Mixing In-Class and Online Learning: Content Meta-Analysis of Outcomes for Hybrid, Blended, and Flipped Courses

Lauren E. Margulieux, Georgia Institute of Technology, l.marg@gatech.edu  
W. Michael McCracken, Georgia Institute of Technology, mike@cc.gatech.edu  
Richard Catrambone, Georgia Institute of Technology, rc7@prism.gatech.edu

Abstract: Over the past 15 years, courses that mix face-to-face and online instructional methods, such as blended, hybrid, and flipped courses, have gained both supporters and skeptics in higher education. Studies that compare mixed courses to face-to-face or online courses have conflicting results: some find improved learning outcomes and some find no significant differences. We contend that these conflicting results are due to inconsistent or vague definitions of hybrid, blended, and flipped. To address this problem, we use the definitions from a recently proposed taxonomy to reclassify studies in the literature. After reclassification, analysis of this literature reveals two main themes that illuminate how mixed instructional methods affect learning outcomes. Courses that use mixed methods can either reduce time in class and maintain learning outcomes or maintain time in class and improve learning outcomes.

Keywords: hybrid, blended, flipped, inverted, content meta-analysis

Introduction
Since 2000, a growing group of educators has been interested in mixing face-to-face and online instructional methods. Mixed method courses, commonly called hybrid, blended, or flipped (which is sometimes called inverted), have both face-to-face and online components. Because they employ pedagogical resources supported by both instructors and technology, they allow students to receive more instruction without substantially increasing the workload of instructors (Banerjee, 2011, U.S. Department of Education, 2010). For example, students in a traditional math course typically receive content in class (i.e., through lectures) and solve problems for homework. In contrast, students in a mixed method math course receive content before class through videos and solve problems during class with instructor feedback. Providing feedback during application activities, like problem solving, has been identified as a critical component of education (National Research Council, 2011) and a service that will keep universities relevant after all content is easily available online for free (Bok, 2006). Many instructors, however, feel that they need to lecture to ensure student understanding of course content. Mixed method courses allow instructors to provide both feedback and lectures, supporting learners when they both receive and apply content.

Much research has been conducted in the past several years to assess the effectiveness of mixed method courses, but the results of that research have been inconclusive: many mixed method courses improved learning and just as many did not. The differences between courses that improved outcomes and those that did not are unclear due to the ill-defined terms used to describe the courses that were evaluated. For example, the term “blended” has been used to describe a course in which students learn content before class and practice applying content in class (Melton, Graf, & Chopak-Foss, 2009) as well as a course in which half of the lectures are delivered in class and the other half are delivered online (Gerlich & Sollosy, 2009). The pedagogy of these courses is different, but they are classified as the same type of course.

The inconsistent definitions make comparing results, replicating experiments, implementing course design, and finding and understanding information from the literature difficult. To address this issue, Margulieux, Bujak, McCracken, and Majerich (2014) proposed a taxonomy that used pedagogically relevant dimensions to define terms. The taxonomy classified the design of courses based on the type of instruction that was delivered (either didactic exposition of content or feedback on application of content) and how that instruction was delivered (either via an instructor or via technology). The taxonomy defines a hybrid as a course that is delivered via both instructor and technology and primarily delivers one type of instruction. It differentiates between lecture hybrids in which instruction primarily delivers content and practice hybrids in which instruction primarily provides feedback. In blended courses, instruction is delivered via both instructor and technology and includes both content exposition and feedback. The most common type of blended course, a flipped blend, delivers content via technology and provides feedback via instructor. Supplemental blends deliver content via instructor and provide feedback via technology, and replacement blends deliver content via both instructor and technology and provide feedback via both instructor and technology. We used these definitions to
reclassify studies of mixed methods courses. Based on these new classifications, the results of the studies were reinterpreted to identify themes in the literature that were previously unrecognized.

**Analysis**
We employed a content meta-analysis methodology. Content meta-analyses systematically aggregate information from a number of studies but use a qualitative approach instead of a quantitative approach (Jeong, Hmelo-Silver, & Yu, 2014). Given the large variations in research methodology and quantitative data sources (e.g., grades on exams, projects, or concept inventories) of the selected papers and inconsistent reporting of key measurements (e.g., sample size), a qualitative approach was more appropriate than a quantitative approach.

**Paper selection**
To find relevant papers, the ERIC, Proquest Education Journals, Academic Search Complete databases, and Google Scholar were queried for permutations of the terms “hybrid,” “blended,” “flipped,” and “inverted” with the terms “class,” “classroom,” “course,” and “learning” in the title or abstract. The title or abstract also had to include “comparison,” “experiment,” “evaluation,” or “performance.” Articles that met these criteria were considered for inclusion. If their abstracts did not mention student “outcomes,” “knowledge,” “achievement,” or “grades,” the articles were excluded. Studies also must have included a control group that was the previous (traditional) version of the course. This analysis includes only research that reported learning outcomes. Much of the research and reviews on mixed method courses have focused on student and instructor perceptions instead of learning outcomes (Ginns & Ellis, 2007), but outcomes are imperative to determine the efficacy of mixed methods. This analysis also focuses on higher education, so only studies of for-credit, higher education courses were included. Measures of learning outcomes must have been equivalent in experimental and control groups.

**Reclassification of studies**
The 49 selected studies (out of 163 considered studies) were reviewed to identify pedagogical components of courses. The designs of the mixed method and traditional (control) courses were coded for how instruction was delivered to students (via an instructor or via technology) and what type of instruction was delivered (exposition to content or feedback during application of content). The difference between course designs was coded for changes in delivery medium, instruction type, and time spent in class. Courses were from a range of domains and mostly from American higher education institutions.

Of 17 courses that were reported as hybrid in the literature, 5 were reclassified as a type of hybrid, 10 as a type of blend, and 2 as other types of courses. Of the 11 courses that were reported as blended, 5 were reclassified as a type of hybrid, 5 as a type of blend, and 1 as another type of course. This array of reclassifications suggests that courses reported as hybrid and blended included several types of courses that differed on the fundamental features of the course. Of the 21 courses that were reported as flipped or inverted, all but 4 were reclassified as a flipped blend (including four with an additional in-class lecture component). Figure 1 summarizes the theme of results for each type of hybrid and blended course. The only type of mixed method course that consistently improved learning outcomes was the flipped blend.

![Figure 1. Mixed method courses categorized by design and split by reported learning outcomes.](image)

**Analysis of differences between mixed method and traditional courses**
To explore why mixed method courses did or did not improve learning outcomes, the differences between mixed method courses and traditional courses were considered.
**Delivery medium**

Of mixed method courses that changed only the delivery medium (e.g., lecture hybrids and replacement blends) from the traditional courses, 79% (15 out of 19) did not report a change in learning outcomes. The four studies that reported improved learning outcomes argued that asynchronous delivery of instruction was beneficial to student learning. This argument is supported by a meta-analysis that compared face-to-face courses, which are inherently synchronous, to synchronous and asynchronous online courses. Bernard et al. (2004) found that students in asynchronous online courses performed slightly better than students in face-to-face courses, and students in face-to-face courses performed slightly better than students in synchronous online courses. They argued that asynchronicity allowed students to receive content at their own pace and to reflect more on their answers while applying that content (e.g., in discussion forums).

The effect of delivery medium was at the center of the educational media debate in the 1990s. In this debate, Clark (1994) took the point of view that media is merely a vessel that delivers information, and properties of the information and the learner are those that affect learning. Taking the opposite view, Kozma (1994) argued that media have many properties that affect learning. The prevailing view today is that different types of media have different affordances. Though some types of media are better suited for some types of learning, several types of media can be equally effective for several different types of learning (Ainsworth, 2006). Therefore, unless an instructor uses an ineffective medium for instruction, media should not have a large effect on learning outcomes.

Though online instruction generally did not improve learning, the lack of a difference in learning due to delivery medium changes is an important finding. If technology can deliver instruction with the same efficacy as instructors, then technology can be used as a resource to either supplement face-to-face instruction or reduce the amount of time students need to be in class. Possible benefits of using technology to supplement instructors include increasing the quality of instruction (by increasing the resources available to students) and the accessibility of instruction (by reducing the amount of class time required for a course). The goals of using mixed methods depend on the instructor, students, course, and institution, but course outcomes should not be negatively impacted by appropriately delivered instruction via technology.

**Type of instruction**

A feature of mixed method courses that made a consistent impact on learning outcomes was type of instruction. Of mixed method courses that added instruction during application of content to traditional courses, 77% (23 out of 30) reported improved learning outcomes. That percentage increases to 88% (23 out of 26) if the four courses that already had feedback during application and simply added more are not included. The majority of classes that added instruction during application and reported improved learning (17 of the 23) were flipped courses. These classes typically have recorded video lectures to be viewed before class and then application activities in class completed in small groups and with a instructor’s (and sometimes teaching assistants’) feedback. Only 4 of the 21 flipped courses did not report improved learning outcomes.

A range of learning theories and frameworks supports the benefits of in-class application activities. Theories based on active learning (the type of learning that requires students to play an active role in education by answering questions, problem solving, etc.) argue that applying content helps students learn more efficiently and deeply. A meta-analysis of 225 studies about active learning in STEM courses found that performance on exams and concept inventories was, on average, 47 standard deviations higher for courses that had some active learning than courses that had traditional lecture only (Freeman et al., 2014). Furthermore, in-class application activities require all (or at least most) students to participate, as opposed to in-class lectures, in which typically only highly motivated students participate (Crouch & Mazur, 2001). Moreover, Black and Wiliam (1998) argue that application activities provide opportunities for students and instructors to interact, allowing students to receive feedback on their grasp of the content and instructors to receive feedback on the efficacy of their instruction. Similarly, experiential learning frameworks are based on students applying content with an instructor present to give feedback. The central belief of experiential learning is that students learn best when they construct knowledge by integrating new content with prior knowledge, and one of the most effective ways to achieve this integration is by allowing students to direct their learning while working through application activities (Hmelo-Silver, 2004).

Feedback during application activities is important for learning. The question for mixed method courses becomes, does this type of instruction need to take place in the classroom? To address this question, the courses that added technology-mediated application were further analyzed. Of the four supplemental blends (i.e., lectures in class and technology-supported application outside of class), two of them reported learning improvements and two of them did not. The two that reported improvements asked students to use technology to practice recurrent skills (skills that are always executed in the same way), such as practicing conjugation for a
language class with vocabulary drills. The two that reported equivalent outcomes asked students to use technology to practice non-recurrent skills (skills that are executed differently depending on the application). In addition, two other courses continued application activities online that started in class (e.g., continued a discussion that started in class), and they both reported improved learning outcomes.

Based on these findings, one tentative conclusion is that technology might effectively support some application activities but not others. Jia, Chen, Ding, and Ruan (2012) argued that technology can support application activities that would be repetitive and time-consuming for an instructor to support. Technology might even be better in these cases because it typically provides feedback more quickly than instructors, leading to higher student satisfaction (Gikandi, Morrow, & Davis, 2011). If a theme can be found in these six studies, it would support Jia et al.’s (2012) argument by suggesting that technology-supported applications are more successful when the applications are repetitive, like practice drills or a continuation of an in-class activity.

For the two studies that did not find learning improvements, both asked students to solve problems (that used non-recurrent skills) with feedback exclusively from a computer program. The nature of instructional support that students received from these programs was unclear, but based on the predominately positive findings from flipped courses and the neutral findings from these courses, it is likely not equivalent to in-class support that students in flipped courses received. In a review that compared human tutoring to computer tutoring, VanLehn (2011) found that answer-based computer tutors (i.e., those that indicate only whether the final answer is correct or not) are less effective than step-based computer tutors and human tutors (i.e., those that indicate whether each step the student took is correct). This difference, VanLehn argues, is due to the granularity of feedback and scaffolding that students receive. Because students using step-based computer tutors or human tutors receive information about each step that they took in the problem solving process, they can identify and repair faulty logic or misconceptions more easily. In their 4C/ID model of complex learning, van Merrienboer, Clark, and de Croock (2002) argue that students need this type of support while learning non-recurrent skills, but not necessarily while learning recurrent skills. The technology used in these courses was likely an answer-based computer tutor and might not have provided sufficient support for non-recurrent skill building.

Time in class
Because of the increased use of technology outside of the classroom, instructors of mixed method courses commonly underestimate the amount of time students will spend on the course, resulting in a more time-consuming course (i.e., “a course and a half”). Though many of the studies in this review did not directly measure time spent on the course outside of class, many did reduce the amount of time students spent in class to accommodate additional coursework outside of class. Nearly half (22) of the studies decreased time spent in class for the mixed method course, and the majority (18 courses, 82%) of these courses did not report improved learning outcomes. These results suggest that courses can reduce time spent in class without negatively impacting learning.

Of the 27 mixed method courses that did not reduce time spent in class, most (85%) reported improved learning outcomes. Though these studies did not reduce class time, approximately half of them reported efforts to keep the workload of the students in the mixed method course equal to that of the students in the traditional course. It is possible, however, that improved learning outcomes are partially caused by a greater workload. Without more research, it is difficult to speculate on the effect size of workload, but these findings suggests that time in class is valuable for learning outcomes. Whether that value comes from face-to-face interactions with the instructor, collaboration with other students, or the culture of the learning environment is much speculated upon by the educational community and in need of additional research.

How mixed methods affect learning
Regardless of instructional method, students generally get the same two types of instruction: course content and application activities. For example, to learn about a law in a physics course, students generally get a lecture and a reading about that law and complete homework problems to ensure that they can apply their knowledge. To learn about an historical event in a history course, students generally get a lecture and a reading about that event and discuss or write about its impact to ensure they can apply their knowledge. Given these general constants, this review analyzed how mixed methods of instruction affected learning. The results suggest that, in general, mixed method courses can maintain learning outcomes for a course while reducing time in class, or they can maintain time in class and improve learning outcomes.

This review found that adding feedback during application of content improved learning outcomes. Courses that improved learning outcomes did not typically add additional application activities, they added only the feedback that students received while completing the application activities. This finding suggests that giving
students application activities is more effective when feedback is provided. In addition, because these courses used multiple methods of instruction, they were not forced to reduce the amount of content covered in didactic lectures when they added feedback.

A common approach to teaching, especially in STEM domains, is to provide only didactic instruction during class and to assign application activities to be completed outside of class. This approach, however, does not typically provide the structure and support that many students need to be most successful when they start applying content (Baeten et al., 2013). Alternatively, experiential learning (e.g., problem-based learning) and constructivism focus on providing feedback during application activities and charging students with identifying and learning the content that they need to know (even if that involves asking the instructor or TA). Proponents of experiential and constructivist learning argue that students have different backgrounds and bring different prior knowledge to the classroom; therefore, the most effective way to help students build new knowledge is to guide them through the process of gathering information and constructing that knowledge rather than providing information based on how the instructor organizes it (Baeten et al., 2013; Hmelo-Silver, 2004; Jonassen, 1999). Research on experiential and constructivist learning methods, however, do not consistently support this argument (e.g., Baeten et al., 2013; Cennamo et al., 2011; Hmelo-Silver, 2004), suggesting that providing instruction through only feedback is not necessarily better than providing instruction through only lectures.

Kirschner, Sweller, and Clark (2006) argue that this inconsistent success of experiential learning is due to high cognitive load of students during application activities. More specifically, they argue that if students are trying to organize new information at the same time that they are attempting to apply that information, then the cognitive load associated with the task is too high to promote learning. Cognitive load can be further taxed when students work in groups, especially if they are expected to provide feedback to their peers (Ching & Hsu, 2013). Instead of starting with application of information, Kirschner et al. (2006) suggest that students should be taught a basic level of information and how to organize it to provide necessary cognitive structure for a topic before starting application activities. This suggestion aligns with Maki and Maki’s (2002) findings that unstructured courses, whether they were face-to-face or online, were less successful than structured courses.

Because students in online learning environments do not necessarily have immediate access to an instructor and therefore cannot ask questions, structure in online learning environments is perhaps more important than in face-to-face learning environments. For example, Xie and Bradshaw (2014) found that, for successful online discussions, well-defined tasks and moderation provided necessary structure. Perhaps for online problem solving tasks, the same type of structure is necessary. We found that courses that used software to facilitate problem solving were more effective than unaided problem solving perhaps because the software did not provide appropriate (in terms of type, quantity, or quality) structure, feedback, or scaffolding during problem solving. As VanLehn (2011) argued, when student receive feedback for each problem solving step that they take, they can more effectively identify incorrect thinking and fix it. In addition, scaffolding is a method used to provide extra structure and support while students are learning to complete application activities, especially with problem solving. It is intended to allow novices to successfully complete activities of which they would otherwise not be capable. For example, if a problem solution involved five steps, a scaffolded problem might provide two of those steps for the learner. As the student becomes better at solving problems, the scaffolding is incrementally removed until learners achieve the solutions independently.

Scaffolding accelerates learning by helping students to complete activities that they would otherwise not be able to do, such as complex or authentic tasks (Hmelo & Guzdial, 1996). Though scaffolding is typically pre-prescribed and static, it can be highly effective in a more dynamic role because the level of support needed differs among learners and providing too little or too much support inhibits learning (Pea, 2004), but few technologies are capable of dynamic scaffolding. Perhaps one of the reasons flipped blends improved learning outcomes is because, during application activities, instructors can provide dynamic scaffolding that gives more support when necessary. If the instructor is readily available to help students, then more challenging application activities (e.g., in terms of transfer or complexity) could be assigned than if students were given those activities to be completed independently. These classes also provided didactic lectures to help students learn and organize content, but the mixed methods of the course allowed this instruction to be delivered at the convenience of the student and did not take class time.

In summary, this review suggests that mixed methods can help educators achieve two main goals. If instructors are trying to improve the learning outcomes of their course, then mixed methods can help by allowing students to receive instruction both inside and outside of class. If instructors are trying to reduce the resources needed to teach a course, then mixed methods can help by providing instruction through technology and reducing time spent in class. By better understanding the potential of mixed method courses, instructors can better employ these methods to fulfill the needs of their students.
References
Ching, Y.-H. & Hsu, Y.-C. (2013). Peer feedback to facilitate project-based learning in an online environment. International Review of Research in Open and Distance Learning, 14(5), 258-276.


Marcey, D. & Brint M. (2012). Transforming an undergraduate introductory biology course through cinematic lectures and inverted classes: A preliminary assessment of the CLIC model of the flipped classroom.


**Acknowledgments**

We would like to thank Keith Bujak for his consultation and Frank Durso, Mark Guzdial, David Majerich, Wendy Rogers, and Wendy Newssetter for critiques on earlier versions of the manuscript.