9 | A comparison of middle and high school classrooms with small enrollment university classes. **(A)** Percentage of instructor collapsed Classroom Observation Protocol for Undergraduate STEM (COPUS) codes for middle school, high school, and university classes with fewer than 30 students ordered by percent Instructor Presenting. **(B)** Percentage student collapsed COPUS codes for middle school, high school, and university classes with fewer than 30 students ordered by percent Student Receiving. Each horizontal bar represents a different class session.

($r = -0.18$, $R^2 = 0.03$, $p < 0.05$), which is considered a small to medium effect size (Cohen, 1988), suggesting that longer class periods had fewer 2-min time intervals that included lecturing.

Laboratory Effect

Another difference between middle school, high school, and university classes is the placement of laboratory activities within the course structure. In middle and high school, laboratory activities are incorporated into the same class periods as other class activities. At the university level, laboratories are often scheduled at separate times and in different locations. Because COPUS is designed to capture observation data in the lecture portion of a course, our data set does not include observations of the laboratory sections. Therefore, an explanation for the instructional differences we observed between educational levels could be that at the university-level we were only focusing on the lecture portion of the classes and, therefore, missing other active-learning activities that are part of the course but taught in the laboratory. To investigate whether or not having a required laboratory influenced the amount of active learning that occurred in the lecture portion of the university classroom, we compared data from university courses that did and did not have a required laboratory section associated with the lecture portion of the course. We saw a similar range of the relative abundance of instructor and student COPUS codes in courses that require laboratory sections and those that do not, suggesting that the presence of the laboratory section of a course is not greatly decreasing the amount of active learning occurring in the lecture section (**Figures 11A–D**). Our data also showed that classes taught with and without required laboratory sections had similar interquartile ranges of percent 2-min time interval values for the Instructor Lec and Student Listening codes (**Figures 11E,F**). A comparison of Instructor Lec and Student Listening medians, using Mann–Whitney *U* Tests, shows that there were no statistically significant differences ($p > 0.05$) between classes taught with and without required laboratory sections for the Instructor Lec and Student Listening codes.

Perceptions of Instruction across Educational Levels

To determine how educators perceive instructional differences across multiple education levels, middle school, high school, and university instructors were sent a survey that asked them to predict which of four graphs showed the correct depiction of how much time on average instructors spent lecturing at each educational level (**Figure 12A**). Graph B, less lecturing in middle school and high school classrooms compared with both first-year and advanced university classes, most closely matches the observation data.

All three groups of instructors (middle school, high school, and university) most frequently selected graphs that showed a shift in the amount of lecturing across educational levels (**Figure 13**). Both middle school and high school instructors most commonly

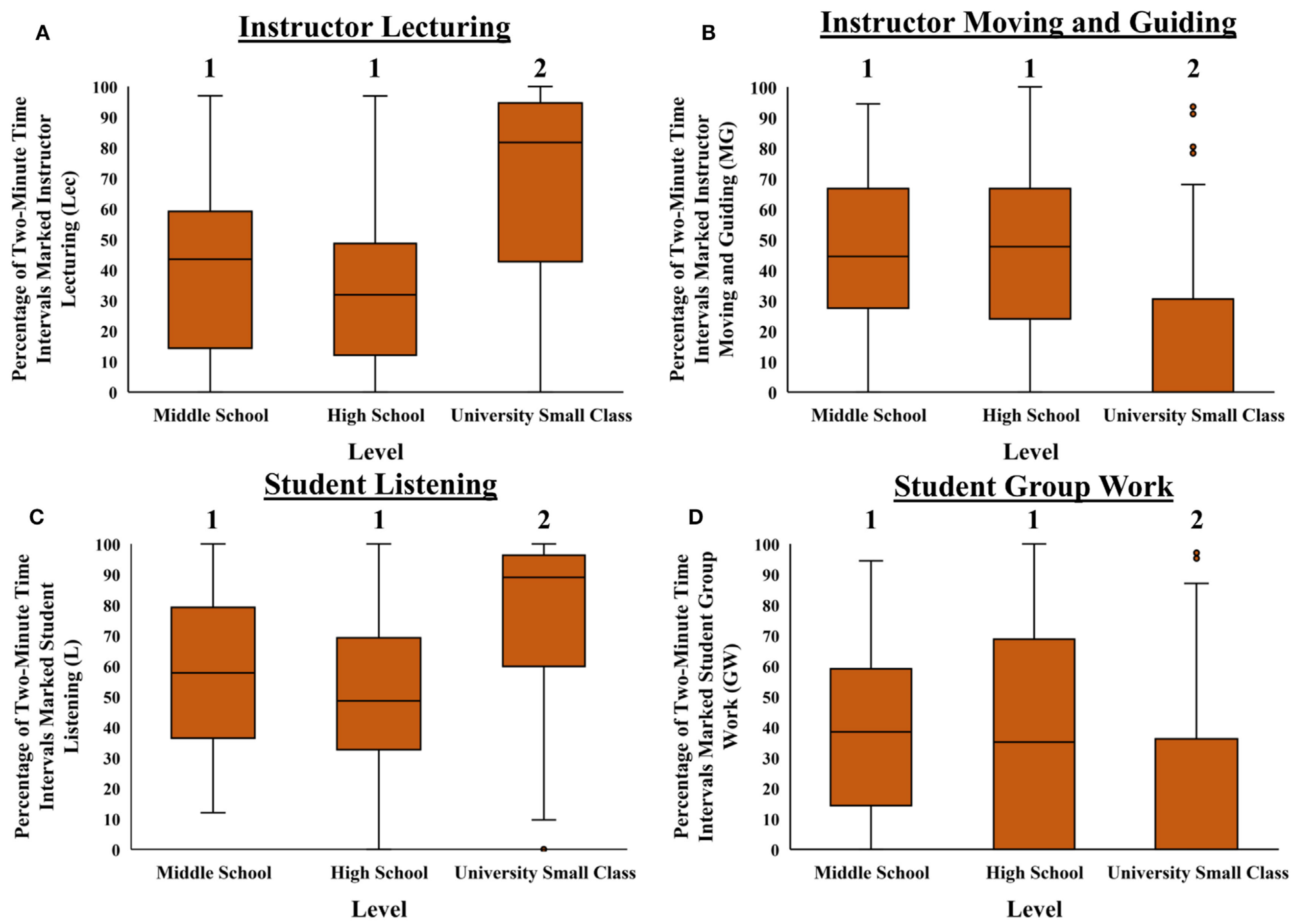

FIGURE 10 | Comparison of the relative frequency of four Classroom Observation Protocol for Undergraduate STEM codes **(A)** Instructor Lecturing (Lec), **(B)** Instructor Moving and Guiding (MG), **(C)** Student Listening (L), and **(D)** Student Group Work (GW) for middle school, high school, and university classes with fewer than 30 students. Box-and-whisker plots showing the median and variation between the four instructional levels. The line in the middle of the box represents the median percentage of 2-min time intervals for the class sessions in each level. Boxes represent the interquartile range, whiskers represent 1.5 times the interquartile range, and data points not included in 1.5 times the interquartile range are shown with dots. Levels labeled with different numbers indicate a significant difference in mean ranks by a *post hoc* Dunn's pairwise comparison (Kruskal–Wallis Test, Instructor Lec $\chi^2 = 43.28$, Instructor MG $\chi^2 = 52.35$, Student Listening $\chi^2 = 40.23$, Student GW $\chi^2 = 18.59$, df = 2, $N = 192$, $p < 0.001$ in all cases; Dunn's pairwise comparisons, $p < 0.001$ for levels labeled with different numbers).

predicted graphs that were not aligned with the observed trend; namely, they tended to select graph A, a gradual increase in the amount of instructor lecturing between each level, and graph D, more instructor lecturing in first-year college classes compared with the other three levels. Because the trends in the data were similar for middle and high school teacher responses, these data were combined in subsequent analyses. University instructors most commonly predicted B, less lecturing in middle school and high school classrooms compared with both first-year and advanced university classes (which most closely matches observation data), and graph D.

The instructors were also asked to explain why they chose a particular graph (**Figure 12B**), and content analysis was used to categorize the responses. Middle and high school instructors who chose graph A, showing an increasing amount of lecturing over all educational levels, most commonly used "personal experience" (47%) as part of their explanation. The instructors drew on experiences both as students and teachers. As one middle school teacher wrote, "From my observations and memories there seems to be an increasing trend to more lecturing and note taking as students progress from middle school to high school to college." Another high school teacher made a clearer distinction by writing, "When I was in class at UMaine in 2002–2004 lecturing was the main teaching method. As a high school teacher now, student exploration is much more prevalent."

Many middle school and high school instructors who chose graph D, showing first-year college classes as having more lecture than the other three levels, pointed to the common difference in class size for these first-year courses (56%). One high school teacher wrote, "First year university classes tend to be very large and held in an area that would be difficult to do anything but lecture." Some viewed the differences in instruction as a result of alternative standards regarding the use of active learning (53%), such as a middle school teacher who wrote, "Active learning is

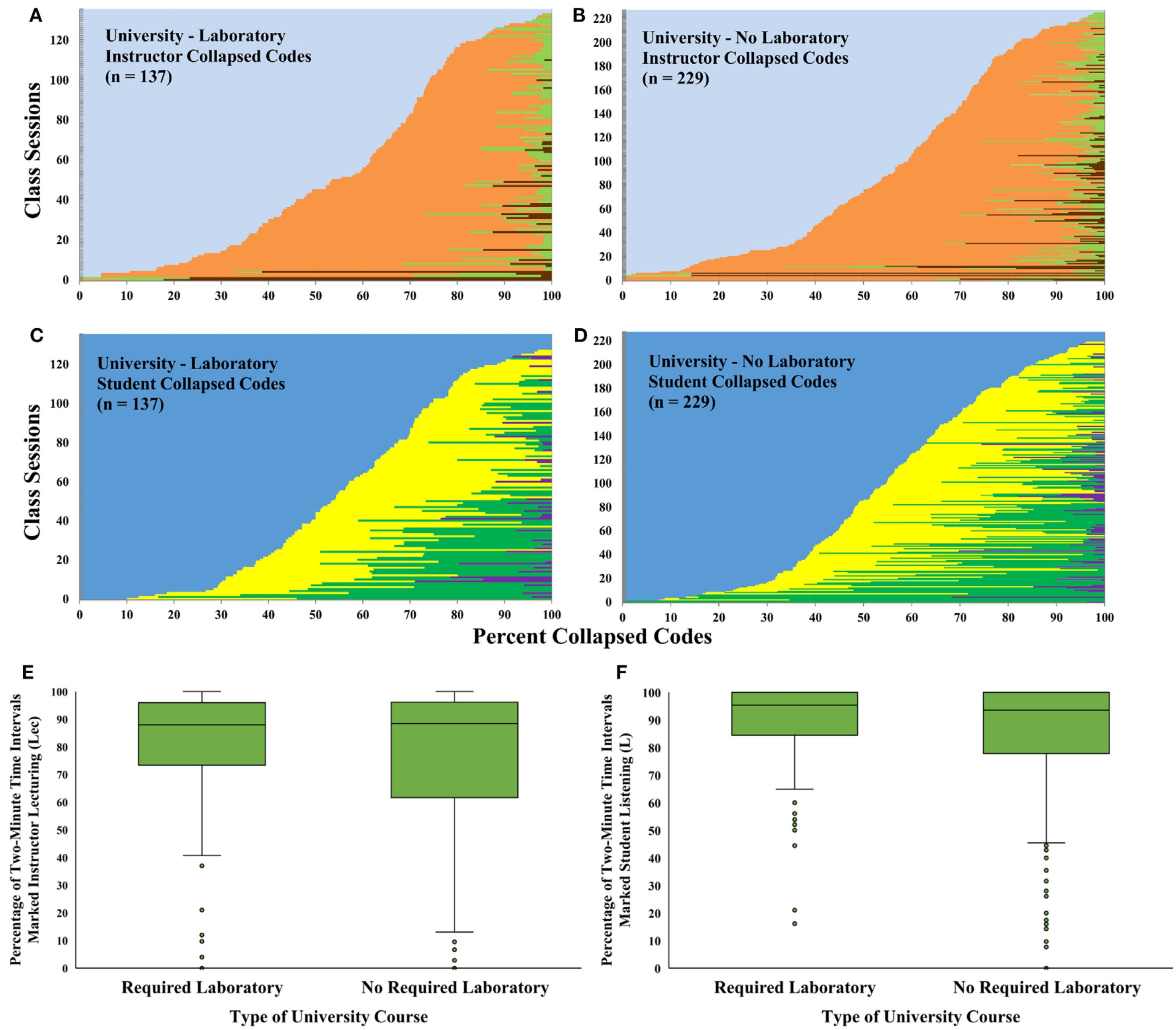

FIGURE 11 | A comparison of university classes with and without a required laboratory section. Percentage of instructor collapsed Classroom Observation Protocol for Undergraduate STEM (COPUS) codes for all university class observations **(A)** with required laboratory sections and **(B)** with no required laboratory section ordered by percent Instructor Presenting. Percentage of student collapsed COPUS codes for all university class observations **(C)** with required laboratory sections and **(D)** with no required laboratory section ordered by percent Student Receiving. Each horizontal bar represents a different class session. **(E)** Box-and-whisker plots showing the median and variation of the COPUS code Instructor Lecturing (Lec) for the two types of university classes. **(F)** Box-and-whisker plots showing the median and variation of the COPUS code Student Listening for the two types of university classes. The line in the middle of the box represents the median percentage of 2-min time intervals. Boxes represent the interquartile range, whiskers represent 1.5 times the interquartile range, and data points not included in 1.5 times the interquartile range are shown with dots. There are no significant differences in Instructor Lec or Student Listening medians for university classes with and without a required laboratory section (Mann–Whitney *U*, Lec χ^2 = 16,526.5, Listening χ^2 = 16,733.0, N = 366, $p > 0.05$ in both cases).

actively encouraged in the middle and high school level as part of our understanding of best practices in pedagogy. University, on the other hand, doesn't require the same level of pedagogical understanding."

Many university instructors who chose graph D pointed to class size (61%) as the effect that the large enrollments of first-year classes have on the instruction, such as one instructor who wrote, "The largest classes are first year university classes and the most traditional way to teach a large class is with lectures." Instructors in this group also stated they had no knowledge of instructional practices at other educational levels (50%): "I have no idea what it would be like in high school, but I would think that first year college classes spend more time lecturing than more senior classes because class size decreases at higher levels."

University instructors who predicted graph B, showing middle school and high school with much less lecturing than both

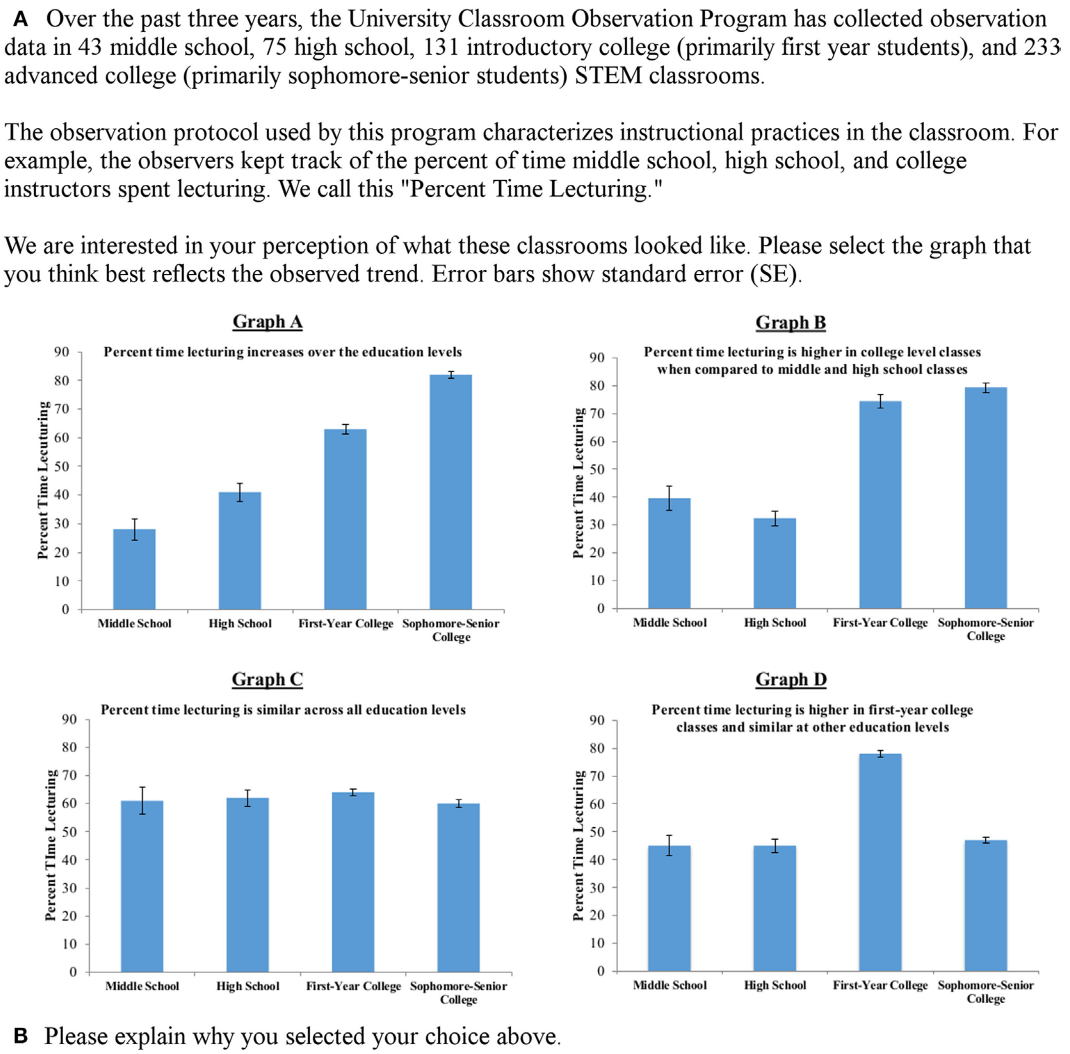

FIGURE 12 | Middle school, high school, and university-level instructors were asked to respond to a survey that included **(A)** a multiple-choice question asking them to predict the Instructor Lecturing code trend we observed with the Classroom Observation Protocol for Undergraduate STEM data and **(B)** an open-response question where they explained their answer choice.

first-year and advanced university classes, most commonly used "active learning" (57%) as an explanation for their choice. Some of these instructors based their reasoning on the perceived needs of students at different levels, such as one instructor who wrote, "I imagine middle and high school students need more hands-on, interactive learning than college-level students." Others invoked the observations they have made in the classroom. A different university instructor explained, "In general, I see more interactive activity happening at the K12 level than the college level."

DISCUSSION

This study is one of the first to compare instructional practices in STEM classrooms across multiple education levels, from middle school through university, using a single observation instrument. Previous studies have used the same instrument to measure instruction in either high school *or* university classrooms (Roehrig and Kruse, 2005; Ebert-May et al., 2011; Yezierski and Herrington, 2011; Lund et al., 2015), but we used the same observation instrument to make direct comparisons across a continuum of educational levels within the same study. Observations conducted with the COPUS instrument show that in middle school and high school classrooms, there is significantly less time dedicated to teacher-centered practices (Instructor Lecturing and Student Listening) and significantly more time dedicated to active-learning practices (Instructor Guiding and Students Working in Groups) compared with university classrooms (**Figures 3–8**). In addition, there are no significant differences in instructional practices or student experiences between first-year and advanced-level university classes. These results show that the largest transition in classroom experiences occurs between high school and first-year undergraduate courses.

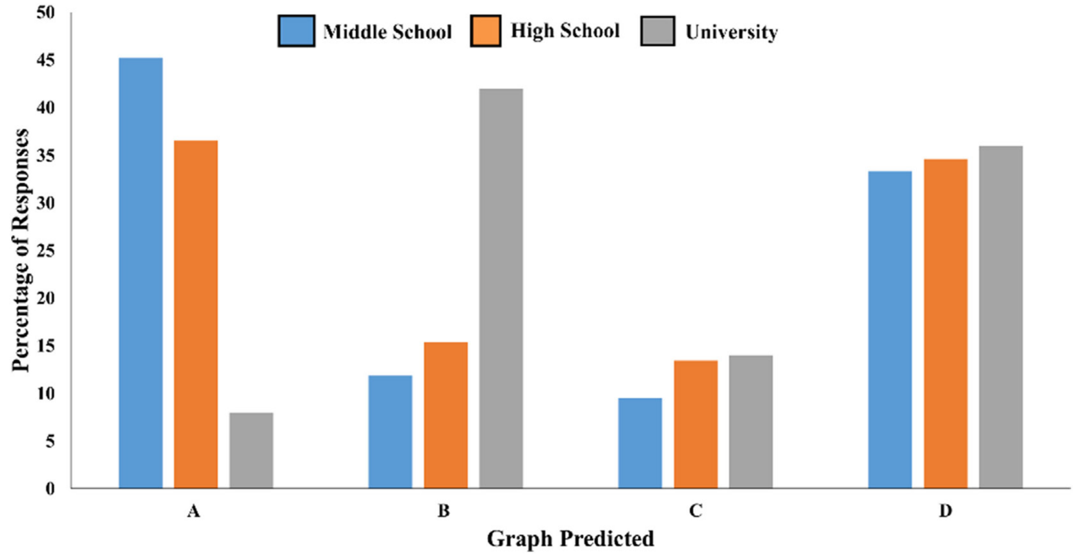

FIGURE 13 | Frequency of responses by graph type for middle school (n = 42), high school (n = 52), and university instructors (n = 50).

Potential explanations for our results include the following: (1) class size effects since university classes are typically much larger than middle school and high school classes, (2) class period length effects, and (3) the fact that laboratory work can be included in middle school and high school class meetings but takes place in dedicated laboratory sections at the university level. We discuss each of these potential explanations below:

Large class size has been reported by faculty as a barrier to implementing active-learning strategies (Henderson et al., 2011; Shadle et al., 2017). Furthermore, the survey data collected as part of this study, which contain explanations of instructor predictions for lecture frequency across multiple educational levels, show that class size was a common explanation for predicted differences in instructional practices across educational levels (**Figures 12** and **13**). However, our observation data reveal that even in small university classes with 30 or fewer students, there is significantly more time dedicated to teacher-centered practices (Instructor Lec and Student Listening) and significantly less time dedicated to active-learning practices (Instructor Guiding and Students Working in Groups) than in middle school and high school classrooms (**Figures 9** and **10**), thereby ruling out this explanation.

Longer classes could provide more opportunities to incorporate active learning into the class period. Our data show that there is a non-significant correlation between class period length and the percentage of time dedicated to lecturing at the middle school, high school, and first-year university levels. At the advanced university level, our data reveal a modest negative correlation between the same two variables, which suggests that having longer time blocks for advanced university classes may enable instructors to use instructional techniques beyond lecture.

Laboratory sections, which are often separate classes at the university level, typically provide opportunities for students to actively engage in doing experiments. Because COPUS is designed for non-laboratory observations, the results here only pertain to the lecture sections of courses, and we are not capturing all the educational opportunities university students engage in during a course. Therefore, it might be predicted that classes with a separate laboratory sections have less active learning in the lecture section because any active components occur in the laboratory. However, we do not find significant differences between the amount of Instructor Lec and Student Listening in university classes that require and do not require a laboratory section (**Figure 11**), which rules out this explanation.

Taken together, these results suggest that the observed differences in instruction between educational levels are not solely a function of class size, class length, or a course structure including a required laboratory section.

One limitation of our study is that the middle school and high school instructors came from a variety of schools and the university faculty came from one institution that is the primary public university in the state of Maine. Future studies should be performed across additional secondary schools and universities to better understand how classroom pedagogies and student transitions are influenced by different school cultures at multiple education levels.

When we surveyed middle school, high school, and university instructors, most were unaware of the instructional differences shown by our findings (**Figure 13**). Other than using class size to explain their predictions, instructors also commonly cited personal experience and/or the perceived amount of active learning at each level as rationale. Our data show that many instructors are unfamiliar with the classroom environments their students are either coming from or heading to in the future. This disconnect represents a barrier to instructional reform aimed at best supporting students as they transition from high school to college. Addressing this issue and developing solutions that target the instructional gap represents an important part of working toward improving retention for undergraduates interested in STEM careers.

How Can We Address the Instructional Gap?

One way to address the instructional gap between high school and first-year undergraduate classes is to promote active learning at the undergraduate level. Due in part to national calls for reform of introductory undergraduate STEM courses (Mervis, 2009; American Association for the Advancement in Science, 2011; President's Council of Advisors on Science and Technology, 2012), many institutions have already begun to implement more active-learning instruction into these courses (Armbruster et al., 2009; Haak et al., 2011; Jensen and Lawson, 2011; Freeman et al., 2014). These changes have led to increases in learning and retention for all students (Freeman et al., 2014), with even greater improvements for traditionally underrepresented minority and first-generation students (Haak et al., 2011; Eddy and Hogan, 2014). Our results show that institutions and instructors could look to high school and middle school classrooms for inspiration on how to begin or continue transforming their introductory courses. In addition, by examining high school and middle school classrooms, university instructors can gain a better understanding of the instructional environment their students most recently experienced.

Given that our results show students experience the greatest shift in classroom experiences between high school and university, institutions and instructors on both sides of the high school to university transition can help students succeed and ultimately persist in STEM degree programs. Due to a number of logistical barriers, connections between instructors at these two levels are rare, but shifting this paradigm could lead to increased instructor awareness and better alignment of instructional practices. One straightforward way to grow connections is through events at which instructors from different levels can meet to discuss common topics, ask questions of one another, and promote a clearer understanding of the types of classrooms students are coming from or heading to in the future. These discussions could be framed around the evidence supporting the use of active learning at the undergraduate level and how it can be effectively used regardless of class size (Resources: http://www.cwsei.ubc.ca/resources/instructor_guidance.htm). Also, because COPUS measures the type of active learning but not necessarily the quality of the teacher–student interactions or educational materials, observation data could be used as a way to start additional conversations about deeper teaching and learning issues.

Classroom observations can provide the basis for another type of productive interaction between teachers at different levels. Specifically, college faculty can observe middle and high school classes and *vice versa*. Previous work has shown that observations can promote change in both the observed instructors and the observers themselves (Cosh, 1998). At the most fundamental level, the feedback received by the instructor based on the observation can lead to an increased awareness of best practices being utilized and areas for future growth. In addition, observing and giving feedback on lessons is helpful for the observer, who can use the opportunity to reflect on their own practices. As a part of the University of Maine's University Classroom Observation Program, middle school and high school teachers observe university faculty and give feedback on specific areas indicated by faculty (Smith et al., 2014). As a result, a subset of the faculty who teach first-year courses have made connections with these teachers and visited high school classrooms. With careful consideration, these types of interactions can be facilitated at multiple levels from individuals to departments to entire institutions. Our group is also beginning to explore long-term professional development activities where groups of university faculty and high school teachers meet regularly to discuss instructional transitions and work toward developing specific approaches that would support student transitions from high school to college STEM instruction.

How Can We Learn More About the Student Experience?

Our results are limited to observation data and instructor perspectives, but additional student surveys could provide useful insight into the perceptions and challenges faced by students as they transition from high school to college. Previous work has used student surveys to document the way groups of students perceive and think about their education. For example, one study investigating what types of expectations students had about pedagogy in college STEM classes found that first-year students expected more active-learning techniques to be used than non-first-year students (Brown et al., 2017). Furthermore, using surveys has been an effective way of measuring student buy-in and engagement with STEM classes (Brazeal et al., 2016; Cavanagh et al., 2016). These studies revealed that students think that active-learning teaching strategies support their learning in class and lead them to engage in more self-regulated learning habits out of class, such as meeting with other students to complete assignments.

To build upon our own findings, student surveys could provide useful information about how students view the transition between high school and university in terms of classroom instruction. For example, these types of student perceptions could inform researchers and instructors alike about which students would be predicted to struggle with the transition to university STEM classrooms. In addition, these surveys would give college students the opportunity to ask questions about the transition, and faculty could be aware of and address these questions in class. Longitudinal studies of how student instructional experiences affect attrition rates and student achievement are also needed to determine the efficacy of increased active learning at the undergraduate level.

CONCLUSION

Our observation-based study of STEM classrooms across multiple educational levels shows that a notable instructional transition occurs between high school and first-year college courses, which cannot solely be attributed to differences in class size, class length, or the presence of dedicated laboratory sections. The shift from more active learning in middle school and high school to classes with more time dedicated to lecture-based instruction at the university level could be contributing to STEM student retention issues. Building on our findings, we propose that future advances in improving retention rates in college STEM majors could be achieved by (1) increasing the amount of interactions between middle school, high school, and university instructors through programs that include classroom observations, (2) developing long-term professional development programs that will work to narrow the instructional gap between high school and university by promoting more active learning in college STEM classrooms, and (3) measuring the efficacy of these programs by tracking the persistence and graduation rates of students who enter universities interested in earning a STEM degree.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the Policies and Procedures for the Protection of Human Subjects of Research, The Institutional Review Board at the University of Maine with informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol were approved by the The Institutional Review Board at the University of Maine.

AUTHOR CONTRIBUTIONS

KA: conceived of study, collected and analyzed data, and drafted manuscript. EB: collected and analyzed data. BC and MS: conceived of study and edited manuscript. JL and EV: collected and analyzed data, and edited manuscript. MRS: conceived of study, discussed data analysis, and edited manuscript. MKS: conceived of study, analyzed data, and drafted manuscript.

ACKNOWLEDGMENTS

This material is based on work supported by the National Science Foundation under grants DRL 0962805 and DUE 1347577, 1712060, and 1712074. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. This work was also supported by a grant from the University of Maine Graduate Student Government. We would like to thank the middle and high school teachers who conducted the university-level observations and Jeremy Smith for his help with the data analysis.

REFERENCES

American Association for the Advancement in Science. (2011). *Vision and Change in Undergraduate Biology Education: A Call to Action*. Washington, DC: American Association for the Advancement in Science.

Armbruster, P., Patel, M., Johnson, E., and Weiss, M. (2009). Active learning and student-centered pedagogy improve student attitudes and performance in introductory biology. *CBE Life Sci. Educ.* 8, 203–213. doi:10.1187/cbe.09

Brazeal, K. R., Brown, T. L., and Couch, B. A. (2016). Characterizing student perceptions of and buy-in toward common formative assessment techniques. *CBE Life Sci. Educ.* 15, 4. doi:10.1187/cbe.16-03-0133

Brockliss, L. (1996). "Curricula," in *A History of the University in Europe: Vol. II, Universities in Early Modern Europe*, ed. H. De Ridder-Symoens (Cambridge: Cambridge University Press), 565–620.

Brown, T. L., Brazeal, K. R., and Couch, B. A. (2017). First-year and non-first-year student expectations regarding in-class and out-of-class learning activities in introductory biology. *J. Microbiol. Biol. Educ.* 18, 1. doi:10.1128/jmbe.v18i1.1241

Campbell, C. M., Cabrera, A. F., Ostrow, M. J., and Patel, S. (2016). From comprehensive to singular: a latent class analysis of college teaching practices. *Res. Higher Educ.* 58, 581–604. doi:10.1007/s11162-016-9440-0

Cavanagh, A. J., Aragón, O. R., Chen, X., Couch, B., Durham, M., Bobrownicki, A., et al. (2016). Student buy-in to active learning in a college science course. *CBE Life Sci. Educ.* 15, 4. doi:10.1187/cbe.16-07-0212

Cleveland, L. M., Olimpo, J. T., and DeChenne-Peters, S. E. (2017). Investigating the relationship between instructors' use of active-learning strategies and students' conceptual understanding and affective changes in introductory biology: a comparison of two active-learning environments. *CBE Life Sci. Educ.* 16, 2. doi:10.1187/cbe.16-06-0181

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Earlbaum Associates.

Cosh, J. (1998). Peer observation in higher education – a reflective approach. *Innov. Educ. Train. Int.* 35, 171–176. doi:10.1080/1355800980350211

Eagan, K., Hurtado, S., Figueroa, T., and Hughes, B. (2014). *Examining STEM Pathways among Students Who Begin College at Four-Year Institutions*. 1–27. Available at: http://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpage/dbasse_088834.pdf

Ebert-May, D., Derting, T., Hodder, J., Momsen, J., Long, T., and Jardeleza, S. (2011). What we say is not what we do: effective evaluation of faculty professional development programs. *Bioscience* 61, 550–558. doi:10.1525/bio.2011.61.7.9

Ebert-May, D., Derting, T. L., Henkel, T. P., Maher, J. M., Momsen, J. L., Arnold, B., et al. (2015). Breaking the cycle: future faculty begin teaching with learner-centered strategies after professional development. *CBE Life Sci. Educ.* 14, 2. doi:10.1187/cbe.14-12-0222

Eddy, S. L., and Hogan, K. A. (2014). Getting under the hood: how and for whom does increasing course structure work? *CBE Life Sci. Educ.* 13, 453–468. doi:10.1187/cbe.14-03-0050

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., et al. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proc. Natl. Acad. Sci. U.S.A.* 111, 8410–8415. doi:10.1073/pnas.1319030111

Garrett, R., and Steinberg, M. P. (2015). Examining teacher effectiveness using classroom observation scores: evidence from the randomization of teachers to students. *Educ. Eval. Policy Anal.* 37, 224–242. doi:10.3102/0162373714537551

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

Henderson, C., Beach, A., and Finkelstein, N. (2011). Facilitating change in undergraduate STEM instructional practices: an analytic review of the literature. *J. Res. Sci. Teach.* 48, 952–984. doi:10.1002/tea.20439

Hora, M. T., Oleson, A., and Ferrare, J. J. (2013). *Teaching Dimensions Observation Protocol (TDOP) User's Manual*. University of Wisconsin-Madison. Available at: http://tdop.wceruw.org/Document/TDOP-Users-Guide.pdf

Jensen, J. L., and Lawson, A. (2011). Effects of collaborative group composition and inquiry instruction on reasoning gains and achievement in undergraduate biology. *CBE Life Sci. Educ.* 10, 64–73. doi:10.1187/cbe.10-07-0089

Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics* 33, 159–174. doi:10.2307/2529310

Lewin, J. D., Vinson, E. L., Stetzer, M. R., and Smith, M. K. (2016). A campus-wide investigation of clicker implementation: the status of peer discussion in STEM classes. *CBE Life Sci. Educ.* 15, 1. doi:10.1187/cbe.15-10-0224

Lund, T. J., Pilarz, M., Velasco, J. B., Chakraverty, D., Rosploch, K., Undersander, M., et al. (2015). The best of both worlds: building on the COPUS and RTOP observation protocols to easily and reliably measure various levels of reformed instructional practice. *CBE Life Sci. Educ.* 14, 2. doi:10.1187/cbe.14-10-0168

Manduca, C. A., Iverson, E. R., Luxenberg, M., Macdonald, R. H., McConnell, D. A., Mogk, D. W., et al. (2017). Improving undergraduate STEM education: the efficacy of discipline-based professional development. *Sci. Adv.* 3, 2. doi:10.1126/sciadv.1600193

Mervis, J. (2009). Universities begin to rethink first-year biology courses. *Science* 325, 5940. doi:10.1126/science.325

Miles, M. B., Huberman, A. M., and Saldana, J. (2013). *Qualitative Data Analysis: A Methods Source Book*. Thousand Oaks, CA: SAGE.

President's Council of Advisors on Science and Technology. (2012). *Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics*. President's Council of Advisors on Science and Technology.

Prince, M. (2004). Does active learning work? A review of the research. *J. Eng. Educ.* 93, 223–231. doi:10.1002/j.2168-9830.2004.tb00809.x

Pryor, J. H., and Eagan, K. (2013). *The American Freshman: National Norms Fall 2012*. Available at: https://www.heri.ucla.edu/monographs/TheAmericanFreshman2012-Expanded.pdf

Rockoff, J. E., Jacob, B. A., Kane, T. J., and Staiger, D. O. (2008). Can you recognize an effective teacher when you recruit one? *Natl. Bur. Econ. Res.* 6, 43–74. doi:10.1017/CBO9781107415324.004

Roehrig, G. H., and Kruse, R. A. (2005). The role of teachers' beliefs and knowledge in the adoption of a reform-based. *School Sci. Math.* 105, 412–422. doi:10.1111/j.1949-8594.2005.tb18061.x

Sawada, D., Piburn, M. D., Judson, E., Turley, J., Falconer, K., Benford, R., et al. (2002). Measuring reform practices in science and mathematics classrooms: the reformed teaching observation protocol. *School Sci. Math.* 102, 245–253. doi:10.1111/j.1949-8594.2002.tb17883.x

Seymour, E., and Hewitt, N. M. (1997). *Talking About Leaving*. Boulder, CO: Westview Press.

Shadle, S. E., Marker, A., and Earl, B. (2017). Faculty drivers and barriers: laying the groundwork for undergraduate STEM education reform in academic departments. *Int. J. STEM Educ.* 4, 1. doi:10.1186/s40594-017-0062-7

Smith, M. K., Jones, F. H. M., Gilbert, S. L., and Wieman, C. E. (2013). The classroom observation protocol for undergraduate stem (COPUS): a new instrument to characterize university STEM classroom practices. *CBE Life Sci. Educ.* 12, 618–627. doi:10.1187/cbe.13-08-0154

Smith, M. K., Vinson, E. L., Smith, J. A., Lewin, J. D., and Stetzer, M. R. (2014). A campus-wide study of STEM courses: new perspectives on teaching practices and perceptions. *CBE Life Sci. Educ.* 13, 624–635. doi:10.1187/cbe.14-06-0108

Van Noy, M., and Zeidenberg, M. (2014). *Hidden STEM Producers: Community Colleges' Multiple Contributions to STEM Education and Workforce Development*. Available at: http://nationalacademies.org/

Watkins, B. J., and Mazur, E. (2013). Retaining students in science, technology, engineering, and mathematics (STEM) majors. *J. Coll. Sci. Teach.* 42, 36–41.

Yezierski, E. J., and Herrington, D. G. (2011). Improving practice with target inquiry: high school chemistry teacher professional development that works. *Chem. Educ. Res. Pract.* 12, 344–354. doi:10.1039/C1RP90041B

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Education

ORIGINAL RESEARCH
published: 07 March 2018
doi: 10.3389/feduc.2018.00008

Use of a Social Annotation Platform for Pre-Class Reading Assignments in a Flipped Introductory Physics Class

Kelly Miller[1]*, *Brian Lukoff*[2], *Gary King*[1] *and Eric Mazur*[1]

[1] Harvard University, Cambridge, MA, United States, [2] Perusall LLC, Brookline, MA, United States

In this paper, we illustrate the successful implementation of pre-class reading assignments through a social learning platform that allows students to discuss the reading online with their classmates. We show how the platform can be used to understand how students are reading before class. We find that, with this platform, students spend an above average amount of time reading (compared to that reported in the literature) and that most students complete their reading assignments before class. We identify specific reading behaviors that are predictive of in-class exam performance. We also demonstrate ways that the platform promotes active reading strategies and produces high-quality learning interactions between students outside class. Finally, we compare the exam performance of two cohorts of students, where the only difference between them is the use of the platform; we show that students do significantly better on exams when using the platform.

Keywords: digital education, flipped classroom, educational software, pre-class reading, physics education research

OPEN ACCESS

Edited by:
Elizabeth S. Charles,
Dawson College, Canada

Reviewed by:
Liisa Ilomäki,
University of Helsinki, Finland
Michael Dugdale,
John Abbott College, Canada

Correspondence:
Kelly Miller
kmiller@seas.harvard.edu

Specialty section:
This article was submitted to Digital Education, a section of the journal Frontiers in Education

Received: *29 September 2017*
Accepted: *19 January 2018*
Published: *07 March 2018*

Citation:
Miller K, Lukoff B, King G and Mazur E (2018) Use of a Social Annotation Platform for Pre-Class Reading Assignments in a Flipped Introductory Physics Class. Front. Educ. 3:8. doi: 10.3389/feduc.2018.00008

INTRODUCTION

Getting students to read the textbook before coming to class is an important problem in higher education. This is increasingly the case as more college classes are adopting "flipped" teaching strategies. A key principle of the flipped classroom model is that students benefit from having access to the instructor (and other peers) when working on activities that, in traditional classrooms, are typically done at home (like problem sets). Moving these activities in class improves student learning as it provides them the opportunity to actively engage with the instructor and each other (Herreid and Schiller, 2013). In a flipped class, the information transfer (traditionally accomplished by the instructor delivering a lecture during class) is moved outside the classroom to a pre-class assignment that students are expected to complete before coming to class. These pre-class assignments typically require students watch a video of a lecture online, or complete a reading. Moving the information delivery out of the classroom allows in-class time to be used for more interactive activities during which students can be actively engaged with instructors and other students.

When students are exposed to the material before class they are better able to follow material in class (Schwartz and Bransford, 1998), they ask more meaningful questions in class (Marcell, 2008), and they perform better on exams (Narloch et al., 2006; Dobson, 2008; Johnson and Kiviniemi, 2009). Students report that one of the most important factors in deciding whether to participate in class is reading the textbook beforehand (Karp and Yoels, 1976). The connection between pre-class reading and in-class participation is particularly relevant in flipped courses that rely on active in-class participation.

As pre-class reading assignments replace lectures in flipped courses and serve as the primary mechanism for information transfer, it is essential that students complete their assignments before class. Even in traditional (non-flipped) college courses, pre-class reading has been shown to be important for student learning and yet 60–80% of students do not read the textbook before coming to class (Cummings et al., 2002; Clump et al., 2004; Podolefsky and Finkelstein, 2006; Stelzer et al., 2009). Clump et al. (2004) studied the extent to which psychology students reported reading their textbooks and found that students only read, on average, 28% of the assigned reading before class and 70% before the exam.

Other studies have looked at how much time students spend reading textbooks, and when they read. Berry et al. (2010) studied pre-class reading habits of undergraduate students enrolled in finance courses across three different universities. They found that 18% of students report not reading the textbook at all, and approximately 92% of students report spending 3 h or less per week reading. Almost half the students (43%) report reading the textbook for less than an hour a week (Berry et al., 2010). The authors also polled the students to find out when they read. Despite instructors' recommendation to read before class, very few students actually do so: just 18% of students report that they frequently read before class; 53% report never or rarely reading the textbook before class (Berry et al., 2010). Podolefsky and Finkelstein (2006) found that only 37% of students report regularly reading the textbook and less than 13% read before class. Instead of reading before class, students reported reading predominately in preparation for exams, to find the answer to a specific question, or to help complete homework (Berry et al., 2010).

There are several reasons why students are not reading before class. Some studies suggest that students do not see the connection between doing well on exams and pre-class reading (Podolefsky and Finkelstein, 2006). To strengthen the connection between pre-class reading and course grades, many instructors have implemented graded, pre-class reading quizzes (Burchfield and Sappington, 2000; Connor-Greene, 2000; Ruscio, 2001; Sappington et al., 2002). "Just in-time teaching" (JITT) is one specific implementation of this way of handling pre-class reading (Novak et al., 1999). With JITT, before class, students are required to answer open-ended questions about the reading online, with one question dedicated to soliciting feedback from students about what aspect of the reading they found most confusing. Instructors can use this feedback to tailor their in-class activities and instruction to the most popular areas of student confusion. Even with grade incentives, however, the rate of pre-class reading compliance is still surprisingly low. Stelzer et al. (2009) reported that even with JITT, 70% of students never or rarely read the textbook before class. Heiner et al. (2014) recorded student pre-class reading compliance in two different classes with a JITT-like implementation of short, targeted readings and associated online reading quizzes (Heiner et al., 2014). They found that 79% of students in one class and 85% of students in the other classes reported reading the pre-class reading assignment every week (or most weeks). While these results are promising, this study (as well as the other mentioned studies on pre-class reading) relied on student-reported responses; Sappington et al. (2002) found that students' self-reported reading compliance is often distorted and invalid.

It is unclear from the literature whether there is a relationship between pre-class reading behavior and in-class exam performance. Podolefsky and Finkelstein (2006) studied the relationship between the frequency with which students report reading and course grades. They conducted this study in three different types of courses: calculus-based physics, algebra-based physics, and conceptual physics. For the calculus and algebra-based courses, they found no significant correlation between course grade and how much students reported reading. For the conceptual course, they found a moderate correlation. Smith and Jacobs (2003) looked at the correlation between time spent reading and course grade for chemistry students and also found no correlation between time spent reading (based on student self-reported data) and course grades. Heiner et al. (2014) found a statistically significant positive correlation between students' exam performance and the frequency with which students completed the online reading quiz. Because much of the research is based on students' self-reported data and because of the lack of consensus about the relationship between pre-class reading and grades, we set out to systematically study this relationship.

The research questions we address are:

(1) What are students' pre-class reading habits on a social learning platform?
(2) Which pre-class reading behaviors are predictive of student in-class exam performance?
(3) What is the efficacy of the platform in promoting student learning?

THEORETICAL FRAMEWORK

It is generally accepted that students understand material better after discussing it with others (Bonwell and Eison, 1991; Sorcinelli, 1991). From the social constructivism perspective, students learn through the process of sharing experiences and building knowledge and understanding through discussion (Vygotsky, 1978). Online learning communities are virtual places that combine learning and community together (Downes, 1999) and provide environments for learners to collaboratively build knowledge. Collaborative learning settings provide students a space to verbalize their thinking, build understanding, and solve problems together (Webb et al., 1995; Crouch and Mazur, 2001).

Online discussion forums have been used successfully as tools to facilitate social interactions and exchanges of knowledge between learners (Rovai, 2002; Bradshaw and Hinton, 2004; Tallent-Runnels et al., 2006). The social constructive theory of learning with technology emphasizes that successful learning requires continuous conversation between learners as well as between instructors and learners (Brown and Campione, 1996). As a result, when designing online learning strategies, educators should create social environments with a high degree of interactivity (Maor and Volet, 2007). The asynchronous nature of online discussion forums allows for discussion between learners and between learners and instructors at any time of day or night, and this is a major advantage over other forms of learning environments (Nandi et al., 2009).

Beyond online discussion forums, collaborative annotation systems have recently been developed and used in education as social learning communities. Online annotation systems are computer-mediated communication tools that allow groups of people to collaboratively read and annotate material online. Many studies have shown that online annotation systems increase student learning across many different educational settings (Quade, 1996; Cadiz et al., 2000; Nokelainen et al., 2003; Hwang and Wang, 2004; Marshall and Brush, 2004; Ahren, 2005; Gupta et al., 2008; Robert, 2009; Su et al., 2010).

PERUSALL: SOCIAL LEARNING PLATFORM FOR READING AND ANNOTATING

Perusall is an online, social learning platform designed to promote high pre-class reading compliance, engagement, and conceptual understanding. The instructor creates an online course on *Perusall*, adopting electronic versions of textbooks from publishers or uploading articles or documents, and then creates reading assignments. Students asynchronously annotate the assigned reading by posting (or replying to) comments or questions in a chat-like fashion.

An instructor view of the course home page is shown in **Figure 1**. The instructor uploads the reading material to the left-hand side of the page (under Documents) and then creates specific reading assignments from these documents which appear in the right panel.

Figure 2 shows what a student sees after opening a reading assignment and highlighting a specific passage on a page in the assignment. A conversation window opens on the right where the student can ask a question or make a comment.

Figure 3 shows a page that has been highlighted and annotated by students. When a student clicks on a specific highlight that highlight turns purple, and the conversation window for that highlight opens on the right.

When a student asks a question about a specific passage, it is automatically flagged with an orange question mark, as shown in **Figure 3**. Other students can respond in an asynchronous conversation.

Perusall also has an integrated assessment tool that provides both students and instructors with constant feedback on how students are engaging with the reading assignments. Finally, *Perusall* has a built-in tool for instructors called the Confusion Report. This report automatically summarizes the top areas of student confusion for instructors so that they can prepare class material that is targeted specifically to the content that students are struggling with the most.

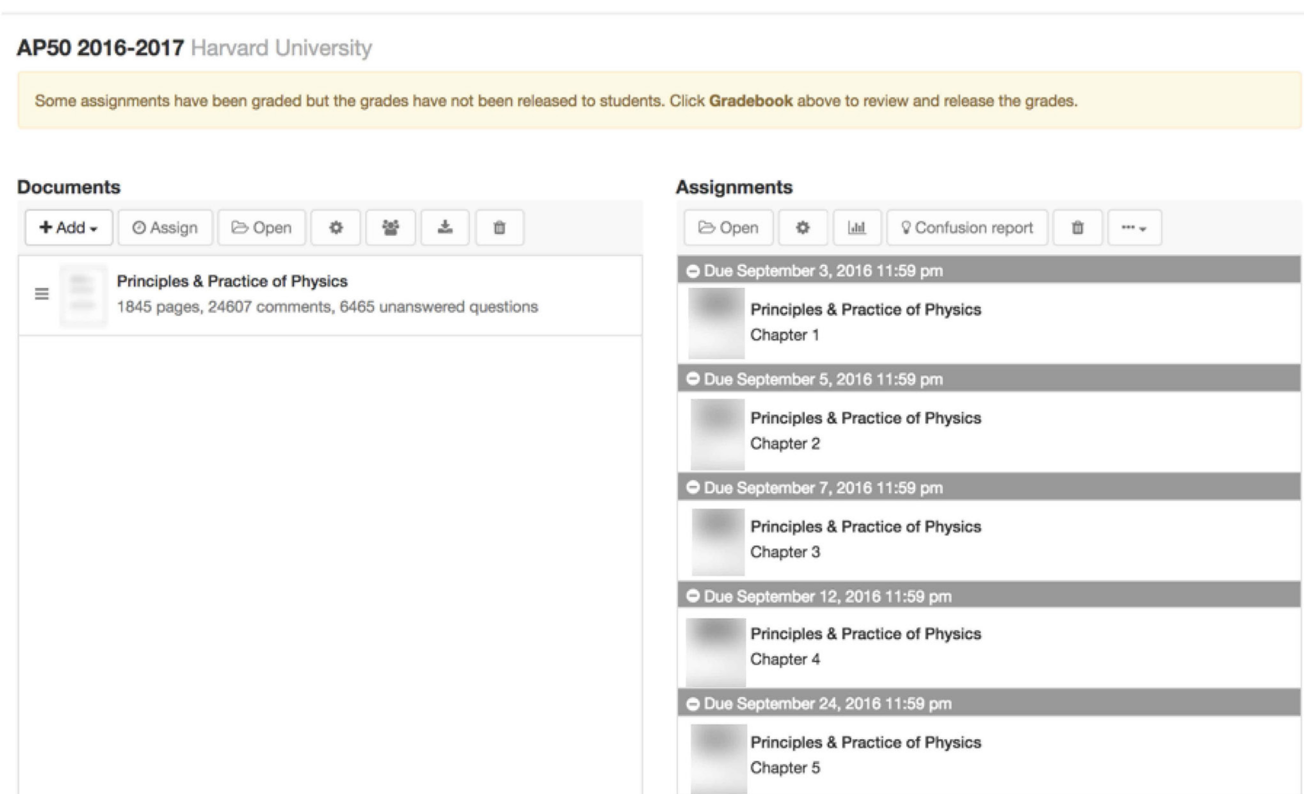

FIGURE 1 | *Persuall* instructor course view.

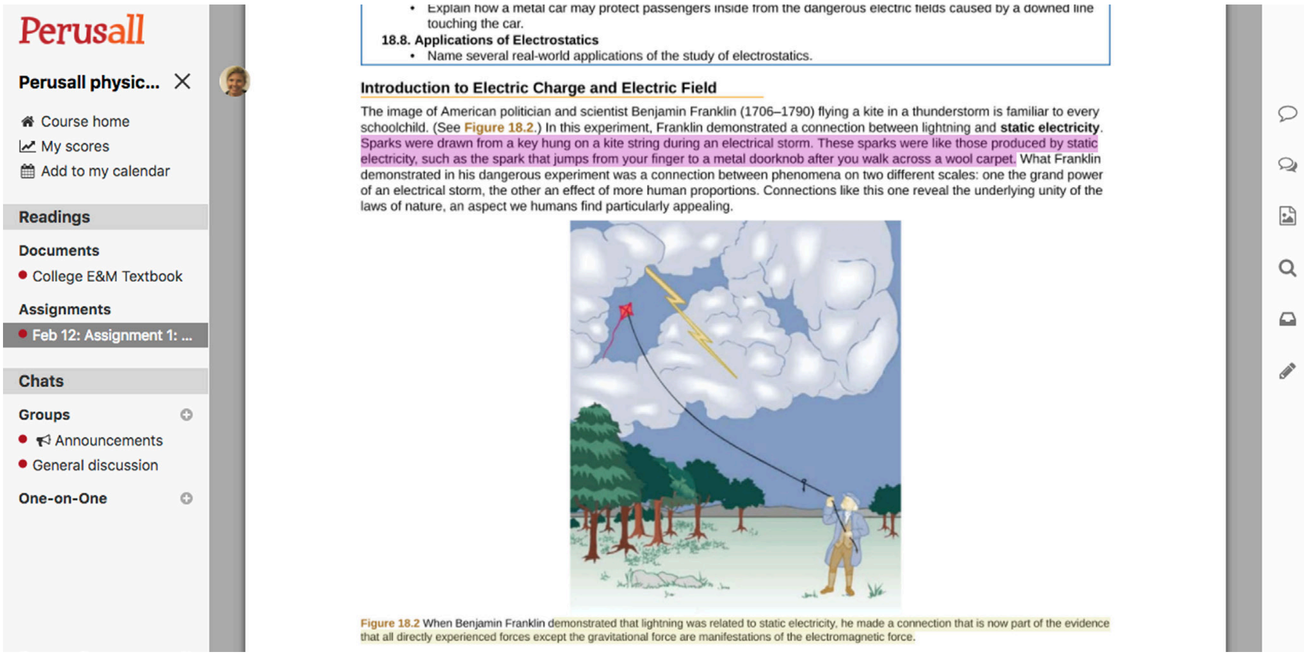

FIGURE 2 | Page of a reading assignment in *Perusall.* (Note: the depicted individual provided written informed consent for the publication of their identifiable image, the image of the textbook book is from OpenStax, University Physics, Volume 1. "Download for free at https://openstax.org/details/books/university-physics-volume-1," no further permission is required from the copyright holders for the reproduction of this material.)

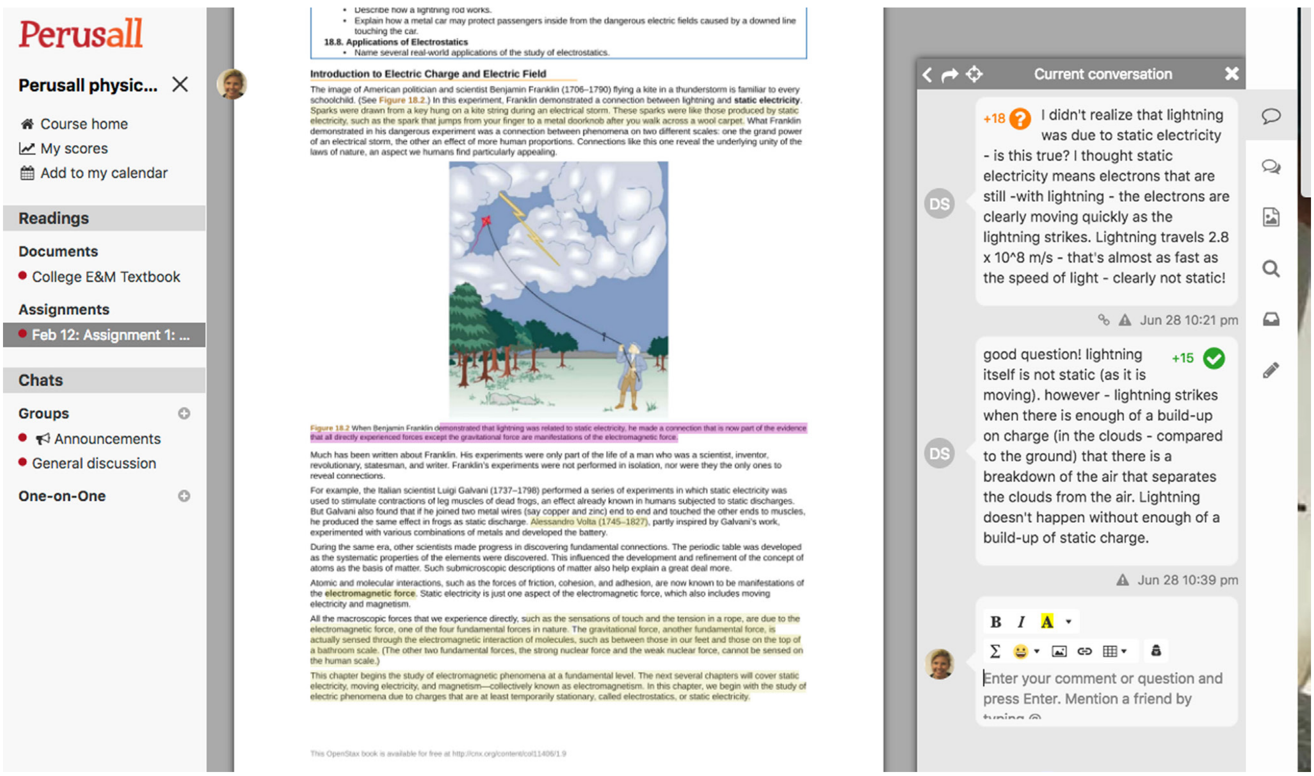

FIGURE 3 | Reading assignment in *Perusall* showing student highlights and annotations (note: the depicted individual provided written informed consent for the publication of their identifiable image, the image of the textbook book is from OpenStax, University Physics, Volume 1. "Download for free at https://openstax.org/details/books/university-physics-volume-1," no further permission is required from the copyright holders for the reproduction of this material).

Social Features

In addition to the basic highlighting and annotating functions, *Perusall* has a number of additional features designed to turn the online reading assignment into a social experience to encourage students to engage with the material and with fellow classmates outside of class. Several features of the software are designed to promote the social aspect of the software.

Sectioning

If the class exceeds 20 students (or another threshold set by the instructor), the software automatically partitions students in the class into groups that function like "virtual class sections." Students can only interact with and see annotations posted by others in their group (as well as any annotations posted by the instructor). This allows students to become more familiar with the other students in their group, and this familiarity helps promote more online interaction. Our prior work demonstrated that when the size of the group is too large, the overall quality of students' annotations decreases (Miller et al., 2016), so these smaller groups prevent students from becoming overwhelmed by an excessively large number annotations and helps keep the overall quality of the interactions high.

Avatars

The avatars of other students and instructors who are viewing the same assignment at the same time appear in the top left hand corner of the screen (**Figure 2**). Being able to see classmates (and instructors) reading the assignment at the same time increases the social connectivity of the reading experience and encourages students to engage more with the reading (through the software).

Upvoting

Students can provide feedback on the annotations made by other students in their section by "upvoting" annotations. There are two types of upvoting in *Perusall*. When students would like to know the answer to a question posed by another student, they can indicate this by clicking on the orange question mark. For example, **Figure 4** shows that three students clicked on the orange question mark button for that question, indicating that they too would like to know the answer. When instructors review questions in *Perusall*, they can pay particular attention to the questions that have been upvoted by other students.

When a student provides a particularly helpful explanation, other students can indicate this by clicking on the green checkmark. In **Figure 5**, five students found the explanation to the initial question to be helpful to their understanding. When students upvote other students' explanations, it helps other students find explanations that are particularly helpful to their conceptual understanding of the reading. Both of these upvoting features are designed to increase and encourage the social component of the online reading assignments and foster a sense of community within the groups.

Email Notifications

Finally, *Perusall* has an email notification feature that is designed to encourage the social interaction aspect of the software even when students are not logged into *Perusall* by letting them know when a classmate has responded to a question or comment they have made (or have clicked the question mark button for). Through the notification, *Perusall* encourages students to continue their conversation about the reading. **Figure 6** shows an example of an email notification that a student receives when a classmate

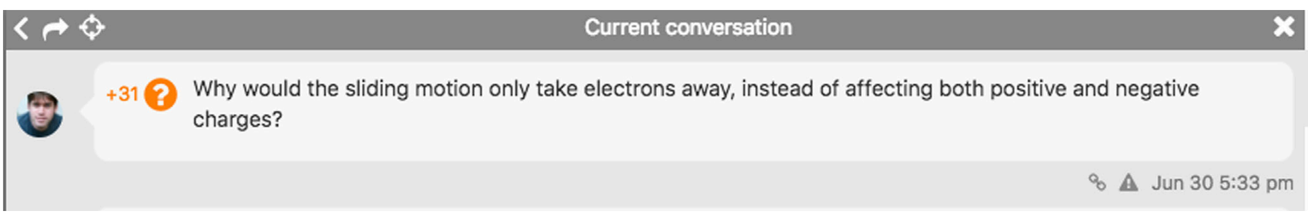

FIGURE 4 | Upvoting of questions in *Perusall* (note: the depicted individual provided written informed consent for the publication of their identifiable image).

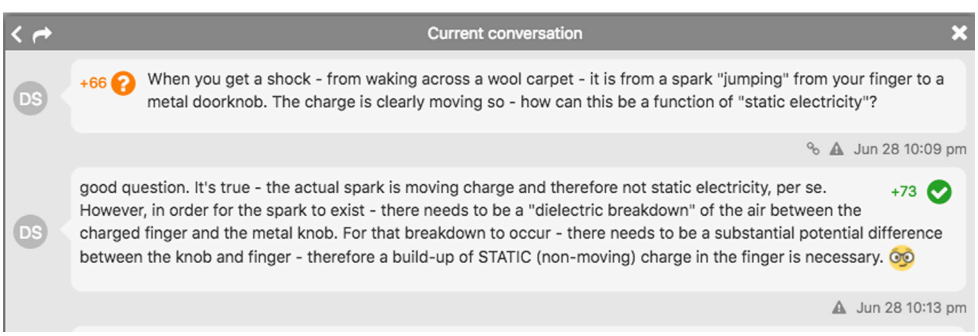

FIGURE 5 | Upvoting of explanations in *Perusall* (note: the depicted individuals provided written informed consent for the publication of their identifiable images).

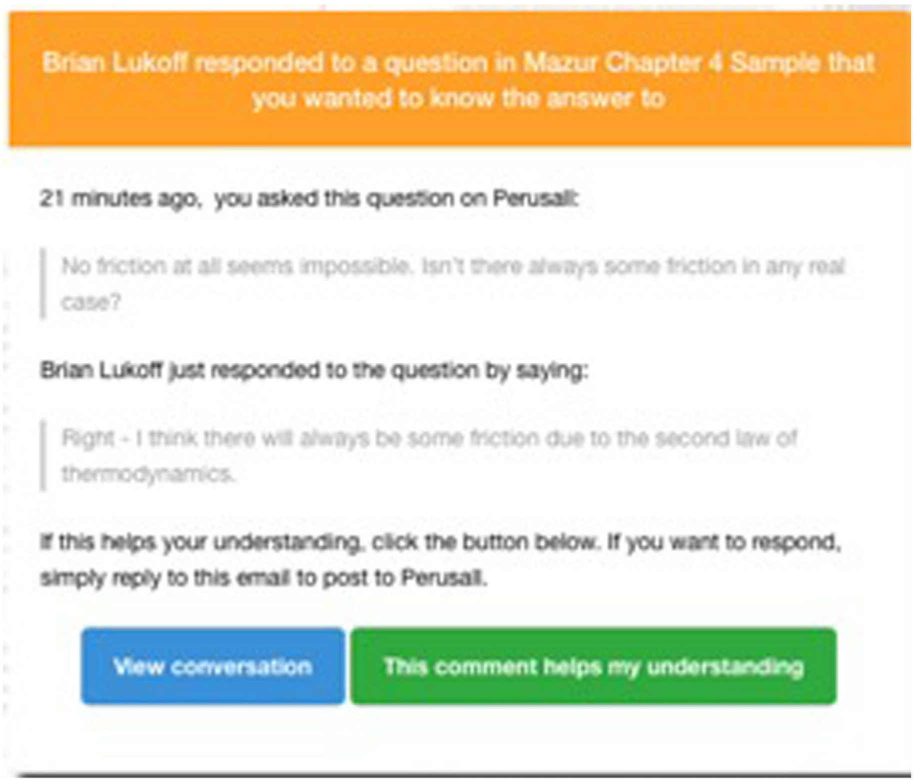

FIGURE 6 | Email notification from *Perusall*.

responds to his/her question. The notification encourages students to re-engage with the reading assignment by viewing the conversation and/or letting the responder know whether the response was helpful to their understanding. Students can reply to the email inside their email client, and the reply is appended directly to the conversation in *Perusall*, as if the student had been online.

Assessment

Perusall has an integrated assessment tool, which automatically evaluates students' participation in the reading assignment and populates an integrated gradebook (**Figure 7**).

The grading algorithm uses four criteria to evaluate a students' collection of annotations for any given reading assignment—timeliness, quantity, quality, and distribution—and students receive an overall score based on all four of these criteria. The grading algorithm uses machine learning to drive desirable student behavior: timely, thorough, and complete reading of the text, with annotations that demonstrate thoughtful interpretation of the subject matter. Students receive a score based on how closely their overall reading and annotating behavior matches behavior that is predictive of success in the classroom.

Instructor Tools

Besides the gradebook and individual reading assignment feedback, which provides important assessment information to both students and instructors, *Perusall* also assists instructors in identifying from the body of annotations the top areas of confusion so they can prepare class material that is targeted specifically at addressing these areas. To this end *Perusall* automatically mines questions that students are asking about a particular reading assignment and, using a topic modeling algorithm, groups questions into three to four conceptual areas of confusion. **Figure 8** shows an example of a confusion report generated for a specific reading assignment. The philosophy behind the confusion report is based on Just-in-Time-Teaching (Novak, 2011), which uses feedback from work that students do at home (like pre-class reading assignments) to inform what is done in the classroom.

RESEARCH METHODS

Participants

The participants in this study were undergraduate students enrolled in an introductory physics course. We collected reading assignment data and exam data from two semesters of the course when *Perusall* was used [spring of 2015 (S15) and fall of 2016 (F16)]. In S15, there were 74 students and in F16 there were 79 students. Students in the S15 course were not the same students as in the F16 course. Due to the fact that the event-tracking feature of the software was not yet developed in S15, most of the analysis focuses on the F16 cohort. As a point of comparison, we also collected exam performance data during the two previous semesters of the course when a different social annotation platform was being used [spring of 2014 (S14) and fall of 2015 (F15)]. There were 72 students enrolled in the S14 course and 75

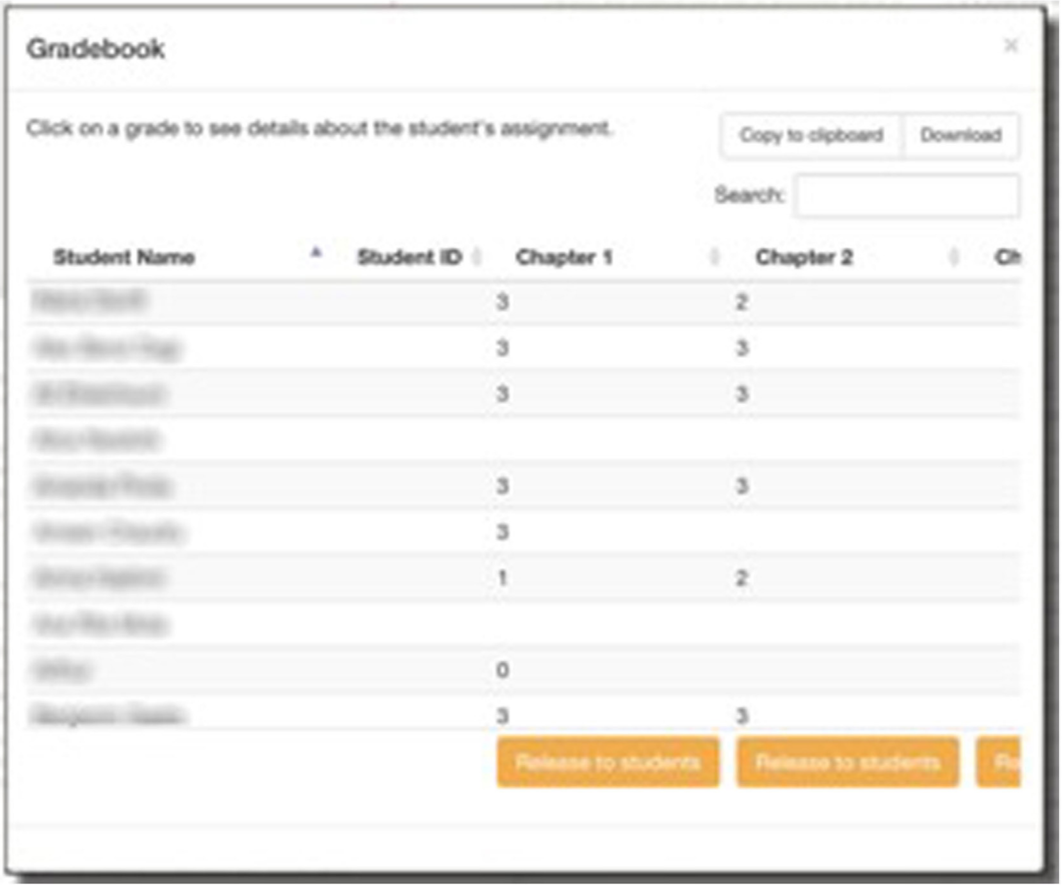

FIGURE 7 | Gradebook in *Perusall.*

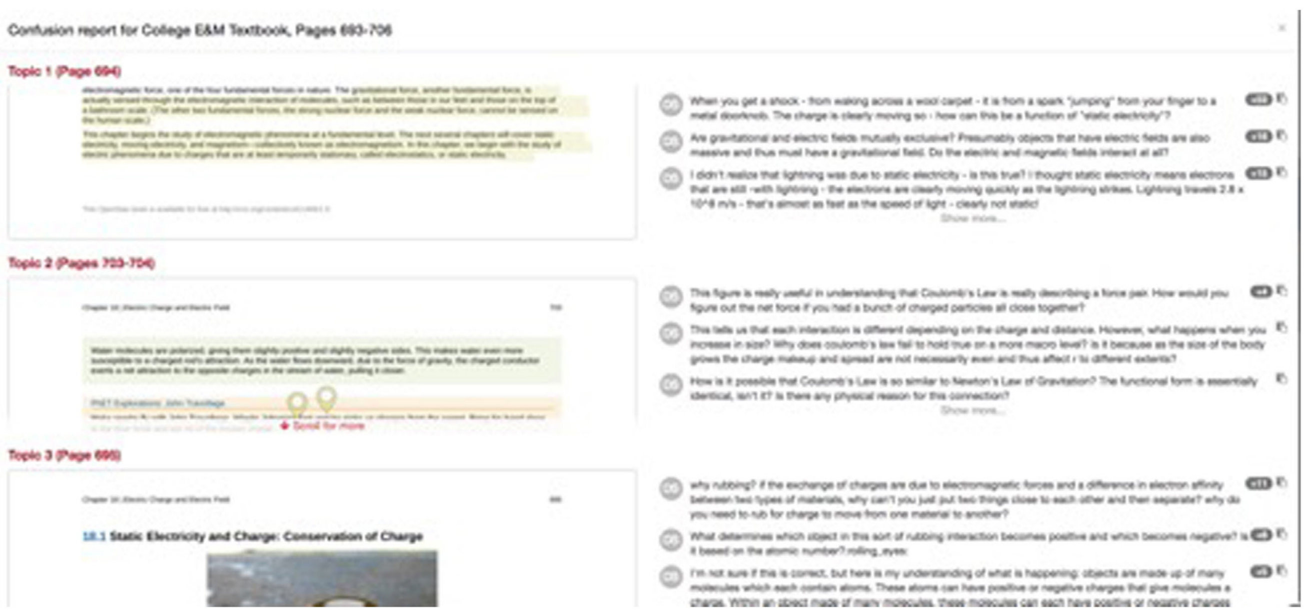

FIGURE 8 | Confusion report in *Perusall*.

students enrolled in the F15 course. The four student populations were very similar in composition. Student populations were comprised of 48–50% premedical students and 50–52% engineering students. All four groups were 53–55% female and consisted of equal ratios of students in their sophomore, junior, and senior years of college. The four groups had a similar level of incoming physics background knowledge, as measured by the average score on the physics conceptual survey administered at the beginning of each semester [*Force Concept Inventory* (Hestenes et al., 1992) for the fall courses and the *Conceptual Survey on Electricity and Magnetism* (Maloney et al., 2001) for the spring courses].

Setting

We conducted this study in the School of Engineering and Applied Sciences at Harvard University in an introductory physics course called Applied Physics 50 (AP50). AP50 is a calculus-based physics course designed for undergraduate engineering students. It is split into two courses; AP50A, a mechanics course taught in the fall, and AP50B, an electricity and magnetism course taught in the spring.

The instructor was the same for all four semesters and the same pedagogy was used each semester. AP50 met twice weekly and each class was 3-h long. In this course, all lectures were replaced by reading assignments in *Perusall* and class time was entirely devoted to active learning. The pedagogy was based on features from both Project-Based Learning (Blumenfeld et al., 1991) and Team-Based Learning (Michaelsen et al., 2002). Students worked in small groups for all aspects of the course, including assessments. There were six different types of in-class activities, each of which designed to help students master the relevant physics and get started on the projects, which were the focal point of the course.

As there were no lectures, students were expected to read the textbook on *Perusall*. By midnight, the night before each class, students were required to complete the pre-class reading assignment by highlighting and annotating an assigned chapter of the textbook, the content of which was the focus of the activities the following day in class. As the class met twice a week, there were typically two chapter-long reading assignments per week. Over the course of each semester there were 17 assigned chapters, with each chapter containing 34 pages on average. To receive full credit for each reading assignment, students needed to enter at least 7 timely and thoughtful annotations per chapter.

Procedure

To evaluate the efficacy of *Perusall* and to study how and when students were using the software, we did three different types of analyses. We first extracted, from *Perusall*, a number of metrics that describe student reading behavior: the amount of time students spent reading, how long before each class students logged on to *Perusall*, and how often they returned to the same reading assignment. We calculated student averages, per reading assignment, for each of these metrics and summarized these

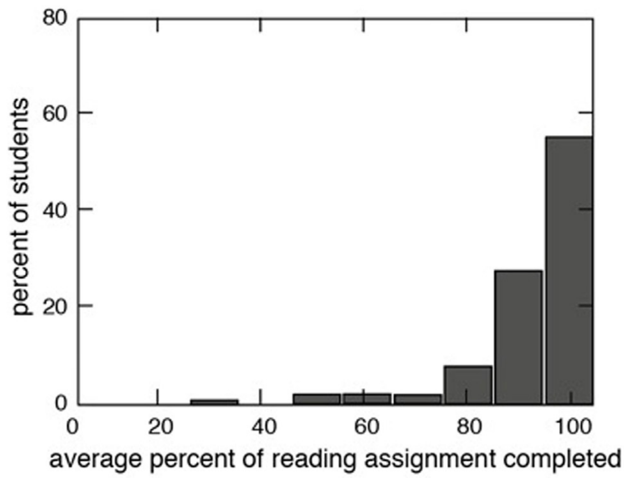

FIGURE 10 | Histogram of the average percent of the reading assignment students complete before class.

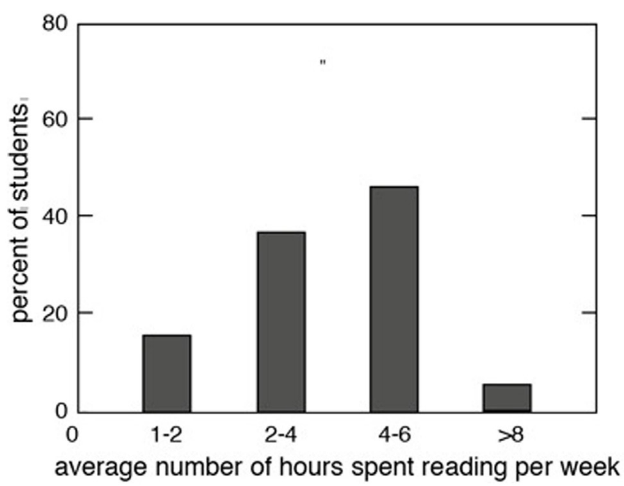

FIGURE 11 | Histogram of the average number of hours students spend on the reading per week.

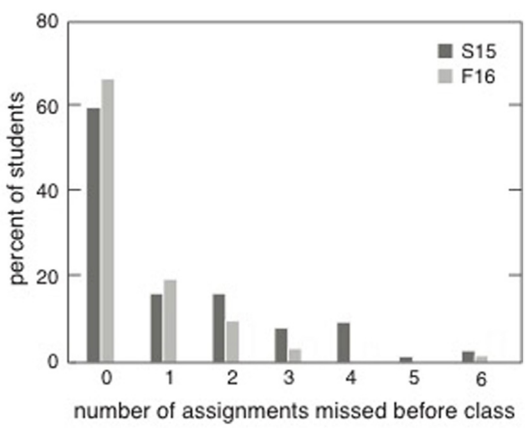

FIGURE 9 | Histogram of the number of assignments students fail to complete before class, in each semester.

(in **Figures 9–13**) as a way of describing how much and when, relative to class, students were reading. Second, we used the statistical software, STATA, to calculate the correlations between specific reading metrics and student exam performance so that we could determine which (if any) types of reading behaviors are predictive of exam performance. Based on these correlations, we used STATA to develop linear regression models to predict exam performance using reading behavior metrics (controlling for physics background knowledge). Third, to study the efficacy of the software in promoting student learning, we conducted a comparative study between two different types of social annotation software platforms. We compared exam performance during two semesters of AP50 when *Perusall* was used (S15, F16) to performance on the same exams during two other semesters (S14, F15) when a different social annotation platform was used. The exams in each of the two fall semesters (F15 and F16) were identical as were the exams in each of the two spring semesters (S14 and S15). To ensure that the students' incoming physics knowledge was the same between the two fall populations and the two spring populations, we used STATA to do a two-sample, *t*-test for equal means (Snedecor and Cochran, 1989).

This study was carried out in accordance with the recommendations of the Institutional Review Board (IRB) at Harvard University, Committee on the Use of Human Subjects. The IRB classified this study as "minimal risk" and, therefore, exempt from requiring written consent from the participants.

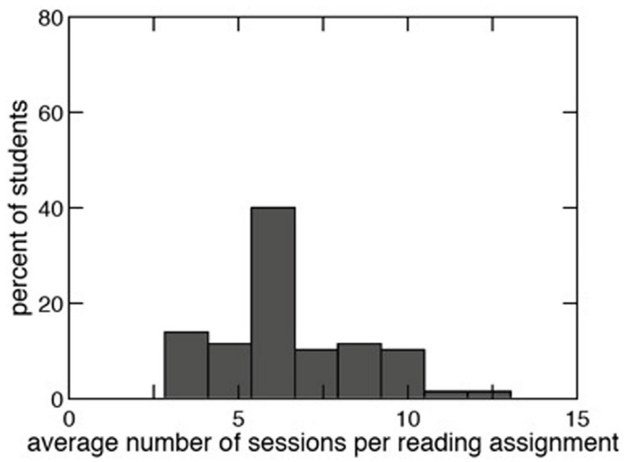

FIGURE 12 | Histogram of the average number of reading sessions per assignment.

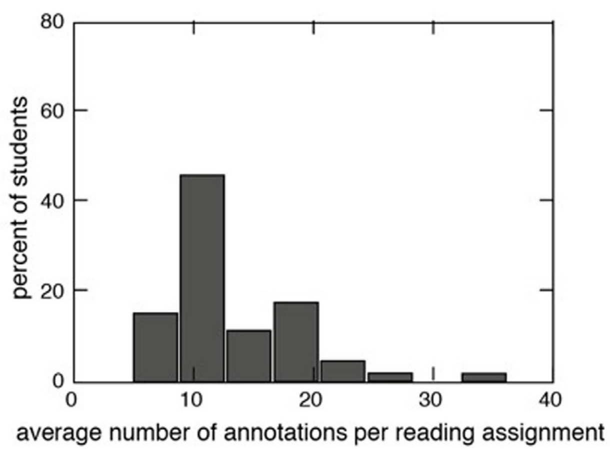

FIGURE 13 | Histogram of average number of annotations entered by students per assignment on *Perusall*.

RESULTS

Students' Pre-Class Behavior on *Perusall*

Figure 9 shows the extent to which students complete pre-class reading assignments over the two semesters that *Perusall* was being used. In each semester, approximately 60% of students completed every one of the 17 reading assignments. **Figure 9** shows that in S15, approximately 90% of students completed all but a couple of reading assignments; in F16, 95% of students completed all but a couple of reading assignments.

Perusall allows us to collect data on how much time students spend on each individual page of a reading assignment. Using these data, we can determine when a student makes it all the way through the assignment. We define a page as "read" when the time spent on that page is longer than 10 s and less than 20 min. We define a student as having completed an assignment by dividing, for each assignment, the number of pages that were read by the total number of pages in the assignment. Based on this metric, we find that 80% of students make it through at least 95% of the reading and that an additional 10% of students make it through 80% of the reading (**Figure 10**).

Using the same data we find that, on average, students spend 3 h and 20 min per week reading on *Perusall* (**Figure 11**).

Figure 12 shows the average number of individual "sessions" students take to complete their reading assignments. We define a session as any cumulative pages read for longer than 10 min with at least 2 h since the previous reading session. On average, students divide their reading of each assignment in seven different sessions.

TABLE 1 | Standardized coefficients for linear regression models predicting average exam performance using the average time students spend reading per chapter and the average number of sessions students break their reading up into as predictor variables and controlling for pre-class physics knowledge (pre-semester FCI).

	MODEL 1	MODEL 2	MODEL 3
N	79	79	79
R^2	0.42	0.43	0.51
	Standardized coefficients		
Constant	−0.03	−0.03	−0.02
Pre-semester physics knowledge	0.65***	0.67***	0.66***
Average time spent per chapter		0.15	−0.16
Average number of reading sittings per assignment			0.41**

(Number of students, N = 79).
****p < 0.001, **p < 0.01.*

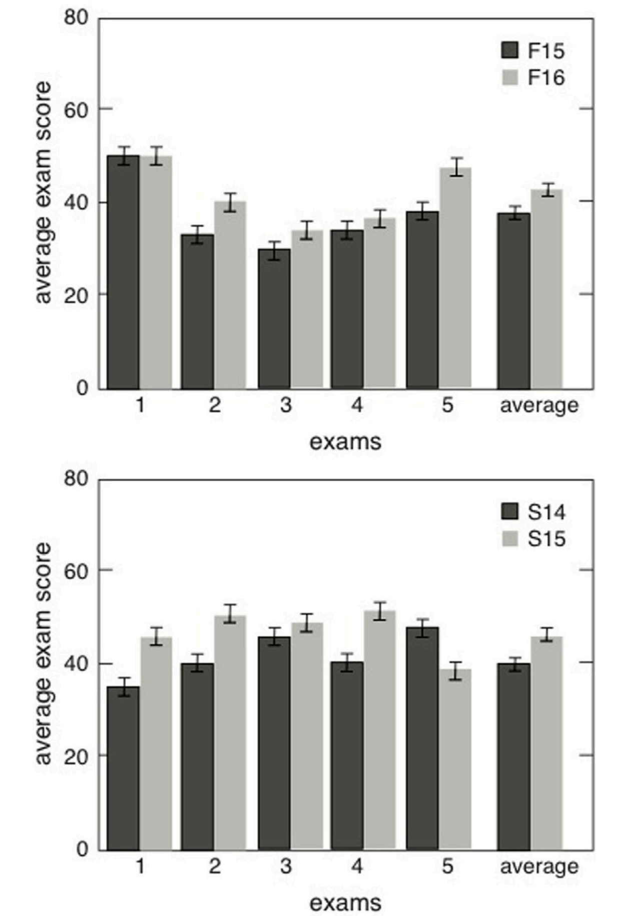

FIGURE 14 | Exam performance comparison between two different cohorts of the same course where the only difference is the use of *Perusall* as the online learning platform. The error bars represent the SEM.

Figure 13 shows the average number of annotations students enter per assignment. Over the course of the semester, students wrote a total of 16,066 annotations on *Perusall*. On average students make 13.3 annotations per assignment—nearly twice the number that the system recommends.

Relationship between Student Reading Behavior and In-Class Performance

To study the relationship between reading behavior and in-class performance, we built a series of linear regression models predicting students' exam performance (averaged over the five exams during the semester) from the reading and annotating metrics previously discussed. These models are presented in **Table 1**. We control for incoming physics background by including students' pre-semester score on the *Force Concept Inventory* (Hestenes et al., 1992) and the *Conceptual Survey on Electricity and Magnetism* (Maloney et al., 2001). We find that students who break up their reading into more sessions do better on in-class exams than students who read in fewer sessions, even when controlling for pre-course physics knowledge and the amount of time students spend reading.

TABLE 2 | Comparison of pre-course conceptual physics survey results between the four semesters of students (S14/F15 when *Perusall* was not used compared to S15/F16 when *Perusall* was used).

	Spring of 2014	Fall of 2015	Spring of 2015	Fall of 2016
Force concept inventory (average score)		17/30 ± 1		14/30 ± 2
Conceptual survey for electricity and magnetism (average score)	8/32 ± 1		7/32 ± 1	

Model 1 shows that we can predict 42% of the variability in students' average exam performance using only their score on the pre-semester *Force Concept Inventory*. If we add the average amount of time students spend reading (model 2) we can predict marginally more (43%) of the variability in exam performance although this difference is not significant. When we add to the model the average number of sessions the students use to complete the reading, we find we can predict almost 10% more of the variability in student exam performance (model 3). Increasing the number of sessions a student completes the reading in by one SD increases average student exam performance by 0.41 of a SD ($p < 0.01$). None of the other reading/annotation metrics are predictive of average student exam performance.

Student In-Class Exam Performance

Finally, we compare two different cohorts of the same course and show that the cohort for which *Perusall* was used to deliver the pre-class reading assignments did significantly better on the same in-class exams compared to students from the previous year when *Perusall* was not used.

Figure 14 shows student exam performance on 10 in-class exams administered over 2 years of AP50 (five in the fall semester and five in the spring semester). While the cohort of students were different over the four semesters, **Table 2** shows that the four groups of students had the same level of incoming physics knowledge at the beginning of each semester (as measured by the *Force Concept Inventory* and the *Conceptual Survey in Electricity and Magnetism*). We conducted two-sample, *t*-tests to confirm that the performance on these conceptual inventories was the same for the two fall groups and for the two spring groups ($p = 0.32$ and $p = 0.36$, respectively).

The only difference in the course between the S14/F15 and S15/F16 was the use of *Perusall*. During the S14/F15 semesters a simpler annotation tool was used to administer the pre-class reading assignments. This annotation tool lacked many of the social and machine learning features of *Perusall*. Students in the S15/F16 semesters scored 5–10% better on all but two of the 10 exams compared to the students from the semesters before when *Perusall* was not being used ($p < 0.05$). Based on a two-sample *t*-test, averaging over all five exams in the fall, students in the class that used *Perusall* scored significantly better than the class that did not use *Perusall* ($p < 0.05$). Students in the fall class that did not use *Perusall* had an average exam score of 38% compared to students in the fall class that used *Perusall* who had an average exam score of 43% (effect size = 0.34). The same is true when we average over all five exams in the spring—students in the class that used *Perusall* outperformed the students from the year before ($p < 0.05$). Students in the spring class that did not use *Perusall* had an average exam score of 41% compared to students in the spring class that used *Perusall* who had an average exam score of 45% (effect size = 0.31).

DISCUSSION AND CONCLUSION

This study explores student pre-class reading behavior on *Perusall*, a social learning platform that allows students to interact and discuss course material online. We find that student completion of reading assignments is substantially higher than what has been reported in the vast majority of the literature. With *Perusall*, 90–95% of students complete all but a few of the reading assignments before class. For comparison, most of the literature reports that 60–80% of students do not read the textbook before coming to class (Cummings et al., 2002; Clump et al., 2004; Podolefsky and Finkelstein, 2006; Stelzer et al., 2009). One study found that, with a JITT-like implementation of pre-class reading, between 80 and 85% of students completed the reading before class (Heiner et al., 2014), but this study was based on student-reported reading data, which has been shown to be unreliable. Using reading data from *Perusall*, we find that 80% of students complete 100% of the reading assignment before class. This percentage, too, is considerably higher than what is reported in the literature: Clump et al. (2004) find that students only read on average 28% of the assigned reading before class.

In addition to higher completion of pre-class reading assignment, we also find that, on *Perusall*, students read for longer than what is reported in the literature. Approximately 92% of students report that they spent 3 h or less per week reading the textbook. On *Perusall*, students spend, on average, 3 h and 20 min per week reading for this one course.

In studying the relationship between reading behavior and in-class performance, we find that the average time spent reading per chapter alone is not predictive of student exam performance. This is consistent with what has already been reported in the literature (Smith and Jacobs, 2003; Podolefsky and Finkelstein, 2006). However, it should be noted that previous studies on the relationship between time spent reading and exam performance have all been based on student-reported data. Our study uses data obtained directly from the *Perusall* platform. We do find, however, that students who break the reading up into more reading sittings perform better on in-class exams than students who read in fewer sittings. This is true even when we control for the amount of time students spent reading, and consistent given the spacing effect, a well-known phenomenon in psychology: material is more effectively and easily learned when it is studied over several times spaced out over a longer time span, rather than trying to learn it in a short period of time (Dempster and Farris, 1990).

Finally, we find that students using *Perusall* perform significantly better on in-class exams than students using a simple annotation tool without some of the social and machine learning features of *Perusall*. We recognize that this result does not indicate causality and must be interpreted carefully given the fact that other factors could be confounding the results. More research needs to be done to pinpoint exactly why students do significantly better using *Perusall*. *Perusall* has many features that the other platform did not have. For example, with *Perusall* assessment is built right into the platform and students get regular and timely feedback. In the other platform, assessment was provided separately by the instructors and so students received sporadic and less targeted feedback. *Perusall* also has many social features (sectioning, avatars, upvoting, email notifications) that are designed to improve the interactions between students. Finally, the Confusion Report makes it easier for the instructor to address main areas of student confusion in class, which both affords better targeting of in-class time to student confusion and allows students to better see the connection between pre-class reading assignments and in-class activities.

We have demonstrated the efficacy of *Perusall* as a social learning platform and have shown that student completion of pre-class reading assignment is substantially higher than what has been reported by other studies. In short, with *Perusall* we are better able to get students to complete reading assignments, and do so in a way—with spaced repetition—that leads to better outcomes. *Perusall*, therefore, is a useful tool for delivering content to students outside class and for building an online learning community in which students can discuss course content and develop understanding. This is particularly important in flipped and hybrid courses or any other course that relies on pre-class reading assignments.

ETHICS STATEMENT

We were covered for this research by Harvard's Committee on the Use of Human Subjects (CUHS).

AUTHOR CONTRIBUTIONS

Several people contributed to the work described in this paper. EM conceived of the basic idea for this work. KM, BL, GK, and EM designed and carried out the study, and KM analyzed the results and wrote the first draft of the paper. All authors contributed to the development of the manuscript.

REFERENCES

Ahren, T. C. (2005). Using online annotation software to provide timely feedback in an introductory programming course. *Paper Presented at the 35th ASEE/IEEE Frontiers in Education Conference*, Indiannapolis, IN. Available at: http://www.icee.usm.edu/icee/conferences/FIEC2005/papers/1696.pdf

Berry, T., Cook, L., Hill, N., and Stevens, K. (2010). An exploratory analysis of textbook usage and study habits: Misperceptions and barriers to success. *Coll. Teach*. 59, 31–39.

Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., and Palincsar, A. (1991). Motivating project-based learning: sustaining the doing, supporting the learning. *Educ. Psychol.* 26, 369–398. doi:10.1080/00461520.1991.9653139

Bonwell, C. C., and Eison, J. A. (1991). *Active Learning: Creating Excitement in the Classroom. 1991 ASHE-ERIC Higher Education Reports*. Washington, DC: ERIC Clearinghouse on Higher Education.

Bradshaw, J., and Hinton, L. (2004). Benefits of an online discussion list in a traditional distance education course. *Turk. Online J. Distance Educ.* 5.

Brown, A., and Campione, J. (1996). "Psychological theory and design of innovative learning environments: on procedures principles and systems," in *Innovations in Learning: New Environments for Education*, eds L. Schauble and R. Glaser (Mahwah, NJ: Erlbaum), 289–325.

Burchfield, C. M., and Sappington, J. (2000). Compliance with required reading assignments. *Teach. Psychol.* 27, 58–60.

Cadiz, J. J., Gupta, A., and Grudin, J. (2000). "Using web annotations for asynchronous collaboration around documents," in *Proceedings of CSCW'00: The 2000 ACM Conference on Computer Supported Cooperative Work* (Philadelphia, PA: ACM), 309–318.

Clump, M. A., Bauer, H., and Bradley, C. (2004). The extent to which psychology students read textbooks: a multiple class analysis of reading across the psychology curriculum. *J. Instr. Psychol.* 31, 227–232.

Connor-Greene, P. A. (2000). Assessing and promoting student learning: blurring the line between teaching and testing. *Teach. Psychol.* 27, 84–88. doi:10.1207/S15328023TOP2702_01

Crouch, C. H., and Mazur, E. (2001). Peer instruction: ten years of experience and results. *Am. J. Phys.* 69, 970–997. doi:10.1119/1.1374249

Cummings, K., French, T., and Cooney, P. J. (2002). "Student textbook use in introductory physics," in *Proceedings of PERC'02: In Physics Education Research Conference*, eds S. Franklin, K. Cummings, and J. Marx (Boise, ID), 7–8.

Dempster, F. N., and Farris, R. (1990). The spacing effect: research and practice. *J. Res. Dev. Educ.* 23, 97–101.

Dobson, J. L. (2008). The use of formative online quizzes to enhance class preparation and scores on summative exams. *Adv. Physiol. Educ.* 32, 297–302. doi:10.1152/advan.90162.2008

Downes, S. (1999). *Creating an Online Learning Community [PowerPoint Slides]*. Edmonton: Virtual School Symposium. Available at: https://www.slideshare.net/Downes/creating-an-online-learning-community

Gupta, S., Condit, C., and Gupta, A. (2008). Graphitti: an annotation management system for heterogeneous objects. *Paper Presented at IEEE 24th International Conference on Data Engineering*, Cancun, Mexico, 1568–1571.

Heiner, C. E., Banet, A. I., and Wieman, C. (2014). Preparing students for class: How to get 80% of students reading the textbook before class. *Am. J. Phys.* 82, 989–996.

Herreid, C., and Schiller, N. (2013). Case studies and the flipped classroom. *J. Coll. Sci. Teach.* 42, 62–66.

Hestenes, D., Wells, M., and Swackhamer, G. (1992). Force concept inventory. *Phys. Teach.* 30, 141–157. doi:10.1119/1.2343497

Hwang, W. Y., and Wang, C. Y. (2004). "A study on application of annotation system in web-based materials," in *Proceedings of GCCCE' 04: the 8th Global Chinese Conference on Computers in Education*, Hong Kong.

Johnson, B. C., and Kiviniemi, M. T. (2009). The effect of online chapter quizzes on exam performance in an undergraduate social psychology course. *Teach. Psychol.* 36, 33–37. doi:10.1080/00986280802528972

Karp, D. A., and Yoels, W. C. (1976). The college classroom: some observations on the meanings of student participation. *Sociol. Soc. Res.* 60, 421–439.

Maloney, D. P., O'Kuma, T. L., Hieggelke, C. J., and Van Heuvelen, A. (2001). Surveying students' conceptual knowledge of electricity and magnetism. *Am. J. Phy.* 69, S12–S23. doi:10.1119/1.1371296

Maor, D., and Volet, S. (2007). Interactivity in professional learning: a review of research based studies. *Aust. J. Educ. Technol.* 23, 227–247. doi:10.14742/ajet.1268

Marcell, M. (2008). Effectiveness of regular online quizzing in increasing class participation and preparation. *Int. J. Scholarsh. Teach. Learn.* 2, 7. doi:10.20429/ijsotl.2008.020107

Marshall, C. C., and Brush, A. J. B. (2004). "Exploring the relationship between personal and public annotations," in *Proceedings of JCDL'04: The 2004 ACE/IEEE Conference on Digital Libraries* (Tucson, AZ), 349–357.

Michaelsen, L. K., Knight, A. B., and Fink, L. D. (eds) (2002). *Team-Based Learning: A Transformative Use of Small Groups*. Westport, CT: Greenwood Publishing Group.

Miller, K., Zyto, S., Karger, D., Yoo, J., and Mazur, E. (2016). Analysis of student engagement in an online annotation system in the context of a flipped introductory physics class. *Phys. Rev. Phys. Educ. Res.* 12, 020143. doi:10.1103/PhysRevPhysEducRes.12.020143

Nandi, D., Chang, S., and Balbo, S. (2009). "A conceptual framework for assessing interaction quality in online discussion forums," in *Same Places, Different Spaces. Proceedings Ascilite Auckland 2009*.

Narloch, R., Garbin, C. P., and Turnage, K. D. (2006). Benefits of prelecture quizzes. *Teach. Psychol.* 33, 109–112. doi:10.1207/s15328023top3302_6

Nokelainen, P., Kurhila, J., Miettinen, M., Floreen, P., and Tirri, H. (2003). "Evaluating the role of a shared document-based annotation tool in learner-centered collaborative learning," in *Proceedings of ICALT'03: the 3rd IEEE International Conference on Advanced Learning Technologies* (Athens), 200–203.

Novak, G. M. (2011). Just-in-time teaching. *New Directions for Teaching and Learning*. 2011, 63–73.

Novak, G. M., Patterson, E. T., Gavrin, A. D., and Christian, W. (1999). *Just-in-Time Teaching: Blending Active Learning with Web Technology*. Upper Saddle River, NJ: Addison-Wesley.

Podolefsky, N., and Finkelstein, N. (2006). The perceived value of college physics textbooks: students and instructors may not see eye to eye. *Phys. Teach.* 44, 338–342. doi:10.1119/1.2336132

Quade, A. M. (1996). "An assessment of retention and depth of processing associated with note taking using traditional pencil and paper and an on-line notepad during computer-delivered instruction," in *Proceedings of Selected Research and Development Presentations at the 1996 National Convention of the Association for Educational Communications and Technology* (Indianapolis, IN), 559–570.

Robert, C. A. (2009). Annotation for knowledge sharing in a collaborative environment. *J. Knowl. Manag.* 13, 111–119. doi:10.1108/13673270910931206

Rovai, A. P. (2002). Building sense of community at a distance. *Int. Rev. Res. Open Distrib. Learn.* 3. doi:10.19173/irrodl.v3i1.79

Ruscio, J. (2001). Administering quizzes at random to increase students' reading. *Teach. Psychol.* 28, 204–206.

Sappington, J., Kinsey, K., and Munsayac, K. (2002). Two studies of reading compliance among college students. *Teach. Psychol.* 29, 272–274. doi:10.1207/S15328023TOP2904_02

Schwartz, D. L., and Bransford, J. D. (1998). A time for telling. *Cogn. Instr.* 16, 475–5223. doi:10.1207/s1532690xci1604_4

Smith, B. D., and Jacobs, D. C. (2003). TextRev: a window into how general and organic chemistry students use textbook resources. *J. Chem. Educ.* 80, 99. doi:10.1021/ed080p99

Snedecor, G. W., and Cochran, W. G. (1989). *Analysis of Variance: The Random Effects Model. Statistical Methods*. Ames, IA: Iowa State University Press, 237–252.

Sorcinelli, M. D. (1991). Research findings on the seven principles. *New Directions for Teaching and Learning*. 1991, 13–25. doi:10.1002/tl.37219914704

Stelzer, T., Gladding, G., Mestre, J. P., and Brookes, D. T. (2009). Comparing the efficacy of multimedia modules with traditional textbooks for learning introductory physics content. *Am. J. Phys.* 77, 184–190. doi:10.1119/1.3028204

Su, A. Y. S., Yang, S. H., Hwang, W. Y., and Zhang, J. (2010). A web 2.0-based collaborative annotation system for enhancing knowledge sharing in collaborative learning environments. *Comput. Educ.* 55, 752–766. doi:10.1016/j.compedu.2010.03.008

Tallent-Runnels, M. K., Thomas, J. A., Lan, W. Y., Cooper, S., Ahern, T. C., Shaw, S. M., et al. (2006). Teaching courses online: a review of the research. *Rev. Educ. Res.* 76, 93–135. doi:10.3102/00346543076001093

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Cambridge University Press.

Webb, N. M., Nemer, K., Chizhik, A., and Sugrue, B. (1995). *Using Group Collaboration as a Window into Students' Cognitive Processes*. CSE Technical Report 404. Los Angeles, CA: National Center for Research on Evaluation Standards, and Student Testing (CRESST).

Conflict of Interest Statement: The authors developed the technology described in this article, mostly at Harvard University. Perusall.com is a commercial product based on this work. The authors are cofounders of Perusall, LLC, the company that runs perusall.com.

frontiers in ICT

ORIGINAL RESEARCH
published: 17 May 2018
doi: 10.3389/fict.2018.00008

Using the TAM and Functional Analysis to Predict the Most Used Functions of an Active Learning Classroom (ALC)

Bruno Poellhuber[1]*, *Samuel Fournier St-Laurent*[2] and *Normand Roy*[1]

[1] Department of Psychopedagogy and Andragogy, Université de Montréal, Montréal, QC, Canada, [2] Learning Support and Pedagogical Development service, Collège Ahuntsic, Montréal, QC, Canada

Since the pedagogical reform undertaken in the field of physics teaching with the SCALE-UP project, research has shown that active learning classrooms (ALC) can lead to substantial gains. The reflection on ALC is now burgeoning, with this area being the number-one focus of university technological investments in 2017. However, even though a kind of ALC standard has emerged (teacher pod at the center of the room, round tables, a projector for each table, etc.), very few researchers actually investigate the precise layout of ALCs and which particular features are the most important from the students' perspective. This is precisely what this study aims to do, relying on the TAM (Technological Adoption Model). The study took place in three colleges in Quebec with ALCs, using a functional analysis approach. In this process, nine functions were identified. A single-item questionnaire was developed around a modified TAM (including interest) and sent to 352 students who rated the frequency of use, utility, interest and ease of use of each of the nine functions. Qualitative data were collected through group interviews with students. Average scores were computed for each construct with the nine functions and they showed satisfactory consistency. Automated text analyses were conducted on the answers to the open-ended question. The results show that from the students' perspective, the most important functions are related to features that facilitate group work (having a team table and using wall surfaces that can support image projections and annotations). Being able to use computers supplied by the college, connect student-owned devices to the team projector and annotate projection surfaces also ranked high. The correlation between frequency of use, interest, utility and ease of use is high and statistically significant. The qualitative data show that having comfortable, movable chairs is also important. The special look and feel of an ALC also seem to make students more comfortable. On a less positive side, some students indicate that visual obstruction is an obstacle in the periods when the teachers lecture in the class. These results may support cost-effective ALC design.

Keywords: TAM, functional analysis, ALC, active learning classrooms, students, classroom layout

OPEN ACCESS

Edited by:
Rob Cassidy, Concordia University, Canada

Reviewed by:
Yotam Hod, University of Haifa, Israel
Tracey Birdwell, Indiana University, United States

***Correspondence:**
Bruno Poellhuber
bruno.poellhuber@umontreal.ca

Specialty section:
This article was submitted to Digital Education, a section of the journal Frontiers in ICT

Received: *30 November 2017*
Accepted: *23 April 2018*
Published: *17 May 2018*

Citation:
Poellhuber B, Fournier St-Laurent S and Roy N (2018) Using the TAM and Functional Analysis to Predict the Most Used Functions of an Active Learning Classroom (ALC). Front. ICT 5:8. doi: 10.3389/fict.2018.00008

INTRODUCTION

The reflection on active learning classrooms (ALC) is burgeoning, but few researchers investigate how to lay out these learning spaces and which particular functions are the most important from the students' perspective. This is precisely what this study aims to do, relying on the TAM (technological adoption model) and a functional analysis approach. From a professor's point of view, it is fairly obvious that a classroom's layout influences the type of pedagogy that can take place in it, facilitating some types and rendering others more difficult: lectures fit perfectly in lecture halls, but cooperative pedagogy is harder to achieve. According to Wesch (2007), a well-known physical anthropologist, the layout of our teaching and learning spaces says a lot about the way we conceive of teaching and learning. The set-up of very large lectures halls common in all North-American universities is such that students are seated very close together. The presenter (professor) stands on the stage, often on a podium which is sometimes next to a large screen for projections. All seats are oriented toward the front and the rows are designed to focus attention toward the front screen and the presenter. Participants in the audience ("students") have little or no room for anything other than a notebook. This set-up implicitly but very clearly communicates a vision that the information will come from an expert who is on the stage, who is worthy of the participants' attention and who will "profess" so they can take notes.

PROBLEM

This very transmissive approach has been challenged in the STEM reform movement in relation to concerns regarding student retention and learning in STEM areas. In 1998, Hake ran a study on 6542 students registered in 62 introductory physics courses, using the Force Concept Inventory (Hestenes et al., 1992) to compare conceptual gains between an "active engagement" condition and a traditional lecture approach. Conceptual gains proved to be significantly higher in the "active engagement" condition (Hake, 1998).

Following these results, the original SCALE-UP project objectives were to improve student learning and attitudes, design "new modes of instruction for large enrolment sections" and develop teaching guides and instructional materials (Beichner and Saul, 1999), but the SCALE-UP project became famous because it was a pioneer in experimenting with classroom layout. It proposed new ways of exploring large classroom layouts to facilitate active learning and collaboration in technology-rich environments. As stated by its originator in an early article, "the primary goal of the SCALE-UP Project is to establish a highly collaborative, hands-on, computer-rich, interactive learning environment in large-enrolment physics courses. We know from extensive educational research that students should collaborate on interesting tasks and be actively involved with the material they are learning" (Beichner and Saul, 1999).

Within a few years, the SCALE-UP project gained speed, and in 2006, about two dozen universities had climbed aboard (Beichner et al., 2007). Over a decade later, over 250 sites inspired by SCALE-UP were in operation in the US, and more than 31 are located in the province of Quebec (Canada). This rapid expansion gave birth to deep, new reflections on how to lay out learning spaces with technology in order to facilitate active learning and collaboration supported by technology. It created a focus on the complex relations between classroom layout, technology, pedagogy and learning in different types of spaces. A project report on the SCALE-UP documented the many advantages of the SCALE-UP model over traditional lectures in lecture halls in introductory physics courses: better conceptual understanding, better course attendance, lower failure rates (better retention), better problem-solving skills (Beichner et al., 2007).

Emerging from a specific SCALE-UP subproject, the TEAL (technology-enhanced active learning) project at MIT was implemented in all MIT introductory physic instructions. "*Technology-enabled active learning* is a teaching format that merges lectures, simulations, and hands-on desktop experiments to create a rich collaborative learning experience" (http://icampus.mit.edu/projects/teal/). The TEAL project went further in the technology enrichment aspect of the pedagogical project, including simulations, visualizations and hands-on experiments in the collaborative learning approach. Both SCALE-UP and TEAL aimed not only to redesign classrooms, but also to redesign instruction, the way the introductory courses were taught, the teachers' roles and the instructional materials. The TEAL project generated conceptual gains similar to those obtained in the SCALE-UP set-up (Dori et al., 2003). Researchers in Quebec replicated these results a few years later (Charles et al., 2011).

In short, in the field of physics teaching, research shows that the student-centered pedagogy used in active learning classes has led to greater conceptual gains than those made with traditional methods (Hake, 1998; Dori et al., 2003; Beichner et al., 2007; Charles et al., 2011), as well as other interesting gains such as lower failure rates (Dori et al., 2003) and better class attendance (Beichner et al., 2007). In these early studies, profound pedagogical changes accompanied the physical changes. From the teacher-student transmission of the material to the student-student interactions with the material that take place in this new environment, many changes are needed. The SCALE-UP report (Beichner et al., 2007) mention of the challenges of course design. Instructional design in this context takes more time than preparing lectures. The design must also take into account the need for students to work in groups and stay engaged in their tasks. The skills required for lecturing are also different from those required to offer adequate cognitive and metacognitive support for the students.

Meanwhile, the evolution of learning technologies, the emergence of low-cost high-performance laptop computers and the birth of the iPad led to an increase in one-to-one initiatives (Bocconi et al., 2013) and brought BYOD (bring your own device) to the fore. BYOD was deemed to be the most important development in educational technology in the 2015 Horizon report (Adams Becker et al., 2017). The advent of BYOD creates a need for accommodation. For example, "University of Scranton leaders assert that BYOD policies will also impact the physical environment of the classroom, and that rigid furniture should be replaced with more flexible workspaces to accommodate the collaboration that mobile apps and other features promote" (Adams Becker et al., 2017, p. 38).

While in the SCALE-UP and TEAL spaces, pedagogy and layout are intertwined in various and complex ways, a recent line of research initiated at the OIT of the University of Minnesota turned its focus on the specific role of classroom layout, using quasi-experimental designs to isolate the classroom factor from the others. Brooks (2011) ran an initial quasi-experimental study in a biology course, keeping all variables constant except for the physical layout of the classrooms. Both sections of the course had the same instructor, were offered in the same time slot (on different days) and relied on the same course material and instructional approach. The students in the active learning classroom (ALC), which had significantly lower ACT scores compared to the students in the traditional classroom, performed as well as them and had the same final grades. These findings suggest that "physical space alone can improve student learning" (Brooks, 2011, p. 725). In a replication of this study with a different instructor in another biology course (Cotner et al., 2013), similar results were obtained.

In a further study, Brooks (2012) used another quasi-experimental design to compare teacher and student behaviors in two sections of the same course, using a systematic behavioral codification grid. This study showed that the classroom layout actually has an effect on the behavior and pedagogy of the instructor. The instructor gave significantly more lectures and significantly fewer group activities in the traditional classroom than in the ALC. In this study, both lectures and team work were linked with student engagement, as measured by the observation of "on task" behaviors. This particular study suggests that room layout does have an impact on the type of pedagogy, a result also obtained by Whiteside et al. (2010), and that both types of pedagogy can lead to on-task behaviors.

In a quasi-experimental ex post facto longitudinal study, Brooks and Solheim (2014) focused on the impact of the pedagogical transformation of a finance course taught in an ALC, supported by a faculty development program. The authors report significant differences in student participation, as well as student grades (for individual assignments and final grades). Other results suggest that it is the active learning pedagogy that is effective in the SCALE-UP project (see Soneral and Wyse, 2017, as well as Stoltzfus and Libarkin, 2016).

There are intricate links between pedagogy, room layout, technology and student outcomes. Whether changes in classroom layouts produce a direct effect on pedagogy is subject to debate, but it does seem that room layout induces or facilitates particular pedagogical approaches and that the greater part of the gains obtained in projects such as SCALE-UP come from the pedagogy rather than from the room layout.

For some, changes in classroom layouts and pedagogical changes should take place simultaneously (Woolner et al., 2012). For example, it seems that teacher-centered approaches are actually less effective in active learning classrooms (Charles et al., 2011). In the context of technology-rich learning spaces, it is also useful to point out that changes in pedagogy are also necessary to effectively use technology (Basque, 2004; Barrette, 2009).

Since the SCALE-UP started, many universities and colleges have picked up the concept and Active Learning Classrooms (ALC) have become somehow standardized, even though there are many variations on the theme. In an ALC, the instructor podium is located in the center of the room, in order to balance interactions with the different student teams. Other features usually found in an ALC are:

- tables for teams of 6 to 8 students (typically round or oval)
- chairs on wheels
- work surfaces on walls such as a projector and wall screen for each team
- a technology-rich environment that provides laptops, computers or tablets to the students, as well as various software programs

The concept of flexibility is now emerging in the literature, usually with the design of flexible learning spaces. For example, at Calgary's Taylor Centre, learning studios, team tables, chairs and even the instructor podium and team projectors are designed to be mobile.

Reflections and experiments with other types of learning spaces are expanding, and we can now see examples of active lecture halls or active labs. McGill University has experimented with these and even set up an active wet lab equipped with advanced biochemistry and chemistry laboratory apparatuses (https://www.mcgill.ca/tls/spaces/classrooms/). The focus is currently also shifting to the learning potential of spaces other than classrooms, and to the student perspective on informal learning spaces such as halls, cafeteria, library spaces, etc. Vo (2015) investigated the factors behind college students' choices of informal learning spaces. Carnell (2017) also focused on informal learning spaces, drawing design principles from students and staff interviews.

While we see many new types of active or flexible learning spaces, few studies actually document how traditional ALC are used (Wilson and Cotgrave, 2016). Research has focused mostly on the teachers' pedagogies in these spaces, and the students' perspective on these spaces has not frequently been taken into consideration. This tendency to focus on faculty perceptions and practices is not consistent with the student-centered practices that are the target of the ALC. The instructors have to design active learning scenarios for these environments, which are usually much more student-centered, and they have more choices than in the more traditional pedagogical scenario. This generally gives them a lot of freedom in the choices of the particular ALC features they will use during team work. Some of the research with students has focused on the choices they make in their personal study spaces (see for example Alphonse-Prescott, 2016) or personal learning environments (Roland and Talbot, 2015).

Very little research has been done on how teachers or students actually use the different features of an ALC and which features are the most important in the design phase. Benoit (2017) is a notable exception. He interviewed both students and instructors on their perceptions of two types of ALC layout on three topics; equipment and technology, learning environment design and interaction. He found that rooms resembling classical ALC layout were more conducive to student-student interactions

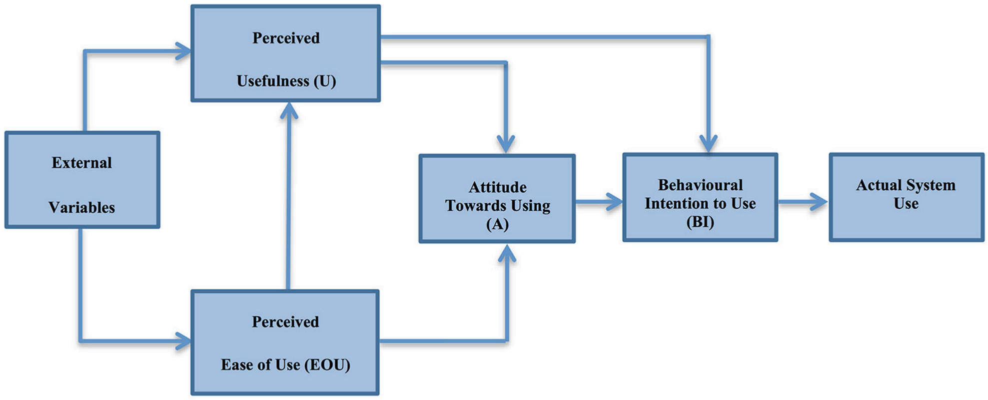

FIGURE 1 | The original TAM (from Davis et al., 1989, p. 985).

and group work, as well as student-instructor interaction. These rooms were also perceived to be more welcoming and more comfortable. Portable whiteboards were the most used technological feature. Concerns about table size and stability were identified.

The costs of setting up an ALC can be quite high, from $100,000 (University of West Florida, 1999) to even $465,000 (McGill Teaching and Learning Services, 2009). Considering the relatively high cost of designing and implementing active learning spaces, identifying crucial aspects of ALC layout is particularly important. In a low-tech SCALE-UP mock-up project, results similar to the original SCALE-UP research were obtained, which suggests that the most expensive technological features of an ALC might not be the most important (Soneral and Wyse, 2017).

The objective of this study is to evaluate the most important functions in the layout and technological choices for active classrooms.

THEORETICAL FRAMEWORK

In order to understand the choices made by student teams to use (or adopt) any particular function, we decided to rely on the technology acceptance model (TAM), because of its simplicity, its applicability to the particular context and its efficiency in predicting the adoption of particular technologies in educational settings (see **Figure 1**).

The TAM has been developed to explain and predict the adoption of technological systems by users, specifically, computer use and software applications. Davis et al. (1989) developed the TAM from the theory of reasoned action (TRA). The TRA predicts the intention to adopt behavior based on a person's beliefs and attitudes. This focus on behavioral intent was incorporated into the TAM to predict the adoption of computers or software. A comparison between the two models revealed a stronger predictive power for the TAM than the TRA between the intention indicators and the actual use of word processing software (Davis et al., 1989).

The TAM has been widely used in research on technology adoption and is one of the most cited models in the literature. It is a simple and effective model for predicting the intention to adopt a technology.

In developing the TAM, Davis (1989) wanted to build better measures for predicting and explaining the use of various technological environments, particularly computer applications. According to the TAM, the intent to use a technological environment such as a computer application depends essentially on its perceived usefulness—"the degree to which a person believes that using a particular system would enhance his or her job performance" (Davis, 1989, p. 320)—and perceived ease of use—"the degree to which a person believes that using a particular system would be free of effort" (Davis, 1989, p. 320).

In this project, we used the TAM in an unusual way. First, we wanted to focus on the adoption of each of the particular functions of an active classroom, rather than look at the adoption of the ALC as a whole, based on the premise that particular ALC functions differ in their usefulness and ease of use. Second, rather than trying to predict actual use from the intent to use (as the TAM is generally used), we had the opportunity to measure adoption through actual use. As done by Ha et al. (2007), McGowan et al. (2012) and others, we proposed to consider the frequency of use of each of the different functions as an indicator of adoption and actual use. The TAM model suggests that the functions perceived as the most useful and the easiest to use would be the most adopted and, hence, the most frequently used. We also proposed to identify the most useful and easiest functions from the point of view of the students, rather than the instructors. Relationships can be established between the TAM and general expectancy-value models of motivation. Perceived usefulness is part of the task value component in the Pintrich model, which also includes intrinsic goals and affects (Pintrich, 2003). We can also pinpoint some conceptual resemblance or at least a relationship between perceived usefulness and self-efficacy. Venkatesh and Davis (1996) actually linked self-efficacy and usability to perceived usefulness. In Eccles's motivational model, interest is part of the task value component (Eccles and Wigfield, 2002). In a previous study, Poellhuber et al. (2013)

showed that interest was a stronger predictor of the adoption of a social networking system than usefulness. This study therefore relies on a modified TAM that focuses on interest in the use of a particular function.

METHODS

This exploratory study relies on the pedagogical value analysis developed by Rocque et al. (1998) and used in other educational contexts (Severin, 2009). This approach adapts the value analysis and functional analysis approach widely used in engineering in order to apply it to educational contexts and developments. It has been frequently used in Quebec to develop innovative products or services aimed at student clienteles with particular needs (see, for example, Chalgoumi, 2011). This approach is deployed in three phases: pre-design, functional analysis and development. The pre-design phase draws on a user needs analysis and on what is actually known in the field, for example, from the scientific literature. It can also be based on a comparative analysis of existing products or services that meet similar needs. In the functional analysis phase, the focus is on the functions that the particular product has to fulfill, while leaving room for creativity on how each particular function can be filled (Rocque et al., 1998). Finally, the development phase is in the hands of developers who develop a prototype based on the identified functions, but keep some freedom in the design.

Context

This particular study is part of a large research project on the conditions of effectiveness of ALC, in which 19 teachers from five Quebec cégeps (junior colleges in the US and Canada) partnered with a university researcher in a design-based study that investigated the conditions of pedagogical practices that were the most conducive to student motivation and engagement. The project started in the winter semester of 2014 and continued until the fall semester 2015. Teacher participation varied from one semester to another, some being in the project for only one semester (not necessarily the first) and others participating for all four semesters. The study focused mainly on the effect of pedagogical practices and conceptions, as well as innovation adoption (St-Laurent et al., 2017) and pedagogical change over time (Fournier St-Laurent and Poellhuber, submitted). The main focus of the project was the pedagogy in the ALC. Early advantages reported by ALC students pertained the pedagogical approaches, technology, collaboration and team productivity and, finally, the classroom layout itself (Poellhuber et al., 2018).

An iterative design-based research (DBR) approach was adopted as the general methodological framework (Brown, 1992). DBR is particularly useful for studies seeking to make both a contribution to theoretical knowledge and usable knowledge applicable to authentic learning situations (Collins et al., 2004; Anderson, 2005). In this study, the researchers and teachers had many opportunities to meet, discuss and work together to develop the learning scenarios to be implemented in the ALC and to interpret the qualitative and quantitative data collected during the project. The instructors were offered training sessions in the ALC to model cooperative scripts, ways to enhance student motivation and engagement, the development of pedagogical scenarios and teamwork management. The teachers were given ample time to discuss both their successes and their failures. This led to a wide variety of the actual scenarios, which were implemented in the ALC.

At the outset of the project, three of the participating colleges did not yet have an ALC and needed to find an effective approach to ALC design. Considering the cost constraints, ALC design came up as a problem that needed to be solved before pedagogical integration could take place. The Cégep regional de Lanaudière in Terrebonne was the last cégep to design its own classroom, and it benefited from the other colleges' designs and from Collège Ahuntic's functional analysis process. College Ahuntsic used a functional analysis approach to design its own ALC, with a project team made up of one educational developer, three teachers and one IT administrator. Using both a literature review and a comparative analysis of existing active learning classrooms (with visits to many ALC spaces), the committee conducted a functional analysis of three particular functions: utility functions, constraint functions and esteem functions. Briefly, utility functions are the main features of a product, which make it useful (e.g., interactive whiteboards can facilitate interaction with digital documents). Constraint functions refer to design limits (e.g., the object may not weigh more than 10 g). Esteem functions are those that make the product attractive (e.g., shiny stickers on a phone). The results of the process at that college were shared with the other participating colleges that were planning to design an ALC.

Sample

Nine teachers from three of the colleges participated at the last semester of the research (fall, 2015). They were teaching in four subjects (physics, French, biology and philosophy) and the particular pedagogical practices deployed in the ALC varied greatly from one teacher to another. A total of 337 of their students answered a midterm questionnaire that had a TAM section on the various functions of their ALC. Of these, 252 answered the TAM part of the questionnaire at the end of the survey.

Data Collection

In this particular case, the location of each ALC supplied many of the constraint functions. College Ahuntsic's committee focused on the utility functions while respecting the identified constraints. It identified nine utility functions that are particularly important in an ALC:

1. Having a team table
2. Using wall surfaces that can support image projections and annotations
3. Using computers supplied by the institution
4. Using tablets supplied by the institution
5. Connecting computers, tablets or other student-owned devices to the team projector
6. Sharing the work of a particular team with the other teams
7. Annotating projection surfaces while working in teams
8. Capturing an image of the work on the team's work surface

9. Capturing and sharing the image of a page or a real object

Each of these functions can be accomplished by a variety of means, which vary widely in cost. For example, functions 2, 7 and 9 can be accomplished by a set-up in which each team has access to a team smartboard (which is the case at Dawson College, in Montréal) for an approximate cost of about $35,000. It can also be accomplished by low-tech whiteboards that serve as a surface for regular projectors. Annotations can then be made on the projections with dry erase pens and the students can take screen captures with their smartphones, for an approximate cost of about $7,500.

This procedure departs from the use of the validated questionnaire developed by Davis comprising four items per subscale, but due to the innovative approach of the modified TAM, the decision was made to use a single-item scale. Wanous et al. (1997) suggest that in particular situational constraints, single-item scales can be as robust as a well-constructed scale. Many studies have demonstrated the reliability of the single-item scale (see Hoeppner et al., 2011; Leung and Xu, 2013). For the purpose of this study, a full scale would dramatically increase the length of the situational questionnaire (nine utility functions by three variables by four items). Each of the nine functions were listed in a table, and for each one of these, students had to rate the frequency of use on a five-point Likert scale, as well as interest (from not at all interesting to highly interesting), utility (from not at all useful to highly useful) and ease of use (from not easy at all to really easy) on a seven-point Likert scales.

To support the single-item questionnaire, one open-ended question asked them whether any other function was important: Are there other functions available in this classroom (e.g., furniture, tools, software) that have helped make this course motivating?

Thirteen semi-structured student group interviews took place during the last semester of the project around several themes, including the physical layout of the ALC. The nine participating teachers in the three colleges were also interviewed. In the interviews, the teachers were invited to comment qualitatively on the importance of these functions and to describe how they would actually be used. Usefulness, interest and ease of use scores were computed, as well as a global modified TAM score.

Analysis

In order to understand the most important functions in the layout and technological choices in active classrooms, we used descriptive analysis for the single-item questionnaire. We used composite items to evaluate the relationships among the constructs. Average scores were computed for each construct with the nine functions and they showed satisfactory consistencies, based on Cronbach's alpha (Frequency of use = 0.72; Interest = 0.84; Usefulness = 0.84; Ease of Use = 0.79). Automated text analyses were conducted on the 99 answers to the open-ended question in the Survey Monkey text analysis machine and revised manually by one coder.

All interviews were audio recorded, transcribed and coded using the QDA Miner qualitative analysis software. The coding grid was developed using a mixed approach relying both on pre-existing categories based on our conceptual framework and on emerging categories (Miles et al., 2013). The coding grid was developed by one researcher and one assistant consensually coding three student interviews. The final coding grid includes 48 codes grouped in nine categories. After stabilization, the coding grid was used independently by both coders on three interviews. Final interjudge agreement on these was 88.7%. Reports were generated on codes and excerpts pertaining to the physical layout of the ALC. A second stage of analysis then took place in order to identify the subjects most frequently discussed concerning classroom layouts. Significant excerpts in relation to the particular functions of an ALC were identified.

Ethics

The project was conducted under an ethics certificate from the Université de Montréal's pluridisciplinary ethics committee (CPER-13-112-D) and from each of the colleges with participating teachers. This study was carried out in accordance with its recommendations with informed consent from all subjects, including those participating in focus groups. All subjects were met by the researchers and gave written informed consent in accordance with the three Canadian Tri-Council guidelines, for both the survey and the group interviews.

RESULTS

In this section, descriptive statistics on the different components of the TAM will be reported first, followed by a correlation table of the TAM sub-scores. The main categories that emerged from the qualitative analysis will then be presented.

Frequency of use roughly represents the level of effective adoption of each function. We can see from **Figure 2** that the team table is by far the function most frequently used by students in the ALC. The other most frequently used functions are wall surfaces (boards), computers supplied by the school, annotations and screen sharing. Real object images and captures are the most rarely used or adopted, with a rating of "rarely" for captures and "never" for tablets supplied by the institution.

Table 1 represents the mean of the interest, utility and ease of use questions for each of the nine functions. It is in descending order by perceived utility, but the order remains the same for perceived interest. If the list is reordered by perceived ease of use, wall boards move to the first place, and the rest of the list remains unchanged.

In the TAM, studies usually show correlation between perceived usefulness, perceived ease of use and attitude. Based on the general composite score, we confirmed the hypothesis that these concepts are highly correlated (**Table 2**), although frequency of use is less correlated with the other concepts.

Automated text analysis shows that chairs, and, more precisely, comfortable chairs on wheels, are an important function omitted from the list (**Figure 3**). It was mentioned by 16 students. The relevance of the use of specialized subject-specific applications by teachers (such as Maple or

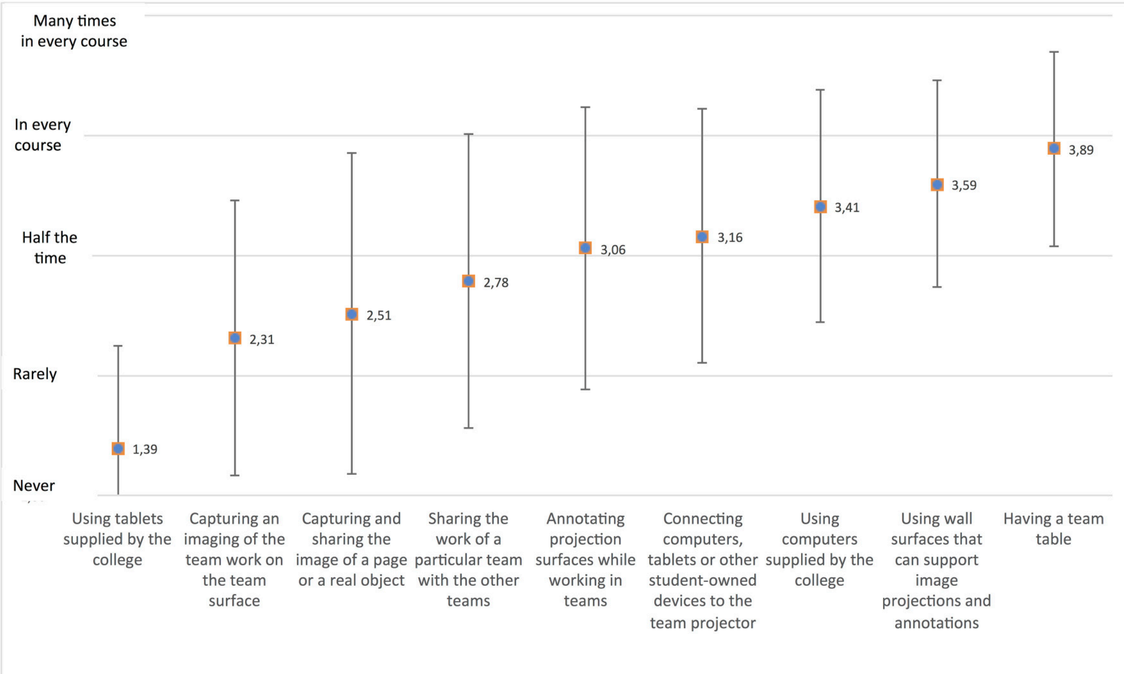

FIGURE 2 | Answer distribution for the frequency of use of the nine different functions.

TABLE 1 | Interest, utility and ease of use means for each function.

	Interest	Utility	Ease of use
Having a team table	6.01	6.12	6.26
Using wall surfaces that can support image projections and annotations	5.75	5.89	6.33
Using computers supplied by the college	5.73	5.82	6.11
Connecting computers, tablets or other student-owned devices to the team projector	5.58	5.67	6.00
Annotating projection surfaces while working in teams	5.35	5.5	5.88
Sharing the work of a particular team with the other teams	5.07	5.23	5.35
Capturing and sharing the image of a page or a real object	4.80	4.85	5.05
Capturing an image of the team work on the team surface	4.53	4.68	5.07
Using tablets supplied by the college	4.48	4.46	4.94

TABLE 2 | Correlation between frequency of use, interest, utility and ease of use.

	Frequency of Use	Interest	Usefulness	Ease of use
Frequency of Use	1	0.48*	0.44*	0.42*
Interest	0.48*	1	0.89*	0.72*
Usefulness	0.44*	0.89*	1	0.75*
Ease of Use	0.42*	0.72*	0.75*	1

*$p < 0.001$.

Geogebra in math) was also highlighted in six comments. The importance of team tables and team wall surfaces was confirmed. Three students also mentioned table colors as an important feature in the classroom layout (in one classroom, each team had a table and wallboard of a different color).

In the three focus groups, the students also insisted on the importance of comfortable and easy to move chairs. This was the first and most important category that emerged from the group interview analysis.

> *It's better on the comfort side. Also, if you want to listen to or work with someone else, you do not have to drag a chair making noise. You just have to roll.* (case 1–7)

> *I think the chair changes a lot of things. I can be at ease and comfortable. When I'm at ease, I am more able to listen. On plastic chairs, you have to sit very straight and you're always moving. With this other chair, we feel well and when we feel well, we can concentrate. It may sound silly, but just being able to move the chair up a little is very handy when someone ahead is blocking your view.* (case 1–8)

> *I really feel it when I have another three-hour course: I have back pain afterwards. On these seats, three hours passes really quickly.* Case 3–12 (R2)

FIGURE 3 | Word cloud generated by a text analysis of 99 open-ended answers.

The second most important qualitative category that emerged from the group interviews (for the part of the interview focused on classroom layout) was a general category related to the special look and feel of the ALC, and the type of "atmosphere" or "climate" it generated.

> *With the atmosphere of this classroom, we are more relaxed than in a normal class. With a lower stress level, you think better and you are more productive…*

> *I like how this class is arranged. It's good. We're comfortable and when the teacher walks around, we just have to rotate our chairs to follow him easily. Not always staying still helps me to focus. It's so different from other classes: the atmosphere is really better.* (case 2–8)

> *I feel it's more welcoming. When you say wooden chair and small individual desk, I think of an exam, that's it. Here it is more family-friendly kind of space. It's like when you go for dinner... at home we have dinner as a family and we talk to each other. It's nice and not stressful.* (case 2–11)

In these group interviews, the students also pinpointed some of the challenges in the ALC set-up. The main challenge that emerged from these interviews is being able to see and hear when the teacher is lecturing. The fact that some students are sitting with their back to the teacher or that the teacher moves around frequently are common problems.

> *If a teacher wants to lecture, it has to be short. When you are at the other end of the class and you try to listen, you see the little PowerPoint or you have to turn towards the other table and then you lose eye contact and you are not focused. If I can't see the teacher talking, I can't see what is being written on the board.* (case 3–10)

For some, the problem is more related to being able to see the teacher than the actual content.

> *It's more the vision that is problematic. I'm often sitting near the wall and the wall board, the third one down, so I cannot see what the teacher is writing.* (case 2–9)

Several students of one particular focus group suggested a more flexible layout in the ALC.

> *For active learning tasks like this, I would rather have a traditional classroom layout with separate desks. When we are working on an active learning tasks, we would only have to move the desks together.* (case 2–10)

DISCUSSION

These results confirm that from the students' point of view, the most important features of an ALC essentially support collaboration within the work environment—round tables, wall projection surfaces—a finding in line with Soneral and Wyse (2017). The importance of supplying very comfortable rolling chairs was the main finding in the qualitative data. The most expensive feature, requiring specific, costly hardware (Sharing the work of a particular team with the other teams), ranked among the lowest.

The fact that round tables are used in every class or many times during every class is not surprising, because classroom layout deeply influences the pedagogy that takes place (Brooks, 2012), but it also ranked first in terms of perceived utility, perceived interest and perceived ease of use.

ALC Technological Environment: Computers and Tablets

The fact that computers supplied by the institution ranked quite high is somewhat puzzling, given that over 90% of college students in Quebec owned a personal laptop in 2011 (Poellhuber and Karsenti, 2012) and that these numbers probably rose. In that earlier study, however, it was found that <50% of students actually brought their computer to the college, for a variety of reasons: fear of theft or accidents and the fact that many teachers actually forbid the use of laptops in their classrooms.

In the tested ALC, one or two computers connected to the projector were available to each team. These computers were easy to connect and were loaded with a variety of software. This result suggests that in planning BYOD approaches, it is still useful to plan to supply a computer or tablet for each team pod.

There is an apparent contradiction in the results concerning tablets supplied by the institutions. Tablets were actually supplied in one of the participating colleges, but no teachers from that particular college took part in the study in the fall 2015 semester. While the frequency of use is very low, the perceived utility or interest of supplying tablets is rated quite high, which is consistent with the result concerning supplying computers.

Relationship Between Ease of Use, Utility and Actual Use in the TAM Model

The connection between ease of use and perceived utility or perceived interest is quite obvious in the correlation tables, so when developing an ALC, the administration should work on both ends: facilitate the use of functions that are less easy to use in order to increase their perceived utility and demonstrate the potential of particular functions in order to increase the students' interest.

From the TAM perspective, in this particular study, the interest component added little, if any, value to the model, since the interest and utility scores were almost identical for all functions. Deeper reflection is required on the links between ease of use and perceived utility. This relationship is already predicted by the TAM but more research could elucidate the process underlying the relationship (e.g., Ease of use Perceived utility or Perceived interest Ease of use), through structural equation modeling, for example (Park, 2009).

Access to Specialized Software

Sharing the screen of a particular team with the whole class is less easy to use than many other features and it is currently done by the teachers themselves, but our observations of class dynamics tend to show that physical proximity is an important vector for inter-team collaboration and sharing.

The relevance of subject-specific software was raised in the students' qualitative comments. The computers supplied by the institution came with pre-installed programs, which are sometimes quite expensive. This converges with other research results in the same project which indicate that the development of subject-specific technopedagogical knowledge is linked to student outcomes (St-Laurent, Poellhuber et al., submitted).

ALC Set-Up

The other salient qualitative category pertains to the special look and feel of the ALC. From the end-users' perspective (students), being in a space that conveys a different kind of atmosphere than a regular classroom seems to contribute to their intellectual and affective comfort, and even their engagement in group activities. This result is similar to what Park and Choi (2014) report: "Students perceived the ALC environment as more inspirational" (p. 749). While exploratory, this result suggests that special attention should also be placed on the esteem function in the functional analysis approach.

Some negative aspects of the ALC set-up also came up in the student group interviews. Many students reported problems when the teacher lectures, mainly related to the difficulty of seeing the teacher clearly and making eye contact or seeing the board the teacher is using (Park and Choi, 2014). Some students suggested a flexible approach to classroom layout rather than a fixed ALC layout. These comments support the flexibility concept which is present in many ALC layouts. Minor changes could also be made in order to make the teacher and the board more visible during lectures. The teacher podium could be shifted away from the center and toward the front and/or a slightly elevated podium could be built for the teacher. Half-rounded tables on wheels could also be used to easily recreate or dissolve the teams.

Cost-Effective Planning of an ALC

These results are good news for institutions that want to invest in ALC, because the most important features of an effective ALC can be designed and implemented at a fairly low cost. The results of **Table 1** offer guidelines for an efficient cost/benefit ALC design. For example, some interesting devices linked to the teacher's computer can project the image of a real object. This type of equipment proved useful in this project because the teachers could easily show a variety of objects (manipulate valuable objects, show a newspaper article, show a problem taken from a book), adding spontaneity and personalization to their presentations. Solutions for this function run anywhere from $100 to a few thousand dollars. Results show that investing a little money to supply one computer per table, loaded with relevant subject-specific software, might also be a good investment.

A traditional classroom with an even floor can be quickly flipped into a basic ALC by adding boards on the walls (approx. $400 each) and grouping individual desks together to accommodate six students. Colour-coded floor stickers can be used to identify the best locations for the desks, should another user move them ($6 per team). More permanent measures (attaching the desks) or a formal agreement between the users of the classroom is another option.

LIMITS

In terms of its limits, this study used single items to measure concepts, an approach that is not mainstream and that is still criticized by some researchers. A social desirability effect is probable, even though the students were reassured that their answers were anonymous.

The value of the TAM for predicting the adoption of particular features by students is also subject to debate. The fact that they use a particular function might be decided more by the instructor's pedagogical scenario than by the students' choice. In some cases, however, it is more likely a joint decision or even a team decision, where for some parts of the work, the team is free to determine which particular tool or function is used. The lower correlation between frequency of use and interest or utility supports this interpretation. A clearer portrait would draw on both teacher and student perceptions and would specify which aspects of use are determined by the teacher's choices and which are determined by the students'. Another limit of this study is that it did not investigate the way each function was mobilized in various pedagogical scenarios. The complex relationship between pedagogy and classroom layout is worthy of future investigation.

One limit of the functional analysis approach is that the particular way a function is made available to users may vary widely as far as ease of use is concerned. If a function is not perceived as easy to use, users will be less likely to use it and will not perceive its potential value. Future research could examine this using the concepts of affordance to investigate how users perceive the educational affordances offered in their environment and how they interact with these possibilities, using, for example, Gibson's ecological approach (John and Sutherland, 2005).

CONCLUSION

This study used a functional analysis approach to identify and prioritize the most important functions of an active learning classroom as linked to actual use by both teachers and learners. This approach offers design teams a high degree of freedom in the choice of how each function will actually be offered.

The most important functions are those can be achieved at moderate cost: the physical layout of the class (tables and chairs), wall boards, the ability to project an image from a device and to annotate that projection, etc. These results can be used to plan the development of active learning spaces in a way that ensures their features will be not only be usable, but actually used. This study clearly conveys the students' point of view on the desired layout of an ALC. From their perspective, planning for some flexibility seems important.

The use of the technology acceptance model was valuable for finding indicators of perceived ease of use, utility and interest that ranked similarly to frequency of use. This similarity suggests that students see the utility of the equipment used in class. Ease of use also points to solutions that could be improved in the future.

Future functional analysis approaches include comfortable chairs, esteem functions and some flexibility. Future studies could therefore explore the differences in the particular ways different functions are offered and draw on both student and teacher perceptions.

While this particular paper focused on ALC design, the whole study examined the sound pedagogical practices that take places in these environments and how to prepare and accompany teachers in adopting them. Planning the design of an ALC is important but it must go hand in hand with the instructors' preparations. The complex relationship between the teachers' pedagogical practices, the physical layout and the technology, on one hand, and the students' motivation and engagement, on the other, should be investigated differently. For example, to gain a better understanding of the decisions to use or not use a particular function that is available in the classroom, a qualitative study could be undertaken to establish the relationship between the uses related to teachers' pedagogical scenarios and the uses developed by student teams.

AUTHOR CONTRIBUTIONS

BP was the main investigator in the research project and designed all aspects of the study, supervising the research at each point. He also wrote the largest part of the article. SF a doctoral student of BP was the co-researcher in the project. He supervised all data collection and participated in the research design. He helped plan the article at a high level and helped validate parts of the text. NR's contribution came at the analysis and article production phase. He conducted many statistical analyses, only a few of which are reported in this particular article. He contributed to the literature review and helped write the methods, results and discussion.

ACKNOWLEDGMENTS

This study was financially supported by an SSHRC partnership development grant (grant number 890-2012-0052), as well as a grant from the PAREA program of the Quebec Ministry of Education (PA2013-012).

REFERENCES

Adams Becker, S., Cummins, M., Davis, A., Freeman, A., Hall Giesinger, C., and Ananthanarayanan, V. (2017). *NMC Horizon Report: 2017 Higher Education Edition.* Austin, TX: The New Media Consortium.

Alphonse-Prescott, A. (2016). *The Physical Learning Environment of Online Graduate Students*. Nova Southeastern University.

Anderson, T. (2005). Design-based research and its application to a call centre innovation in distance education. *Can. J. Learn. Technol.* 31. doi: 10.21432/t26k60

Barrette, C. (2009). Métarecherche sur les effets de l'intégration des TIC en pédagogie collégiale. *Revue internationale des technologies en pédagogie universitaire/Int. J. Technol. Higher Educ.* 6, 18–25. doi: 10.7202/1000008ar

Basque, J. (2004). En quoi les TIC changent-elles les pratiques d'ingénierie pédagogique du professeur d'université? *Int. J. Technol. Higher Educ.* 1, 7–13. doi: 10.18162/ritpu.2004.52

Beichner, R. J., and Saul, J. (1999). "Student-centered activities for large-enrolment university physics (SCALE-UP)," in *Proceedings of the Sigma Xi Forum on the Reform of Undergraduate Education* (Minneapolis, MN), 43–52.

Beichner, R. J., Saul, J. M., Abbott, D. S., Morse, J. J., Duane, L., Allain, R. J., et al. (2007). "Research-based reform of university physics," in *Physics,* ed E. Redish and P. Cooney (College Park, MD: American Association of Physics Teachers), 1–42.

Benoit, A. (2017). Monitoring implementation of active learning classrooms at lethbridge college. 2014-2015. *J. Learn. Spaces* 6, 14–25. Available online at: http://libjournal.uncg.edu/jls/article/view/1417

Bocconi, S., Kampylis, P., and Punie, Y. (2013). Framing ICT-enabled Innovation for Learning: the case of one-to-one learning initiatives in Europe. *Eur. J. Educ.* 48, 113–130. doi: 10.1111/ejed.12021

Brooks, D. C. (2011). Space matters: the impact of formal learning environments on student learning: impact of formal learning environments on student learning. *Br. J. Educ. Technol.* 42, 719–726. doi: 10.1111/j.1467-8535.2010.01098.x

Brooks, D. C., and Solheim, C. A. (2014). Pedagogy matters, too: the impact of adapting teaching approaches to formal learning environments on student learning. *New Direct. Teach. Learn.* 2014, 53–61. doi: 10.1002/tl.20085

Brooks, D. D. (2012). Space and consequences: the impact of different formal learning spaces on instructor and student behavior. *J. Learn. Spaces* 1. Available online at: http://www.partnershipsjournal.org/index.php/jls/article/view/285

Brown, A. L. (1992) Design experiments: theoretical and methodological challenges in creating complex interventions in classroom settings. *J. Learn. Sci.* 2, 141–178. doi: 10.1207/s15327809jls0202_2

Carnell, B. S. (2017). Connecting physical university spaces with research-based education strategy. *J. Learn. Spaces* 6, 1–12. Available online at: http://discovery.ucl.ac.uk/1561146/

Charles, E. S., Lasry, N., and Whittaker, C. (2011). *Scaling Up Socio-Technological Pedagogies*. PAREA Research Report.

Chalgoumi (2011). *Balises Pour l ' Intervention Avec les Technologies Auprès des Élèves Qui ont des Incapacités Intellectuelles*. Thesis, Montréal, QC: Université de Montréal.

Collins, A., Joseph, D., and Bielaczyc, K. (2004). Design research: theoretical and methodological issues. *J. Learn. Sci.* 13, 15–42. doi: 10.1207/s15327809jls1301_2

Cotner, S., Loper, J., Walker, J. D., and Brooks, D. C. (2013). " It's Not You, It's the Room"—Are the High-Tech, Active Learning Classrooms Worth It? *J. Coll. Sci. Teach.* 42, 82–88. doi: 10.2505/4/jcst13_042_06_82

Davis, F. D., Bagozzi R. P., and Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Manage. Sci.* 35, 982–1003. doi: 10.1287/mnsc.35.8.982

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* 13, 319–340. doi: 10.2307/249008

Dori, Y. J., Belcher, J., Bessette, M., Danziger, M., McKinney, A., and Hult, E. (2003). Technology for active learning. *Mater. Today* 6, 44–49. doi: 10.1016/S1369-7021(03)01225-2

Eccles, J. S., and Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annu. Rev. Psychol.* 53, 109–132. doi: 10.1146/annurev.psych.53.100901.135153

Ha, I., Yoon, Y., and Choi, M. (2007). Determinants of adoption of mobile games under mobile broadband wireless access environment. *Inform. Manage.* 44, 276–286. doi: 10.1016/j.im.2007.01.001

Hake, R. R. (1998). Interactive-engagement versus traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses. *Am. J. Phys.* 66, 64–74. doi: 10.1119/1.18809

Hestenes, D., Wells, M., and Swackhamer, G. (1992). Force concept inventory. *Phys. Teacher* 30, 141–158.

Hoeppner, B. B., Kelly, J. F., Urbanoski, K. A., and Slaymaker, V. (2011). Comparative utility of a single-item versus multiple-item measure of self-efficacy in predicting relapse among young adults. *J. Subst. Abuse Treat.* 41, 305–312. doi: 10.1016/j.jsat.2011.04.005

John, P., and Sutherland, R. (2005). Affordance, opportunity and the pedagogical implications of ICT. *Educ. Rev.* 57, 405–413. doi: 10.1080/00131910500278256

Leung, S. O., and Xu, M. L. (2013). Single-item measures for subjective academic performance, self-esteem, and socioeconomic status. *J. Soc. Serv. Res.* 39, 511–520. doi: 10.1080/01488376.2013.794757

McGowan, B. S., Wasko, M., Vartabedian, B. S., Miller, R. S., Freiherr, D. D., and Abdolrasulnia, M. (2012). Understanding the factors that influence the adoption and meaningful use of social media by physicians to share medical information. *J. Med. Int. Res.* 14:e117. doi: 10.2196/jmir.2138

Miles, M. B., Huberman, A. M., and Saldana, J. (2013). *Qualitative Data Analysis*. Thousand Oaks, CA: Sage.

Park, E. L., and Choi, B. K. (2014). Transformation of classroom spaces: traditional versus active learning classroom in colleges. *High. Educ.* 68, 749–771. doi: 10.1007/s10734-014-9742-0

Park, S. Y. (2009). An analysis of the technology acceptance model in understanding university students' behavioral intention to use e-learning. *J. Educ. Technol. Soc.* 12, 150–162.

Pintrich, P. R. (2003). "Motivation and classroom learning," in *Handbook of Psychology, Vol. 7, Educational Psychology,* eds W. M. Reynolds and G. E. Miller (Hoboken, NJ: John Wiley & Sons), 103–122.

Poellhuber, B., Duclos, A.-M., Fournier St-Laurent, S., and Moukhachen, M. (2018). Avantages et défis des classes d'apprentissage actif au collégial selon les enseignants et les étudiants : les résultats d'une première itération d'une recherche de type ≪ design-based ≫. *Formation et Profession* 26, 7–25. doi: 10.18162/fp.2018.387

Poellhuber, B., and Karsenti, T. (2012). *Les Habitudes Technologiques au Cégep: Résultats d'une Enquête Effectuée Auprès de 30 724 Étudiants*. Research Report. Montréal, QC: Centre de recherche interuniversitaire sur la formation et la profession enseignante (CRIFPE).

Poellhuber, B., Roy, N., and Duclos, A.-M. (2013). *A Mixed Methods Analysis of Closed-Gate Academic Social Network Usages Based on the TAM Model*. San Francisco, CA: Dans Actes du colloque de l'AERA.

Rocque, S., Langevin, J., and Riopel, D. (1998). L'analyse de la valeur pédagogique au Canada: méthodologie de développement de produits pédagogiques. *La valeur des produits, procédés et services* 76, 6–11.

Roland, N., and Talbot, L. (2015). *L'environnement personnel d'apprentissage : un système hybride d'instruments*, *21*. Available online at: http://sticef.univ-lemans.fr/num/vol2014/20-roland-epa/sticef_2014_NS_roland_20.htm

Severin, P. P. (2009). *L'analyse Fonctionnelle, de la Méthode aux outils Course Handouts from the Course "Biotechnologies Santé Environnement." Lycée des métiers de l'hôtellerie et du tourisme de Toulouse.* Available online at: http://asset.free.fr/IMG/pdf/Analyse_fonctionnel_complet_pub.pdf

Soneral, P. A., and Wyse, S. A. (2017). A SCALE-UP mock-up: comparison of student learning gains in high- and low-tech active-learning environments. *Cell Biol. Educ.* 16:ar12. doi: 10.1187/cbe.16-07-0228

St-Laurent, S. F., Poellhuber, B., and Moukhachen, M. (2017). Liens entre le modèle CBAM et l'approche d'enseignement dans le contexte de l'adoption d'une classe d'apprentissage actif par des enseignants au postsecondaire| Relationships between the CBAM Model and the approach to teaching inventory in the adoption of the active learning classrooms by postsecondary teachers. *Can. J. Learn. Technology/La Revue Canadienne de l'apprentissage et de La Technologie*, 42. doi: 10.21432/T2S32X.Available online at: http://www.cjlt.ca/index.php/cjlt/article/view/27528/20234

Stoltzfus, J. R., and Libarkin, J. (2016). Does the room matter? active learning in traditional and enhanced lecture spaces. *CBE Life Sci. Educ.* 15, ar68. doi: 10.1187/cbe.16-03-0126

Venkatesh, V., and Davis, F. D. (1996). A model of the antecedents of perceived ease of use: Development and test. *Dec. Sci.* 27, 451–481. doi: 10.1111/j.1540-5915.1996.tb01822.x

Vo, N. T. B. (2015). *College Students' Choice of Informal Learning Spaces*. Columbia: Doctoral dissertation, University of Missouri.

Wanous, J. P., Reichers, A. E., and Hudy, M. J. (1997). Overall job satisfaction: how good are single-item measures?. 82, 247–252.doi:10.1037/0021-9010.82.2.247

Wesch, M. (2007). *A Vision of Students Today*. Available online at: https://www.youtube.com/watch?v=dGCJ46vyR9o

Whiteside, A., Brooks, D. C., and Walker, J. D. (2010). *Making the Case for Space: Three Years of Empirical Research on Learning Environments. EDUCAUSE Quarterly*. Available online at: http://www.educause.edu/ero/article/making-case-space-three-years-empirical-research-learning-environments

Wilson, H. K., and Cotgrave, A. J. (2016). Factors that influence students' satisfaction with their physical learning environments. *Struct. Survey* 34, 256–275. doi: 10.1108/ss-01-2016-0004

Woolner, P., McCarter, S., Wall, K., and Higgins, S. (2012). Changed learning through changed space: when can a participatory approach to the learning environment challenge preconceptions and alter practice? *Improving Sch.* 15, 45–60. doi: 10.1177/1365480211434796

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in ICT

ORIGINAL RESEARCH
published: 10 January 2018
doi: 10.3389/fict.2017.00030

Check for updates

Flipped Classroom in Organic Chemistry Has Significant Effect on Students' Grades

*Caroline Cormier** *and Bruno Voisard*

Department of Chemistry, Cégep André-Laurendeau, Montréal, QC, Canada

The flipped classroom as a form of active pedagogy in postsecondary chemistry has been developed during the last 10 years and has been gaining popularity with instructors and students ever since. In the current paradigm in science, technology, engineering, and mathematics education, it is widely recognized that active learning has significant positive effects on students' grades. Postsecondary organic chemistry is a difficult course for students, and the traditional way of teaching does not foster students' active involvement. Implementation of active pedagogy could increase students' achievement in this course. However, few quantitative data are available on the impact of active pedagogy in general, or flipped classrooms in particular, on learning in organic chemistry at a postsecondary level. Thus, in this study, we evaluated the gain on final grade scores in organic chemistry after implementing a flipped classroom approach to promote active learning in this course. We encouraged students to be active by having them watch educational videos before each class and then having them work during class time on problems that focused on applying the concepts presented in the videos. Exams were the same as those completed by students in the traditional classrooms of our college. In an *a posteriori* analysis of our students' grades, we compared final grades in traditional classrooms (control group, $N = 66$) and in flipped classrooms (experimental group, $N = 151$). The sample was stratified in three categories depending on students' academic ability in college, from low-achieving to high-achieving students. Our results show that students in the experimental group have significantly higher final grades in organic chemistry than those in the control group, that is, 77% for students in the active classroom vs. 73% in the traditional classroom ($p < 0.05$). The effect was the greatest for low-achieving students, with final scores of 70% in the active classroom compared with 60% in the traditional one ($p < 0.001$). This difference in performance is likely due to students spending more time solving problems in a flipped classroom rather than having the questions assigned to them as homework.

Keywords: flipped classroom, organic chemistry, higher education, active learning, educational video

Edited by:
Elizabeth S. Charles, Dawson College, Canada

Reviewed by:
Kelly Anne Miller, Harvard University, United States
Lisa McDonnell, University of California, San Diego, United States

***Correspondence:**
Caroline Cormier
caroline.cormier@claurendeau.qc.ca

Specialty section:
This article was submitted to Digital Education, a section of the journal Frontiers in ICT

Received: *29 October 2017*
Accepted: *19 December 2017*
Published: *10 January 2018*

Citation:
Cormier C and Voisard B (2018) Flipped Classroom in Organic Chemistry Has Significant Effect on Students' Grades. Front. ICT 4:30. doi: 10.3389/fict.2017.00030

INTRODUCTION

Organic chemistry has always been considered a difficult topic (O'Dwyer and Childs, 2017). Some authors attribute this to the new and non-familiar tasks organic students are required to perform (for example, drawing and interpreting tridimensional molecules on a two-dimensional surface, or predicting the products of a reaction based on the nature and reactivity of the reactants) and because organic chemistry is usually very fast paced due to the large quantity of topics to be covered in one semester (Fautch, 2015). Around the world, general chemistry is typically a prerequisite for enrollment in organic chemistry. However, topics studied in these two courses are very different. While general chemistry largely relies on mathematical analysis, organic chemistry focuses more on the relationship between structure and reactivity (Halford, 2016), and on more difficult intellectual tasks that are more prone to alternative conceptions (Rushton et al., 2008; McClary and Bretz, 2012). Reinforcing understanding of organic chemistry's relationships and tasks requires practice. Traditional teaching approaches address this concern through out-of-class homework exercises and reserve in-class time for lecturing. Conversely, the flipped-class approach uses in-class time for reinforcement and moves the lecturing out of class.

As remarked by Lasry et al. (2014), from a strictly economic standpoint, the most expensive resource in a classroom is the teacher. But this resource is not always used the most economically: when a teacher merely lectures from the textbook, his or her greater asset, helping the students to actively construct and apply their knowledge, is not employed. Following this consideration, several educators have undertaken the task of encouraging active engagement from their students during class time. Reviewing several studies conducted in science, technology, engineering, and mathematics (STEM) education; Freeman et al. (2014) concluded that any active pedagogy in STEM education improves students' grades. Their definition of active pedagogy is very wide and includes class activities as diverse as "occasional group problem solving, worksheets or tutorials completed during class, use of personal response systems with or without peer instruction, and studio or workshop course designs" (Freeman et al., 2014, p. 1). Their observation implies that science educators should concentrate their efforts on modifying their lesson plans and having students be more active in their learning. This, as they suggest, can be done in many ways.

Most organic chemistry educators are in favor of having students become more active in class. However, with a full syllabus, it can be difficult to free class time to do so. The flipped classroom solves this problem.

Several authors describe the flipped classroom, its purpose and the way it is implemented by teachers. As defined by Smith (2013), "flipping the classroom, at its simplest, involves pushing lecture material outside the classroom as a form of homework or other pre-class preparation, leaving more time in class for interactive or engaging exercises" (p. 607). There is a real challenge in implementing the flipped classroom, which is the necessity to integrate work at home and work in class into a pedagogically sound teaching approach that fosters the best learning outcome for students.

In a flipped classroom, direct instruction is moved outside of class (Flipped Learning Network, 2014), usually by assigning videos as homework. The flipped classroom is not merely distance education. Class time is crucial, and activities done in class are central in the approach. Therefore, implementing the flipped classroom involves having the instructor "redesign the curriculum so that the videos watched before class are integrated into each class with active learning pedagogies" (Albert and Beatty, 2014, p. 422). Moreover, "the practice of flipping involves activities pre-class, in-class, and post-class" (Estes et al., 2014), which should be designed by the instructor to form a coherent, engaging and effective pedagogical approach.

Abeysekera and Dawson (2015) define more precisely the flipped classroom as follows: a "set of pedagogical approaches that (1) move most information-transmission teaching out of class; (2) use class time for learning activities that are active and social and (3) require students to complete pre- and/or post-class activities to fully benefit from in-class work" (p. 3). Students receive the content in advance, generally through educational videos that they view at home, and then they are asked to perform higher-order learning activities in class while the teacher can help them instead of lecturing to them (Smith, 2013; O'Flaherty and Phillips, 2015). Benefits of the flipped classroom are multiple: students at home can pause and rewind the videos and are therefore less likely to fall behind than during a live lecture; class time is no longer passive; the teacher is available to guide students when they encounter difficulties, increasing their chances of persevering; and students receive feedback from teachers immediately, improving their self-awareness and confidence (Horn, 2013).

Feedback was reported by Hattie and Timperley (2007) as 1 of the 5 most effective factors influencing achievement in school, based on a review of 12 meta-analyses of almost 200 studies. They proposed that feedback is so effective because it helps "reduce the gap between current and desired understanding" for students (Hattie and Timperley, 2007, p. 86). These authors suggest that important aspects of feedback are thus to provide the students with a clear goal (answering the question "Where am I going?"), an appreciation of their current understanding (answering the question "How am I going?") and to design specific challenging problems as targets, or to set with them a target of greater automaticity in completing problems (answering the question "Where to next?").

Flipped classrooms are being used more and more as a pedagogical approach in higher education (O'Flaherty and Phillips, 2015). As reported by researchers, most implementations of the flipped classroom are occurring in STEM education (Roehling et al., 2017). However, not all STEM disciplines are equally aware of the effectiveness of this approach and studies on the impact of the flipped classroom on grades are still relatively few in number (Ryan and Reid, 2016). A review of 28 studies about the use of flipped classrooms in higher education (O'Flaherty and Phillips, 2015) reports implementation of flipped classrooms in a wide variety of disciplines, most of them in STEM, namely, in health sciences courses (15 out of 28), in applied sciences (6 out of 28), and in pure sciences (only 2 out of 28, 1 in chemistry and 1 in mathematics), the rest of the studies being in humanities and social science education.

Research results show that students generally seem to appreciate the flipped approach (O'Flaherty and Phillips, 2015). McNally et al. (2016) studied the correlation between the appreciation of flipped classrooms and grades. They observed an improvement in grades with flipped "endorsers" and flipped "resisters," pointing toward the interpretation that "preferences alone may not be the most informative aspect on which to evaluate a flipped classroom environment" (McNally et al., 2016, p. 292). Several studies conducted on the evaluation of flipped classrooms in the past years concentrated mostly on students' appreciation of the approach (Critz and Wright, 2013; Butt, 2014; Yeung and O'Malley, 2014; Young et al., 2015) and did not evaluate other aspects of its potential effectiveness.

The one study in chemistry education from the aforementioned review (O'Flaherty and Phillips, 2015) was conducted in two Physical Chemistry courses in the UK (Yeung and O'Malley, 2014). All the lectures were filmed in screencasts lasting between 20 and 40 min. Problems were submitted each week to students, who had to work them on their own and could get help from professor during in-class optional workshops. At the end of the semester, students were questioned on their appreciation of the approach. While most of them reported having preferred the flipped classroom to a traditional course (the preference for the flipped classroom was around 80% with a response rate of around 50%), some students still report that the screencast videos were not as engaging as a live lecture and that they did not allow for students to ask questions.

A few other studies were conducted in undergraduate college chemistry education with flipped classrooms. Ryan and Reid (2016) implemented flipped classroom in general chemistry. They questioned students on their appreciation of the approach as Yeung and O'Malley (2014) did before but designed their study to be able to measure academic improvement as well. One part of a student cohort was enrolled in a traditional, lecture-based course, while the other part was enrolled in the same course using a flipped classroom approach. The two populations took the same standardized test at the beginning of the study and another version of the same test at the end of the semester. The setting of the flipped classroom in that study included educational videos of screencast PowerPoint slides to be viewed before class, and cooperative activities conducted during class time, but no traditional lectures at all, while the control group met in class for traditional lectures during the entire semester. Authors reported a significant improvement of academic grades for the lowest-achieving category of students in the flipped classroom setup, but no statistical difference for the entire population studied was observed. They interpreted this finding as follows: "Our results are consistent with the idea that active learning holds particular benefits for students who are capable but less well prepared" (Ryan and Reid, 2016, p. 21). Furthermore, they noted a significant diminution of withdrawal rates from 23% in the control course to only 6% in the flipped classroom, as a likely result of students being more engaged by the setting of the flipped classroom. This seems to be contrary to what was observed by Yeung and O'Malley (2014), who reported less engagement with the flipped classroom. This difference might be explained by the fact that Yeung and O'Malley's flipped setting was only long screencasts of lectures without a particular device developed to use during class time. On the other hand, Ryan and Reid described discussions and activities conducted during class time to complement shorter videos to be watched pre-class. The integration of activities might be the reason why students felt more engaged in the latter setting.

In organic chemistry, Christiansen (2014) conducted a very small-scale study with one group of seven students in a flipped classroom and one group of six students in a traditional classroom. Flipped-class students were required to watch screencasts of the PowerPoint presentation of the lecture at home before the course and class time was used to work problems in groups with the help of the professor when needed. To encourage students to watch the videos, a quiz was included at the beginning of every other class. No difference was noted between the experimental and the control groups, although this is perhaps due to the very small number of participants.

Also in Organic Chemistry, Mooring et al. (2016) studied the effect of flipping a large-enrollment university course on students' grades and attitude toward the course. They had students watch pre-class videos and answer pre-class online quizzes. In class, students worked in small groups on worksheets. Each class started with a traditional lecture of around 20 min. Students also had to participate in weekly out-of-class exercise sessions with teacher assistants. The authors reported an increase in A and B scores in the flipped course when compared with historical data. However, this study was conducted with only one instructor, and the sample consisted of only one group in one semester. Also, the out-of-class sessions with teaching assistants in the flipped setting replaced online homework in the traditional classroom. The authors warn us that it is impossible to disentangle the effect of these sessions from the effect of the flipped classroom in itself. Also, students' results in a standardized exam were not significantly different between flipped instruction and traditional instruction.

Overall, very few studies have been conducted in chemistry, much less in higher education chemistry and organic chemistry. Most studies report flipped-class use in high schools (Fautch, 2015). Still students seem to appreciate this method of teaching and more and more educators around the world are beginning to implement this approach. It is surprising that so few studies have been conducted in chemistry since the popularization of the flipped classroom approach owes a great deal to two chemistry teachers, Bergmann and Sams (2012), who published a book recounting their experience of developing and implementing flipped classrooms for high school chemistry teaching. An ever-increasing number of educators have followed their example since then in a very wide array of disciplines.

The current lack of data demonstrating the effectiveness of the flipped classroom have drawn some criticism, with authors asking the research community to provide actual data before spending time and resources on its implementation. Abeysekera and Dawson (2015) describe this issue with lucidity: "flipped classroom approaches are being adopted with much enthusiasm despite the paucity of specific evidence about their efficacy" (p. 10). The keen interest in flipped classrooms is sometimes motivated by budget preoccupations in countries where universities may "see the flipped approach as a means of delivering cost-effective, student-centered curricula in the face of increasing student numbers" (O'Flaherty and Phillips, 2015, p. 86).

While budget considerations are always a concern, this was not our motivation for the implementation of a flipped classroom approach. We based our choice on pedagogical reasons, recognizing that flipped classrooms have the potential to make the students more active, to free class time for significant activities, to offer more flexibility for students to learn at their own pace, and to increase student responsibility toward learning. Research also pointed toward the fact that it might be an effective pedagogical approach albeit with insufficient data on its efficacy, particularly in higher education and in chemistry (Ryan and Reid, 2016).

Following the cautionary notice provided by O'Flaherty and Phillips (2015), "one of the greatest obstacles [is] related to staff capacity to design, implement and evaluate the effectiveness of their flipped classrooms" (p. 94), we decided to verify the impact of the flipped classroom approach we implemented in our organic chemistry class on our students' grades. This led to the formulation of the following research question:

> Have our organic chemistry students' grades improved since we implemented the flipped classroom?

This research question follows one of the calls for research from Abeysekera and Dawson (2015), who suggest that quantitative studies should be conducted to evaluate the impact of small-scale interventions, to answer the question, "what is the efficacy of the flipped classroom approach in this discipline, this classroom, with these students?" (p. 11). Data analyzed were students' grades before and after the implementation of the new pedagogy.

In addition to this main question, we pursued a second question in this study:

> Did the students appreciate the flipped classroom we implemented?

However, before answering this, it is necessary to provide a full account of our pedagogical approach to the flipped classroom, since several types of flipped classroom exist, and their differences do not reside only in the use of educational videos.

This study is different from previous studies since it reports on flipped classrooms in higher education organic chemistry in a small-enrollment course with different instructors through four successive semesters. It also compares students' grades in traditional and flipped classroom by considering a measure of academic ability as a moderating factor.

MATERIALS AND METHODS

This research reports from an *a posteriori* analysis of data collected during the normal course of our teaching of organic chemistry to verify if the flipped classroom we implemented had a significant effect on students' grades. In this section, we first describe the learning environment in which this research falls, and then we present the method employed to answer the research question.

Learning Environment

Quebec's Colleges and Organic Chemistry

This study was conducted in a postsecondary college in Montreal, Canada. In Canada, education is under the responsibility of the provinces (CICIC, 2017). Quebec, the province where Montreal is located, has a unique postsecondary system. All students in Quebec must obtain a 2-year college diploma prior enrolling in university. Colleges offer both 2-year pre-university diplomas and 3-year vocational diplomas. This study was conducted in the pre-university 2-year science program. Furthermore, education in this college is conducted in French, the first language in the province of Quebec.

Organic chemistry is an optional course in the science program. About two-thirds of science students select it, since it is a prerequisite for several health and pure science university programs in Quebec. Students enroll in this course during their third semester, after having studied General Chemistry in the two previous semesters.

The organic chemistry course as designed in our college consists of 5 h/week of class time split into one block of 2 h and one of 3 h. Lab periods, lasting 2 h, take place during the 3-h block approximately once every 2 weeks.

Organic chemistry taught in Quebec's colleges is very similar, in terms of content and difficulty level, to what is taught in undergraduate programs elsewhere in the world. Results collected in this study could therefore be of interest for educators outside of Quebec.

Traditional Classroom vs. Flipped Classroom

Class time in the traditional (control) setting used to be devoted to lectures either supported with PowerPoint presentations or printed course notes and a textbook. These were interspersed with some professor-led exercises on the board and some exercises that students could practice, for which the professor provided the answer. Then, as homework, students were assigned end-of-chapter exercises and problems from a textbook to consolidate their knowledge.

As Jensen et al. (2015) explained in their paper about the comparison of traditional and flipped classrooms in a university biology course, the difference between these two types of pedagogy is principally the moment the students are first in contact with new subject matter and the platform through which this first contact is made. In our traditional setting, students first learned about new topics in class, in the presence of the professor. At this moment, students would be engaged toward the material, they would explore the contents and the professor would explain to help facilitate learning (Jensen et al., 2015). After class, students in the traditional classroom would be asked to apply their knowledge to novel situations, that is, by practicing textbook problems at home. They would be evaluated during summative examinations, but no formal formative assessment was included in the course and homework were not graded. Students were responsible to verify if they were able to complete textbook problems and to see the professor outside of class time for any questions about the course content.

In the traditional as well as in the flipped settings, the layout of the classroom was a traditional seating arrangement, with tables facing the board and grouped in two- or three-table pods.

In our organic chemistry course in a flipped classroom, the engagement, exploration, and explanation phase would occur before class time, through a series of video on an online platform. Then, during class time, students would participate in face-to-face activities to apply the new knowledge. After class, students were still assigned end-of-chapter exercises and problems as homework. These homework exercises were not verified nor graded by the instructors. The three moments of the flipped classroom setting (pre-class, in-class, and post-class) are described in the following sections.

Pre-Class: Videos

Students enrolled in the flipped classroom organic chemistry course were required to watch videos before coming to class. Typically, three to five videos were assigned each week, for a total video time of 30 min. The videos are of four types: theory, exercises, laboratory techniques, and software use. Our teaching staff, consisting of 2 professors, with the sporadic help of 2 other professors and 1 laboratory technician, prepared all 75 videos during the Fall semester of 2013. Most videos are about theory (57 videos), principally showing one or both professors in front of a white board, explaining concepts to the camera or to each other and noting key concepts or examples on the board; a small number of videos are rather screencasts of a PowerPoint presentation with a voice-over by one of the professors (see **Figure 1**).

Videos were shot with special attention to their length, which was kept as short as possible. Mean video length is in fact 6 min and 13 s (SD = 2:19). Guo et al. (2014) conducted an empirical study about features of videos used in massive open online courses on students' engagement and observed that normalized engagement stayed high with videos up to 6–9 min long, but that it dropped significantly with longer videos. Each of our videos was constructed around one topic, allowing students to find the topic they were looking for easily.

Most videos prepared for this organic course also have a "pause-solve-resume" feature. That is, professors would suggest an exercise, and invite students to press pause, solve the exercise on paper, then resume the video for the solution. This feature is rather low-tech, considering the abundance of interactive tools to segment videos for this purpose (for instance, http://EDpuzzle.com). However, using the low-tech version of a "pause-solve-resume" feature was less time consuming for the professors and allowing students to get immediate feedback on their understanding and ability to solve simple problems on new content. As was reported in the literature, students are less engaged in the outside-of-class activities of a flipped classroom if these activities lack interactivity or feedback (O'Flaherty and Phillips, 2015). The "pause-solve-resume" feature can provide a minimum of interactivity and feedback to students.

Students were encouraged to take notes while watching the videos, and to note any unresolved questions they had. These notes and questions were then used in class, as explained in the next section. However, no incentive was used to ensure students would watch the videos nor did we verify if they did. Since students would have to use their notes to complete classroom activities, we noticed when a student had not watched the videos. However, it rarely occurred as the semester progressed since students rapidly learned that not watching the videos would impede the work they would be able to do during class time.

In-Class: Questions, "Portfolio" Exercises, and Micro-Lectures

In the flipped classroom, the course flow was constructed as follows: first, professors would answer students' questions about the videos they watched before class, for periods ranging from a few minutes to 15–20 min, depending on the number of questions from students. Second, students would work on a sheet of exercises, called "portfolio exercises," brought to class by the professors. These exercises were a direct application of the topics covered in the videos. Students were encouraged to work in pairs and to ask the professors questions whenever they needed help. This practice, having students work in class on face-to-face activities was recommended by Strayer (2012) as a means to strengthen and apply students' understanding of more formal notions seen in videos. Depending on the length of the portfolio exercises, 15–30 min would be devoted to this activity, at the end of which students were asked to give them back to the professors for a formative assessment. Several types of exercises were designed. Practicing organic nomenclature, reaction mechanisms, and forms of drawing molecular structures were part of them.

Typically, after the portfolio exercise, a micro-lecture would be given by the professors on a subject that was not covered in the videos. In fact, some topics were intentionally reserved for micro-lectures, often because they were more difficult or needed a subtler understanding (Sweet, 2014). For example, the explanation of the factors used to predict if a chemical reaction would undergo a nucleophilic substitution or an elimination mechanism was given in a micro-lecture in class, with several examples and the possibility for students to ask questions immediately. These micro-lectures were variable in length, typically lasting between 20 and 30 min.

After the micro-lecture, another portfolio exercise sheet would be distributed to students about topics covered in the micro-lecture. These exercises would then be completed, handed in to the professors and formatively assessed. Approximately one portfolio exercise sheet was thus distributed every class hour.

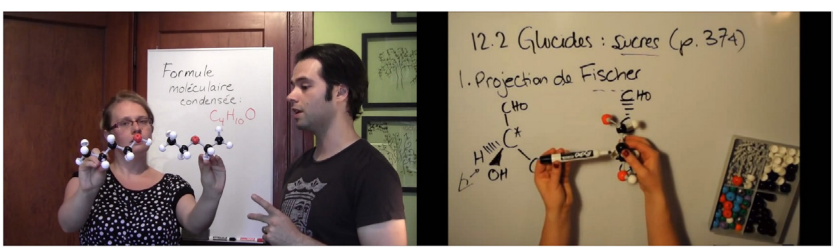

FIGURE 1 | Screenshot of two videos filmed for this flipped classroom. Left panel: an example of a theory video; right panel: an example of a solved exercise.

Between classes, professors formatively assessed the portfolio exercises by indicating where students had made mistakes, but without giving the right answer. Portfolio exercises were handed back to students at the beginning of the next class. Students then had to put the portfolio exercises into their portfolio (a cumulative report) and, outside of class time, they were required to correct their mistakes. Prior each exam, students handed the corrected portfolio in for grading. Only remaining mistakes lost them points. A very small mark was allotted to the portfolio (5% in total), but it was sufficient for students to comply with these requirements. The formative assessment of portfolio exercises provided prompt feedback to students and the format allowed them to make mistakes without being punished. The portfolio was seen by students and professors as a learning tool and not as an evaluation tool. Students are more likely to benefit from the approach if their professor integrates assessment into the design of the flipped classroom (McNally et al., 2016).

As reported by Jensen et al. (2015), this flipped setting allowed students to receive more explanation since they are provided with answers to their questions about the videos at the start of the course and to their other questions during exercise time. Furthermore, it allows a phase of evaluation of knowledge that is not possible with traditional classroom, the immediate feedback the students receive while applying their new knowledge in class.

Post-Class: Consolidation Exercises

After class, teachers suggested exercises in the textbook for students to continue practicing the problems worked on in class and to consolidate their knowledge. Students were autonomous in these exercises, and their completion was not verified during class. These exercises resembled the portfolio exercises and since the textbook was also written by the professors of several sections of the experimental sample (Voisard and Cormier, 2013), they were relevant to the topics studied and adequate to the level of the course.

Co-Teaching

Some classes taught in the flipped classroom were also taught by two professors in co-teaching. This co-teaching consisted of both professors being present during class time, alternatively answering students' questions, giving micro-lectures, and helping students during portfolio exercises. The experimental group was therefore of two types: of the seven classes taught in flipped classroom for this study, four were co-taught while the remaining three were taught by a single professor.

Co-teaching was done on a volunteer basis, meaning the extra amount of class hours were not considered in the teachers' remuneration. However, since the workload of implementing a flipped classroom approach can be demanding (O'Flaherty and Phillips, 2015), co-teaching, particularly in the numerous hours spent preparing videos, was greatly appreciated by both professors.

Research Method

We compared grades in organic chemistry in our college before and after the implementation of the flipped classroom. This was done by an *a posteriori* analysis so the actual evaluation was conducted after students had completed the course either in traditional or flipped settings. As noticed by O'Flaherty and Phillips (2015) in a large review of studies on flipped classroom in higher education: "the majority of articles evaluated student outcomes by comparing an existing course taught in a traditional manner with a course imbedding a flipped class" (p. 89). Several authors have used historical data to find the effect of the flipped classroom, in particular Ryan and Reid (2016) in higher education chemistry. This research approach was also used in this study, where historical data were used as the control sample to which outcomes of the flipped classroom were compared.

Population

Since we worked with two consecutively enrolled populations of students, the sample is the entire population of organic students between 2012 and 2014. **Table 1** presents the two groups we compared. In total, 74 students were enrolled in the control sample, but 8 of them (10.8%) withdrew during the semester. Students who withdrew were not included in the analysis since the reason of their withdrawal was not documented. Similarly, 13 students from the experimental sample withdrew from the course (7.9%) and were not included in the analysis.

Class size was similar between traditional and flipped classrooms: the traditional sample of 66 students was distributed into 3 classes of 22 students on average, while the flipped sample of 151 students was distributed into 6 classes of 25 students on average.

The composition of both samples regarding the sex of the students is slightly different with proportionally more women being enrolled in the experimental sample. This difference in composition is, however, not statistically significant (Pearson's $\chi^2 = 1.093$, $p = 0.296$).

The last information presented in **Table 1** is the *R*-Score means of each sample. This measure of academic ability will be explained in the next section.

Organic Chemistry Grades and *R*-Scores

Quantitative data collected for this study are students' organic chemistry final grades. The final grade is on 100 points, the passing grade being 60%. These grades include theoretical evaluation (exams) for 65% of the total ponderation and laboratory evaluation (lab reports and lab exam) for 30% of the total ponderation. The remaining 5% is allotted to either a group homework in the traditional setting or the portfolio in the

TABLE 1 | Description of control (traditional) and experimental (flipped classroom) samples.

	Control sample	Experimental sample
Years of data collection	2012	2013–2014
Number of students	74	164
Number of groups (classes)	3	6
Number of withdrawals	8 (10.8%)	13 (7.9%)
Number of students included in analysis	66	151
Gender	56.1% F, 44.0% M	63.5% F, 36.4% M
R-Score mean (SD)	27.6 (3.96)	27.0 (3.86)
R-Score median	27.5	27.2

flipped setting. This difference is somewhat minor (only 5%) between the traditional and flipped classrooms, thus the final grades can be compared.

Since historical data were compared with data after the implementation of the flipped classroom, no randomization of samples could have been done. As suggested by Ryan and Reid (2016), the non-randomization of samples was taken into account with a measure of ability of each sample. These authors had both samples take the same pretest and thus demonstrated their equivalence. In the case of this study, since it was not possible to have past students take a test, we chose to compare the control sample (traditional teaching) with the experimental sample (flipped classroom) based on students' academic ability as measured by *R*-Score.

The *R*-Score "is the instrument of choice for analysis of all applications to university programs" (BCI, 2017) in the province of Quebec. It is an improved *Z*-Score in the sense that it considers the group strength and the group dispersion, making it a robust measure of a student's academic ability. It is calculated for every college student, at the end of every semester by college Academic Dean's offices. For this study, *R*-Scores were used as calculated by our college.

Although the theoretical maximum *R*-Score is 50, it is virtually impossible to get such a number. *R*-Scores above 30 are considered "high" and might lead students to be admitted into limited enrollment university programs such as Medicine or Dentistry (BCI, 2017). In addition to students' final grades in all courses they were enrolled in each semester, *R*-Score calculation considers the strength of the group and the dispersion of the group (as measured with high school grades of all students in their groups, for each course). This measurement can therefore allow universities to sort students based on their academic performance, with no regard for the college students were enrolled in. *R*-Scores are also used locally, in colleges, to evaluate the mean academic strength of group classes, for program evaluation purposes, for example. *R*-Scores have a very high correlation factor to all college chemistry courses grades ($r = 0.873$, $n = 229$, $p < 0.001$), including organic chemistry. For this reason, *R*-Scores are used in this study as a measure of academic ability. *R*-Scores are calculated by Quebec's ministry of education, the *Ministère de l'Éducation et de l'Enseignement supérieur*. *R*-Scores used for this study were obtained by the authors from the Academic Dean's office of the college where the study was conducted.

As shown in **Table 1**, mean *R*-Scores for both samples are 27.6 and 27.0. This difference is not statistically significant ($t = 0.951$, $p = 0.343$), thus the two samples can be considered equivalent in terms of academic ability in college, albeit the control group had a slightly better *R*-Score average.

Ryan and Reid (2016) divided students of each sample (traditional and flipped) into three bins of equivalent ability for further analysis upon each of these bins. Following that example, we divided students of the traditional teaching sample and students of the flipped classroom into three academic achievement categories: low achievers, average students and high achievers. Composition and average ability of these bins are presented in **Figure 2**.

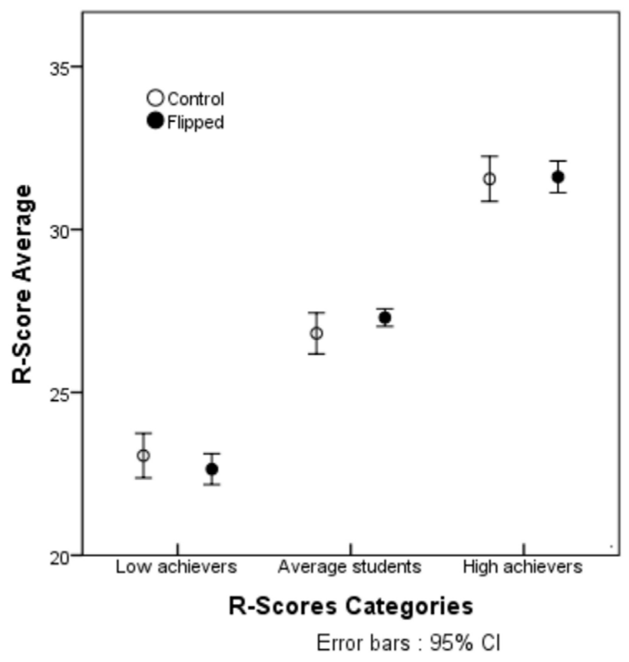

FIGURE 2 | Composition of each *R*-Score category in flipped classroom ($N = 151$) and in traditional (control) classroom ($N = 66$).

Students' Appreciation

At the end of the semesters of fall of 2013, spring of 2014 and fall of 2014, students were questioned on their appreciation of the pedagogical approach in organic chemistry. Students were sent an email containing a link to an anonymous electronic questionnaire containing nine items (Likert-scale and open ended) regarding their appreciation of the pedagogical formula and the videos and probing them on the number of hours devoted to the course outside of class. Examples of items (translated from French by the authors) are presented below:

- What is your appreciation of the course? [I liked it very much; I liked it; I somewhat disliked it; and I hated it].
- What is your appreciation of the flipped format, that is, watching videos before class and working on portfolio exercises in class? [I liked it very much; I liked it; I somewhat disliked it; and I hated it].
- How many required videos do you usually watch prior class? [All of them; most of them; only a few of them; and none of them].
- What type of video do you prefer? [open-ended item].
- How many hours do you spend on organic chemistry material outside of class time each week in average? [0; 1–2; 3–4; and > 4 h].

This questionnaire that was devised as a means of getting feedback from students for a new pedagogical approach was not sent to the control sample, for which data were collected the year before its implementation.

The questionnaire was answered on a volunteer basis, since no control was exerted on the students and that students who chose to answer did so anonymously. The electronic survey was left open for 1 week after inviting students by email to answer it. Students answered it outside of class time.

Constant Parameters between Traditional and Flipped Classrooms

Two elements were kept constant between the course taught traditionally and by flipped teaching. First, exams were kept the same, with year-to-year slight modifications, to prevent cheating. For example, for one version of an exam, students had to draw the mechanism of an esterification reaction between methanol and acetic acid, while in another version, they were asked to draw a mechanism for the same reaction between ethanol and propionic acid. The same knowledge is necessary to answer both problems, making the exams sufficiently similar for the students' grades to be compared. These exams included items on nomenclature and isomerism, drawing of organic molecules and reaction mechanisms, designing of synthesis schemes, and formulating explanations of properties of matter based on molecular structure.

Second, the same laboratory curriculum was used in both settings, with the same laboratory exam. Lab experiments were based on the practice of synthesis, purification and characterization of organic compounds.

Ethical Considerations

Results collected for this study did not include students' identification, but only their *R*-Score, organic chemistry grade, and their sex. Data were provided by the institution's admission service through a list of file numbers, from the admission database. No analysis necessitated students' identification. Since the analysis of data was done *a posteriori* on data present in a database, and no students' identification was collected nor used, no approval from an ethics committee was required for that type of study, as being the analysis of an archival record. Appreciation questionnaires were answered anonymously and on a volunteer basis. Students were informed that their answers might be used for publication, but that no information that might identify them would be collected nor disclosed.

RESULTS

Quantitative results regarding grades in organic chemistry prior and after the implementation of the flipped classroom are presented in this section, followed by qualitative results of the students' appreciation of this pedagogical approach.

Quantitative Results: Grades in Organic Chemistry

By comparing the traditional classroom and flipped classroom in organic chemistry, we first observed that the latter led to statistically better grades for the overall sample. Indeed, in **Table 2**, the overall results show that flipped students had a grade average of 77.1% in organic chemistry compared with 72.9% for the control sample, even though both samples showed no difference in academic achievement as measured by their *R*-Score, as presented earlier. The effect size of this difference is, however, small, as measured by Cohen's d ($d = 0.32$) (Cohen, 1988).

Correlation between organic grades and *R*-Scores is still very high for both the traditional (control) sample and the experimental (flipped) sample: for the traditional sample, Pearson's correlation coefficient between these two variables is 0.827 ($n = 66$, $p < 0.001$), and for the flipped sample, the coefficient is 0.662 ($n = 151$, $p < 0.001$). Note that the correlation coefficient is smaller in the flipped classroom sample. This might be explained by the result that will be presented in the next subsection, which is that not all students' grades increased with the same magnitude.

Difference in Grades for Low Achievers, Average Students, and High Achievers

To further the analysis, we then proceeded to disaggregate results to verify if subgroups of our sample benefited differently from the flipped classroom approach. For this purpose, we analyzed the three bins of students separately based on their *R*-Score, namely, low achievers, average students, and high achievers. Note that these subsamples had similar composition regarding their academic ability.

This analysis led to the most striking results from this study, presented in **Figure 3**. Lowest-achieving students are the ones presenting the largest difference between control and experimental settings, having their grade average going from around the 60% success threshold mark to almost 10% above of it with the flipped classroom, with a large effect size ($d = 0.94$). Difference is also significant for average students, who show a grade average of 7% higher in flipped classroom than in the control group (69 vs. 76%), with a moderate-to-large effect size ($d = 0.73$); high achievers have also slightly better grades in

TABLE 2 | Students' mean grades in organic chemistry in control (traditional classroom) and experimental (flipped classroom) groups, depending on their level of achievement in college.

	Entire sample		Low achievers		Average students		High achievers	
	N	**Mean grade (SD)**	***N***	**Mean grade (SD)**	***N***	**Mean grade (SD)**	***N***	**Mean grade (SD)**
Traditional classroom	66	72.9 (14.7)	21	59.8 (11.0)	18	68.5 (10.6)	27	86.0 (6.6)
Flipped classroom	151	77.1 (11.2)	51	69.8 (10.4)	55	75.6 (8.6)	45	87.1 (7.1)
t Score	2.053*		3.663**		2.868*		0.651	
Effect size	Small effect (Cohen's $d = 0.32$)		Large effect (Cohen's $d = 0.94$)		Moderate effect (Cohen's $d = 0.73$)		No significant difference	

*Statistically significant ($p < 0.05$).
**Statistically significant ($p < 0.005$).

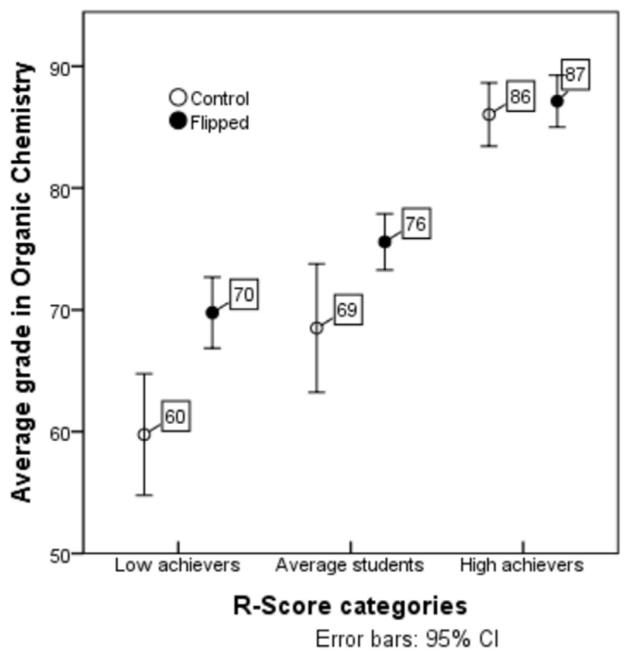

FIGURE 3 | Differences in organic chemistry grades for students in flipped classroom (*N* = 151) vs. students in traditional (control) classroom (*N* = 66).

flipped classroom (86 vs. 87%); however, that difference is not statistically different for this subgroup.

Difference in Withdrawal from the Course

As stated earlier, only students who completed the course we considered for analysis in this study. We observed that a smaller percentage of students withdrew from the flipped classroom groups (7.9%) as from the traditional groups (10.8%), the difference being, however, not statistically significant ($\chi^2 = 0.527$, $p > 0.05$). A similar difference in lower withdrawal rates in the flipped classroom was observed by Ryan and Reid (2016), but with a much larger effect size. These authors explained this difference by their flipped format being more engaging to all students than a traditional lecture in a large (300 seats) lecture hall. Due to our groups being much smaller (around 25 students per group) than those in the study by Ryan and Reid, the increase in engagement may not have been as great between traditional and flipped settings as the one these authors observed.

No Effect for Co-Teaching

It was also important to verify if co-teaching, as used in four flipped classes of the experimental sample, influenced students' grades. We therefore analyzed three subsamples with a one-factor ANOVA: traditional teaching (no co-teaching) in a traditional classroom, traditional teaching (no co-teaching) in a flipped classroom, and co-teaching in a flipped classroom. Since one condition is missing from the design (co-teaching in a traditional classroom), it is impossible to conclude with certainty on the impact of co-teaching with the results collected from this study. Still, results show that co-teaching did not significantly influence grades in the flipped classroom: grade average without co-teaching was 78.5%, and grade average with co-teaching was 75.7%, the difference not being statistically significant. Grade average without co-teaching and without flipped classroom (traditional classroom), on the other hand, was significantly lower at 72.9%. Further research would be necessary, but from the results available now, we can suggest that co-teaching in a flipped classroom as we applied it in this particular setting does not significantly influence students' grades.

Yet, other reasons for wanting to practice co-teaching in a flipped classroom environment might still exist and will be explained in the discussion.

Qualitative Results: Appreciation

Students were questioned on their appreciation of the pedagogical approach in organic chemistry through an anonymous questionnaire. Only data from flipped classroom students are available, but even so, some results are interesting enough to be noted even if no comparison can be made with the control group.

Eighty-eight students responded to the online questionnaire anonymously, after the end of the semester upon email invitation by their professors. The questionnaire was sent to the 99 students who were enrolled in the course at the beginning of the semester and did not withdraw before the end. The high response rate (89%) makes it possible to believe that the answers obtained are representative of the experimental sample of this study.

General Appreciation

When asked if they liked the flipped classroom approach, 83% of the students answered positively (either "I liked it" or "I liked it very much"), which is comparable to results from other studies on general appreciation of this pedagogical approach (Smith, 2013). Since a part of the sample was taught by a pair of professors in co-teaching, we also asked students if they appreciated the co-teaching and 97% of them responded positively. We expected the perception of the flipped classroom to be somewhat lower than the perception of co-teaching, as the first involves more work from the students than simply being taught by two professors. It should be noted, however, that the most popular approaches are not necessarily the most effective: indeed, co-teaching, which is highly appreciated, has had no effect on student grades as seen in the previous section.

Most Preferred, Least Preferred Aspects of the Flipped Classroom

Two open-ended questions asked students to comment on the aspects of the course, the most and least preferred. Answers were grouped under categories, and number of occurrences in each category is presented in **Table 3**. Not all students provided answers to these questions, leading to an unequal total of occurrences. The most frequently mentioned preferred aspect is the flipped classroom in general (21 answers). When adding up all the positive aspects of the flipped classroom, we observe that 47% of the positive comments regarded that particular type of pedagogy. On the contrary, the flipped classroom was only mentioned three times as the least preferred aspects of the course. But by adding up all the least preferred aspects relating to the flipped classroom

TABLE 3 | Number of occurrences of most preferred and least preferred aspects of the organic chemistry course in the flipped classroom format as answered by students.

Most preferred aspect	*N*	Least preferred aspect	*N*
About the flipped classroom			
Flipped classroom (in general)	21	Flipped classroom (in general)	3
Questions are answered in class	7	Workload	8
Face-to-face exercises (portfolio)	4	Face-to-face exercises (portfolio)	2
Videos	4	Videos	2
		Missing: summary of videos at the beginning of class	5
		Missing: printed notes	4
		Missing: more homework	1
		Clarity of PowerPoint used in micro-lectures	1
About the professors or the co-teaching			
Professor	13	Professor	2
Co-teaching	10	Co-teaching	3
About other topics			
Classroom atmosphere	4	In class: pace (too fast), time on exercises (too short)	7
Subject matter	8	Subject matter	1
Lab curriculum or lab reports	3	Lab curriculum or lab reports	9
"I liked everything"	2		
Total of answers	76	Total of answers	48

as we implemented it (type of work outside of class, face-to-face activities, etc.), the total percentage of the flipped classroom being the least preferred aspect of the course was 54%. Students were therefore writing a lot about the flipped classroom for these two questions, either as it being their most preferred or least preferred aspect of the course. This is understandable: for all students, it was their first time being enrolled in a flipped course. Since it probably appeared very different from the lectures they were used to, they had many comments to formulate on the topic. It is also noteworthy to mention that several students cited one aspect of the flipped classroom as being their most favorite of the course and another aspect as being their least favorite, for example, from two given students:

Student 1 favorite aspect: "*Professor availability in class allows for more individual attention to each student.*[1]"
Student 1 least favorite aspect: "*Too much time needed to prepare for class.*"
Student 2 favorite aspect: "*Videos are nice, since they allow me to learn at my own pace (I can pause or rewind if I did not understand).*"
Student 2 least favorite aspect: "*The problem is, if videos are not watched prior to the next class, I would feel lost (I don't always have time to watch all videos) […].*"

Those two students appreciated the general format of the flipped classroom, but some aspects or requirements from it were seen more critically from them.

One of the most frequent negative aspects of the flipped classroom mentioned by students was indeed the workload, an aspect that was also reported by other studies on the flipped classroom (Lage et al., 2000; Mason et al., 2013). Another item of the appreciation questionnaire was specifically about the amount of work outside class hours. We asked the students to estimate the average number of hours of work devoted to the course outside class (possible answers being less than 2 h/week, between 2 and 4 h/week, and more than 4 h/week). Most students answered that they devoted between 2 and 4 h/week at home to the course, which does not exceed the expected 3 h/week as prescribed in the course syllabus and remains the same in the traditional classroom. It is possible that students who complained about the workload in the course felt like they were working more than they used to in other science courses, which may be desirable if they were used to working less than the expected number of hours.

Some negative aspects mentioned by students are about elements that might be missing from our implementation of the flipped classroom, namely, the lack of a form of video summary at the beginning of class, of printed notes for students to fill out either at home while watching videos or in class during micro-lectures, and even the lack of graded homework. These elements were suggested by students as possible ameliorations to the course and could be interesting to consider in the future.

Other very frequent negative aspects regarded the pace of the course (seven answers) and the laboratory curriculum or lab reports (nine answers). These aspects are not related to the new implementation of the flipped classroom, since the pace is always perceived as rather fast in organic chemistry (Fautch, 2015), and the laboratory curriculum was not modified in the flipped classroom implementation. Comments from students for these aspects will therefore not be discussed here.

Degree of Preparedness prior Class

Students were questioned on their assiduity in watching videos before class. 72% of the students declared watching all videos before class, and a further 25% said watching almost all of them, so 97% of the students who answered watched all or almost all assigned videos before coming to class. Note that we chose not to check if the students had seen the videos through online or classroom tests, but most of them still seem to have done the preparatory work.

We did not verify the degree of preparedness of students because videos are available on an open online platform (YouTube) on which students do not need to register. Pedagogical platforms that can host videos, such as Moodle, for example, include tools to verify the completion of pre-class work by students, but were not used in this study.

Appreciation of Videos

We filmed different types of videos, as explained earlier, and in the anonymous appreciation questionnaire, we asked students what type of videos they preferred. As this item was open-ended, students' answers were grouped into categories. The most preferred videos, as evoked by 19 students (30% of all answers), were those where the professors appear on screen, as shown on the left panel of **Figure 1**. Typically, these videos were filmed by a camera placed on a tripod in front of a white board, and in which the professors were discussing with each other, taking notes on the board

[1] Student comments are translated from French by the authors.

and explaining how to solve problems. Most of the videos are of this type, which seemed to have been appreciated by students. The flexibility of the format, the naturalness of presentation, and the fact that students knew the professors might have influenced them in their preferences for these videos.

The second most preferred type of videos mentioned by 13 students (21%) showed the professors solving exercises on the board. In this type of video, the professors were sometimes inside the camera range and other times outside of it, only their arm being shown (see right panel of **Figure 1**). These videos were probably favored because the professors modeled the procedure to solve problems by types, and that the procedure can afterward be practiced by students, either by a "pause-solve-play" feature in the video itself or by suggested exercises in the textbook. No comments concerned the "pause-solve-play" feature of some videos, but it might explain why some students cited the exercise videos as their favorites.

A lot fewer students preferred screencast of a PowerPoint presentation (five students, 8%), showing that seeing the professor is more important than seeing a presentation. A possibility to add personalization to a screencast could be to overlay a video of the professor in the corner of the screencast, thus mitigating the risk of the video being less engaging to the students (Awad et al., 2017).

DISCUSSION

Effectiveness of the Flipped Classroom in Organic Chemistry

Our results point toward the fact that the flipped classroom, as implemented in our course, had a significant effect on learning in organic chemistry, since students' grades improved with that pedagogy as compared with traditional teaching. Very similar exams were answered in the control and the flipped samples, and academic ability of students was controlled in both samples. The only modified factor was the type of pedagogy.

The actual pedagogical device designed for the implementation of the flipped classroom probably has a lot to do with this effectiveness. Indeed, pre-class, in-class, and post-class activities were all integrated to foster mastery in the subject matter. Pre-class videos were short and presented a "pause-solve-resume" feature as means to keep students engaged and reduce cognitive load, as they could go through them at their own pace (Abeysekera and Dawson, 2015).

It is important to remember that we chose not to check that the students had seen the videos by online or classroom tests. The responsibility of being well prepared for class thus lied with students, and even if this responsibility might not be have been taken as seriously by everyone, it was thought to improve self-discipline and the development of students' self-regulation skills (Adnan, 2017, p. 2).

The success of flipped classroom in general is probably dependent on students' self-directing learning skills (Estes et al., 2014). Since our study was conducted in a second-year course, students might have already developed such skills. The same results might not have been obtained by a similar implementation of the flipped classroom with younger students, for instance with first-year general chemistry students. However, with a similar pedagogical device, Ryan and Reid (2016) actually did not see overall improvement of grades in flipped general chemistry. One of the reasons might be the younger students' lack of self-regulation.

We therefore believe that it is not necessary, for second-year college students, to use coercive means to ensure that they watch the videos. The pedagogical device should be sufficient to make them feel that watching the videos is useful and necessary prior coming to class because of the face-to-face activities required from them.

It is recognized that several low-achieving students have less motivation than higher achievers (Horn, 2013) and that that lack of motivation might negatively impact their engagement in pre-class activities. However, our results show the best outcomes for low achiever students. Our interpretation of these results goes in the same direction as a comment made by chemistry professor Christopher J. Cramer, who explained the success of flipped classrooms by the willingness of students to watch videos, more so than reading the textbook prior coming to class. Professor Cramer says: "We're tricking the students into spending twice as much time on the material as they would have otherwise" (C. J. Cramer, reported by Arnaud, 2013). Low achievers, who might spend very few hours outside class doing homework in a traditional setting would have time, in a flipped setting, to do portfolio exercises in class. For some of them, the portfolio exercises might be the only ones they would do, which would still be several more than what they would do in a traditional setting. This extra time and extra practice likely explain the 10% increase of low-achieving students' grades in the flipped classroom.

Some authors propose that the effect researchers observe on students' grades in flipped classrooms probably has more to do to the active-learning setting of the class than the flipped setting in itself (Jensen et al., 2015). This might also be the case in this study. However, we simply implemented the flipped classroom as a model of active-learning environment. Other environments could have been considered, and this study only reports on that one.

Also, our flipped classroom had a higher structure than our traditional classroom, that is, students were graded more frequently, through the portfolio activities, and they had more time to talk to each other and to the professor during class. Increase in structure in the active-learning classroom was reported to have a different influence on some subpopulations, especially first-generation students (whose parents did not go to college) (Eddy and Hogan, 2014). It was proposed that this kind of classroom setting was beneficial because of extra time students devoted to course material and because of a culture of community in class instead of a competitive environment (Eddy and Hogan, 2014). Our classroom setting might also have benefited from these two factors. We did not collect data about parents' schooling for this study, but that could be an interesting question to pursue further.

All students would benefit from a higher time devoted to chemistry outside of class time. We did not verify how many hours students were doing homework in the traditional setting, so we cannot evaluate if they spend more time watching pre-class videos and practicing textbook exercises in the flipped classroom format, but we can postulate that this was the case. In the traditional classroom, students would still have to work the same textbook exercises but did not have to do any pre-class activities.

We were concerned that students in the flipped classroom would therefore have to spend too much time outside of class on our course so we questioned them on that topic in the online questionnaire. For a 5-h college course in Québec, students are expected to work 3 h outside of class each week, as specified by the Ministry of Education program (MELS, 1998, 2000). In average, students in the flipped classroom actually declared devoting around 3 h to the organic chemistry course each week, thus meeting that expectancy. Students may not spend as much time on homework in a traditional setting due to a lack of engagement. The time on task was probably increased with the flipped classroom, but the students still did not exceed the time requirements in the tasks we assigned them. The way our class was designed simply encouraged students to meet the required number of hours expected in the curriculum.

Furthermore, flipped classrooms can improve engagement and motivation because the learning environment they provide are more likely "to satisfy student needs for competence, autonomy and relatedness" (Abeysekera and Dawson, 2015, p. 7).

Another aspect of the effectiveness of our style of implementation of the flipped classroom was probably due to the opportunity, in each class, to provide formative assessment to the students through the portfolio exercises. Prompt feedback is one of the principles of good practice in undergraduate education (Chickering and Gamson, 1987) as it can help students to situate themselves in the learning of the content and it can help instructors monitor individual progression. The way our feedback was designed helped students answer the three questions suggested by Hattie and Timperley (2007): students can answer "Where am I?" by trying portfolio problems, which were designed for the students to be answered right after watching videos. By trying the problems, they had a feedback on the appropriateness of their note taking while watching videos, and by listening to other students' questions at the start of class, they had an appreciation of their level of understanding compared with other students. The question "How am I going?" was answered with the formative assessment of the portfolio exercises, as well as the direct feedback the professors gave students during the portfolio period in class. Finally, the question "Where to next?" was answered by some portfolios exercises that integrated notions from different chapters, for example, chemical synthesis problems. Indeed, videos and regular portfolio exercises were mostly compartmentalized by chapter, and the integration of these chapters constituted a unique challenge to the organic chemistry course. Consequently, by helping students answer these three questions, our flipped classroom approach built on the effective feedback model suggested by Hattie and Timperley (2007).

Students' Appreciation of the Flipped Classroom

Students had a very positive general impression of the flipped classroom course, but still several had critiques regarding certain aspects, mostly concerning the workload it implied. It is interesting to note that Yeung and O'Malley (2014), in their study of flipped classroom in physical chemistry, found that the principal advantage of the flipped classroom as reported by students was the flexibility this format offered to students, and that overall, students were less satisfied with the flipped classroom as with traditional teaching. Conversely, our students reported several positive aspects of the flipped classroom, such as the ease of receiving answers to their questions and the convenience of receiving regular formative assessment through the portfolio. This observation points toward the fact that it is not a single aspect of the flipped classroom (such as the flexibility it offers) that might be sufficient for students to develop an overall positive perception of this pedagogical approach. Our students, when asked if they liked the flipped classroom, considered all the aspects of the approach we set in place to form their opinion. This points toward our understanding of the flipped classroom as not merely a mode of distance education. Its most distinguishing feature, the videos, is not its most important aspect. Rather, it is an entire pedagogical approach developed around the ideal of the most effective use of in-class and homework time.

Some students reported not liking the flipped classroom, that is, 17% of students answered that they somewhat disliked it (12 out of 88) or very much disliked it (3 out of 88). This result cannot be related to student grades because of the anonymity of the appreciation questionnaire. However, results from other studies can shed light on this observation. McNally et al. (2016) classified students in a flipped classroom as either "flipped endorsers" or "flipped resisters." They reported that "although differences were found between those who endorse and those who resist flipped teaching environments (particularly in their expectations of higher education courses and engagement), this differentiation based on preferences did not correspond to differences in their final grades in a flipped course" (McNally et al., 2016, p. 292). This can explain why we saw such an increase in grades even if 17% of the students did have a rather negative impression of the approach. Since several students reported that the reason for not liking the flipped classroom was the extra workload it necessitates, it is possible that these supplementary hours spent preparing for class would be hours not spent on organic chemistry in a traditional setting.

This can further explain why low-achieving students were the ones benefiting the most from the flipped classroom. Indeed, as reported by Enfield (2013), low achievers are the most likely to report that watching videos outside of class takes too much time, with 42.9% of the bottom-third of their sample mentioned it to be too long, compared with 27% for the entire sample. It might point toward the fact that low-achieving students, who habitually spent less time working on the material at home, are the ones who find the workload heavier than usual and benefit the most from the flipped classroom approach.

Positive Aspects of Teaching in a Flipped Classroom Environment

We decided to try the flipped classroom approach back in 2013 because we saw its potential to free class time by pushing a part of traditional lectures outside of class. To this day, we still view this as the principal advantage of this type of pedagogy. We first thought that this free time could be used to have students practice problem solving while we would be there to help them if need be, contrary to traditional homework. This was exactly what we did, but we did not realize that this free time meant much more than just having students practice in class. Smith describes very aptly what this time is also used for: "much more time was available for explanation, interaction, and conveyance of insight than had been in the past" (Smith, 2013). We now have more time to explain concepts in detail, to present relationships between notions and to provide concrete life examples to increase the relevance of studied topics for students.

Furthermore, we noticed a really significant difference in the time students spent in our office during office hours. Even if this observation is somewhat anecdotal, it is still relevant. When we taught in a traditional setting, we used to receive students during office hours to answer their questions, help them with homework, etc. During a normal week, around five to seven students would come, for a total of 2–3 h of individual consultation each week. By offering more of these interactions in class using the flipped approach, virtually no students come during our office hours anymore. This is a real advantage to all students, since some of them are not comfortable or motivated enough to come talk to their professor outside of class time. Now, all students can ask their questions during class time and benefit from others' questions. This might also be a factor explaining why low-achieving students benefit the most from this approach, since that type of student seldom used to come to office hours.

Downside for Teaching in a Flipped Classroom Environment

The principal disadvantage of teaching in a flipped classroom environment is the enormous amount of time its implementation necessitates, which was also reported by other instructors and researchers (Enfield, 2013; O'Flaherty and Phillips, 2015).

As seen with this study, co-teaching did not influence students' grades. It might then be seen as an investment that is not worth the time needed. Yet this is what allowed the implementation of the flipped classroom. We found early on that working as a team on the design of the pedagogical device, which includes videos but also all the in-class activities, as well as formative assessment of these activities, can alleviate the heavy workload needed.

It should be noted that the investment in time is only necessary during implementation. During the subsequent years, a minimum amount of time was necessary to further the bank of videos since most of them were filmed already, and several office hours were then freed since most students were asking their questions during class time. Office hours were therefore used for formative assessment and improving in-class activities, for example.

Recommendations for Teaching

O'Flaherty and Phillips (2015) deplored, in their review of the literature on flipped classrooms in higher education, that several authors reported on the positive results of the flipped classroom without providing design recommendations for its implementation in small undergraduate classes. Based on the results of this study, we are providing such guidelines (Awad et al., 2017) for educational videos to be used in a flipped classroom environment, with the following main features:

- Keep videos short (6 min).
- Use informal tone of voice, speak enthusiastically.
- Address your students directly to engage them.
- Use signalization (e.g., subtitles) on screen.
- Keep videos simple, avoid complex background or music.
- Provide "pause-solve-resume" features within videos to have students apply their knowledge immediately.

The results of this study also provide the opportunity of suggesting the following recommendations that focus specifically on the need of adequacy between pre-class and in-class activities:

- Every video watched pre-class must be used in an in-class activity.
- Some class time must be reserved to answer students' questions about videos.
- In-class activities must necessitate or encourage collaboration and interaction between students, and between professor and students.
- In-class activities must be devised in a way that allows verification of the completion of pre-class activities.
- Coercive measure to verify completion of pre-class activities might not be necessary with already self-regulated students but probably are with younger/less self-regulated students.
- "Redoing" lectures that were seen in video should be avoided, at the risk of students stop watching videos prior class over time.
- A significant portion of class time should be devoted to active-learning activities (not to lecture).
- In-class activities must be achievable without the express help of the instructor, but the instructor should be available to provide help on demand.
- In-class activities should be the opportunity of giving formative assessment or other form of feedback to help students monitor their progression.
- Faculty should work as a team to implement the flipped classroom, since a lot of time will be necessary.

Other researchers provide relevant recommendations, some of them are noted here:

- Redesign the course to foster active learning, for example, by selecting topics for classroom discussions (Albert and Beatty, 2014).
- Foster students' participation by creating incentives (Albert and Beatty, 2014).
- Explain the flipped classroom model to your students to diminish their resistance (Albert and Beatty, 2014; Estes et al., 2014).
- Flip the entire course (McNally et al., 2016).

The implementation of the flipped classroom as we suggested through this article respects all seven principles for a good undergraduate education as listed by Chickering and Gamson (1987): "1. Encourages contact between students and faculty 2. Develops reciprocity and cooperation among students. 3. Encourages active learning. 4. Gives prompt feedback. 5. Emphasizes time on task. 6. Communicates high expectations. 7. Respects diverse talents and ways of learning" (p. 3).

Recommendations for Future Research

This research considered students' grades and questioned them on their appreciation of the course, but since the questionnaire was anonymously answered, no correlation between grades and appreciation could be measured. Future research could concentrate on elucidating this point, as suggested by O'Flaherty and Phillips (2015): "future research should consider the relationship of other indicators of student engagement in the flipped class (not just examination scores)" (p. 94). Moreover, the effect of the flipped classroom could be evaluated in a true experimental setting, with randomized attribution of students in control and experimental samples. The difficulty of working with a control sample in parallel to a flipped classroom would be the leaking of videos that probably would occur if they were hosted on a public platform such as YouTube. This aspect of design should be considered if such an evaluation would be envisioned.

ETHICS STATEMENT

Academic Dean office of Cégep André-Laurendeau approved the study. Students were informed that their course would be evaluated for effectiveness. Results collected for this study did not include students' identification, but only their *R*-Score, organic chemistry grade, and their sex. Data were provided by the institution's admission service through a list of file numbers. No analysis necessitated students' identification. Appreciation questionnaires were answered anonymously and on a volunteer basis. Students were informed that their answers might be used for publication, but that no information that might identify them would be disclosed.

AUTHOR CONTRIBUTIONS

CC and BV both conducted the research and taught the classes described in the article. They analyzed the data, discussed results, and drew conclusions. They both participated in the creation of the qualitative instrument and in the data collection. CC wrote the paper and BV reviewed and expanded it.

ACKNOWLEDGMENTS

The authors wish to thank Asma Zaghbani, François Arseneault-Hubert, and Marc Amyot for their help with filming some videos and the office of the Academic Dean of Cégep André-Laurendeau for their help to collect the data. They also wish to thank Brett Fischer for his help with the manuscript.

REFERENCES

Abeysekera, L., and Dawson, P. (2015). Motivation and cognitive load in the flipped classroom: definition, rationale and a call for research. *Higher Educ. Res. Dev.* 34, 1–14. doi:10.1080/07294360.2014.934336

Adnan, M. (2017). Perceptions of senior-year ELT students for flipped classroom: a materials development course. *Comput. Assist. Lang. Learn.* 30, 204–222. doi:10.1080/09588221.2017.1301958

Albert, M., and Beatty, B. J. (2014). Flipping the classroom applications to curriculum redesign for an introduction to management course: impact on grades. *J. Educ. Bus.* 89, 419–424. doi:10.1080/08832323.2014.929559

Arnaud, C. H. (2013). Flipping the chemistry classrooms. *Chem. Eng. News* 91, 41–43. doi:10.1021/cen-09112-educ

Awad, E. W., Brouillette, Y., Cormier, C., and Turcotte, V. (2017). *Planifier, réaliser et diffuser des vidéos éducatives: lignes directrices et suggestions à l'intention des enseignants*. Profweb. Available at: http://www.profweb.ca/publications/dossiers/planifier-realiser-et-diffuser-des-videos-educatives-lignes-directrices-et-astuces-pour-les-enseignants

BCI. (2017). *The R Score: What It Is and What It Does*. Montréal, QC: Bureau de coopération internationale. Available at: http://www.bci-qc.ca/wp-content/uploads/2017/09/03-R_Score_what-it-is_what-it-does_BCI-july12-2017.pdf

Bergmann, J., and Sams, A. (2012). *Flip Your Classroom: Reach Every Student in Every Class Every Day*. Eugene, OR: International Society for Technology in Education.

Butt, A. (2014). Student views on the use of a flipped classroom approach: evidence from Australia. *Bus. Educ. Accredit.* 6, 33–43.

Chickering, A. W., and Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *Am. Assoc. Higher Educ. Bull.* 39, 3–7.

Christiansen, M. A. (2014). Inverted teaching: applying a new pedagogy to a university chemistry class. *J. Chem. Educ.* 91, 1845–1850. doi:10.1021/ed400530z

CICIC. (2017). *Ministries/Departments Responsible for Education in Canada*. Available at: https://www.cicic.ca/1301/Ministries-Departments-responsible-for-education-in-Canada/index.canada

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Critz, C., and Wright, D. (2013). Using the flipped classroom in graduate nursing education. *Nurse Educ.* 38, 210–213. doi:10.1097/NNE.0b013e3182a0e56a

Eddy, S. L., and Hogan, K. A. (2014). Getting under the hood: how and for whom does increasing course structure work? *CBE Life Sci. Educ.* 13, 453–468. doi:10.1187/cbe.14-03-0050

Enfield, J. (2013). Looking at the impact of the flipped classroom model of instruction on undergraduate multimedia students at CSUN. *TechTrends* 57, 14–27. doi:10.1007/s11528-013-0698-1

Estes, M., Ingram, R., and Liu, J. C. (2014). A review of flipped classroom research, practice, and technologies. *Int. HETL Rev.* 4. Available at: https://www.hetl.org/a-review-of-flipped-classroom-research-practice-and-technologies/

Fautch, J. M. (2015). The flipped classroom for teaching organic chemistry in small classes: is it effective? *Chem. Educ. Res. Pract.* 16, 179–186. doi:10.1039/c4rp00230j

Flipped Learning Network. (2014). *Definition of Flipped Learning*. Available at: http://www.flippedlearning.org/domain/46

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., et al. (2014). Active learning increases student performance in science, engineering and mathematics. *Proc. Natl. Acad. Sci. U.S.A.* 111, 8410–8415. doi:10.1073/pnas.1319030111

Guo, P. J., Kim, J., and Rubin, R. (2014). "How video affects student engagement: an empirical study of MOOC videos," in *Proceedings of the First ACM Conference on Learning Scale Conference* (New York, NY), 41–50.

Halford, B. (2016). Is there a crisis in organic chemistry education? *Chem. Eng. News* 94, 24–25.

Hattie, J., and Timperley, H. (2007). The power of feedback. *Rev. Educ. Res.* 77, 81–112. doi:10.3102/003465430298487

Horn, M. B. (2013). The transformational potential of flipped classrooms. *Educ. Next* 13, 78–79.

Jensen, J. L., Kummer, T. A., and Godoy, P. D. D. M. (2015). Improvements from a flipped classroom may simply be the fruits of active learning. *CBE Life Sci. Educ.* 14, ar5. doi:10.1187/cbe.14-08-0129

Lage, M. J., Hall, G. J., and Treglia, M. (2000). Inverting the classroom: a gateway to creating an inclusive learning environment. *J. Econ. Educ.* 31, 30–43. doi:10.1080/00220480009596759

Lasry, N., Dugdale, M., and Charles, E. (2014). Just in time to flip your classroom. *Phys. Teach.* 52, 34–37. doi:10.1119/1.4849151

Mason, G. S., Shuman, T. R., and Cook, K. E. (2013). Comparing the effectiveness of an inverted classroom to a traditional classroom in an upper-division engineering course. *IEEE Trans. Educ.* 56, 430–435. doi:10.1109/TE.2013.2249066

McClary, L. M., and Bretz, S. L. (2012). Development and assessment of a diagnostic tool to identify organic chemistry students' alternative conceptions related to acid strength. *Int. J. Sci. Educ.* 34, 2317–2341. doi:10.1080/09500693.2012.684433

McNally, B., Chipperfield, J., Dorsett, P., Del Fabbro, L., and Frommolt, V. (2016). Flipped classroom experiences: student preferences and flip strategy in a higher education context. *Higher Educ. Int. J. Higher Educ. Res.* 73, 281–298. doi:10.1007/s10734-016-0014-z

MELS. (1998). *Science (200.B0) Preuniversity Program, College Education*. Québec, QC: Gouvernement du Québec.

MELS. (2000). *Les prescriptions ministérielles et l'élaboration d'un programme défini en objectifs et standards*. Gouvernement du Québec. Available at: http://collections.banq.qc.ca/ark:/52327/bs41437

Mooring, S. R., Mitchell, C. E., and Burrows, N. L. (2016). Evaluation of a flipped, large-enrollment organic chemistry course on student attitude and achievement. *J. Chem. Educ.* 93, 1972–1983. doi:10.1021/acs.jchemed.6b00367

O'Dwyer, A., and Childs, P. E. (2017). Who says organic chemistry is difficult? Exploring perspectives and perceptions. *EURASIA J. Math. Sci. Technol. Educ.* 13, 3599–3620. doi:10.12973/eurasia.2017.00748a

O'Flaherty, J., and Phillips, C. (2015). The use of flipped classrooms in higher education: a scoping review. *Int. Higher Educ.* 25, 85–95. doi:10.1016/j.iheduc.2015.02.002

Roehling, P. V., Root Luna, L. M., Richie, F. J., and Shaughnessy, J. T. (2017). The benefits, drawbacks, and challenges of using the flipped classroom in an introduction to psychology course. *Teach. Psychol.* 44, 183–192. doi:10.1177/0098628317711282

Rushton, G. T., Hardy, R. C., Gwaltney, K. P., and Lewis, S. E. (2008). Alternative conceptions of organic chemistry topics among fourth year chemistry students. *Chem. Educ. Res. Pract.* 9, 122–130. doi:10.1039/B806228P

Ryan, M. D., and Reid, S. (2016). Impact of the flipped classroom on student performance and retention: a parallel controlled study in general chemistry. *J. Chem. Educ.* 93, 13–23. doi:10.1021/acs.jchemed.5b00717

Smith, J. D. (2013). Student attitudes toward flipping the general chemistry classroom. *Chem. Educ. Res. Pract.* 14, 607–614. doi:10.1039/C3RP00083D

Strayer, J. F. (2012). How learning in an inverted classroom influences cooperation, innovation and task orientation. *Learn. Environ. Res.* 15, 171–193. doi:10.1007/s10984-012-9108-4

Sweet, D. (2014). Microlectures in a flipped classroom: application, creation and resources. *Mid West. Educ. Res.* 26, 52–59.

Voisard, B., and Cormier, C. (2013). *Chimie organique*. Anjou, QC: Les éditions CEC.

Yeung, K., and O'Malley, P. J. (2014). Making "the flip" work: barriers to and implementation strategies for introducing flipped teaching methods into traditional higher education courses. *New Dir. Instit. Res.* 10, 59–63. doi:10.11120/ndir.2014.00024

Young, T., Bailey, C. T., Guptil, M., Thorp, A., and Thomas, T. (2015). The flipped classroom: a modality for mixed asynchronous and synchronous learning in a residency program. *West. J. Emerg. Med.* 15, 939–944. doi:10.5811/westjem.2014.10.23515

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in ICT

ORIGINAL RESEARCH
published: 24 August 2018
doi: 10.3389/fict.2018.00022

Boundary Crossings Resulting in Active Learning in Preservice Teacher Education: A CHAT Analysis Revealing the Tensions and Springboards Between Partners

*Thérèse Laferrière**

Studies on Teaching and Learning, Laval University, Québec, QC, Canada

We present a meta-analysis of a twenty-year long university-school partnership in which pre-service teachers collaborated with cooperative teachers and peers during practicums in innovative programs that featured active learning. The partnership evolved as a design experiment. Papers presented at conferences but never submitted to a research journal were revisited applying cultural-historical activity theory (CHAT) to understand the dynamics at play, and especially boundary crossing, within the university-school partnership's activity in terms of motive/object, instruments, community, roles, and rules/policies. We point to tensions that manifested contradictions of different levels between activity systems as the innovation unfolded. Suggestions for boundary crossing when field experiences are part of an undergraduate program are made.

Keywords: innovation, cultural-historical activity theory (CHAT), university-school partnership, online collaborative platforms, authentic problem, collaborative reflective practice, knowledge building, virtual community

OPEN ACCESS

Edited by:
Elizabeth S. Charles,
Dawson College, Canada

Reviewed by:
Cristina Prieto,
Universidad de Salamanca, Spain
Frank De Jong,
Aeres University of Applied Science, Netherlands

Correspondence:
Thérèse Laferrière
Therese.Laferriere@fse.ulaval.ca

Specialty section:
This article was submitted to Digital Education, a section of the journal Frontiers in ICT

Received: *12 December 2017*
Accepted: *02 August 2018*
Published: *24 August 2018*

Citation:
Laferrière T (2018) Boundary Crossings Resulting in Active Learning in Preservice Teacher Education: A CHAT Analysis Revealing the Tensions and Springboards Between Partners. Front. ICT 5:22. doi: 10.3389/fict.2018.00022

INTRODUCTION

Teaching and learning in the digital era are taking many forms and shapes. Our own journey began over 20 years ago as we engaged in the exploration of the possibilities of the Internet to support university-school partnerships dedicated to active learning. A process of co-design began, first among a few teacher educators engaged in a national research network, each attempting to locally develop a university-school partnership centered on ICT integration. At one of the sites, researchers invited a school district superintendent to invest in a school that could become a lighthouse for other schools regarding the uses of information and communication technologies (ICTs). After conducting a need assessment with families, a large urban secondary school introduced a one-to-one laptop program that emphasized project-based learning. Researchers engaged in collaborative action research with school practitioners on ICT integration and effective use, selected and supervised student teachers interested in doing a 5-week or a 15-week practicum in the program.

Wanting student teachers to reflect on their teaching practice in a collaborative manner, teacher educators designed a virtual community of support and communication using two web-based platforms (Virtual-U's VGroups and, later, Knowledge Forum). Successive cohorts of students became virtually linked to one another as incoming cohorts accessed the contributions of previous ones and added their own contributions. Contributions were the results of onsite/online interaction for understanding a pedagogical problem that captured their interest.

This successful case of active learning is reported through an analysis of the university and school partners' activity systems (motive/object, instruments, community, roles, and policies). The first sections provide contextual and conceptual background and the methodology that led to successful use of ICTs in this higher-education case. The latter sections present a meta-analysis of previous research papers related to this case with the goal of identifying boundary crossings and resulting tensions and springboards between partners' activity systems. Suggestions are made for boundary crossing when fieldwork, as a form of active learning, is part of undergraduate programs.

BACKGROUND

In the mid-nineties, university-based teacher educators themselves had to uncover what could be the added value of information and communication technologies (ICTs) for teaching and learning. Even today, the challenge is still there: digital technologies develop rapidly; meta-analyses emphasize that pedagogy remains the critical factor (Tamim et al., 2011; Means et al., 2013); implementation factors such as training and support need to be considered when assessing the effectiveness of ICT interventions (Archer et al., 2014); and "learning is best supported when the student is engaged in active, meaningful exercises via technological tools that provide cognitive support" (Schmid et al., 2014, p. 285).

At the onset of our own use of ICTs, the assumption was that they could enable active learning, one of Chickering and Gamson (1987) widely accepted seven principles for improving practice in undergraduate education. Bracewell et al. (1998) revisited Schwab (1973) four commonplaces of the educational situation—someone teaching something to someone in a given context—in the following terms: A learner learning something, under the guidance of a teacher, in a given context. This reformulation acknowledged the control given to the learner in a context, especially one supportive of onsite/online human interaction. Dirckinck-Holmfeld and Sorensen (1999) stressed the importance of viewing collaborative learning as a holistic process that is taking place in a context—a community of practice. At the time, computer-supported collaborative learning (CSCL) was still in its infancy (Roschelle, 1992; Koschmann et al., 1994).

In teacher education, Schön (1983) book on the reflective practitioner, building on Dewey (1925/1989, 1934/1989, 1938) understanding of experience and reflection, was highly influential (Baird, 1992; Zeichner and Liston, 1996; Boud and Walker, 1998). Schön distinguished reflection-in-action from reflection-on-action. He defined the former as "a reflective conversation with the situation" (p.163), and referred to reflection-on-action as an activity occurring before or after practice. Kolb (1984) model of experiential learning also emphasized reflection on experience. Co-designing a virtual community of support and communication with pre-service teachers[1], we were encouraged by our early results, and found ourselves in agreement with Blanton et al. (1998) who suggested the adoption of a socio-constructivist pedagogical framework to guide teacher educators in the use of telecommunications. A few years earlier, collaborative reflective practice on ill-defined problems for science teachers had been suggested as best practice by Desouza (1994).

At our university, the education of reflective practitioners was the primary aim of the four-year teacher education program. Collaborative reflective practice with school-based teachers engaged in innovative practice with ICTs was appearing most desirable. The working hypothesis put forward by our research team was that reflection on practice and knowledge building, supported by an online collaborative platform, could be highly relevant for the education of pre-service teachers doing practicums in the partner school, and as incoming practitioners of teaching and learning in the digital age.

The partner school was in the early stages of integrating ICTs in teaching and learning. Pedagogies such as cooperative learning and project-based learning were part of teacher professional development activities offered by the school district, and some teachers were doing their Master's Degree at our university. The school adopted an incremental approach, admitting the first year 60 students that were beginning secondary school. The administration hired two elementary school teachers, considering that their teaching practices were more attuned to the classroom processes they wanted to see being installed, namely ICT use, teamwork and project-based learning. On Year two, there were four classrooms instead. In 2002, the program was spreading over the 5 years of high school, and has since proven to be sustainable.

METHODOLOGY

We engaged in a design experiment, a methodology developed to create and evaluate educational innovations (Brown, 1992; Collins, 1992). Researchers adopting this methodology give to intervention special attention, and several research iterations are usual (design-based research, Collins et al., 2004; Zheng, 2015). We were also influenced by Engeström (1987, 2011) activity theory framework and formative interventions as they are, like design-based research, especially suitable when innovation is concerned. Engeström's framework is used to focus on tensions/contradictions between an activity system's main elements or between activity systems. It served to collect manifestations of tensions as data. It is through the identification of tensions/contradictions, and their resolution, that innovation occurs within the activity of a group or a community. We present here the basic constituents of the intervention conducted over the years by the author of this paper who was the pre-service teachers' supervisor during their practicums. She was also a researcher on ICT integration in teaching and learning.

Intervention

Participants (or primary activity systems' actors) involved preservice students, cooperative teachers, and teacher

[1] A R&D project of the TeleLearning Network of Centres of Excellence (TL-NCE, Canada, 1995-2002).

educator(s)/supervisor(s). The school district administrators and personnel, the school principal and other school teachers, and parents also formed other activity systems in interaction with the primary activity systems considered in this paper.

Volunteer Participation

A pre-service teacher (PST) cohort doing a practicum in one-to-one laptop classrooms (OLC) was composed of five to eight participants (PST-OLC). This option attracted more volunteer students than available places. Selection interviews were conducted for matching pre-service teachers with cooperative teachers. What a practicum in one-to-one laptop classroom entails (teamwork, self- and peer-regulated learning, collaborative project-based learning or inquiry, and, sometimes, knowledge building) is hereafter presented.

Student Engagement With Authentic Problems

For authentic problems to lead to socio-cognitive knowledge, the learning environment must be designed to this end (Bransford et al., 1999; Savery and Duffy, 2001). Being a pre-service teacher in a one-to-one laptop classroom had, and still has, its load of challenges (e.g., understanding the curriculum in depth; teaming up with the cooperative teacher; knowing less than classroom students about software in use; moving from a teacher-centered to a student-centered approach and to a learning community model regarding participation in the classroom; managing students' use of laptops during lectures).

Collaborative Reflective Practice

Miettinen (2000) wrote: "It is the failure and uncertainty of the primary experience that gives rise to reflective thought and learning" (p.65). Shireen-Desouza and Czerniak (2003) defined collaborative reflective practice as follows: "A voluntary effort of the part of teachers in a school to share and critique idea about teaching, to reflect upon one's teaching and students' learning, formulate aims and goals about the curriculum through collaboration, and also take responsibility for their actions and consequences of their actions" (p. 77). Yoon and Kim (2010) showed the advantage of collaborative reflection to enhance individual reflection. For a PST-OLC, entries in an individual journal for reflective practice were replaced by contributions in an online forum. As genuine engagement was sought, there was no requirement for posting a specific number of contributions per week. Though participation in the forum was mandatory, a pre-service teacher had the option of opting out during the trimester. Over the years only one of them, who was encountering serious difficulties, chose to use a journal for reflective practice.

Focus on Ill-Defined Problems

The university-based teacher educator guided the PST-OLCs toward identifying one or a few practical problems for which there was no simple or clear definition or solution. They were invited to collaborate for reaching a better collective understanding of the problem and also for co-influencing their individual teaching practices. The university-based teacher educator, and also some school-based teacher educators, provided references, cases, and other forms of advice.

Seamless Onsite/Online Interaction

Given that the pre-service teachers of the PST-OLCs were all doing their practicums in the same school, they had their own room for individual work or exchange with others. Cooperative teachers were sometimes present in this room but pre-service teachers often met with them elsewhere. Seminars with the university-based teacher educator were conducted in that same room. At times, a teacher attended. There were, therefore, plenty of opportunities onsite to engage conversation on problems of practice. Online interaction was encouraged for leaving traces of one's thinking and building on one another's thinking on problems first discussed onsite. Cooperative teachers had access to the forum on the web-based platform.

Collaborative Knowledge Building

Scardamalia (2002) knowledge-building principles (e.g., real ideas and authentic problems; improvable ideas; collective cognitive responsibility for a community's advancement of knowledge) were highly relevant. Each PST-OLC was called to become a knowledge building community, and to leave the results of their collaborative inquiry on the platform. For the 2002-2012 period, such results were available in the form of a virtual tour, developed by one or two participant(s) who were then hired as research assistants, and required to seek validation of the tour from other pre-service teachers before posting it online.

Applying Wenger (1998) concepts of shared repertoire and regime of competence, pre-service teachers' learning and knowledge building artifacts were to contribute to the conceptualization of the teaching practice in a networked classroom. As an exercise of legitimate peripheral participation (Lave and Wenger's, 1991), incoming cohorts had to do an online practicum, that is, the reading/visioning of three virtual tours and/or, for the years 2013–2016, of the PST-OLCs' forum contributions themselves, and they had to write an individual reflective statement on the value of such an activity prior to their practicum in a one-to-one laptop classroom. During the trimester, they could search the platform, using keywords, for previous contributions made by participants of previous cohorts on a problem they were collaboratively reflecting upon in an attempt to advance their individual and collective knowledge and practice as well as the knowledge and practice of the virtual community as a whole.

The second design cycle (2013-2016) that replaced the first one (2002–2012) grew out of necessity given the fact that the new version of the online platform (Knowledge Forum) did not include the contents of the previous database. Therefore, the 2012 PST cohort and the 2013 PST cohort had to switch platforms to do the three required virtual tours. When the supervisor asked the 2013 PST cohort permission to make all their contributions accessible to future cohorts instead of only those part of a virtual tour, they accepted. No more virtual tours were developed.

Research

Socio-technical designs[2] for effective uses of ICTs in teaching and learning were at the heart of our research program, and especially those integrating collaborative platforms. Our own use was enhanced through a number of research iterations. In the first iteration, the Internet was used for bridging university and school practices in teacher education, and patterns of connection were identified (information exchange, coordination of teaching practices, and joint inquiry) (Laferrière et al., 1997). In the second iteration, the notion of a networked community helped integrate the connections that were taking place between the university and the school (Laferrière et al., 1998). In the third iteration, the research narrowed on the activity of the networked community of learners, meaning the online interaction between pre-service teachers (Collins et al., 2000). In iteration four, the research effort expanded to document the connections between networked communities—the university-school partnerships were inspired by the Holmes Group's (1990) professional development school model (PDS)[3] that emphasized (1) practice teaching, (2) professional development, and (3) collaborative research (Laferrière, 2001; Breuleux and Laferrière, 2004). Researchers also studied pre-service teachers' online discourse with regard to content and process: project-based learning (Laferrière et al., 2002); argumentation procedures (Campos et al., 2003); teaching and learning in a networked classroom (Laferrière et al., 2013, 2016).

For this research work, we applied Engeström's cultural-historical activity theory framework Engeström (1987, 2015) to the papers[4] mentioned in the preceding paragraph for conducting, in an illustrative manner, a meta-analysis[5] of the university and school partners' activity systems' components: motive/object, tools/instruments, community, roles, and rules/policies are examined. For innovation to occur, two activity systems must minimally compose the unit of analysis (Engeström, 2001). Most enduring tensions within and between activity systems' constituents and those created by emerging activity systems are pinpointed. Such tensions manifested more basic contradictions at different levels:

- Level 1: contradiction within the same component of an activity system (L1c)
- Level 2: contradiction between components of an activity system (L2c)
- Level 3: contradiction between an established and an emerging activity system (L3c)
- Level 4: contradiction between the new activity system and its neighboring activity systems (L4c)

For each tension identified, the level of contradiction it could reflect is indicated (L1c, L2c, L3c, and L4c).

For activity systems to evolve, boundary crossing reflected in moving beyond traditional roles and in the co-construction and adoption of new models is key (Akkerman and Bakker, 2011). This case study is attentive to such moves.

RESULTS

Partners' Shared Object: Innovation With ICTS

Since the start (1995), the motive of the founding and sustaining partners' (dean, school superintendent, school principal, one-to-one laptop program (OLP) teachers (including cooperative teachers), university-based teacher educators, pre-service teachers, one-to-one laptop program (OLP) learners, and parents) activity had been and remained innovation in teaching and learning with ICTs (**Figure 1**). For the school system, innovation was primarily pursued through the goal of initiating, developing, and sustaining the one-to-one laptop program. For the university, innovation focused on the preparation of pre-service teachers for work in the digital age through PST-OLCs' reflective practice and knowledge building with the support of a collaborative platform. In 2002, this university-school partnership had shrunk to one university-based teacher educator (Tension, L2c) but the one-to-one laptop program had grown, applying a school-within-a-school model, and was offered to all grade levels (Secondary 1 to Secondary 5). In the most recent years, a former member of two PST-OLCs, who is now a professor at the Faculty of Education, is introducing reflective practice and knowledge building on a collaborative platform for all pre-service students preparing to teach at the elementary level.

Partners' Tools and Instruments

Both university-based and school-based partners accessed the Internet at a high level for information and communication purposes, for teaching and for active learning (e.g., the design of practicums for pre-service teachers and the design of a one-to-one laptop program for school students). (**Figure 1**). Advanced collaborative platforms for active learning were the choice of the university-based teacher educators involved in the partnership, while teachers and school learners of the one-to-one laptop program were interested in software diversity and valuing open access ones (Tension, L4c).

Partners' Communities

Each Fall or Winter trimester, at the core of the university-based community was the PST-OLC (pre-service teacher cohort) and the university-based teacher educator involved in the OLP. On campus, they interacted with university peers and colleagues (**Figure 1**), explaining what the OLP was about and its *raison*

[2]Socio-technical design is a concept borrowed from the Tavistock Institute for Human Relations in London that goes back to the '40s.

[3]This was the Holmes Group's strategy for fostering innovation within pre-service teacher education programs as well as within local schools. PDSs caught the attention of the National Council for Accreditation of Teacher Education (NCATE, Washington).

[4]Studies connected to these papers were carried out in accordance with the recommendations of the Social Sciences and Humanities Research Council, SSHRC, Canada. Protocols were approved by Laval University's Ethics Committee.

[5]This meta-analysis focuses on the author herself who reflected on her own experience as she revisited those previous papers to which she had contributed. She is grateful to two university colleagues, Stephane Allaire and Christine Hamel, who previously worked in the partner school as students and research assistants. They validated the analysis.

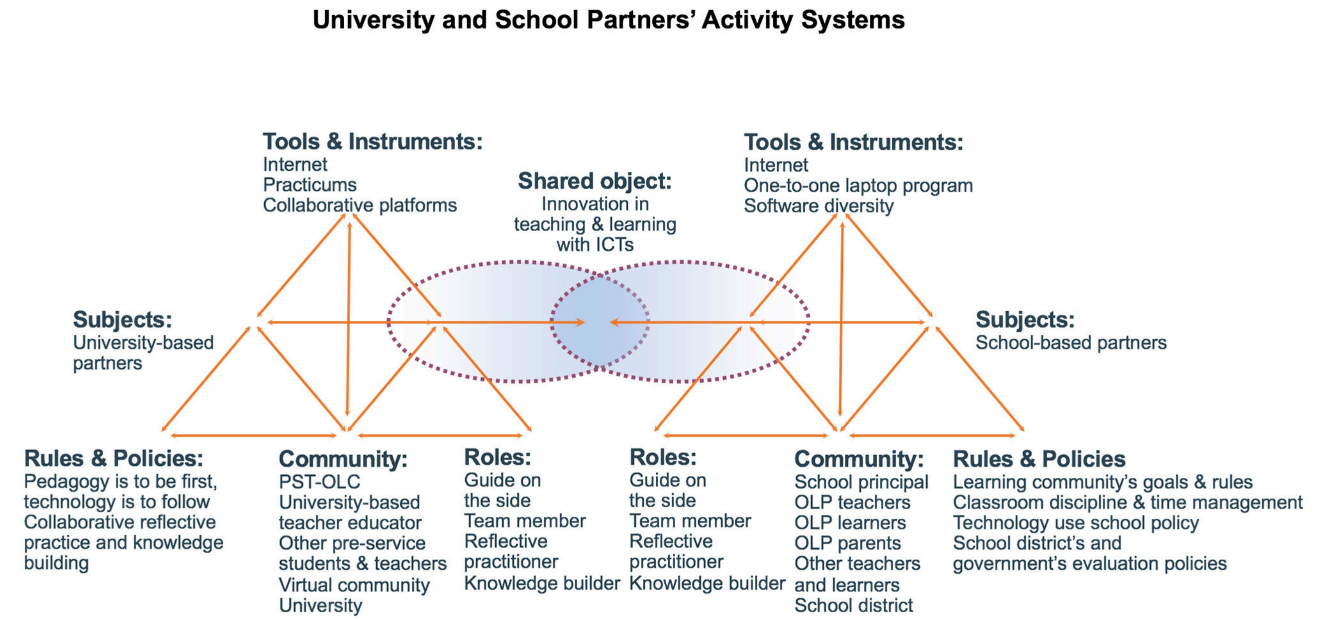

FIGURE 1 | University and School Partners' Activity Systems represents two interconnected systems of activity and the complex dynamics within and between each of them.

d'être (Tension, L1c). OLP teachers had to do the same within their own school-based community, with the help of the school principal (Tension, L1c). Meanwhile, the repertoire of the virtual community of support and communication, composed of all the contributions of previous PST-OLCs, including those of the university-based teacher educator and of some OLP teachers, was underused during and after the practicums. Onsite and online communications with OLP parents, who had chosen to register their child (children) to the OLP and bought his/her laptop, were frequent.

Partners' Roles

The university-based teacher educator, the OLP teachers and the pre-service teachers all had to learn to be "guides on the side" (**Figure 1**). The university-based teacher educator's requirement that pre-service teachers' write personal learning projects, ahead of the practicum but after three or four visits in an OLP classroom, generated insecurity (Tension, L3c). They were instructed to refer to the template provided by the Faculty of Education's Placement Office only when getting short of ideas regarding the planning of their practicum. OLP teachers working with pre-service teachers were often present in the classroom compared to other cooperative teachers. They also favored teamwork more often (**Figure 1**). Moreover, they were learning, and letting pre-service teachers know it, when to instruct and when to give OLP learners control of their use of time when working individually or in teams (Tension, L1c). They liked the learning community model but often found themselves having to take central stage in the classroom (Tension, L1c). A few of them used Knowledge Forum, and considered the students of their classroom as knowledge builders. They worked in teams with other teachers and engaged in collaborative reflective practice and knowledge building although they did not use a collaborative platform. They published individual webpages (Tension L1c). Pre-service teachers were welcomed at all teacher meetings. Having little conceptual and experiential knowledge of active learning and lacking deep understanding of the curriculum, pre-service teachers had a lot to learn. For instance, in the classroom, they leaned toward teacher-centered project-based learning, giving students the freedom to choose the "how" and, rarely, the "what" to be studied (**Figure 1**). They struggled with aligning the curriculum goals, pedagogical intents, and results (Tension, L2c). Nonetheless, some guided classroom students in the use of Knowledge Forum even when students tended to think that other software tools were "cooler" (Tension, L2c). On the whole, pre-service teachers found ways to contribute to the conceptualization of teaching in a networked classroom, that is, when all own a laptop connected to the Internet. Almost half of these pre-service teachers are now OLP teachers.

Partners' Rules and Policies

Pre-service teachers were advised by outsiders (university teachers and peers, and family members with teaching experience) to the one-to-one laptop community (OLC), to spell out, as they introduce themselves to a classroom, the rules they wanted to apply. That was contrary to the thinking of the OLP teachers and the university-based teacher educator who were favoring the learning community model (**Figure 1**): learning goals were to be established with the classroom, and rules were to derive from them (L4c). Pre-service teachers did not want to lose control of the classroom, an implicit rule they perceived was important (Tension, L1c). For instance, they did not want classroom students to break the school policy with regard to the use of the computer (Tension, L2c). Being in touch with what was going on in the classroom, including on screens, while scaffolding a student or a small group of students, was expected of them (Tension, L2c). Working individually or in groups, classroom students were not always on-task and, sometimes, disturbed others. Pre-service teachers had to act. Another difficulty regarded learning assessment. At the beginning of the OLP, the school district had loosened up its evaluation policies but over time they tightened them up (Tension, L4c). At the government level, shortly after recommending the OLP as an exemplary case regarding learning assessment practices to the Organization for Economic Co-operation and Development (OECD, 2005), less emphasis was put on the acquisition of competencies, and OLP teachers and pre-service teachers felt the pressure of assessing rote knowledge in preparation of provincial exams (Tension, L4c).

Pre-service teachers also had to meet the expectations of the university-based teacher educator with regard to giving attention to pedagogy first and technology after (Tension, L2c), and engaging in collaborative reflective practice (onsite/online) and knowledge building (especially online) (**Figure 1**). To construct and maintain a joint problem space (Roschelle and Teasley, 1995; Fischer et al., 2013) was not easy for each PST-OLC, and for the university-based teacher educator as well (Tension, L3c). Pedagogical concepts such as socio-cognitive conflict and positive interdependence required deeper understanding. Ill-defined problems were for instance: How to interact with classroom students in ways that will allow for an authentic question to arise and engage them into a collaborative inquiry? Which technology would best support this or that learning activity? When to release students' agency, and for how long? How to organize and manage a networked classroom? A PST-OLC could search the collaborative platform and refer to the contributions of previous PST-OLCs having work on the same or a similar problem but such an action was not mandatory.

DISCUSSION

Under the lens of cultural-historical activity theory (CHAT), which serves as the theoretical underpinnings of this case study that grew out of a meta-analysis of previous unpublished papers, one gets a systemic view of what innovation in the classroom, supported by ICTs, entails, and especially when active learning is on the agenda. Active learning was enacted through reflective practice and knowledge building supported by a collaborative platform. The ill-defined problems that pre-service teachers struggled with when learning to teach in one-to-one laptop classrooms were brought forth during collaborative reflective practice and knowledge building. In her work with the PST-OLCs, the university-based teacher educator experienced the same problems being pinpointed (e.g., how to be a guide on the side; how to engage students in project-based learning; how to exercise control). While each PST-OLC had access to the repertoire of the virtual community, they nonetheless needed to engage in their own meaning negotiation over such problems as a way to face the internal L1c and L2c contradictions they were experiencing.

These ill-defined problems do not appear to be that different also from the ones that post-secondary teachers face, inside and outside the classroom, when engaging students in active learning. For instance, student engagement into active learning require that they venture into a more active role, and some resist such role modification (Parent, 2017). When this happens, the teacher's emerging activity system enters in contradiction with the student's well-established activity system (L3c). At such a time, the partners (teacher and student) need to find a shared object in order to move forward.

In spite of the fact that with the school's partners activity system advanced collaborative platforms were not very popular, pre-service teachers were presented Virtual-U's VGroups and, later, Knowledge Forum for collaborative reflective practice and knowledge building. There was an obvious lack of coherence between the two activity systems but OLP teachers and the university-based teacher educator respected one another's boundaries, and accepted this L4c contradiction. In the end, only a few teachers and pre-service teachers had referred to the knowledge-building principles and made use of Knowledge Forum. It may be inferred that the use of similar instruments would have deepened pre-service teachers' experience with the same instruments, and, therefore, their use for active learning purposes.

The university-based and the school-based partners belonged to different communities, each with its beliefs and ways of thinking and doing. The experiential approach that led to sending pre-service teachers to emerging one-to-one laptop classrooms, and favored the use of advanced collaborative platforms went against the grain of the mainstream activity of the Faculty of Education, and, introduced, therefore, another L3c contradiction. While active learning was voiced, only a few professors enacted it with undergraduate students (L1c). An even smaller number showed interest in advanced collaborative platforms (L4c). Similarly, most pre-service teachers seemed to underestimate the value of active learning (L1c). But not the parents of the OLP learners (L4c). In a few words, the emerging activity system was installing a contradiction between the old and the new (L3c). The CHAT framework emphasizing that the resolution of tensions/contradictions leads to innovation, one gets a sense here of the boundary crossing that was required from the university system's actors. This activity system is more complex than the activity system of a school. Even when a school decides to implement a school within-a-school model, which adds to the complexity of its activity, the emerging activity system kept expanding (e.g., the number of teachers involved) while the emerging university-based activity system did not expand beyond one PST-OLC engaging in reflecting practice and knowledge building with the support of a collaborative platform during the Fall and Winter trimesters.

Being a guide on the side is more of a self-effacing role than being the sage on the stage, and requires a capacity to face the unknown as students take more active roles (e.g., generating questions and problems, searching for information, engaging in project-based learning, collaborative inquiries, and in knowledge building). It may not be what prospective teachers have in mind when choosing this profession, and, if so, their expectations are in contradiction with the expectations for life and work in the digital age (Pellegrino and Hilton, 2012). Teaching beliefs and educational systems in place, including university professors'/lecturers' and students' expectations of their role, are key factors to work with for innovation and change in education, and these raise L1c, L2c, L3c, and L4c contradictions that will need to be overcome. The task will not be easy given that teachers' and students' roles become more complex than conventional ones when active learning is enacted. Technology seems to add to, rather than diminish, this complexity.

Learning to release students' agency without losing control, to negotiate behavioral rules with students that will allow for the learning objectives to be met, to scaffold student learning, and to proceed fairly in assessing individual and group learning are requirements of an active learning pedagogical approach. It requires boundary crossing within the university activity system and between university and school activity systems. Students of each of these activity systems also are facing a steep learning curve as they are required to exercise agency when they operate in less scripted learning environments, negotiate their different representations of an ill-defined problem and seek knowledge and action convergence with their peers. As pointed by Dede (2017), students must be prepared to reinvent themselves. Will these emerging practices transform into new rules and policies at the institutional level? Applying cultural-historical activity theory (CHAT), one may foresee that such an emerging activity system is bound to bring more tensions/contradictions between the old and the new ways of being a teacher and a student in post-secondary education. CHAT has a methodology for interested administrators and teachers to address such contradictions and bring about, in an informed and consensual manner, effective models, namely the Change Laboratory (Engeström, 1987, 2015; Virkkunen and Newnham, 2013).

CONCLUSION

We presented a case of active learning that stands out by its duration, and its systemic nature. It featured pre-service teachers learning to teach in networked classrooms with their cooperative teachers and university-based teacher educators who fostered their active learning by using, among others, collaborative platforms to support reflective practice and knowledge building. CHAT was used to provide a sense of the dynamics at play in such innovation. However, this study has limits with regards to the way CHAT was used for analytical purposes. For instance, many units of analysis, each involving two different activity systems with their respective subjects who participated in the university-school partnership, could have been analyzed. Contradictions, as manifested by identified tensions, could have been understood at a much deeper level with a fuller application of the theory and the Change Laboratory as its related methodology.

Nonetheless, the results illustrate what is at stake when post-secondary teachers venture into engaging students in active learning. In this case, it was done through reflective practice and knowledge building using a collaborative platform. It is our way to prepare pre-service teachers for teaching and learning in the digital era, and to work with students that will have to demonstrate future skills that still remain to be completely uncovered.

Given the breadth and length of this innovation that fostered active learning, we formulate four suggestions, and the contradiction level (L1c, L2c, L3c, L4c) they address, for the boundary crossing of one's activity system when field experiences or practicums are part of an undergraduate program:

- A student who wants to evolve and thrive in the digital era will find him-herself advantaged by registering for elective courses or programs that promote active learning through the use of digital tools and resources, and, among others, collaborative platforms (L1c).
- A post-secondary teacher who wants to engage students in active learning will find him-herself advantaged by taking the role of a designer, or of a design researcher, proceeding through iterative cycles by collecting data that will inform his or her practice (L2c).
- A post-secondary teacher who wants to engage students in authentic problem setting and solving will find him-herself advantaged by being part of a partnership where both partners have agreed on a shared object toward which to direct their respective activity forward (3c).
- A post-secondary institution who wants to contribute at most advanced levels at cultural, societal and economic levels will find itself at advantage by spelling out to prospective students that active learning is expected of them (L4c).

CONSENT PROCEDURE

University students were informed that the innovation they were part of was part of a research program. Participation was on a voluntary basis. University students read and signed the consent form.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and approved it for publication.

FUNDING

This paper was financed by successive SSHRC grants (Social Sciences and Humanities Research Council, Canada).

ACKNOWLEDGMENTS

The author would like to thank Alain Breuleux and Robert Bracewell (McGill University) and Gaalen Erickson (University of British Columbia) for their collaboration in the early years of this endeavor as well as all preservice teachers who participated in this endeavor and cooperative teachers.

REFERENCES

Akkerman, S. F., and Bakker, A. (2011). Boundary crossing and boundary objects. *Rev. Educ. Res.* 81, 132–169. doi: 10.3102/0034654311404435

Archer, K., Savage, R., Sanghera-Sidhu, S., Wood, E. E., Gottardo, A., and Chen, V. (2014). Examining the effectiveness of technology use in classrooms: a tertiary meta-analysis. *Comput. Educ.* 78, 140–149. doi: 10.1016/j.compedu.2014.06.001

Baird, J. R. (1992). "Collaborative reflection, systematic inquiry, better teaching," in *Teachers and Teaching from Classroom to Reflection,* eds T. Russell, and H. Munby (Bristol: The Falmer Press), 33–48.

Blanton, W. E., Gary Moorman, G., and Woodrow, T. (1998). Telecommunications and teacher education: a social constructivist review. *Rev. Res. Educ.* 23, 235–275. doi: 10.3102/0091732X023001235

Boud, D., and Walker, D. (1998). Promoting reflection in professional courses: the challenge of context. *Stud. High. Educ.* 23, 191–297. doi: 10.1080/03075079812331380384

Bracewell, R. J., Breuleux, A., Laferrière, T., Benoit, J., and Abdous, M. (1998). *The Emerging Contribution Of Online Resources And Tools To Classroom Learning And Teaching.* Available online at: http://www.tact.fse.ulaval.ca/ang/html/review98.html

Bransford, J. D., Brown, A. L., and Cocking, R. R. (1999). *How People Learn: Brain, Mind, Experience, and School.* Washington, DC: National Academy Press.

Breuleux, A., and Laferrière, T. (2004). "University-school partnerships: online ethnography," in *Paper Presented at the Annual Meeting of the American Educational Research Association (AERA),* (San Diego, CA).

Brown, A. L. (1992). Design experiments: theoretical and methodological challenges in creating complex interventions in classroom settings. *J. Learn. Sci.* 2, 141–178. doi: 10.1207/s15327809jls0202_2

Campos, M., Laferrière, T., Messas, and Allaire. (2003). "Analysing argumentation procedures of electronic conference transcripts of student teachers: a conceptual tool," in *Paper Presented at the Annual Meeting of the American Educational Research Association (AERA),* (Chicago, IL).

Chickering, A. W., and Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *Am. Assoc. High. Educ. Bull.* 39, 3–7.

Collins, A. (1992). "Toward a design science of education," in *New Directions in Educational Technology*, eds E. Scanlon and T. O'Shea (Berlin: Springer-Verlag), 15–22.

Collins, A., Joseph, D., and Bielaczyc, K. (2004). Design research: theoretical and methodological issues. *J. Learn. Sci.* 13, 15–42. doi: 10.1207/s15327809jls1301_2

Collins, A., Laferrière, T., Grimmett, P., and Cole, A. (2000). "Networked communities of learners: the TL∗NCE teacher education approach to teacher education reform," in *Paper Presented at the Symposium Innovations in Teacher Education Programs in Canada. Canadian Society for the Study of Education's Special Session at the Annual Meeting of the American Educational Research Association (AERA),* (New Orleans, LA).

Dede, C. (2017). Students must be prepared to reinvent themselves. Commentary. Ed Week, December 11th (2017).

Desouza, S. (1994). "Do science teachers intend to engage in collaborative reflective practice?" in *Paper Presented at the Annual Meeting of the National Association for Research in Science Teaching* (Anaheim, CA).

Dewey, J. (1925/1989). Experience and nature: The Later Works of John Dewey. Vol.1 Edited by Jo Ann Boydston: Carbondale & Edwardsville: Southern Illinois University Press.

Dewey, J. (1934/1989). *Art as experience. Vol. 10. Edited by Jo Ann Boydston*. Carbondale & Edwardsville; Southern Illinois University Press.

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

Dirckinck-Holmfeld, L., and Sorensen, E. K. (1999). "Distributed computer supported collaborative learning through shared practice and social participation," in *Proceedings of Computer Support for Collaborative Learning (CSCL)* (Stanford, CA), 136–141.

Engeström, Y. (1987). *Learning by Expanding: An Activity-Theoretical Approach to Developmental Research*. Helsinki: Orienta-Konsultit.

Engeström, Y. (2001). Expansive learning at work: toward an activity theoretical reconceptualization. *J. Educ. Work* 14, 133–156. doi: 10.1080/13639080020028747

Engeström, Y. (2011). From design experiments to formative interventions. *Theory Psychol.* 21, 598–628. doi: 10.1177/0959354311419252

Engeström, Y. (2015). *Learning by Expanding: An Activity-Theoretical Approach to Developmental Research. 2nd Edn.* New York, NY: Cambridge University Press.

Fischer, F., Kollar, I., Stegmann, K., and Wecker, C. (2013). Toward a script theory of guidance in computer-supported collaborative learning. *Educ. Psychol.* 48, 56–66. doi: 10.1080/00461520.2012.748005

Holmes Group (1990). *Tomorrow's Schools*: *A Report of the Holmes Group*. Erickson Hall, East Lansing, MI.

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

Koschmann, T. D., Myers, A. C., Feltovich, P. J., and Barrow, B. S. (1994). Using technology to assist in realizing effective learning and instruction: a principled approach to the use of computers in collaborative learning. *J. Learn. Sci.* 3, 227–264. doi: 10.1207/s15327809jls0303_2

Laferrière, T. (2001). "Learning to teach in networked PDSs," in *Paper Presented at the Symposium Learning to Teach in the Networked Classroom Through Collaborative Inquiry. Annual Meeting of the American Educational Research Association (AERA),* (Seattle, WA).

Laferrière, T., Breuleux, A., Willinsky, J, Wolfson, L., Erikson, G., Legault, F., et al. (1998). "TeleLearning: knowledge-building in networked communities and the shaping of teacher's professional knowledge," in *Symposium at the Annual Meeting of the American Educational Research Association (AERA),* (San Diego, CA).

Laferrière, T., Hamel, C., and Allaire, S. (2013). "Learning to teach in a networked classroom: a systemic approach," in *Paper Presented at the Annual Meeting of the American Educational Research Association (AERA),* (San Francisco, CA).

Laferrière, T., Hamel, C., Hamel, M.-D., Perreault, C., and Allaire, S. (2016). "Sustaining innovation in a networked classroom: the contribution of pre-service students," in *Paper Presented at the Annual Meeting of the American Educational Research Association (AERA),* (Washington, DC).

Laferrière, T., Legault, F., and Breuleux, A. (2002). "Project-based learning in a fully networked classroom," in *Paper Discussion at the Annual Meeting of the American Educational Research Association (AERA),* (New Orleans, LA).

Laferrière, T., Willinsky, J., Gervais, F., Laberge, C., and Legault, F. (1997). "Patterns of connection and community," in *Symposium Presented at the Annual Meeting of the American Educational Research Association (AERA),* (Chicago, IL).

Lave, J., and Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation.* Cambridge: Cambridge University Press.

Means, B., Toyama, Y., Murphy, R., and Baki, M. (2013). The effectiveness of online and blended learning: a meta- analysis of the empirical literature. *Teach. Coll. Rec.* 115, 1–47.

Miettinen, R. (2000). The concept of experiential learning and John Dewey's theory of reflective thought and action. *Int. J. Lifelong Educ.* 19, 54–72. doi: 10.1080/026013700293458

OECD (2005). *Formative Assessment Improving Learning in Secondary Classrooms: Improving Learning in Secondary Schools*. Paris: OECD.

Parent, S. (2017). *L'engagement D'enseignants, La Variation De L'engagement D'étudiants Sur Une Base Trimestrielle Et La Présence De Conditions D'innovation En Situation D'enseigner Et D'apprendre Avec Le Numérique Au Collégial / Teacher Engagement, Variability of Student Engagement on a Trimestrial Basis and Presence of Innovation Conditions for Teaching and Learning With Digital Tools and Resources at the College Level*. Doctoral dissertation, Laval University.

Pellegrino, J. W., and Hilton, M. L. (eds.) (2012). "Education for life and work: developing transferable knowledge and skills in the 21st century," in *Committee on Defining Deeper Learning and 21st Century Skills; Center for Education; Division on Behavioral and Social Sciences and Education; National Research Council, U.S.A.* (Washington, DC).

Roschelle, J. (1992). Learning by collaborating: convergent conceptual change. *J. Learn. Sci.* 2, 235–276. doi: 10.1207/s15327809jls0203_1

Roschelle, J., and Teasley, S. D. (1995). "Construction of shared knowledge in collaborative problem solving," in *Computer-Supported Collaborative Learning,* ed C. O'Malley (New York, NY: Springer-Verlag), 69–97.

Savery, J. R., and Duffy, T. M. (2001). *Problem Based Learning: An Instructional Model and its Constructivist Framework. CRLT* Technical Report No. 16-01. Bloomington, IN: Indiana University.

Scardamalia, M. (2002). "Collective cognitive responsibility for the advancement of knowledge," in *Liberal Education in a Knowledge Society,* ed B. Smith (Chicago, IL: Open Court), 67–98.

Schmid, R. F., Bernard, R. M., Borokhovski, E., Tamim, R. M., Abrami, P. C., Surkes, M. A., et al. (2014). The effects of technology use in postsecondary education: a meta-analysis of classroom applications. *Comput. Educ.* 72, 271–291. doi: 10.1016/j.compedu.2013.11.002

Schön, D. (1983). *The Reflective Practitioner*. New York, NY: Basic Books Inc.

Schwab, J. J. (1973). The practical 3: translation into curriculum. *Sch. Rev.* 81, 501–522. doi: 10.1086/443100

Shireen-Desouza, J. M., and Czerniak, C. M. (2003). Study of science teachers' attitudes toward and beliefs about collaborative reflective practice. *J. Sci. Teach. Educ.* 14, 75–96. doi: 10.1023/A:1023653911246

Tamim, R. M., Bernard, R. M., Borokhovski, E., Abrami, P. C., and Schmid, R. F. (2011). What forty years of research says about the impact of technology on learning: a second-order meta-analysis and validation study. *Rev. Educ. Res.* 81, 4–28. doi: 10.3102/0034654310393361

Virkkunen, J., and Newnham, D. S. (2013). *The Change Laboratory: A Tool for Collaborative Development of Work and Education*. Rotterdam: Sense Publishers. doi: 10.1007/978-94-6209-326-3

Wenger, E. (1998). Community of Practice. *Learning, Meaning and Identity*. Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9780511803932

Yoon, H. G., and Kim, M. (2010). Collaborative reflection through dilemma cases of science practical work during practicum. *Int. J. Sci. Educ.* 32, 283–301. doi: 10.1080/09500690802516538

Zeichner, K. M., and Liston, D. P. (1996). *Reflective Teaching: An Introduction*. New Jersey, NJ: Lawrence Erlbaum Associates, Inc.

Zheng, L. J. (2015). A systematic literature review of design-based research from 2004 to 2013. *J. Comput. Educ.* 2, 399–420. doi: 10.1007/s40692-015-0036-z

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in ICT

ORIGINAL RESEARCH
published: 19 June 2018
doi: 10.3389/fict.2018.00012

Check for updates

Change Process of Two Postsecondary Teachers in the Early Adoption of an Active Learning Classroom

Samuel Fournier St-Laurent[1] *and Bruno Poellhuber*[2]*

[1] *Service de Soutien à l'Apprentissage et du Développement Pédagogique, Collège Ahuntsic, Montreal, QC, Canada,*
[2] *Department of Psychopedagogy and Andragogy, Université de Montréal, Montreal, QC, Canada*

There is a growing interest for specialized classrooms, termed active learning classrooms (ALC), which are designed to facilitate the use of active learning methods and information and communication technologies (ICT) by students. Thanks to pioneering studies such as SCALE-UP, there is a better understanding of the benefits of these classrooms and the pedagogy taking place in them. Teachers accustomed to traditional classes have to change many aspects of their pedagogy in order to reap the benefits of the ALCs, however. The purpose of this research is to gain a better understanding of the adoption process of an ALC by teachers and how its adoption modify teaching preferences and practices. Relying on an in-depth case study methodology founded on interviews and questionnaires about the adoption of innovations (CBAM), Approaches to Teaching Inventory, technopedagogical competencies and collaborative, competitive or individual teaching preferences, this article describes the cases of two teachers who used an ALC over a three-semester period. The results show that the teachers develop their courses quickly, with an emphasis on the active learning aspects of their pedagogy rather than on ICT integration, and that there are a lot of personal and management concerns. When the pedagogical changes are stabilized, the teachers retained their personal concerns about the innovation and were highly motivated to collaborate with other ALC users. Finally, apparently minor increases in student-centered teaching approaches result in significant pedagogical changes when they are studied qualitatively. These changes did not lead to a reduction in teacher-centered teaching approaches, suggesting that a significant portion of teacher-directed activities remain.

Keywords: active learning, active learning classroom, pedagogical change process, adoption, student-centered practices, cooperation

Edited by:
Rob Cassidy,
Concordia University, Canada

Reviewed by:
Alain Breuleux,
McGill University, Canada
Liisa Ilomäki,
University of Helsinki, Finland

Correspondence:
Bruno Poellhuber
bruno.poellhuber@umontreal.ca

Specialty section:
This article was submitted to Digital Education, a section of the journal Frontiers in ICT

Received: *28 December 2017*
Accepted: *18 May 2018*
Published: *19 June 2018*

Citation:
Fournier St-Laurent S and Poellhuber B (2018) Change Process of Two Postsecondary Teachers in the Early Adoption of an Active Learning Classroom. Front. ICT 5:12. doi: 10.3389/fict.2018.00012

INTRODUCTION

Several postsecondary educational institutions in Quebec were inspired by the American project SCALE-UP (*student-centered activities for large enrollment undergraduate programs*; Beichner et al., 2007). SCALE-UP emerged from major changes that take place in STEM education in the United States. It aimed to improve student learning by integrating collaborative, hands-on learning activities with abundant use of information and communication technologies (ICT) in large enrollment programs where the use of amphitheater is frequent. Even though the changes made to pedagogy were important, SCALE-UP became most famous for presenting a rationale, practical applications, and demonstrating positive impacts of a classroom layout adapted for collaborative work and ICT use: the active learning classroom (ALC).

A body of research specific to ALCs is emerging. The research methods often include groups of students in traditional settings as a control condition (Dori and Belcher, 2005; Beichner et al., 2007; Charles et al., 2011). With regard to the students, the results are encouraging: increased conceptual understanding (usually double), higher success rates (double to sextuple), higher attendance (80–90%), and other positive outcomes associated with motivation.

The results of SCALE-UP and similar projects (such as the TEAL project) showed that classroom layout goes hand-in-hand with pedagogy even though early research doesn't discriminate between the effects due to pedagogical changes and those due to room layout. Charles et al. (2013) focused on the relationship between the pedagogy and the type of classroom layout. They reported that the type of pedagogy used (traditional or active learning) may have a different impact when used in ALCs or traditional classrooms, with lecturing actually faring worse in an ALC than in a traditional classroom. The authors were also the first to explore the relationship between teachers' beliefs about their role and student learning in the ALC setting. Results suggest that both the teachers' beliefs and the pedagogical approaches used in an ALC can influence learning outcomes. To make the most of the potential advantages of an ALC, a pedagogical change process must take place for many teachers who are typically more used to lecturing. In a sense, ALC layout, technologies and tools offer different pedagogical affordances that teachers may or may not use to influence the students' learning and academic success.

While the pedagogical benefits of ALC are more documented, the implementation of ALCs are generally time-consuming and costly. The general objective of investing in these classrooms is to facilitate active learning and group collaboration using ICT. However, both pose particular challenges to teachers and represent innovations in pedagogical practices that takes a long time to implement. Beichner et al. (2007) describe the change process for departments in terms of years. The process of adopting an ALC is not only costly in terms of equipment, but in time and energy for teachers as well.

CEGEPs are postsecondary colleges exclusive to the province of Quebec in Canada. They offer general (2 years) and specialized (3 years) programs in an educational system where undergraduate degrees and secondary school are both 1 year shorter. In this network, Kingsbury (2012) reported the appearance of eight ALCs in 2012, with the number rising to over 30 in 2014 (CLAAC, 2014). The proliferation of ALCs in CEGEPs can be explained in part by the interest of these institutions for active learning and integration of ICT to improve student academic success. Nonetheless, the rapid appearance of ALCs combined with the possible link between learning outcome and teachers' approaches raise concerns as to how this innovation is adopted. Additionally, Brooks (2012) showed that classroom layout does induce changes in pedagogical practices, with traditional classrooms generating more lectures and ALC generating more group activities. The CEGEP context offers a good opportunity to explore the impacts of ALCs' use of early adopters.

In this study, we seek to gain a better understanding of the adoption process of an ALC by teachers and how its adoption modify teaching preferences and practices. We rely on a "thick description" that provides interpretive depth (Spiegelberg, 1978). The study's objective is to describe the cases of two teachers (selected from a larger group) who made the most significant changes toward ALC-oriented pedagogy over a three to four semesters period.

CONCEPTUAL FRAMEWORK

This study takes place in a postsecondary education setting with recently acquired ALCs. In our attempt to better understand the adoption of the ALCs as an innovation, we selected key theoretical concepts from the teachers' perceptions, beliefs or practices linked to active learning and ICT integration. It should be noted that these concepts are linked to self-reported indicators, as we chose to avoid direct monitoring of practices at this stage. Additionally, we chose concepts that have already been adopted in education and are accompanied by validated instruments.

Preferences Regarding the Type of Instructional Methods and Teaching Approaches

Active learning is a broad term often presented in opposition to lectures or other types of "traditional instruction" (Prince, 2004). In practice, active learning refers to several instructional methods grounded in active pedagogies, such as problem-based learning (Barrows, 1996; Savery and Duffy, 1996), project-based learning (Blumenfeld et al., 1991), peer learning (Crouch and Mazur, 2001), and various collaborative and cooperative techniques (e.g., pause, jigsaw, pyramidal scripts–see Howden and Kopiec, 2000 as an example). Productive failure is a more recent method that could also qualifies as an active learning method: the initial failure part focusing on collaborative research and attempts to solve a problem (Kapur et al., 2010). Among these methods, two salient types emerge from the descriptions given by authors: student-centered learning and collaborative learning.

Student-Centred Learning

The difference between active learning and traditional instruction generally lies in the role of the students in these different situations. The role of students during lectures is mostly to receive the knowledge given by the teacher and take notes. While some students may be active and engaged in lectures, it is useful to contrast the relative passivity of listening to a speech with the active role required of students in instructional methods such as problem-based learning (where they have the responsibility to research new ideas, collect data, analyse problems, and more), cases studies, and cooperative learning. Students are also active in many other ways: they may act on the work of other students

(Macpherson, 2007), develop a product (Barron et al., 1998), attempt to solve a very difficult (or impossible) problem (Kapur, 2012). In summary, active learning is associated with the idea that the students are required to be more active, through the tasks they have to accomplish in the associated instructional methods.

When students take an active role, the role of the teacher changes accordingly. The teacher no longer acts as the main intermediary between the students and the material (Bonwell and Eison, 1991). Students may not have the required level of cognitive and metacognitive strategies to assume these new responsibilities, however, so the teacher has to guide them in choosing and applying the appropriate strategies (Hmelo-Silver, 2004; Gijbels et al., 2005). This is an important role in the context of active learning, with an effect on student performances (Yukselturk and Bulut, 2007).

Added to this list of new responsibilities is the need for instructional design that provides proper scaffolding and technical support for the students (Laffey et al., 1998). In problem-based learning, for example, students learn through ill-structured problems that have multiple acceptable or correct answers. They must explore many solutions (and much material) to find the one that seems best (Hmelo-Silver, 2004).

We can illustrate the shift in responsibilities of the teacher and the students with a gradient. This image can be found in the taxonomy of Chamberland et al. (2006): each instructional method has a relative position on a "control of learning" continuum ranging between a totally teacher-controlled point (teacher-centered) and a totally student-controlled point (student-centered). The teacher-controlled end refers to activities where the teacher has complete control over the activities, such as the pace of learning and the material shown to the students. Lectures are a good example of the teacher-centered method. At the other end, the students have more freedom to explore, determine the pace of learning and choose their strategies. Bonwell and Sutherland (1996) also presents a similar approach to describe the nuances of teaching methods associated with active learning.

The teacher- vs. student-centered opposition is also found in the Approaches to Teaching Inventory model (Trigwell et al., 2005) which is based on a list of strategies adopted by teachers at the university level. This inventory provided the basis for a short questionnaire with two scales: student-centered and teacher-centered. It offers a useful tool for appraising the relative position of the teacher's approach on a continuum. It was used in one ALC study with six teachers (Charles et al., 2011). Even though the number of cases was small in this study, the students obtained higher conceptual gains as their teacher self-reported more student-centered approaches.

Collaborative Learning

Another central aspect of active learning is collaboration among the students, which typically ranges from teams of two people (e.g., in peer learning) up to 12 (Wilkerson, 1996). Collaboration and cooperation can be seen as learning in a team of students who are working toward a common goal, although it is sometimes useful to make a distinction between collaboration and cooperation, to take into account the potential effects of specific roles, contributions and hierarchy among the team members (Dillenbourg, 1999; Kirschner, 2001). It can also be described through comparison with two other types of interactions that students have with each other: competition and individual work (absence of interaction).

Johnson et al. (1998) refer to the early work of Koffka, Lewin and Deutsch in the 1900s and 1940s to describe cooperation as the result of interdependence structures among students: cooperation occurs when one student's success depends on the success of their teammates, through task, and reward structures. Slavin (1996) mentions a good example of reward structures in group contingency, where rewards are given to a group of students if every member reaches a specific goal. Another type of interaction—competition—results from negative interdependence or contexts where the success of one student depends on the failure of another (e.g., single winner in a tournament). Finally, students are likely to work individually when there is a lack of interdependence.

One aspect of the teacher's role in an ALC is to design contexts in which students will work together efficiently. Interdependence offers a practical objective for instructional design, since the literature offers examples of task and reward structures that foster positive interdependence. The analysis of teachers' beliefs about collaboration offers a general perspective on what motivates the choice of learning activities. It is an alternative to direct monitoring of changes in the number and quality of collaborative activities put in place by teachers.

Technopedagogical Competencies

ALCs usually offer a wide range of technologies, from laptops to systems designed to share multimedia content among groups of students. In this technology-rich environment, teachers are likely to design activities where students will use technology to learn. As was the case for active learning, to be used effectively, integrating technology demands some changes in pedagogy (Conseil Supérieur de l'éducation, 2000; OCDE, 2008).

There are several models to describe ICT integration by teachers. One popular model is Technological Pedagogical And Content Knowledge (TPACK). It places ICT integration at the intersection of three kinds of knowledge required of teachers: content, pedagogy, and technology. TPACK does not focus on adding technology to the teacher's existing pedagogy, but rather on the harmonious merging of the three components of interest. TPACK is useful for illustrating the key components of an instructional strategy.

While many models focus on the pedagogical integration of ICT, the approach used here to describe ICT integration is through the pedagogical skills needed to integrate ICT, for example, those identified by the International Society for Technology in Education (ISTE) in 2008. This approach offers a broad view of the possible changes teachers may implement when using an ALC, without focusing on specific equipment or applications.

The work of the ISTE and the technological pedagogical skills suggested by the TPACK, among other references, inspired the development of a framework of technopedagogical competencies for teachers in the Quebec college network (Bérubé and Poellhuber, 2005). This model is founded on a broad review of international models of ICT integration and professional development, followed by interviews and validation with local experts. The model identifies four areas where teachers have to develop technopedagogical competencies; (1) communication and collaboration; (2) informational competencies; (3) instructional design (lesson planning, implementation and evaluation); and (4) production of educational resources. This model is anchored in a socioconstructivist perspective that fits well with the use of an ALC.

Adoption of an Innovation

The previously identified scales give little information about the possible concerns, challenges, and motivation factors for pedagogical change. Adopting an ALC entails complex interactions between equipment, pedagogy, and classroom layout.

One model that is frequently used in the context of pedagogical and technology adoption is the concern-based adoption model (Hall et al., 2006; Hord et al., 2006; George et al., 2013). It rests on the idea that the adoption of an innovation is first a process of professional change for the teachers. Furthermore, the users' perceptions determine what can be done to help them adopt the innovation. A key aspect of CBAM is the profiles of user concerns about the innovation being studied. The "self" concerns refer to informational and personal stages, where users have general awareness about the innovation and perhaps some doubts or questions about the effects of the innovation on themselves. The "task" concerns are directly related to the management stage, where users may have issues with regard to organizing and scheduling. The "impact" concerns are related to the consequence, collaboration and refocusing stages, which respectively refer to interest in the possible impacts of the innovation on the students, interest in cooperating with other users in the use of the innovation, and focus on exploring new ways to use the innovation (or even replace it).

The CBAM also shares similarities with models of pedagogical ICT integration through its Level of Use (LoU) branching interview. By asking questions about the use of the innovation in a specific order, the interviewer can quickly determine whether a teacher is using the innovation (first branching), what kind of changes the teacher has made to use the innovation (second branching), whether collaboration is occurring with other users of the innovation to generate student-oriented changes (third branching), and whether major changes are planned (fourth branching). For example, using an innovation (positive for first branching) and making personal, teacher-oriented changes to use it is labeled "mechanical use." This means the teacher focuses most of their effort on short and day-to-day use of the innovation.

The CBAM can complement indicators related to the teacher's role in active learning and ITC integration, since it can explain the changes observed. In the previous example, the teacher operating at the level of mechanical use may also have management concerns and report improved ICT competency. He may therefore focus on using new technology and equipment to alleviate management problems (e.g., distribution of material, time management, better monitoring of the students' work).

METHOD

The research team is composed of researchers, teachers, and pedagogy professors from Université de Montréal and five CEGEPs (postsecondary colleges with pre-university and technical programs). For the purpose of this article, a multi-case approach was used: each case was treated individually and compared with the other cases. The description of each case is based on the key aspects of ALC use proposed earlier (approaches to learning, teaching preferences, technopedagogical competencies, and adoption of an innovation). Data was collected each term, using questionnaires, and individual interviews with the teachers. The total project duration was four terms, although some of the teachers recruited in this study participated for only three consecutive semesters.

Teachers

Although the results of this article mainly focus on two teachers, they were selected from a group of 13 CEGEP teachers teaching five different subject matters (literature, mathematics, physics, biochemistry, and philosophy) in three different CEGEPs. The teachers were initially assigned to an ALC by the administrative service at their CEGEP (in one CEGEP, the classroom was reserved for a specific subject matter). All invited users agreed to participate in this study. To ensure at least minimal use of the classroom, all the teachers committed to use it for at least 50% of their classroom time. Activities done in a laboratory setting (physics, biology) were excluded from the calculation.

Classrooms

Each teacher used one of three classrooms, each located in a different CEGEP. The cases described in his study took place in two different classrooms. The classrooms contained seven to eight permanently fixed tables large enough to accommodate teams of up to six students and equipped with electric and media connections (electricity, VGA, and internet). Each team was allowed to use their own team projector or TV screen. One white board and at least two laptops were also available for each team. The teacher's desk was either located in the center of the class (in one case) or included in the ring-shaped disposition of the tables around the room. Interactive whiteboards linked to the teacher's desk were available in two classrooms. The teachers reported that they were mostly used to present material, however: the students rarely or never interacted with this equipment.

Questionnaire

All the teachers answered a questionnaire at the beginning of each semester and at the end of the project. The questionnaire examined four dimensions: teaching preferences, approaches to teaching, technopedagogical competencies, and adoption of an innovation. It was an adaptation of the Stages of Concern questionnaire, a CBAM tool that is used to determine the relative intensity of each of the seven stages of concern. The teachers answered the questions based on their agreement with the items, using a Likert scale of 5 or 7 points.

The items for the scales on teaching preferences and approaches to teaching were only available in English, so they had to be translated to the native language of the teachers, French. This was done using a cultural transvalidation procedure where each question was translated into French by a professional translator. Another translator then did a back translation. The original and final questions in English were compared. The questionnaire was read by five teachers working at the same CEGEP as the participants, to ascertain the clarity of the questions. After minor adjustments, the questionnaire was distributed to nearly 900 teachers from the participating CEGEP. A total of 128 teachers answered. The data were then used to examine the reliability and factorial validity of the translated scales.

In addition to these dimensions, demographic questions (e.g., age, years of experience in education) and two open questions about the advantages and challenges of using an ALC were added to the questionnaire. The final version contains 127 questions and takes around 20 min to answer.

Approaches to Teaching

Active learning is linked to instructional methods in which students take an active role in researching, organizing and analyzing knowledge. Accordingly, the teachers take less responsibility in the dissemination of knowledge and greater responsibility in providing cognitive process support for the students. For the teachers, this shift in responsibilities can be depicted on a continuum between a teacher-centered approach to teaching and a student-centered approach to teaching. If teachers see their role predominantly as the source of knowledge, their position on the continuum is toward the teacher-centered end.

Trigwell and Prosser (2004) offered a practical tool for assessing teachers' approaches with regard to these two dimensions. The Approaches to Teaching Inventory questionnaire was first developed with 58 university professors. The inventory of strategies adopted by the professors and their underlying intentions were organized and validated on two scales: teacher-focused and student-centered. In a second article, more items were added to the inventory (Trigwell et al., 2005). The participants answered 22 items using a five-point Likert scale.

The questionnaire was translated to French and validated. The final version contained eight items in the teacher-focused scale (alpha = 0.733) and nine items in the student- centered scale (alpha = 0.833).

Teachers' Preference

As stated earlier, active learning is closely related to collaborative learning. The teacher's preferences in this regard can be useful in understanding the potential impact of ALC adoption. Slavin and other authors offer a model which clearly separates cooperation (working together), individual work, and competition (working against others) (Slavin, 1996). For this purpose, a questionnaire from Owens and Barnes (1992) was used to determine the teachers' preferences in these three dimensions. The original questionnaire is composed of 33 items divided into the three dimensions of interest.

After validation, the final questionnaire contained seven items for the individual dimension (alpha = 0.65), seven items for the competitive dimension (alpha = 0.73), and nine items for the collaborative dimension (alpha = 0.82).

Technopedagogical Competencies

For this project, we chose to address ICT integration by the teachers' appraisal of their own technopedagogical skills.

A questionnaire was developed and validated based on Bérubé and Poellhuber's model (2005) and used in a previous unpublished study. It allows teachers to report how they perceive their own pedagogical ICT integration skill. For this project, questions were added for the "collaboration" and "use of specialized resources" dimensions. Exploratory factorial analysis of the original 30 items during the validation phase of the questionnaire yielded three scales:

1. Choice of instructional methods (5 items, alpha = 0.788).
2. Use of ICT for creation and collaboration in active learning (12 items, alpha = 0.895).
3. Use of resources related to field of study (8 items, alpha = 0.846).

CBAM

Two CBAM tools were used in this study. The Stages of Concern (SoC) questionnaire is composed of 35 statements aimed to determine the teacher's level of concern about using an ALC related to seven stages of concern: unconcerned, informational, personal, management, consequence, collaboration, and refocusing. ALC use was defined as the general use of the classroom, including active learning pedagogy, and ICT. The data were analyzed and presented as recommended in the SoC guide (George et al., 2013). The second CBAM tool, Level of Use, was used in the interviews.

Individual Interviews

At the end of each term, the teachers were invited to an interview. Questions were based on the CBAM Level of Use tool which, as the authors state, "breaks use and nonuse into several levels" (Hall et al., 2006). LoU gives indications about the extent to which the ALC is used by a teacher. The levels are (0) nonuse, (1) orientation and acquiring information about the innovation, (2) preparation for the first use, (3) mechanical use focusing on short-term efforts, (4A) routine use where few changes are made, (4B) refinement to increase the impact on students, (5) integration and collaboration with other users, and (6) renewal. One question was added to clarify the opportunities for collaboration for teachers within the project and with the researchers. Another assessed the perceived impact of the ALC on their own work and on student learning.

A qualitative analysis was conducted using two coding lists. The first was a list taken from the Levels of Use, which allowed the coding team to identify segments linked to one of the seven levels of the LoU tool. The second coding list was designed using a mixed approach (Miles and Huberman, 1994). Codes were first listed based on the main items of the project's conceptual framework. New codes were added by two researchers who read the material after the first term. Each term, an inter-coder agreement was made on a sample composed of 20% of the transcripts to be coded. The coders were the same and the percentage of agreement was always between 82 and 89%.

Ethics

The project was conducted under an ethics certificates from the Université de Montréals' pluridisciplinary ethics comittee (CPER-13-112-D) and from each of the colleges with participating teachers. This study was carried out in accordance with its recommendations with informed consent from all subjects. All subjects were encountered by researchers and gave written informed consent in accordance with the canadian three coucil guidelines, both for the survey and the interviews.

RESULTS

For this case study, we selected two cases (1 and 2) whose numerical indicators over the course of the project showed the most changes toward student-centered approaches, collaborative preferences, and high technopedagogical competencies. During the selection process, priority was placed on cases with the most change in student-centered ness, since it was a factor of interest in two previous ALC studies. Interestingly, the two cases identified using this rule were also the two cases that showed the most change toward collaborative preferences and they were among the top four teachers in terms of positive change in their perception of their skills. **Table 1** shows the change in perception between the last semester and first semester for all thirteen cases.

For each case, we first present a summary of the teacher's numerical change indicators over the project. The SoC profile is also shown. The quantitative data are linked to segments of the interviews conducted with the teachers to highlight factors that contributed to their adoption of the ALC.

Case 1

Case 1 is the teacher who showed the greatest positive changes in the following scales: collaboration preferences, student-centered approach to teaching and technopedagogical competencies (see **Table 2**). When he joined the project, he was mid-career (10 to 20 years). He had some previous formal training in pedagogy (less than 15 university credits) and showed great interest in the use of technology with students in the ALC. He prepared and gave the same course for three terms in the ALC and usually had three to four groups of 30 to 40 students each semester.

During the semester prior to his participation in the study, he attended an activity given by another teacher in the ALC. At that time, he saw the difference between a simple group assignment, where the teacher gives work to students and then sits at his desk, and the active learning setting, where the teacher guides the students' cognitive processes and the team engagement. The importance of models as a source of inspiration was underlined several times during the interviews.

> *I did not have models to show me how this works and how we work in this kind of classroom. On the other hand, lecturers are the models we always have seen.*

For this teacher, his early experiences in the ALC were influenced by a need he felt to plan something new and innovative for each class. This pressure quickly led to fatigue, frustration, and the accumulation of small failures.

> *I had the feeling that since I was there, I had to use every piece of equipment and that everything about my planning had to fit perfectly with the tools. Otherwise, I would have failed.*
>
> *Each time I was, like, "I need to do something new." Of course, I was trying something that wasn't perfectly ready. So it was rarely a success.*

This personal pressure to innovate and use the equipment was found in other cases. Many reasons were offered. For example, the cost of the classroom and the fact that it was made available for them in the context of a special project made them feel privileged. As such, they felt a certain level of performance was somehow expected of them. Another example given by the teachers was the perception that the students expected something special.

At some point during the semester, this teacher stopped creating new activities and concentrated on some models that worked well. He then began working to improve these.

> *So I repeated it four times and, as I said, there was no longer this pressure that I had to do something new. I think the students liked it and I found my place.*

Approaches to Teaching

Despite a small 0.33 increase in the student-centered subscale on the ATI (4.78 to 5.11), this teacher is one of the two cases who showed the greatest increase for this indicator. His goal was primarily to reduce lectures by replacing them with collaborative activities. During the first semester, emphasis was placed on the variety of these activities, but this set a design pace that was difficult to maintain and led to activities that were less successful. Furthermore, with this level of variety, he felt that the students were getting lost in the instructions. Toward the end of the first semester, he chose fewer models of activities that he could then work to improve.

> *You need to create habits. Then the students know what to do and ask fewer questions.*

A similar change was made with the classrooms. During the first semester, he maintained access to a traditional classroom in which lectures were sometimes offered. After the midterm, he decided to stay in the ALC, mainly for practical reasons: students occasionally ended up in the wrong classroom. The students reacted positively to this decision, saying the ALC was more comfortable, attractive, and fun.

Despite these positive comments from the students, he was uncomfortable giving lectures in the ALC. During the second semester, he shared his concerns with students.

TABLE 1 | Change in teachers's perceptions between the last and first semester of ALC use for all cases.

Indicator	1	2	3	4	5	6	7	8	9	10	11	12	13
APPROACHES TO TEACHING													
Teacher–centered	0.89	0.37	0.25	0.75	0.25	−0.13	0.87	1.50	−2.50	−1.00	−0.37	0.25	−1.00
Student–centered	0.33	0.33	0.33	−0.11	−0.11	−0.11	−0.11	−0.22	−0.55	−0.55	−0.56	−0.89	−1.00
TEACHING PREFERENCES													
Individual	−0.97	0.07	0	−1.17	−0.33	−0.83	0.50	−0.67	−0.67	−0.17	0.50	0.33	−0.84
Collaboration	0.57	0.86	−0.14	0.43	0.43	0.15	0.14	0.43	0.29	0	0	0.28	−0.43
Competition	0.53	0.29	0.14	−0.15	−0.29	0.28	−0.28	0.14	−0.85	1.14	1.43	0.86	−0.43
TECHNOPEDAGOGICAL COMPETENCIES													
Choice of methods	0.20	0	0	−0.80	−0.60	−0.20	0	0.20	0	−0.40	−1.40	0.20	0.40
Creation/Collaboration	0.83	0.95	0.33	−1.42	−0.49	−0.58	−0.64	0.25	0.08	0.34	−0.17	−0.75	−0.17
Resources	0.88	0.75	0.13	−0.38	−0.50	0	0.12	0	0	−0.75	0.75	−0.63	0.88

A negative value indicate a lower value at the final semester.

TABLE 2 | Case 1: Change indicators before and after three semesters.

Indicator	Before	After
APPROACHES TO TEACHING		
Teacher-centered	3.25	4.14
Student-centered	4.78	5.11
TEACHING PREFERENCES		
Individual	3.17	2.20
Collaboration	5.86	6.43
Competition	5.33	5.86
TECHNOPEDAGOGICAL COMPETENCIES		
Choice of methods	4.00	4.20
Creation/Collaboration	1.50	2.33
Resources	1.75	2.63
CBAM		
LoU	4B	4B

TABLE 3 | Case 2: Indicators of change before and after three semesters.

Indicator	Before	After
APPROACHES TO TEACHING		
Teacher-centered	3.88	4.25
Student-centered	4.11	4.44
TEACHING PREFERENCES		
Individual	3.60	3.67
Collaboration	4.57	5.43
Competition	3.57	3.86
TECHNOPEDAGOGICAL COMPETENCIES		
Choice of methods	3.40	3.40
Creation/Collaboration	1.55	2.50
Resources	2.75	3.50
CBAM		
LoU	3	4B

> *In fact, the students said no! It's not a problem. I explained to them that they looked less engaged during lectures and that they did not seem to know where to look. They said they were not.*

He welcomed these comments and later mentioned that he was less nervous about giving short lectures in the ALC.

A notable challenge to his new role as teacher surfaced during the second semester. Once the students were actively engaged in teamwork, it was difficult to stop them in order to give further instructions or small, lecture-like interventions. Even when the students stopped, their attention was not focused on the teacher. He found the solution in a routine in which a short lecture was given at the beginning of an activity. Later, the students had access to complete instructions for the activity. The teacher visited each team to offer theoretical support or special instructions. These small adjustments to the activity design offered a new way to fulfill his role.

> *When you plan lectures, you can adjust as you go and fill the time easily. Now, there are more activities to plan and more teamwork. There is less space. In fact, I believe it is a different way to plan courses.*

The new routine eventually ended up giving his students more time to accomplish their learning tasks and more control over their learning. It is important to mention that he remained critical about the changes.

> *I see students take notes and pay attention when we discuss the solutions at the end. I see them take pictures. But I do not know how well they organize this information. Yes, we have more interactions, but have they improved their retention of what was discussed?*

During the third semester, he reported fewer changes, but he occasionally engaged the students in a new routine where they work on a problem and present their solution to the group.

Teaching Preferences

For this teacher, the collaborative teaching preference rose from 5.86 to 6.43, the competitive preferences also rose from 5.33 to 5.86, and the individual preferences decreased from 3.17 to 2.20.

During the first term, this teacher often designed activities with a cooperative work component. Briefly, each team was assigned a portion of the material to be covered and was responsible for sharing their work with others. Dividing the labor in this way also helped him cover more material. This advantage was a strong incentive to favor collaboration.

> *I slowly discovered that to use this classroom efficiently, I had to make the most of the fact that students were divided into teams. I find it very interesting to divide the material between the teams and bring them together at the end.*

With some activities well established at the end of the first term, he focused on team management. During the second term of the project, he questioned his system of randomly forming teams. This subject was covered during a meeting between the researchers and teachers: early results indicated that many students preferred to choose their own teammates. Other teachers also mentioned trouble forming perfectly balanced teams.

> *At the beginning of the course, it [random assignment] is fun because the students meet new people. Once they have worked together, they are reluctant to change because they have already established a team dynamic. When they sit together, it is because they want to work together. So at that point, I stop randomly assigning students to teams.*

Early in the project, he mentioned problems with student engagement in group work. One problematic situation was students who disengaged from the work. In this case, he tried to find structures, offer support, and adjust instructions. The teams were also asked more often to present their work to the group.

> *Last term, the students received specific tasks in their team and it worked well. I don't know why, but this year, I did not distribute the tasks. I feel the students worked less.*

Another problem observed was students who were so engaged that they did not stop when the teacher had to give a general message. As a solution, he designed activities with minimal interruptions.

> *Once it starts, if you had the bad idea of planning a small lecture to explain something...forget about it… too difficult. Eventually they listen, but you really have to take over.*

From the beginning, this teacher saw the positive impacts of collaboration on his own work (saving time by dividing the work). He also dealt with challenges in team management by seeking alternatives and by changing his pedagogy. This type of positive experience with group work aligns with the corresponding increase in the preference for collaboration.

Technopedagogical Competencies

This teacher's personal perception of his skills showed the greatest change in the use of ICT for collaboration/creation (1.50 to 2.33) and specialized resources (1.75 to 2.63). The interviews revealed two salient contexts of ICT use: the use of videos and the collaborative tool Google Docs.

To explain the increase in the use of specialized resources and creation, it is relevant to mention that in the first semester, this teacher developed specialized videos so the students could review the course content before coming to class. These videos were initially part of a flipped class approach, but he did not formally pursue this idea in the following semesters. The flipped class approach was maintained for a limited number of activities. He concluded his first semester by saying that no other significant ICT integration was made other than having the students use computers to look for information on the web.

> *No, I did not make major changes. I abandoned the exploration of some technologies because I had no idea what the other teachers were doing with them.*

During the second semester, he tried the Google Docs application, effectively replacing Microsoft Word in activities where students had to write texts which were later presented to the class. This application later played an increasing role in keeping traces of the students' work and as public notes to prepare for exams. Google Docs may have contributed significantly to his increase on the collaboration/creation scale. He also began to use specialized applications to manage time (e.g., public stopwatch), but all these changes took quite some time to develop and implement.

> *Give yourself some time. After two semesters, I begin to feel ready to try more complicated things. This is a lot of change and you have to give yourself a chance.*

One observation that almost every teacher in this project made about ICT integration is the difficulty of effectively managing the computers and other electronic equipment. For this teacher, these difficulties can be mitigated by adopting a routine, which he tried to instill in his students.

> *Of course, if you use this place once, you will take a lot of time explaining to the students what they have to do, where to get the computer and how to install everything. But once they do it, they know what to do and there will be fewer problems the next time.*

Unlike many other teachers, he did not report ICT issues as distractions.

> *No, not really. It is when I lecture that I notice students doing something else. They are openly on Facebook...and not embarrassed about it.*

Except for developing videos, which replaced lectures on theoretical concepts, this teacher seemed to focus on other aspects of ALC use during the first semester. After establishing a routine, he began to replace some aspects of his activities with

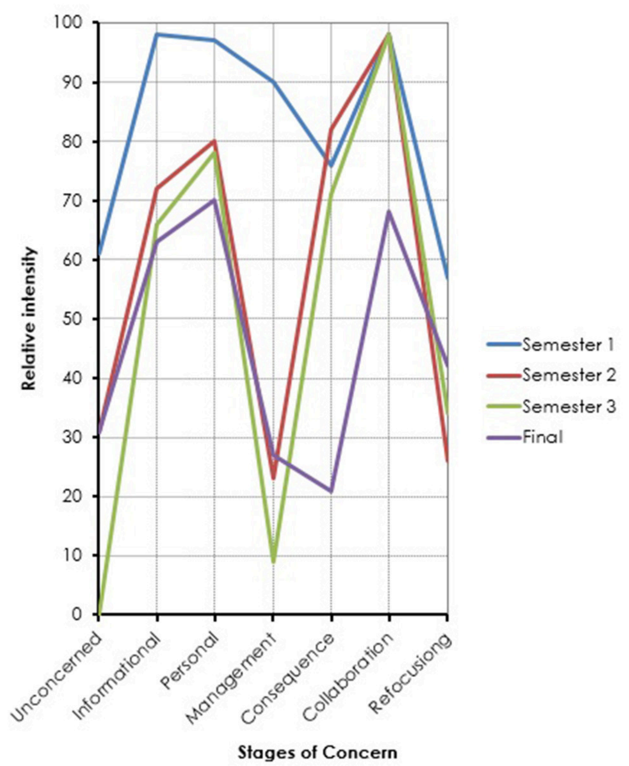

FIGURE 1 | Stages of Concern profiles for Case 1.

new technology, such as Google Docs. The lack of major issues and the increase of the perceived skills suggest that further change is possible in the future.

Adoption of the Innovation

The SoC profiles in **Figure 1** show the relative levels of concern for Case 1 at the beginning of each semester and at the end of the project.

This teacher's initial profile indicates relatively strong personal concerns that lasted for the duration of the project. The first semester is also characterized by a peak in the informational stage, meaning that this teacher was actively researching the innovation. In the first semester, management concerns were also high, but they quickly dropped afterwards.

In the second and third semester, the profiles are similar. One notable observation is the personal stage, which is higher than the informational stage. This represents a theoretical pattern called negative one-two split. According to the SoC guide, it indicates that this teacher may have personal doubts about the innovation that interfere with his interest in knowing more about it (George et al., 2013).

Of course, the authors also call for prudence in this kind of interpretation. This result was not supported by comments from the interview, as negative aspects of the ALC were always linked to management problems, such as lack of student preparation, time management, amount of effort to invest in course preparation. One exception was a comment presented earlier about the way the students organize their notes and whether they remember what is discussed in class. We can see from the interview, however, that this teacher remained critical of his pedagogical choices and sought to improve student learning.

The peak observed in the collaboration stage can be linked to comments made in the interview about collaboration with other teachers. He showed interest in collaboration, but he could not find local colleagues to collaborate on the use of ALCs. In fact, the ALC in his institution was new and only a handful of teachers used it for more than one lesson. According to the SoC guidelines, a high informational stage and a high collaboration stage "suggests a desire to learn from what others know and are doing, rather than a concern for leading the collaboration." This explanation is supported by previous comments about his interest in observing other teachers and learning how they use the ALC and ICT. This teacher also tried to reach others by giving oral communications about ALC use during the third semester. He was also visited by several colleagues who wanted to observe a typical lesson in the ALC. These actions may well have been the first step toward future collaboration.

Analysis of the interviews with regard to the LoU revealed that he stayed at level 4B throughout the project. This level corresponds to refinement: the teacher varies the use of the innovation to increase its impact on the students. He did not reach the next level because he did not collaborate with others to use the innovation.

In summary, Case 1 rapidly designed activities and video resources during the first semester. Afterwards he focused on improving teamwork effectiveness and integrated more ICT in his activities. Collaboration was a need that was not filled for this teacher. One possible way to help him is to provide him with examples of ALC use by other teachers.

Case 2

Case 2 is the teacher who demonstrated the greatest change in teaching preferences for the collaborative indicator (see **Table 3**). He was also among the teachers with the greatest changes in the technopedagogical competencies subscales. When he joined the project, he was in mid-career (10 to 20 years) in education. He had no academic base of pedagogy and showed great interest in the use of the ALC as a way to generate new experiences with his students. He prepared two courses over his three-semester experience and usually had three groups of 25 to 35 students. Prior to this project, this teacher designed another course in the same subject with the intention of using it in the ALC. Departmental assignments did not allow him to use his work. It should be noted that despite these efforts, he described himself as a teacher who mainly uses lectures in class. He also had several concerns about ICT use before the project.

When asked about the main advantages of the ALC, this teacher always made positive comments about the layout and the fact that his classroom is different. He refers to the ALC as a source of creativity for developing new activities for students.

Approaches to Teaching

As in Case 1, this teacher's results increased for both student-centered and teacher-centered scales. The small 0.33 increase in the student-centered scale was the highest increase observed.

During the first term, he focused on testing activities similar to those he had previously designed. He also planned new ones. Typical activities began with a lecture and were followed with a teamwork period where students gathered specific information and did calculations to answer a problem. In contrast with the ill-structured problems usually found with problem-based learning, his activities mostly required a single correct answer from the students.

After the first term, this teacher showed enthusiasm about the classroom. He reported that the ALC fostered new ideas and changed the teacher-student dynamic. Instead of the teacher trying to make the students do things, they began to raise questions themselves.

> *I could give the same activity in another classroom, but this one stimulates me. There is still work to do, but this place motivates me to design interesting activities. This year, I have done more and I have plenty of ideas for the future.*

The boards on the walls are a notable example of equipment that allows for a different way of thinking about the tasks students will do.

> *So I would draw a graph in front of them. But in the ALC, they all have their own boards, so I like to project the image of a grid and they draw their own graph.*

Once this base had been established after the first term, he focused on the design of his existing activities by replacing parts where he was still lecturing. The teacher also saw these improvements as a way to reduce the time the students spent listening. Listening is perceived to be more difficult for students in an ALC.

> *Students still have some trouble listening and I think there is still room to cut back on my lectures. I want to plan more teamwork.*

More time allotted to collaborative learning meant that the students had more time for discussions. The perception of these discussions was generally positive:

> *The students look happy when they do these activities. They discuss and negotiate...why...how…how did you get the answer...this answer makes no sense…*

He also reported that he let the students discuss a problem instead of readily giving the information, which is in line with the new role of teacher in an ALC.

> *I just said something to a group of students and the team at the next table piped up. I didn't say a word, just listened and heard them out.*

During the second term, new ideas also came from the students. Their favorite movies, music, and hobbies became the starting points for new contexts for the problems to be solved.

> *He told me he liked that very much. So I spent 20 hours designing an activity on it. It took me so much time to research the subject that I didn't work on the actual design as much as I hoped.*

He also mentioned an increasing interest in designing activities that look like games or allow students to study problems in fictional, yet entertaining contexts.

During the third term, he introduced music in some activities. This was an interesting change since in a traditional lecture setting, music would be seen as an auditory interference for the transmission of knowledge.

> *So I put on music to go with the subject of their activity. There was a calm sort of mood in the classroom. It was fun to put on a bit of music to enjoy the activities.*

Teaching Preferences

The results for this indicator improved the most on the collaboration scale (4.57 to 5.43), with a small increase observed on the competitive scale (3.57 to 3.86). The relatively low score for the collaborative scale during the first semester (4.57, vs. 5.98 for the mean of the cases) may be explained by comments about the fact that in the new setting, the students interacted among themselves more and were less inclined to listen as they did before.

> *Students rapidly develop a sense of complicity among themselves and less with me.*

There were problems with students who came to class unprepared and slowed the progress of their team. He felt that no matter what action was taken, they remained disengaged. This belief did not change during the project.

Despite these concerns, this teacher continued to dedicate more time to collaborative activities during the second semester. It should be noted, however, that the description of the activities revealed similar patterns. This observation will be further explored later.

> *This time I did more group work. I had a small script and I did many similar activities with it.*

In general, this teacher mentioned many situations where collaborative work was beneficial. Both skilled and less skilled students seemed to enjoy a positive impact from the ALC. Teamwork was often associated with perceptions of increased engagement:

> *The students sometimes seem apart and not very active when they sit at a table. When we begin group work, they get close, they explain things to each other. I find this interesting.*

Technopedagogical Competencies

As in Case 1, the Case 2 results show an increase in the use of resources (2.75 to 3.50) and creation/collaboration (1.55 to 2.50). There was no change in the choice of methods scale between the beginning and the end of the project.

The main discoveries this teacher made during the project were Google Docs and Google Spreadsheet. He had had the opportunity to use these resources in a professional context prior to the project. These online and collaborative tools greatly improved support, including support from other teams.

> *I go everywhere in Google Docs and I quickly get to know my students. If they are stuck, nothing stops them from looking at the work of other teams.*

The use of Google's collaborative tools is probably the main reason for the perceived improvement of informational and communication skills, since they are the only collaborative ICT this teacher said he used in his courses. Other applications and equipment were considered, but not tried.

> *I do not think I used technology much. The students often used their computers to look for information. I did not take the time to use surveys and I stopped thinking about other tools. I am not too familiar with them because there is no way to give individual feedback to the students.*

When designing activities, he also drew on many subject-matter resources that the students could use to solve problems in class. He was not afraid to explore new resources, even in front of the students:

> *Anyway, I am older than them. Naturally, they have more computer skills than I do. Sometimes, if I'm stuck, a student helps me. It makes them so proud!*

Like many teachers in the project, he reported that ICTs were sometimes a source of distraction for students. It should be noted that the ALC layout made it more difficult for this teacher to notice disengaged students:

> *As for engagement, the problem in this classroom is that I cannot see everything. If a student is playing with his phone, it is harder to see.*

Adoption of the Innovation

The SoC profiles in **Figure 2** show the relative levels of concern for Case 2 at the beginning of each semester and at the end of the project.

The SoC profiles in **Figure 2** show that personal concerns remained high relative to the other concerns. The profile at the beginning is similar to the model of nonusers proposed by the authors of the SoC guide, except for the relatively small increase in the collaboration and refocusing stages (George et al., 2013).

The high collaboration stage indicates an interest in coordinating and cooperating with other users of the innovation, while the refocusing stage is associated with exploring the more general benefits of the innovation and the possibility of replacing

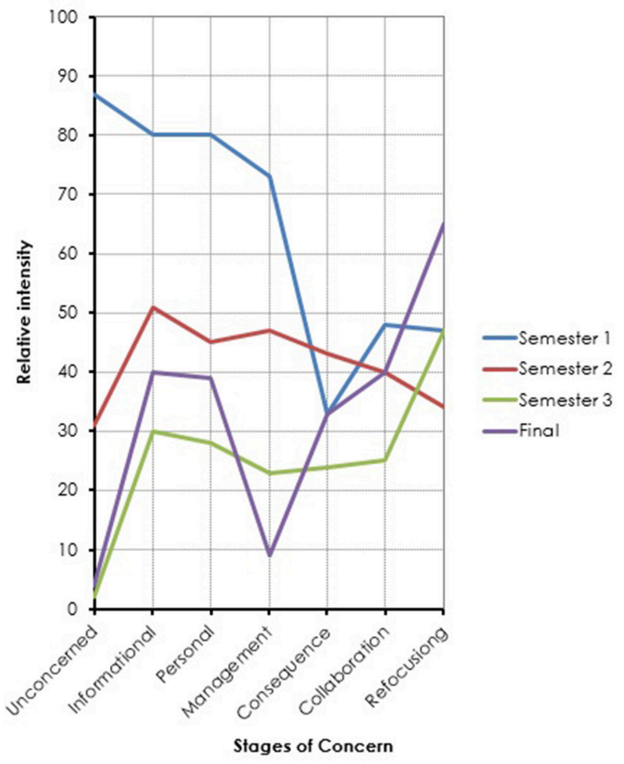

FIGURE 2 | Stages of Concern profiles for Case 2.

it or making major changes to it. While collaboration was encouraged in this project, we found the level of the refocusing stage surprising for a new user of the innovation. In the second term, the profile is similar to an intermediate state of adoption, although the informational stage is lower in the theoretical profile.

As users gain experience, they usually become less concerned about the personal impacts of the innovation, while in the later stages, consequence and collaboration increase. This change did not occur in the third semester and final profiles. Instead, this teacher's personal concerns remained relatively high and we see an increase in the refocusing stage. As shown in the SoC guide, this tailing up at the last stage may indicate resistance to the innovation.

To find signs of doubt that would explain the step rise for the refocusing stage, we can return to the teacher's collaborative preferences. They increased between the beginning and the end of the project and the reported amount of time dedicated to collaborative learning increased. However, the answers to questions about typical activities and possible changes in the way activities were designed revealed that no particular structure for collaboration had been put in place.

> *They do not have to work in teams. In fact, they can do the problems alone, but if they join a team, they must collaborate.*

In short, this teacher was interested in collaboration and used it more often, but there were no design features in place to promote collaboration. With regard to ICT use, plenty of resources were available to the students. He also used technologies himself, but with the exception of Google Docs, his activities were not designed so that students will use technology.

It should also be noted that this teacher invested a lot of effort in lesson design, including a course that could not be given in the ALC. Outside the context of the interview, he mentioned fears of losing access to the ALC once the research project ended. These factors may have contributed to his hesitation to invest further effort in development.

With regard to the LoU aspects of the interviews, the coded segments indicated a level 3 for the first semester and 4A for the remaining semesters. Level 3 refers to mechanical use: the user focuses on the short-term use of the innovation and changes are usually made to meet his needs rather than those of the students. In short, this teacher was probably in a survival state during the first semester. Level 4A is considered routine use: few changes are made and little effort is put into improving the use of the innovation to achieve impact for the students. The teacher's use of music and the students' interests is interesting—only two teachers in the project used ambient music—but in this case, it did not balance the lack of changes in the general activity model.

> *No! I think my activities and approaches work and I intend to use them again this semester.*

In summary, Case 2 was inspired by the classroom and showed positive changes in student- centered ness and collaboration, including collaboration with ICT. He invested a lot of effort in the development of new activities during the first semester but did not explore his design options much. The high refocusing stage in the SoC profile may be explained by awareness of the need to make changes in the pedagogy. In this case, combined interest in collaboration and information could mean that he did not clearly see how these changes could be made. One way to help him would be to propose modifications to his existing activities or, as in Case 1, provide him with examples of ALC use by other teachers.

DISCUSSION

The two cases described in this study are those whose indicators progressed the most toward the idea of adopting student-centered approaches to teaching, collaborative preferences and technopedagogical competencies. The common aspects of this change can be found in the description of the cases.

Develop and Stabilize

In each participating institution, the ALCs were the only places specially designed to facilitate active pedagogies and the use of ICT. These were unique, special and often expensive places. The first semester of use for this innovation revealed a sense of performance for one teacher and a burst of creativity for the other. In both cases, the first semester was associated with a significant phase of development of new learning activities. Case 1 shows us a danger in this rapid expansion, that is, testing too many different teaching methods. This approach requires a lot of effort and the activities include many aspects that the teacher had not had the opportunity to test before. In addition, the many changes in student tasks from one activity to another and the sheer variety of required learning tasks can become confusing for them. After discovering this problem, the teacher chose to focus on fewer types of activities that he repeated and refined. Creating routines seemed to offer some stability, for both the teacher and his students, which is seen in a drop in the management concern on the SoC subscale. It was from this stability that he initiated changes to create a better structure that fostered collaboration and explored new ICT technologies. For the other teacher, the design effort seems to have focused more on the diversity of problems and situations than on the use of different teaching methods. In both cases, informational, personal and collaboration concerns were high, indicating an open mind about ALC use by other teachers.

Development of Active Pedagogies Before ICT

The changes made in the ALC first focused on the use of more active learning methods. Apart from the use of computers to search for information, ICT use was limited on the student side during the first semester for Case 1. Case 2 also integrated Google Docs in the first semester, which Case 1 did in the second semester. Case 1 justified this limited integration of ICTs by the fact that using an ALC requires a lot of changes and that he needed to give himself some time. As both teachers have an interest in collaboration, the integration of ICTs could involve collaboration with other ALC users in the future.

Decrease of Lectures

Once the initial development phase is complete, teachers can improve activities to maximize the impact on the students or keep the business models already developed and use them more often in one semester. A feature that both these teachers shared was to focus on replacing lectures with known models. This feature is most evident in the third semester, where little change occurred in the overall form of activities. Rather than mentioning an interest in selecting teaching methods appropriate to the knowledge that students were supposed to learn, the teachers instead described their design efforts as a replacement for lectures. The absence or weak progression of indicators for the choice of methods supports this observation.

The relatively low level of the consequence stage in the SoC questionnaire results could mean that the teachers were so busy or concerned with the design of their activities or the development of their ICT resources (such as videos) that they were less concerned about the changes needed to maximize the impact on student learning. In short, after a considerable initial phase of development, the teachers seem to have continued their adoption of ALC by focusing on the proportion of learning activities in which active pedagogies were used. This strategy appears to have worked well to the extent that the teachers reported changes in their role as teachers, progressing toward a student-centered approach.

High Scores for Personal Concerns

We observed that personal concerns remain high in the various stages of concern. According to the SoC guide, the profile of a more experienced user shows a decrease in personal concerns, but the teachers mentioned the significant investment of time required to design the activities. Additionally, the use of active pedagogies limited their room for maneuver in what they could do in class. They also mentioned that students engaged in an activity were difficult to stop. In fact, both teachers developed activity models where they gave a presentation at the beginning of the course. Their rapid adoption of active methods and discovery of new limits to their role justify their uncertainties and personal doubts about the innovation.

High relative intensity in the early stages may also be related to concerns for collaboration. Case 1, in particular, revealed this link. He explained that he gained a better understanding of his role as a guide thanks to the example of another teacher. He also mentioned that he could integrate ICT if he saw ideas from other teachers. Unfortunately, neither case collaborated with other ALC users. The fact that the ALCs are new and the lack of a collaborative structure among the teachers in the participating institutions (e.g., community of practice) reduced the opportunities for collaboration. Another participant in the project also mentioned that she had the most time to interact with other ALC users during the teachers' strike days: a strike of a few days took place during the course of this study and it was at that moment that she had the most discussions about her practices.

Increase in the Teacher-Centered Approach

The increase in teacher-centered teaching approach scale scores seems difficult to explain, since the student-centered approach also increased. Although the teachers said they provided more time for teamwork activities, it should be noted that they both set aside time for lectures at the beginning of lessons and they both experienced a major change from their previous teaching approaches. Case 1 outlined his strategy for disseminating content to each team and voiced doubts about the students' ability to keep track of their discussions. Case 2 emphasized individual feedback as a limitation with ICT integration. Interestingly, both teachers had different scores for approaches to teaching based on the setting they were teaching in (Lindblom-Ylänne et al., 2006). A portion of their courses were less affected by the pedagogical changes and their indicators show that they did not fully adopt the innovation. The increase in teacher-centeredness may be related to the portion of their courses given in a traditional lecture format. Maintaining a teacher-centered role may be justified in some lessons that were less affected by the changes after three semesters.

Strengths and Limitations

Numerical indicators combined with interviews helped to clarify several lines of explanation related to the adoption of the ALC. Notably, the interviews showed that extensive changes can be made in pedagogy with small changes in the corresponding indicators.

Case selection is a limit in this study, in that quantitative results do not necessarily reveal users whose practices have evolved the most: there were no systematic observations in class. This study is also limited by the fact that the data are self-reported, so desirability phenomena may have come into play.

Although the descriptions of the two cases share several similarities, this study only describes the experience of two teachers who previously had limited experience with active pedagogies. They were also the first to use a classroom which was unique in their institution: a truly frontier experience for them. To participate in the project, the teachers had to use the ALC beyond a minimum threshold (50% of theoretical lessons), which could have influenced the innovation adoption process, especially during the first semester where several activities had to be implemented to attain the threshold.

CONCLUSION

The purpose of this article was to describe the adoption of an ALC by two teachers whose individual scores with regards to ALC use changed the most positively toward student-centeredness, collaboration and high technopedagogical competencies. These two cases were selected from a sample of 13 teachers offering courses in five different subject matters in three different institutions. Quantitative and qualitative data were used to describe the teachers' adoption process over a period of three semesters.

Both teachers were motivated to develop new activities during the first term, despite the efforts required. After some time, activity models were reused multiple times with the objective of transforming lectures into teamwork activities. Most of the pedagogical changes involved active learning, rather than ICT integration. While the teachers mentioned several uses of technologies, the students mostly used computers to look for information and Google Drive to collaborate. Elevated informational and personal concerns for the use of an ALC indicate that the teachers may not have been comfortable about the change in their role. The increase in teacher-centered approaches supports the idea that the change in the teacher's role is not straightforward. The interest in collaboration in this setting suggests that it would be useful for them to see concrete examples of ALC use in which a teacher assumes solely the role of guide. Teachers may also simply be given more time to adapt to what seem to be complex and demanding changes.

With regard to the results and limitations of this study, teachers who attempt to use an ALC for the first time could aim to develop a routine with their students by implementing a few activity models they are comfortable with. Observing and collaborating with other teachers should be encouraged. Institutions and professionals who collaborate with teachers can facilitate such collaboration. They may consider establishing a community of practice for ALC users (e.g., such as SALTISE in the CEGEP network) or contributing to digital collections of sample activities done in ALC settings.

This study adds to the emerging research on the impact of the ALC on teachers' pedagogy. While the cases share some similarities related to the early phase of development and general priorities, there are differences in their adoption process, especially with the CBAM. Therefore, keeping a variety of indicators could be considered in future research. It could also be useful to verify whether the increase in teacher-approaches and the high level of personal concerns are specific to these cases or a common adoption stage for early ALC users with limited experience in active learning.

AUTHOR CONTRIBUTIONS

SF a doctoral student of BP was the co-researcher in the project. He supervised all the data collection, participated in the research design and conducted the quantitative and qualitative data analysis. He wrote the largest part of the article. BP provided the intellectual leadership and designed all aspects of the study, supervising the research at each point. He also helped plan the article at a high level and validated its writing.

ACKNOWLEDGMENTS

This research received financial support from the PAREA program of the Quebec Ministry of Education (PA2013-012) and from an SSHRC partnership development grant (grant number 890-2012-0052).

REFERENCES

Barron, B. J. S., Schwartz, D. L., Vye, N. J., Moore, A., Petrosino, A., Zech, L., et al. (1998). Doing with understanding: lessons from research on problem- and project-based learning. *J. Learn. Sci.* 7, 271–311.

Barrows, H. S. (1996). Problem-based learning in medicine and beyond: a brief overview. *N. Dir. Teach. Learn.* 1996, 3–12. doi: 10.1002/tl.37219966804

Beichner, R. J., Saul, J. M., Abbott, D. S., Morse, J. J., Duane, L., Allain, R. J., et al. (2007). "Research-based reform of university physics," in *Physics*, eds E. Redish and P. Cooney (College Park, MD: American Association of Physics Teachers), 1–42.

Bérubé, B., and Poellhuber, B. (2005). *Un Référentiel de Compétences Technopédagogiques. Destiné au Personnel Enseignant du Réseau Collégial*. Montréal, QC: Collège de Rosemont.

Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., and Palincsar, A. (1991). Motivating project-based learning: sustaining the doing, supporting the learning. *Educ. Psychol.* 26, 369–398.

Bonwell, C. C., and Sutherland, T. E. (1996). The active learning continuum: choosing activities to engage students in the classroom. *N. Dir. Teach. Learn.* 1996, 3–16. doi: 10.1002/tl.37219966704

Bonwell, C., and Eison, J. (1991). *Active Learning: Creating Excitement in the Classroom. ASHE-ERIC Higher Education Report*. Washington, DC: The George Washington University.

Brooks, C. (2012). Space and consequences: the impact of different formal learning spaces on instructor and student behavior. *J. Learn. Spaces* 1.

Chamberland, G., Lavoie, L., and Marquis, D. (2006). *20 Formules Pédagogiques. Formules Pédagogiques*. Québec City: Presses de l'Université du Québec.

Charles, E. S., Lasry, N., and Whittaker, C. (2011). *Scaling Up Socio-Technological Pedagogies: PAREA Report*. Montréal, QC: Dawson College.

Charles, E. S., Lasry, N., and Whittaker, C. (2013). L'adoption d'environnements sociotechnologiques comme moteur de changement pédagogique. *Pédagogie Collégiale* 26, 4–11. Available online at: http://aqpc.qc.ca/sites/default/files/revue/Kingsbury-25-3-2012.pdf

CLAAC. (2014). *CLAAC: Les classes d'apprentissage actif* (Retrieved 12July, 2017). Available online at http://claac.org/

Crouch, C. H., and Mazur, E. (2001). Peer Instruction: ten years of experience and results. *Am. J. Phys.* 69:970. doi: 10.1119/1.1374249

Dillenbourg, P. (ed.). (1999). "What do you mean by 'collaborative learning'?," in *Collaborative-Learning: Cognitive and Computational Approaches* (Oxford: Elsevier), 1–19.

Dori, Y. J., and Belcher, J. (2005). How does technology-enabled active learning affect undergraduate students' understanding of electromagnetism concepts? *J. Learn. Sci.* 14, 243–279.

Conseil Supérieur de l'éducation (2000). *La Formation du Personnel Enseignant du Collégial: Un Projet Collectif Enraciné Dans le Milieu*. Sainte-Foy: Conseil Supérieur de l'éducation.

George, A. A., Hall, G. E., and Stiegelbauer, S. M. (2013). *Measuring Implementation in Schools: the Stages of Concern Questionnaire (SEDL)*. Austin, TX: SEDL.

Gijbels, D., Dochy, F., Van den Bossche, P., and Segers, M. (2005). Effects of problem-based learning: a meta-analysis from the angle of assessment. *Rev. Educ. Res.* 75, 27–61. doi: 10.3102/00346543075001027

Hall, G. E., Dirksen, D. J., and George, A. A. (2006). *Measuring Implementation in Schools: Levels of Use (SEDL)*. Austin, TX: SEDL.

Hmelo-Silver, C. E. (2004). Problem-based learning: what and how do students learn? *Educ. Psychol. Rev.* 16, 235–266. doi: 10.1023/B:EDPR.0000034022.16470.f3

Hord, S. M., Stiegelbauer, S. M., Hall, G. E., and George, A. A. (2006). *Measuring Implementation in Schools : Innovation Configuration*. Austin, TX:SEDL.

Howden, J., and Kopiec, M. (2000). *Ajouter aux Compétences. Enseigner, Coopérer et Apprendre au Post-Secondaire*. Montréal, QC: Chenelière/McGraw-Hill.

Johnson, D. W., Johnson, R. T., and Smith, K. A. (1998). Cooperative learning returns to college: what evidence is there that it works? *Change* 30, 26–35. doi: 10.1080/00091389809602629

Kapur, M. (2012). Productive failure in learning the concept of variance. *Instr. Sci.* 40, 651–672. doi: 10.1007/s11251-012-9209-6

Kapur, M., Dickson, L., and Yhing, T. P. (2010). Productive failure in mathematical problem solving. *Instr. Sci.* 38, 523–550. doi: 10.1007/s11251-009-9093-x

Kingsbury, F. (2012). Le projet scale-up: une révolution pédagogique qui nous vient du sud. *Pédagogie Collégiale* 25, 37–44.

Kirschner, P. (2001). Using integrated electronic environments for collaborative teaching/learning. *Res. Dialog. Learn. Instr.* 2, 1–9. doi: 10.1016/S0959-475200021-9

Laffey, J., Tupper, T., Musser, D., and Wedman, J. (1998). A computer-mediated support system for project-based learning. *Educ. Technol. Res. Dev.* 46, 73–86. doi: 10.1007/BF02299830

Lindblom-Ylänne, S., Trigwell, K., Nevgi, A., and Ashwin, P. (2006). How approaches to teaching are affected by discipline and teaching context. *Stud. High. Educ.* 31, 285–298. doi: 10.1080/03075070600680539

Macpherson, A. (2007). *Cooperative Learning Group Activities for College Courses: A Guide For Instructors*. Surrey, BC.

Miles, M. B., and Huberman, A. M. (1994). *Qualitative Data Analysis, An expanded Sourcebook*. Londres: Sage Publications.

OCDE. (2008). *Les Grandes Mutations Qui Transforment l'Éducation*, Centre pour la recherche et l'innovation dans l'enseignement Ed.

Owens, L., and Barnes, J. (1992). *Learning Preference Scales: Handbook and Test Master Set*. Melbourne, VIC: Australian Council for Education Research Ltd.

Prince, M. (2004). Does active learning work? A review of the research. *J. Eng. Educ.* 93, 223–231. doi: 10.1002/j.2168-9830.2004.tb00809.x

Savery, J. R., and Duffy, T. M. (eds.). (1996). "Problem based learning: an instructional model and its constructivist framework," in *Constructivist Learning Environments: Case Studies in Instructional Design* (Englewood Cliffs, NJ: Educational technology Publications), 135–148.

Slavin, R. E. (1996). Research on cooperative learning and achievement: what we know, what we need to know. *Contemp. Educ. Psychol.* 21, 43–69. doi: 10.1006/ceps.1996.0004

Spiegelberg, H. (1978). *The Phenomenological Movement. A Historical Introduction*, *2nd Edn*. The Hague: Martinus Nijhoff.

Trigwell, K., and Prosser, M. (2004). Development and use of the approaches to teaching inventory. *Educ. Psychol. Rev.* 16, 409–424. doi: 10.1007/s10648-004-0007-9

Trigwell, K., Prosser, M., and Ginns, P. (2005). Phenomenographic pedagogy and a revised approaches to teaching inventory. *Higher Educ. Res. Dev.* 24, 349–360. doi: 10.1080/07294360500284730

Wilkerson, L. (1996). Tutors and small groups in problem-based learning: lessons from the literature. *N. Dir. Teach. Learn.* 1996, 23–32. doi: 10.1002/tl.37219966806

Yukselturk, E., and Bulut, S. (2007). Predictors for student success in an online course. *Educ. Technol. Soc.* 10, 71–83.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.