CrossMark

Development of an instructional design model for flipped learning in higher education

Jihyun Lee[1] · **Cheolil Lim**[2] · **Hyeonsu Kim**[2]

Abstract In response to pedagogical challenges in higher education, blended learning has become a prevalent practice in colleges and universities. Flipped learning (FL) represents a newly emerging form of blended learning, where students individually watch online lectures prior to class and then engage in classroom learning activities interacting with peers and instructors. Although the conceptual framework of FL may be intuitively appealing, its design and implementation involve considerable intricacy. The purpose of this study was to develop a FL design model for higher education that could systematically guide instructors or designers in creating an appropriate blend of individualized online lectures and collaborative face-to-face learning activities. Using an established method for model development research, a theoretically constructed initial model was iteratively improved and underwent internal and external validation through model usability tests, expert review, and field evaluation. The implementation of an algebra class in a Korean university designed by the final model resulted in significant increases in the maturity of mathematical views, quality of reflections, and satisfactions of students. Features of the model are discussed, along with theoretical and practical implications and suggestions for further research.

✉ Cheolil Lim
  chlim@snu.ac.kr

  Jihyun Lee
  leeji1@snu.ac.kr

  Hyeonsu Kim
  midori@snu.ac.kr

[1] School of Dentistry, Seoul National University, 101 Daehak-ro, Jongro-gu, Seoul 03080, South Korea

[2] Department of Education, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, South Korea

🖄 Springer

Introduction

Colleges and universities have made ongoing calls for innovation in higher education, where for decades knowledge has been transmitted mainly through traditional lectures, despite continuous quality concerns (Arum et al. 2012). Numerous studies indicated that students do not learn critical, creative thinking, or complex reasoning skills during intellectually unstimulating classroom experiences (e.g. Arum and Roska 2011; Garrison and Kanuka 2004; McLaughlin et al. 2014). Since the early 2000s, blended learning (BL)—a convergence of face-to-face (F2F) classroom learning with technology-mediated online learning—has emerged in response to such challenges and become a prevalent pedagogical practice (Ginns and Ellis 2007; Graham et al. 2013). Typically implemented as a combination of synchronous classroom lecture and asynchronous text-based discussion forum (Graham 2006; Kerres and de Witt 2003), BL is intended to offer a more interactive and flexible learning environment for students that fosters core values of higher education such as advanced research and scholarship or critical and creative knowledge claims (Garrison and Vaughan 2008; Garrison and Kanuka 2004; Kerres and de Witt 2003; Lee and Jang 2014).

A crucial issue in the design of BL experiences is designing the right blend of learning online and F2F (Allen et al. 2007; Garrison and Kanuka 2004; Gedik et al. 2013; Hofmann 2002). Blending the two modes of learning does not mean simply combining the two or laying one on top of the other, as frequently occurs, but rather integrating the two by maximizing the advantages of each environment. This sort of complementary relationship requires recreating and reorganizing nearly the entire learning experience, however. The typical mix of blended learning involves adding supplementary online discussion activities to the classroom lecture. Furthermore, such F2F lectures have remained unchanged leaving learners passive recipients, and online activities have not engaged learners as intended (Arum and Roska 2011; Garrison and Kanuka 2004). Many researchers have critiqued that such a mix achieves neither the tremendous benefit of effective human interaction in the classroom nor the advantages of online systems and technological affordances for individualized pacing and flexible access to learning resources (Huon et al. 2007; López-Pérez et al. 2013; Lou et al. 2006; Orton-Johnson, 2009). The current improper mix of blended learning activities, then, calls for a better blend.

As a newly emerging type of BL, flipped learning (FL) represents a means to such a better blend. FL combines asynchronous online lectures that individual students study outside of class with F2F classroom learning activities in which students interact with peers and instructors (Bergmann and Sams 2013; Bishop and Verleger 2013; Goodwin and Miller 2013; Hamdan et al. 2013; Herreid and Schiller 2013; Mason et al. 2013; McLaughlin, et al., 2014; Milman 2013; Strayer, 2012). The driving questions to suggest FL were "when students are struggling and seeking help from their instructor or peers?", and "how instructors can help students while optimizing the technological benefits of an online learning environment and the in-person benefits of human interactions in the classroom?" (Bergmann and Sams 2012). Classroom time is devoted to discovering and sharing ideas with one-to-one assistance, scaffolding, and inspiration, all made feasible by offloading content delivery onto the online lectures that are better at visual representations and self-pacing (Bush 2013; Goodwin and Miller 2013). In other words, FL saves for human instructors to do what only they can do best, and leaves to the technology to do what it can do best. This way of blending posits that neither component has a

supplementary role, but rather each is a core complementary constituent of the learning experience as a coherent whole (Stannard 2012).

Yet, although the conceptual framework of FL may be intuitively appealing, there is considerable intricacy in its design and implementation, and practically infinite design variability in diverse contexts (Bergmann and Sams 2013; Garrison and Kanuka 2004; Overmyer 2012). Instructors and designers may find designing activities for a flipped learning classroom to be a daunting challenge for several reasons. First, online lectures must be designed so that their instructional and technical qualities are adequate enough to prepare students for more engaging classroom activities. Yet, even if the lectures are of good quality, students may not view or comprehend them, perhaps due to distractions online, a lack of self-regulation, or inadequate live support by the instructor (Bergmann and Sams 2012; Hamdan et al. 2013; Milman 2013). Second, the design of F2F requires greater preparation and deliberation in order to provide the interactive learning experiences that achieve the traditional values of higher education (Arum et al. 2012; McLaughlin et al. 2014). Most importantly, the design needs to create a tight link between what students do in class and what they do at home (Stannard 2012), a link that is not addressed by existing models that prescribe separate design processes for online lecture and classroom learning activities. Thus, though FL offers many advantages, its potential may be compromised by inadequate design and implementation, resulting in a corresponding lack of improvement in learning practices and outcomes.

The purpose of this study was to develop a FL design model to systematically and effectively guide designers or instructors throughout the design and implementation process of FL at the higher education level. This FL design model is intended to result in an appropriate blend of individualized online and collaborative F2F learning activities that coherently harmonize meaningful learning experiences for specific needs and contexts. Models make possible a systematic approach to design in that they intentionally lead designers to balance considerations of varied critical factors (Daly et al. 2012). Models incorporate both theoretical and empirical research in related fields (Dick 1997; Smith and Boling 2009). Further, the iterative process of evaluation and modification leads to improvement of practice, pointing to the potential of design to stimulated continued progress in a larger context (Daly et al. 2012; Smith and Boling 2009). The FL design model developed for this study is prescriptive by nature. It proposes when and how particular elements of FL should be selected and designed in a learning arrangement. Up to now, few prior publications on BL or FL have investigated closely the full design process (Halverson et al. 2014). The process for constructing and validating an FL design model reported in this study, and its implications for improvement course-level practice and educational innovation are thought to contribute to the growing literature on this meaning approach to learning.

Review of literature

Studies on FL began to appear around 2012, after FL practitioners had begun to share anecdotal reports with other (Goodwin and Miller 2013). Even if FL takes a fresh approach to learning by redefining class time and the role of technology, however, the pedagogy underlying FL is not entirely novel (Milman 2013). Numerous distance education and e-learning researchers have studied online lecture, just as those researching collaborative/cooperative learning, inquiry-oriented learning, or problem-based learning have

studied F2F learner-centered learning. Team-based learning of Michaelsen, Knight, and Fink (2004) even proposed similar format to FL, but with more emphasis on the F2F part. Both lines of literature are relevant to FL because its success depends on defining the right blend of online and F2F sessions and arranging each ingredient in such a way as to build an integrated optimum learning experience (Halverson et al. 2014; Singh 2003).

The development of instructional design models specific to FL may provide invaluable assistance in the instructional design process. The steps in the instructional design process are based on knowledge of what creates a successful design product, knowledge that is primarily derived from related theoretical implications (Reigeluth 1992; Richey 1986). The BL literature and the burgeoning body of FL literature provide such insights, and can inform the development of a design model for FL. BL studies can inform decisions about the best blend of the two FL components, while FL studies can offer detailed and practical implications for design within each of the components.

Studies on blended learning

Although typical forms of BL include a mix of face-to-face instructor-led lectures and supplementary online activities (Garrison and Kanuka 2004; Ginns and Ellis 2007; Kember et al. 2010; Kerres and de Witt 2003), BL scholars and practitioners have long sought to identify the optimal blend of the two learning modes. According to a thematic synthesis of BL research performed by Halverson et al. (2014), publications on BL from 2000 through 2011 mostly questioned instructional design issues like models, strategies, and course structure.

The didactical 3C model developed by Kerres and de Witt (2003) involves a framework specifying three components of BL and corresponding learning tasks. The three components of BL were: (1) a *content component* that makes factual learning content available to learners, (2) a *communication component* that provides interpersonal interactions for more complex or arguable learning tasks between learners or learners and instructors, and (3) a *constructive component* that facilitates learners' active engagement in most complex learning tasks. In this 3C model, the three ingredients are arranged according to the complexities of the learning tasks, and the weight of each ingredient depending on the learning objectives.

Several theories have addressed how these components of the 3C model can best be arranged using available media. One is media richness theory of Daft and Lengel (1984) and Rice (1992), which categorizes media into richer media (media with richer sensory input, immediate feedback or personal focus, for example, face to face meeting) and less rich media (media with less rich sensory input, delayed feedback or shared focus, for example, online discussion board), and learning tasks into equivocal tasks and unequivocal tasks. Media richness theory proposes that equivocal learning tasks with multiple or conflicting interpretations or solutions are better supported by richer media, while unequivocal tasks with low haziness (such as learning factual knowledge) may be allocated to less rich media.

Another theory that informs the allocation of BL components and tasks involves the cost of media/communication (Hollingshead, 1996; Kerres and de Witt 2003). Cost, then, is seen as related to the dependency of learning on time and space (asynchronous or synchronous), and the directionality of communication (uni-directional transmission or bi-/multi-directional interaction) in the learning process. In other words, the more dependent learning is on time and location, and the more directions in which interactions occur in the learning experience, the greater the cost of media or communication. Thus,

face-to-face classroom learning would be a more expensive mode of communication, and would be recommended for intensively challenging, high-benefit learning activities.

Media synchronicity theory of Dennis and Valacich (1999) considers the dependency of time and location and different modes of communication, and proposes a match of synchronous and asynchronous settings and two communication processes that lead to either conveyance or convergence. That is, in a synchronous environment where learners work together with a shared focus at the same time and place, learning should be designed to facilitate a convergence process that reaches a common understanding. Conversely, in an asynchronous environment where learners work and shared information regardless the constraints of time and place, learning is more reasonably designed for conveyance, that is, the exchange of information. Table 1 summarizes the blending arrangement that related theories suggest.

Studies on flipped learning

The FL research literature may be divided into studies with conceptual or empirical discussions. The conceptual studies concern definitions, comparisons with traditional approaches, educational rationales, and subject appropriateness of FL (e.g. Bergmann and Sams 2013; Bishop and Verleger 2013; Bush 2013; Mason et al. 2013; Milman 2013), while the empirical studies typically report on the effectiveness of actual cases of FL instruction and include implications for design that can guide potential practitioners (e.g. Covil et al. 2013; Herreid and Schiller 2013; Slomanson 2014; Strayer 2012; Talbert 2012).

The term FL derives from Baker's (2000) phrase "*the classroom flipped*" (p. 9). More recently, Bergmann and Sams (2012) described FL in this way: in this way: what is traditionally done in class is done at home, and what is traditionally done as homework is done in class. Learners begin at home, learning with video lectures or screencasts by

Table 1 Blending suggestions in related BL theories

	Face-to-face	Online
Didactical 3C model (Kerres and Witt 2003)	Communication component that provides interpersonal interactions for more complex or arguable learning tasks Constructive component that facilitates learners to actively engage in most complex learning tasks	Content component that makes factual learning content available to learners
Cost of communication (Hollingshead, 1996; Kerres and Witt 2003)	Dependent on time and location Higher cost Bi-/multi-directional interaction	Less dependent on time and location Lower cost
Media richness theory (Daft and Lengel 1984; Rice, 1992)	Intensive high benefit challenging learning activities Equivocal learning tasks with multiple interpretations	Uni-directional transmission Less intensive, less challenging learning activities Unequivocal tasks with low haziness
Media synchronicity theory (Dennis and Valacich 1999)	Synchronous environment Convergence process for learners' shared focus and a common understanding at the same time	Asynchronous environment Conveyance process for the exchange of information

themselves prior to the class, and then engage in enriching activities that help them apply content at a deeper level (Bergmann and Sams 2013; Collins et al. 2001; Covil et al. 2013; Gannod et al. 2008; Lage et al. 2000; Strayer 2012). This way of blending maximizes the advantages of two modes of learning environment. A great strength of the online learning environment is that it can readily provide visual representations like space simulations or comparisons and allow for self-pacing (Bush 2013; Goodwin and Miller 2013). The F2F learning environment can be undoubtedly optimal media for group work, idea sharing with one-to-one assistance, immediate scaffolding, and individual inspiration. Thus, by flipping or inverting the allocation of time and locations for learning, FL combines two types of learning activities—active problem-solving learning activities and direct instruction/mastery learning activities (Bishop and Verleger 2013)—while making both approaches core components of a coherent and meaningful learning experience (Stannard 2012).

Most FL literature stresses that the most critical issue in making FL successful is structuring online and F2F learning experiences so that each component coherently supports the other at the macro-course level and micro-lesson level (Ginns and Ellis 2007). Most researchers commenting on the macro-course level emphasize the need for a much tighter link between the online and F2F modes of learning (Bush, 2013; Covil et al. 2013; Mason et al. 2013; Stannard 2012; Strayer 2012). In order to achieve a desirable macro-structure, the lecture part of learning should be kept to a minimum and the inquiry, discovery, and application parts of the learning process should be allocated maximum time (Stannard 2012). Technology should be used for what it can do better than humans, saving for instructors what only humans can do (Bush 2013). These recommendations largely correspond with recommendations in the BL literature. In addition, the FL literature specifically addresses the need for a consistent course structure with thorough descriptions of the course and design rationales in the syllabus, as well as an introductory orientation that addresses learners' potential fear or resistance to a new method (Herreid and Schiller 2013; Mason et al. 2013; McLaughlin et al. 2014). Assessment of FL courses should include a variety of instruments designed to evaluate higher-order thinking skills like critical and creative thinking skills and problem-solving skills along with content mastery (McLaughlin et al. 2014). Assessment also should provide learners with multiple ways to demonstrate their understanding (Milman 2013).

Studies investigating the design of online sessions contain three important concerns: (1) physical features of online lectures, such as length, speed, and auditory quality (Bush 2013; Goodwin and Miller 2013; Hattie 2008; Khan, 2012; Mason et al. 2013; Smith and McDonald 2013); (2) content features of online lectures, such as proper allocation of online portions, interactivity, and clarity (Bush 2013; Goodwin and Miller 2013; Milman 2013; Shim, 2013; Smith and McDonald 2013); and (3) logistic features outside of the online lecture, such as formative evaluation, scheduling, and after-activity (Bush 2013; Mason et al. 2013; Slomanson 2014; Talbert 2012). In order to retain the learners' attention the online lecture should be less than twenty minutes (Mason et al., 2013; Smith and McDonald 2013) and even less than ten minutes if it is a static screencast that does not show the instructor's face (Khan 2012). Further, creating one clip per topic is better for further reorganizing and modularizing (Smith and McDonald 2013). The clip should have a function for controlling speed (Bush 2013; Goodwin and Miller 2013; Hattie 2008) and the audio quality should eliminate auditory distractions (Mason et al., 2013; Smith and McDonald 2013). In terms of content, online lectures should contain a basic overview or prerequisite content for an upcoming unit (Milman 2013), including a clear introduction and summary of main points (Smith and McDonald 2013).

The critical feature of a successful FL online lecture is its interactivity. Some researchers have suggested embedding interactive activities such as stop-think-answer, interactive graphs, or optional synchronous Q & A for this purpose (Bush 2013; Goodwin and Miller 2013; Shim 2013). In order to improve the quality of online lectures both instructionally and technically, formative evaluation of the resulting instructional product with potential learners and following revisions of the design are invaluable (Slomanson 2014). Among the most crucial advice emerging from FL practitioners is that it is important to design a speculative schedule for studying online content in order to help learners come to class prepared (Mason et al. 2013; Talbert 2012). Summarizing activities or self-checkup quiz also can help to ensure their understanding of online instruction (Bush 2013; Slomanson 2014).

Designing the F2F portion of FL can be much more demanding because it is the core element of FL. Relevant studies primarily concern: (1) initial activities such as verification quizzes and mini-lectures (Roehl et al. 2013; Strayer 2012; Talbert 2012); (2) main learning activities in terms of task features and the instructor's role (Bergmann and Sams 2012; Goodwin and Miller 2013; Mason et al. 2013; Talbert 2012); and (3) external features such as after-activities, ground rules, and classroom culture (Covil et al. 2013; Hamdan et al. 2013; Smith and McDonald 2013; Talbert 2012). One suggestion for F2F sessions is starting with a quick verification quiz to check learners' understanding of online content (Roehl et al. 2013; Talbert 2012), or conducting a 5–10 min Q and A session instead. These initial activities are a very important part of the FL design because successful engagement in those activities as entry tickets for the classroom activities can support the link between online and F2F session, and also promote learners' motivation to come to class prepared. Furthermore, such initial activities enable the instructor to catch students' lack of understanding or misconceptions before the class activity begins. Mini-lectures can then address the identified issues, and be geared to the level of the class. The relative ratio of lecture to class activity depends on the radicality of the instructor's approach of the course: the more radical the approach, the higher the ratio of class activity to lecture. Though one researcher cautioned against entirely eliminating lectures or re-teaching (Talbert 2012), another suggested that mini-lectures should gradually be reduced as learners become more familiar with the FL structure (Strayer 2012).

The main F2F learning activity can be designed around tasks or problems. Studies provide guidance on the optimal features of the classroom tasks and problems of the FL. First, they should represent the most difficult components of learning in the course, and accordingly, may require an instructor's guide, feedback, or peer learner collaboration (Talbert 2012). The tasks also should be authentic and intrinsically rewarding as learners apply what they learned in the online lecture (Mason et al. 2013). The instructor should not only provide direct and immediate feedback and correct learners' misconceptions (Goodwin and Miller 2013); but also respond to their social and emotional needs (Bergmann and Sams 2012). As external design factors, after class activities can be designed to encourage further reflection that consolidates learning (Covil et al. 2013). Some practitioners also recommend setting clear ground rules or appropriate boundaries, and creating a positive engaging culture for the successful implementation of F2F sessions in FL (Hamdan et al. 2013; Smith and McDonald 2013; Talbert 2012). Table 2 summarizes the micro-level elements/issues to be considered when designing FL from previous studies.

Table 2 Micro-level design elements/issues for designing FL from previous studies

FL part	Design elements/issues
FL online part	Physical feature design
	Length; speed control; audio quality
	Content feature design
	Appropriateness; interactivity; clarity
	External feature design
	Formative evaluation of design; study Schedule; after activity
FL F2F part	Initial activity design
	Verification quiz; mini-lecture
	Main learning activity design
	Task; instructor's role
	External feature design
	After activity; ground rule; culture
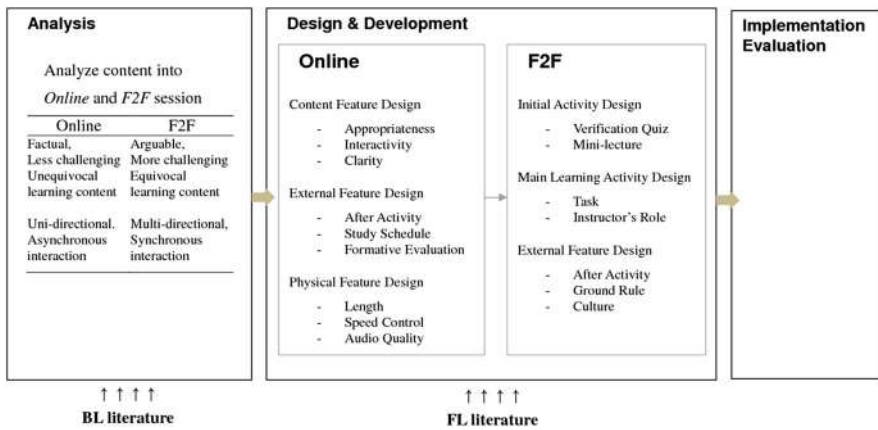
Fig. 1 Synthesis of relevant findings from BL and FL literature

Synthesizing Findings from the BL and FL Literature

As an initial synthesis of literature, a tentative model can be formulated as in Fig. 1. The BL and FL literature are mapped onto ADDIE (Analysis, Design, Development, Implementation, and Evaluation) process. BL literature constitutes the Analysis step, and each of online and F2F FL literature was arranged in the Design and Development step, and sequenced online session followed by F2F session.

Methods

An ID model is a systematic tool that assists designers in understanding related instructional variables and/or guides them through the design process (Lee and Jang 2014). Research on ID models may be classified into three different types: model development, model validation, and model use. This study concerned the first two of these, and largely

followed the research methodology for model development and model validation explicated by Richey and Klein (2007), who noted that ID models may be developed through theoretical or practical means, or both. Theoretical approaches work by synthesizing related literature, while practical approaches utilize simulated design tasks or real-life design projects. The FL design model in this study was developed from a synthesis of relevant literature, simulated design tasks, and real-life design project data.

Further, a developed model may be validated. The validity of a model refers both to the appropriateness of the model components and the usefulness of the model with respect to purpose (Barlas 1994). ID model validation is a carefully planned process of collecting and analyzing empirical data to (1) provide support for each component of the model, or (2) prove its usefulness in practice (Richey 2005). ID model validation can be performed either internally or externally, or through the use of both methods. *Internal model validation* is a validation that addresses the integrity and its usefulness of a model (Richey and Klein 2007). The integrity of the model refers to how valid the components or processes of a model are; and the usefulness of the model points to how effectively the model "assists designers in understanding related instructional variables and/or guides them through the process of analyzing, designing, developing, implementing, and evaluating instructional products" (Lee and Jang 2014, p. 744). Prevalent methods of internal validation are expert reviews and model usability tests. *External model validation* deals with the effects of using the model: the quality of the ID products it creates and the benefits of these products for learners, clients, or organizations. Questions that might be asked in the course of an external validation include: "To what extent does the resulting instruction meet learner needs, motivate learners, or satisfy clients?" or "To what extent do changes occur in the learners' or organization's performance or learning?" (Richey 2005; Richey and Klein 2007). However, such outcome issues may be influenced by a variety of factors such as instructor variables, learner characteristics, or organizational priorities or policies. Typical methods used for external validation are field evaluations or controlled tests. The FL design model in this study underwent internal validation through expert reviews and model usability tests, and external validation through field evaluations.

Procedure

Construction of the initial model

The first step of the study involved mapping specific design details found in the BL and FL literature onto ADDIE (Fig. 1), and presenting this synthesis to an FL instructional design team at a Korean university prior to the start of the fall 2013 semester. The team consisted of one university professor and three teaching assistants, who were chosen based on their position, field, and years of experience (See Table 3 for information on participants). The team refined the synthesis of findings into an initial model that served as the basis for the design of the course implemented in the semester.

Internal validation I: model usability test

The refinement of the initial model led to the second model that the instructional design team used to design and implement a semester-long college calculus course for mathematics education majors at a Korean university. As the semester progressed, the team continued to develop both online learning content and F2F learning activities. We collected the team members' reflections on the FL design process via individual interviews. The

Table 3 Participants for internal validation of the FL design model

Participant	Role	Position	Expert field	Years of teaching experience
Usability test				
A	Instructor	Professor	Mathematics education	21 years
B	Teaching assistant	Doctoral student	Mathematics education	3 years
C	Teaching assistant	Master student/secondary school teacher	Mathematics education	4 years
D	Teaching assistant	Master student/secondary school teacher	Mathematics education	8 years
Expert review				
E	Instructor/FL instructional design	Professor	Physics	10 years (3 FL courses)
F	Instructor/FL instructional design	Professor	Mathematics education	25 years (2 FL courses)
G	Instructor/FL instructional design	Professor	Electronic engineering	12 years (2 FL courses)
H	Instructor/ instructional design	Professor	Educational TECHNOLOGY	10 years (1 FL courses)
I	Instructor/FL instructional design	Professor	Mathematics Education	16 years (1 FL courses)

interview questions concerned conceptual and actual design tasks, such as, "Does this FL design model appropriately reflect actual design practice?"; Are there any step that are hard for designers to follow?; and "Can you elaborate on the component design process in a more detailed manner?" Insights gained from this first usability test were then used to create a revised second FL design model.

Internal validation II: model usability test

The second FL design model also underwent a the second model usability test involving members of the original design team and additionally recruited three experts who have experience in the design of FL (see Table 3 for information on these participants). These individuals completed a model usability questionnaire modified from Tracey's (2001) model development study, and additional questions added for this study. The items on the model usability test concerned on the usability of overall FL design model, specific stages of the model, and resulting products. The responses to the questionnaire resulted in revisions to the second FL design model and, ultimately, the final FL design model for this study.

Internal validation III: expert review

The internal validity of the final model was tested using experts' reviews. Total of five experts who have theoretical expertise and practical experience in the FL design were

asked to answer questions on validity, as elaborated into explicability and comprehensibility (for addressing the integrity of the model components); and usability and generality (for addressing the usefulness of the model) of FL design model (See Table 1 for information on experts). Four of the experts were asked to review the model through individual F2F interviews, and one expert was asked to do it via email. The final content validity index (CVI) and inter-rater agreement (IRA) were calculated and reported. The CVI is a measure of how valid the model is, and the CVI of an item is calculated by dividing the number of experts with positive ratings (i.e. 3 or 4) by total number of experts. A CVI higher than 0.80 is recommended (Davis 1992). The IRA is a measure of the reliability of their ratings and overall agreement among experts (Lynn 1986). The IRA of a scale is calculated by dividing the number of items with an item IRA over 0.80 by the total number of items. The CVI and IRA represent the integrity of a model, that is, the validity of the components of a model.

External validation

The data used for external validation included (1) pre- and post-semester scores from a survey of changes in student views about mathematics (VAMS) (Carlson, 1999); (2) the reflection journal scores; (3) a class survey of learners' satisfaction and follow-up interviews with selected students.

The VAMS The VAMS measures learners' epistemological beliefs and attitudes about math by having them choose from two sentences that represent opposing views about math. For the purposes of this study, we included in the survey only items dealing with epistemological beliefs and dropped other items that did not fit our research objectives. We also added two items related to authentic learning tasks and collaborative learning, both previously validated in a study by Rasmussen et al. (2006) (See Table 4). The responses to the survey were coded on a five-point scale and analyzed using SPSS 21.0. The internal consistency of the converted scale in Cronbach alpha was 0.79.

Reflection journals The instructor of the FL algebra course also asked students to record their reflections after each class. The students then submitted their reflection journals to the teaching assistants, who analyzed the content qualitatively according to the three types of reflective thinking identified by (Mezirow 1990a, b): content reflection, process reflection, and critical reflection. These types served as a priori categories for qualitative content. The reflections also were evaluated quantitatively for completeness and punctuality.

Learner satisfaction survey and follow-up interviews Learners' satisfaction data were collected through a class survey in the seventh week of class. In the fifteenth week of class, three students who had reported high satisfaction and two who reported low satisfaction in the seventh week were interviewed one-to-one using questions such as, "What do you find most helpful in the FL class?" "What do you find most challenging in the FL class?" and "What do you think is the difference between the FL class and the traditional lecture-centered class?"

Table 4 Views about mathematics (VAMS) test items

Appendix.   *Views About MathematicS (VAMS) Test Items*

(1)	(2)	(3)	(4)	(5)	(6)
Only (a)	**More (a)**	**Equally**	**More (b)**	**Only (b)**	**Neither**
Never (b)	than (b)	(a) and (b)	than (a)	Never (a)	(a) nor (b)

		VAMS Items	1	2 3 4 5	6
1		What I mostly liked in the college algebra course was : (a) In-class lecture (b) Learning alone or with peers out of class		...	
2		After I see a solution to a mathematics problem that I got wrong: (a) I discard my solution and learn the one I saw. (b) I try to figure out how the solution I saw differs from mine.		...	
3		3. In solving mathematics problems, graphing calculators or computers help me: (a) understand the underlying mathematical ideas. (b) obtain numerical answers to problems.		...	
4		For me, the relationship of mathematics courses to everyday life is usually: (a) easy to recognize. (b) hard to recognize.		...	
5		The role of a mathematics teacher is to: (a) show me how to work specific problems. (b) guide me in learning to solve problems		...	
6		Scientists use mathematics as: (a) a tool for analyzing and communicating their ideas. (b) a source of factual knowledge about the natural world.		...	
7		College algebra course should engage students with learning with (a) conceptual problems (b) application problems with practical context		...	
8		When I experience a difficulty while studying mathematics: (a) I immediately ask for help, or give up trying. (b) I try hard to figure it out on my own or with others.		...	
9		Reasoning skills that are taught in mathematics courses can be helpful to me: (a) in my everyday life. (b) if I were to major in mathematics or a related field.		...	
10		A major goal of mathematics instruction is to: (a) impart information. (b) equip students to solve problems independently.		...	
11		When studying mathematics in a textbook or in course materials: (a) I memorize it the way it is presented. (b) I make sense of the material so that I can understand it.		...	
12		Graphing calculators or computers: (a) bring new methods for solving mathematics problems. (b) speed up problem solving using established methods.		...	
13		In a mathematics course, it is important: (a) to understand main concepts (b) to understand how concepts can be utilized	+	...	
14		How well I do on mathematics exams depends on how well I can: (a) recall material in the way it was presented in class. (b) do tasks that are somewhat different from ones I have seen before.		...	
15		When solving a challenging mathematics problem, a mathematician: (a) makes many incorrect attempts. (b) moves directly to a correct solution.		...	
16		After I go through a mathematics course and feel that I understand them: (a) I can solve related problems on my own. (b) I have difficulty solving related problems.		...	
17		Using graphing calculators or computers: (a) increases my interest in studying mathematics. (b) is a waste of time.		...	
18		For me, doing well in mathematics courses depends on: (a) how much effort I put into studying. (b) how well the teacher explains things in class.		...	

Participants

The study involved three sets of participants whose feedback was sought at different times in the model construction and validations. The first set of participants was composed of four members of an instructional team (one university professor and three teaching assistants) who served as model users. The instructional design team took part in the model usability test and provided ratings on the general model usability, specific stages of the model, and the resulting products. The instructor was a professor with twenty-one years of experience in the field of mathematics education. The teaching assistants, one doctoral student and two masters students (who concurrently were employed as secondary school math teachers), had from three to eight years of teaching experience.

The second set of participants included five professors from US and South Korean universities who provided the expert review in the second test of internal validity. These five experts were recruited based on their theoretical expertise and experience in designing and implementing FL for at least one course. Their fields of study were physics, mathematics education, electronic engineering, and educational technology. Their experience in designing FL ranged from one to three classes, as indicated in Table 3.

The third set of participants in the study included 18 college students enrolled an algebra course developed from the study's initial FL design model and taught by the first set of participants, the instructional team. The algebra course was offered in the department of mathematics education of a South Korean university in the fall 2013 semester. The group of students contained eight females (44%) and ten males (56%), and all were freshman mathematics education majors.

Results

Initial FL design model

The initial FL design model was developed from the synthesis of BL and FL literature depicted via the ADDIE process (Fig. 1). Design suggestions from the BL studies were incorporated and fleshed out in the Analysis step, the main task of which involved allocating content into online or F2F sessions. Suggestions from the FL literature were mainly embedded in the Design and Development steps along online or F2F sessions. Sub-ID activities that also were suggested in the literature and generic instructional design tasks were added and arranged along the two tracks of the ADDIE model. After analyzing learning goals, content, learners, and technological environment, the content features (such as content sequence and hierarchy and interactivity), and external feature (such as quiz, study schedule, formative evaluation) of online sessions were designed, and then were developed as video clips created from slides and graphics, for which shooting, editing, and revising occurred throughout the formative evaluation. The design of the F2F sessions centered on the design of initial and main learning activities, for which related worksheets, quizzes, and an optional instructor's guide were developed. The resulting F2F sessions were implemented and evaluated, which provided data for further revisions of instructional products. The initial design model described above is shown in Fig. 2.
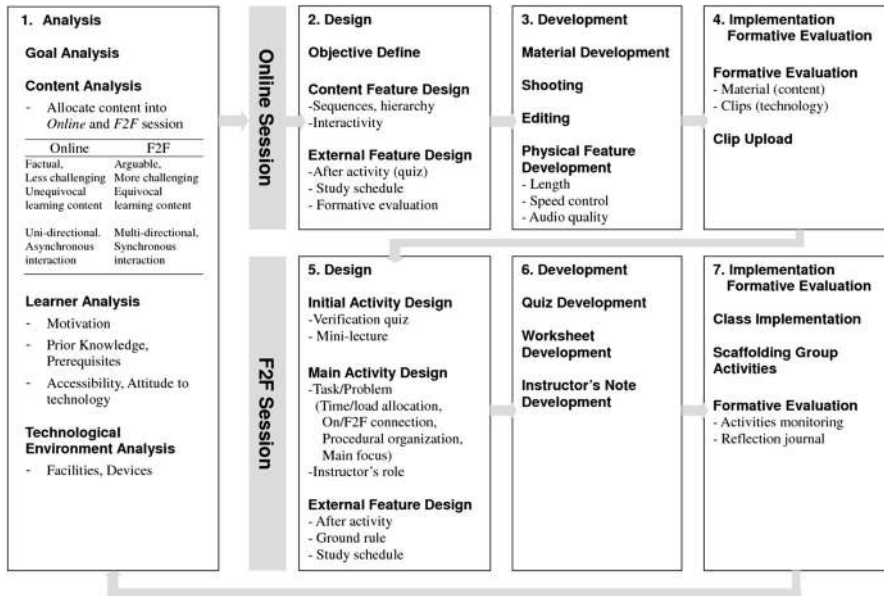
Fig. 2 The initial FL design model

Second design model

The initial FL design model was first validated via two rounds of model usability tests. The results of the first model usability test led to revisions that were incorporated in the second FL design model. In the process of reflecting on the design and implementation process and sharing their insights, the instructional team helped identify two major ways the model might be improved. The first related to course level design and lesson level design; these were seen as needing to be divided since the instructional designers usually do different tasks for each level of design. At the course level, the analysis of goal, content, learner, and environment, and the content outline and instructional strategy design are performed at the macro-level. At the lesson level, however, the objectives of each lesson must be analyzed and online and F2F sessions must be designed and developed at the micro level. The second improvement inspired by the feedback of the instructional team was to include a formative evaluation step in the center of the model in order to reflect how the actual FL design process happens. An iterative revision process proceeds after a set of online and F2F sessions is designed and implemented. The observation and students' feedback data collected during implementation of a lesson promotes the modification of the lesson or the next lessons.

Final FL design model

The second model usability test brought about revisions that resulted in the final FL design model. The feedback of the three members of the instructional team and five experts on the usability of the second FL design model resulted in a number of improvements, the first of which was at the course and lesson levels, where rapid prototyping processes should be represented. At the lesson level, this rapid prototyping also should be present in the design
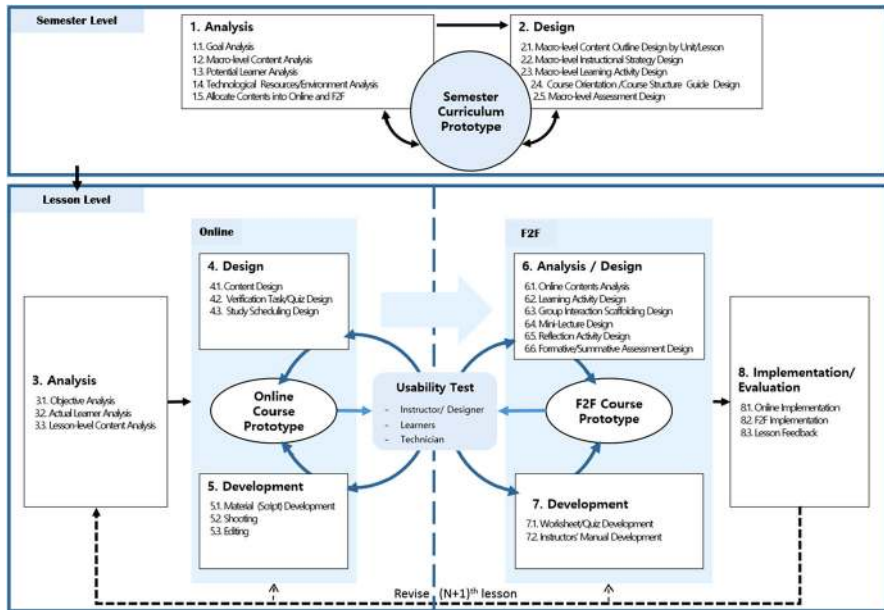
Fig. 3 The final FL design model

of F2F and online sessions. Another improvement related to the need for an explicit textual and visual explanation of the flow of the design process, because the complexities of the model may cause instructional designers to feel frustrated about tracking the procedure that the model suggests. The third and final improvement was a response to the need to describe as clearly as possible the implicit assumptions and usage scenarios of the model in order to prevent possible misunderstandings. The outcome of this last iteration, the final version of the FL design model and model assumptions, is shown in Fig. 3. Figure 4 provides a textual description of the final FL design model.

Finally, the FL design model can have three different usage scenarios at the lesson level. In the first usage scenario, a full set of online sessions is developed, and then a full set of F2F sessions is developed. After development, implementation begins. In the second scenario, a full set of online sessions is developed and each F2F session is developed for immediate implementation. In the third scenario, pairs of online and F2F sessions are developed for immediate implementation.

Final evaluation of final FL design model

Internal validation (expert review)

The panel for the expert review consisted of five professors from a variety of disciplines (educational technology, math education, electrical engineering, physics, and foreign language education), all with experience in the design of FL and theoretical expertise with FL. They were asked to evaluate the validity of final FL design model by providing ratings on the validity, explicability, usability, generality, and comprehensibility of the final model. Mean scores ranged from 2.8 to 3.8 on a scale of 4 to 1, with 4 indicating "strongly agree" and 1 indicating "strongly disagree". The content validity index (CVI) and inter-

rater agreement (IRA) was higher than 0.80 for all items, indicating that the validity of the model is acceptable (Davis 1992; Lynn 1986), and experts' evaluations were mostly in agreement about the usefulness of the model. They were especially positive about the separated course and lesson level design in the higher education context, and the inclusion of a usability test step by which the FL design could be tailored to the particulars of a
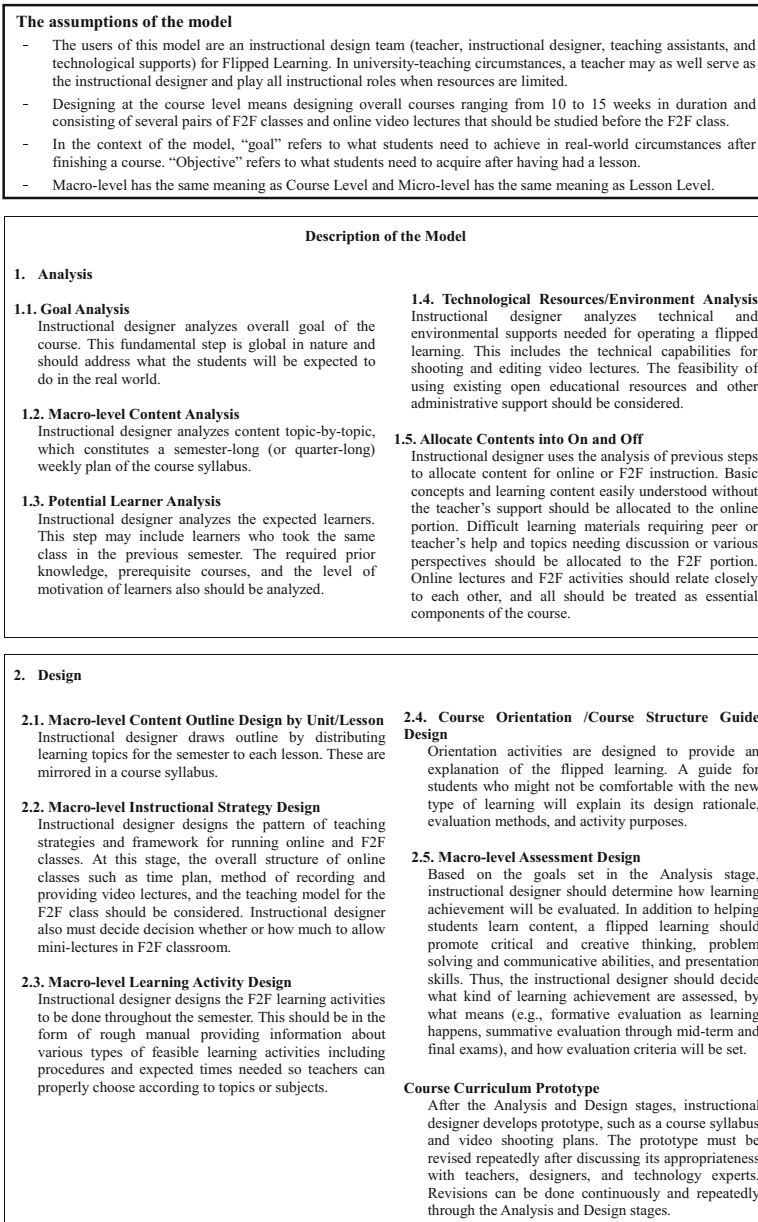
The assumptions of the model
- The users of this model are an instructional design team (teacher, instructional designer, teaching assistants, and technological supports) for Flipped Learning. In university-teaching circumstances, a teacher may as well serve as the instructional designer and play all instructional roles when resources are limited.
- Designing at the course level means designing overall courses ranging from 10 to 15 weeks in duration and consisting of several pairs of F2F classes and online video lectures that should be studied before the F2F class.
- In the context of the model, "goal" refers to what students need to achieve in real-world circumstances after finishing a course. "Objective" refers to what students need to acquire after having had a lesson.
- Macro-level has the same meaning as Course Level and Micro-level has the same meaning as Lesson Level.

Description of the Model

1. Analysis

1.1. Goal Analysis
Instructional designer analyzes overall goal of the course. This fundamental step is global in nature and should address what the students will be expected to do in the real world.

1.2. Macro-level Content Analysis
Instructional designer analyzes content topic-by-topic, which constitutes a semester-long (or quarter-long) weekly plan of the course syllabus.

1.3. Potential Learner Analysis
Instructional designer analyzes the expected learners. This step may include learners who took the same class in the previous semester. The required prior knowledge, prerequisite courses, and the level of motivation of learners also should be analyzed.

1.4. Technological Resources/Environment Analysis
Instructional designer analyzes technical and environmental supports needed for operating a flipped learning. This includes the technical capabilities for shooting and editing video lectures. The feasibility of using existing open educational resources and other administrative support should be considered.

1.5. Allocate Contents into On and Off
Instructional designer uses the analysis of previous steps to allocate content for online or F2F instruction. Basic concepts and learning content easily understood without the teacher's support should be allocated to the online portion. Difficult learning materials requiring peer or teacher's help and topics needing discussion or various perspectives should be allocated to the F2F portion. Online lectures and F2F activities should relate closely to each other, and all should be treated as essential components of the course.

2. Design

2.1. Macro-level Content Outline Design by Unit/Lesson
Instructional designer draws outline by distributing learning topics for the semester to each lesson. These are mirrored in a course syllabus.

2.2. Macro-level Instructional Strategy Design
Instructional designer designs the pattern of teaching strategies and framework for running online and F2F classes. At this stage, the overall structure of online classes such as time plan, method of recording and providing video lectures, and the teaching model for the F2F class should be considered. Instructional designer also must decide decision whether or how much to allow mini-lectures in F2F classroom.

2.3. Macro-level Learning Activity Design
Instructional designer designs the F2F learning activities to be done throughout the semester. This should be in the form of rough manual providing information about various types of feasible learning activities including procedures and expected times needed so teachers can properly choose according to topics or subjects.

2.4. Course Orientation /Course Structure Guide Design
Orientation activities are designed to provide an explanation of the flipped learning. A guide for students who might not be comfortable with the new type of learning will explain its design rationale, evaluation methods, and activity purposes.

2.5. Macro-level Assessment Design
Based on the goals set in the Analysis stage, instructional designer should determine how learning achievement will be evaluated. In addition to helping students learn content, a flipped learning should promote critical and creative thinking, problem solving and communicative abilities, and presentation skills. Thus, the instructional designer should decide what kind of learning achievement are assessed, by what means (e.g., formative evaluation as learning happens, summative evaluation through mid-term and final exams), and how evaluation criteria will be set.

Course Curriculum Prototype
After the Analysis and Design stages, instructional designer develops prototype, such as a course syllabus and video shooting plans. The prototype must be revised repeatedly after discussing its appropriateness with teachers, designers, and technology experts. Revisions can be done continuously and repeatedly through the Analysis and Design stages.

Fig. 4 Description of the final FL design model

3. Analysis

3.1. Objective Analysis
Instructional designer analyzes lesson level-objectives. These might be described considering the items used in the goal analysis stage.

3.2. Learner Analysis
More specific learner analysis should be conducted for the each lesson. The required prior knowledge, prerequisite courses, and the level of motivation of learners should be analyzed. Learners' accessibility to personal computer should also be assured.

3.3. Lesson-Level Content Analysis
The learning content of each class is analyzed. Related learning resources are analyzed at this stage and the hierarchy is analyzed in light of the relative difficulties of the learning content.

4. Design

4.1. Content Design
Instructional designer decides on the specific learning content to be delivered by the online video lecture. Video-lecture continuities are designed, along with pop-up quizzes or self-checking activities inserted within the video lecture to check learner understanding.

4.2. Verification Task/Quiz Design
Instructional designer designs quizzes or activities to check overall understanding of video lectures. Quiz checking could be conducted at the opening of F2F class or the end of online class. Summary assignments and question making activities can be designed as well.

4.3. Study Scheduling Design
Instructional designer sketches study schedule for online learning based on information on learner characteristics derived from learner analysis. A schedule to complete study of online learning material before F2F class sessions should be designed.

5. Development

5.1. Material Development
Various materials that can be used in the online video clip, including presentation material, images, and tables, are developed. Teaching plan and script are developed to give overall direction.

5.2. Shooting
Instructional designer (teacher) removes visual and audio noises and checks audio recording status. Teacher then shoots self and develops video lecture. Open learning resources like OCW or MOOCS can be used in a mixed way. Explicit learning objectives should be provided at the opening of video lecture and summary should be done at the end. Important concepts must be repeated and simple reading of a script should be avoided. When real learners are in front of teacher, it might help for the teacher to overcome unfamiliarity with video shooting.

5.3. Editing
Instructional designer edits segments video lecture and learning materials. Considering typical attention span of learners video lecture should be shorter than 20 minutes and be segmented by topics.

Online Course Prototype
At this stage, already designed and developed online learning prototypes are reviewed in terms of achieving learning objectives and revised as needed. The results of usability testing by teacher, instructional designer, several learners, and technology experts can provide feedback about the prototype. They must check to see if learning material covers learning objectives, if it is overloaded, or if technological limitations exist. Allocation issues should be reviewed in light of both F2F and online portions of the class, and their alignment must be examined. This work can be done repeatedly and continuously in design and development stages.

6. Analysis /Design

6.1. Online Content Analysis
In order to design an F2F session that is tightly link to online counterpart, the online content is analyzed. The online content should be prerequisite for the F2F activities.

6.2. Micro-level Learning Activity Design
Instructional designer designs learning activities for F2F classroom. Decisions are made concerning the procedure, time, and proportion of opening quiz and following activities as well as main class activities. The unique goal of the F2F activities is to make different knowledge applications among peers and instructor visible and to promote interactions.

6.3. Group Interaction/Scaffolding Design
Activity groups and group interaction and scaffolding are designed. Decisions are made about in-group interaction and inter-group interaction in light of overall course design. The scaffolding methods to facilitate group interaction also are considered and designed.

6.4. Mini-Lecture Design
The instructional designer designs mini-lecture to be used at the opening of class to remind students of online video lecture previously watched and to make linkages between the two. When used midway through class, it can help adjust learning activities; if used at end of class, it can help learners to organize what they have learned and enable them to connect it with upcoming learning materials.

6.5. Reflection Task/Assignment Design
Reflection assignment is provided after F2F classroom activities to help learners consolidate what they have learned individually. Also can be done following the same patterns on a regular basis.

6.6. Formative/Summative Assessment Design
Instructional designer designs formative assessment and summative assessment. Teacher and learner can interactively complete formative assessment during offline activities, for which assessment plan and evaluating criteria need to be designed. Assignments and mid-term and final exams are also planned.

Fig. 4 continued

7. Development **7.1. Worksheet/Quiz Development** Worksheets and quizzes to be used in the classroom are developed. Worksheets should be developed in light of the hierarchy of learning content and its order, expected time spent, and students' workload. The online learning content should be reflected in the quiz so that once learners have studied it they are able to move forward. **7.2. Instructors' Manual Development** If several teaching assistants and facilitators are involved, they collaboratively develop learning activities, interactions, scaffolding, mini-lectures, and teaching plans. The role of teachers as facilitators of interaction-provoking activities should be made explicit in the learning activities emphasized points in the learning activities.	**F2F Lesson Prototype** The teacher, instructional designer, and several learners review the prototype of F2F learning activities and worksheets to see if they are appropriate to the learning goals and to check if they can be implemented as planed without an excessive overload. They should be also compared with the online learning prototype, and allocation issues on-off alignment should be examined. This process can be conducted repeatedly and continuously in the design and develop stages.
8. Implementation/ Evaluation **8.1. Online Implementation** Developed online learning materials are uploaded on time according to a schedule so that learners do not have problem studying online.	**8.2. F2F Implementation** F2F classroom activities are conducted. **8.3. Next Lesson Feedback** Suggestions for improvement drawn from implementtation are integrated into the design of the next class.

Fig. 4 continued

course. Additionally, they felt strongly that a detailed description of the model needed to be included for novice instructors who might otherwise find it difficult to effectively utilize the model. In addition, they believed that guidelines for allocating learning content to online or F2F sessions would be very helpful.

External validation (field evaluation)

A 15 week algebra course was designed according to the model. The topics of a traditional algebra class were assigned throughout the fifteen weeks, with a mid-term exam in the eighth week and a final exam in the fifteenth week. Each week, two clips with 20 min lectures recorded by the instructor were uploaded on the online course management system for students. In addition, students had a 75 min F2F class each week. Each F2F class started with a quiz on the lecture clips, followed by a 10–15 min mini-lecture. For approximately the next hour, students then engaged in group problem-solving activities with three to four problems provided by the instructors. The difficulty of these problems gradually increased over the class time. Students first discussed the problem in teams, and then participated in a whole class discussion. Students also were assigned to submit a reflection journal within the class day.

After implementing the FL course designed according to the final FL design model, the various data indicate how resulting instruction impacted students' views of mathematics (VAMS), student reflections about the FL design of the class, and students' satisfactions in regard to the course overall.

VAMS (views about mathematics survey) Although the normality assumptions of *Shapiro–Wilk* statistics were met (p = 0.20, and p = 0.61, respectively), both parametric and non-parametric statistics were calculated due to small sample size (n = 17). Both Wilcoxon signed rank test and paired *t* test indicated that participants showed significantly higher scores in the post-VAMS (*M* = 3.59, *SE* = 0.085) than in the pre-VAMS (*M* = 3.28, *SE* = 0.061), ($Z = 2.33$, $t(17) = 2.774$, $p < 0.05$, $r = 0.56$).

Studies suggest that students with low VAMS scores usually have naïve views toward math and knowledge whereas students with high VAMS scores have advanced mathematical views and epistemological beliefs closer to those of experts (Hofer 2004). Students in the former group regard knowledge as a solid, absolute, structured, separate, and stable truth. To the contrary, students in the latter group regard knowledge as a variable and provisional agreement that is complex and constructed by interaction, and these students tend to have better mathematical problem-solving ability (Higgins 1997; Verschaffel et al. 1999) and greater ability to synthesize information than naïve viewers (Muis, 2008). Research also has indicated that the VAMS scores rarely increase much in a short period like a single semester (Carlson, 1999; Chris et al. 2006). In this study, then, the meaningful difference between students' pre- and post-VAMS scores after just a single semester suggests that the FL design model had a very positive impact on the development and implementation of the FL algebra course.

Reflection journals   Over the fifteen weeks of the FL course, the teaching assistants discovered qualitative improvement in the level of students' reflective thinking. At the start of the semester, students' reflective journals mostly contained summaries of learning content and activities, and mentions of the novelty of the video clips and comments about the difficulties of online and F2F activities, much in the way of a diary. These thoughts were mostly content and process reflection. As students approached the end of the semester, they described their own ways of critical understanding what had not been adequately discussed during the class activities, and even provided suggestions for FL implementation such as the optimum number of class activities or ideas about group composition, which can be regarded as critical reflections. The journals also were assessed using a combined score for the quality of the reflection (70%) and punctuality of submission (30%), and these scores increased overall throughout the semester, as indicated in Fig. 5.
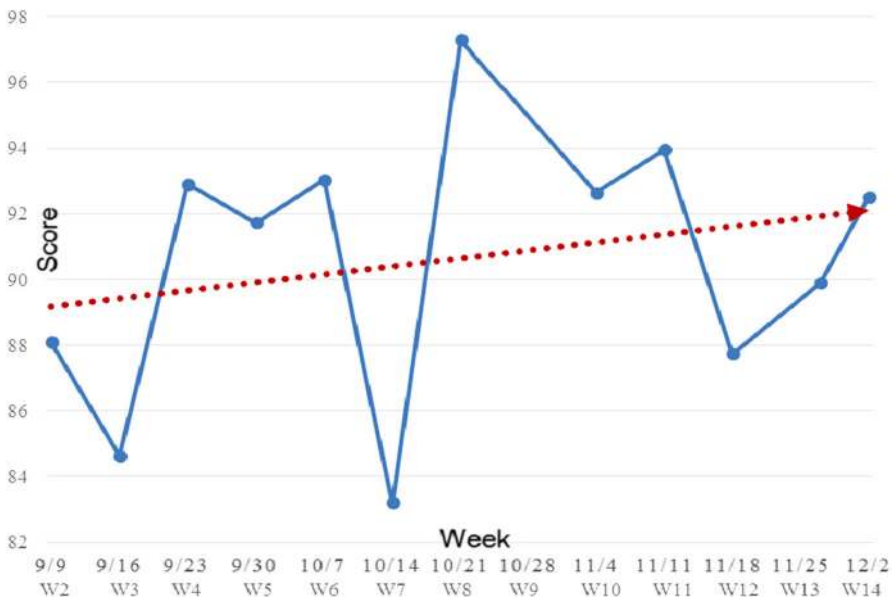
Fig. 5 Average combined score changes of weekly reflection journal throughout the semester

Student satisfaction   In the seventh week of the algebra course, a short-answer mid-term survey was conducted to measure learners' satisfaction. Students' responses were somewhat mixed. Some students complained that their workload was too heavy due to the lecture clips, discussion participation, and reflection journals. Other complaints involved the limited time for problem solving during the F2F sessions and the inaccessibility of instant Q & A when learning with the lecture clips. Students also commented on the positive impacts of FL on their learning. Many valued being able to pause and repeat the clips when they did not understand a specific part or needed a while to reflect or think deeper. Students also responded very positively to the increased opportunities for discussion and active engagement in the overall learning process, and they reported enhanced motivation for learning itself.

At end of the semester, three of the most satisfied learners and two of the least satisfied learners on the mid-term survey were asked to participate in in-depth interviews. The results showed that all students, including the least satisfied, expressed great satisfaction with FL designed by the FL design model. Students reported that they appreciated the opportunities for active participation, including the interaction between them and the other learners and the instructor. Four of the five learners interviewed (A, B, C, and E) reported that they felt motivated by the activities and engaged in deep understanding:

I think group work was the best part here. When I asked other friends about what I couldn't understand, I found at least one of us understood it well. Hearing this, I learned, and at the same time, I could practice explaining things that I understood well. That made me feel organized and I loved it. (Learner A)

In my group, there were some smart guys, and I felt more comfortable asking them questions than asking the teacher. Of course, they knew me [and my needs] better than the teacher and so they could explain it better. (Learner C)

Thanks to the group work, I think I studied harder this semester than last semester. Last semester, I rarely studied after the mid-term exam. However, this semester, in order to participate in discussions in class, I had to bring some basic knowledge to the classroom. I had to watch the video lectures and study the subject before the class. That was the main difference from the last semester. (Learner B)

I believe the key to success is students' participation. I myself became engaged in the process of understanding during group discussions. Just watching lectures doesn't guarantee understanding and they do not always give clear explanations about what I haven't understood. Even when what I couldn't understand was not a big part, [my lack of understanding] but was eliminated in group discussions. I loved the group work. (Learner E)

Those interviewed also mentioned the benefits of the online lectures, particularly because of their repeatability. In the previous semester, the length, breadth, and rapid progression of mathematical proofs gave students little time to stop and think, and they often lost pace as a consequence, and frequently giving up in the end. In contrast, the ready availability of the online lectures in the course studied enabled students to pause, rewind, and re-think.

Last semester, the teacher couldn't finish the class in time because of too much content to cover, and some students dozed off or gave up. However, when watching the video lectures we don't doze off and we can take a nap if it is really needed and

ask the teaching assistants questions in class. That was the good part. There was a synergistic effect when combining the video lecture and offline class. (Learner C)

First of all, the video lectures were helpful for reviewing the study materials when preparing for exams. They could be replayed anytime, and [helped me] to master calculus anyway because calculus requires incremental understandings. I felt I had to master theories by myself but it gave chances to think deeply about the subject. (Learner B)

One of the critiques of the FL class was the increased workload—even among students who admitted it was good for them, like Learner D. Learner E, however felt that the total burden of work was not greater, but was distributed more evenly than when the focus of learning centered on mid- and final exams.

It should be a four-credit course with this much homework. There were lots of small homework assignments and they counted for 5% of our total grade. We were evaluated multi-dimensionally. That was a really good part of it. It was possible because it was flipped classroom. So, it went past rote learning and went deep into the subject. More subjects were introduced, and students were more interested in them. (Learner D)

I didn't study that much for the mid- and final exams. Actually, I didn't need to…'cause I already had studied and mastered the content. Last semester, I usually stayed up nights for the exams because I had to start to review and recall the content from the beginning. (Learner E)

All the students, including some who complained in the mid-term survey, said they would take a FL course again if one were available in the next semester.

I would take it again because it was fun. I confess I am one who dozes off a lot in class, but I never did in this one. (Learner D)

If I could choose between a normal class and a flipped learning class, I would definitely choose a flipped one. (Learner E)

Discussion and conclusion

This study was a developmental investigation of a process for constructing and validating a flipped learning design model. Through an iterative process of review and revision, the FL design model was formatively improved and internally validated. The implementation of the final model was shown to result in meaningful increases in students' maturity of mathematical views and epistemological beliefs, reflections, and satisfactions. Related issues that emerged during the process will be discussed below.

Distinctive features of the final FL design model

Two distinctive features of the final FL design model would worth further discussion: the macro and micro two-level approach and the integrative design of online and F2F sessions. First, in taking both a macro-view of the course level, and a micro-view of each lesson level, the FL design model creates a more systematic approach to FL design. At the macro-level, course goals are defined, which in turn guides the design of macro-level content and

micro-level objectives and content. The macro-design corresponds to the syllabus design, and the micro-design corresponds to the lesson plan design, by which the model reflects the design context of higher education. This sort of two-level approach for designing FL has been suggested in the FL literature, though without any explicit recommendations. In fact, the review of FL literature revealed its clear division into studies concerning either macro-level or micro-level design elements. Macro-level studies tended to concern issues like content allocation and a consistent structure throughout the lessons of a course (Bush 2013; Covil et al. 2013; Mason et al. 2013; Stannard 2012; Strayer 2012), an introductory orientation to the overall design rationale (Herreid and Schiller 2013; Mason et al. 2013; McLaughlin et al. 2014), and course outcomes and assessment at course end (McLaughlin et al. 2014; Milman 2013). By contrast, micro-level studies concerned the length of lessons (Mason et al. 2013; Smith and McDonald 2013), speed control of lessons (Bush 2013; Goodwin and Miller 2013; Hattie, 2008), audio quality of lecture clips (Mason et al. 2013; Smith and McDonald 2013), strategies for interactivity (Bush 2013; Goodwin and Miller 2013; Shim 2013) the study schedule of online lectures (Mason et al. 2013; Talbert 2012), in-class verification quizzes (Roehl et al. 2013; Talbert 2012), mini-lectures (Strayer 2012; Talbert 2012); and in- and after-activity design of F2F sessions (Covil et al. 2013; Goodwin and Miller 2013; Mason et al. 2013). The apparent distinction between the two lines of FL studies indicates that design tasks at the macro- and micro-level are best modeled separately, as the final FL design model in this study proposes.

Secondly, the final model starts with a collective analysis step dealing with both online and F2F sessions, and proceeds to the formative design and development of the online lessons. The design and development of F2F lessons start with another analysis specifically intended for the design of F2F activities. Results from the common implementation and evaluation are to be reflected in the analysis step of the next lessons. These model changes can be interpreted as working towards the right blend, that is, the optimal integration of the two parts of FL. The common analysis, implementation, and evaluation steps will undoubtedly improve the amalgamation of the online and F2F pair by aligning the two sessions with a shared lesson objective and by allocating content into on and off, thus maximizing the advantages of each environment.

These strategies are consistent with recommendations derived from related FL literature. Most FL studies have stressed that online and F2F learning experiences should coherently support each other in order to create a successful FL experience (Bush 2013; Covil et al. 2013; Mason et al. 2013; Stannard 2012; Strayer 2012). This kind of coherent support means that the two FL components have a complementary rather than supplementary relationship. Put simply, without the online sessions, F2F activity should not be effective; and without the F2F sessions, learning with online lectures should not be complete or have enough depth. Conducting an integral evaluation of online and F2F learning, in particular, can increase the congruency of all the learning experiences in a FL course. This is because such an evaluation reveals the extent to which each portion of the design as well as the entire design optimally contributes to the achievement of course and lesson goals and objectives. Another analysis for designing F2F activities represents an effort to increase the tight integration of the two parts, and to design the F2F sessions more deliberately.

Model specificity

The final FL design model created in this study evolved gradually, and more making it possible to create a specified guide to the FL design process with model assumptions, usage

scenarios, and step-by-step descriptions. Model specificity can be achieved both by specified model assumptions and by specified step components. The model assumptions and usage scenarios specify the target users, scope, and design context for which the final FL design model should be used. The step-by-step descriptions provide more precise guides that are useful when designers make decisions on the actual design.

However, in general, the more specific a model gets, the narrower the application of the model becomes. Downes (2003) referred to this dilemma when asserting that design requires specificity but specificity is incompatible with reusability and general application. The task of finding the best balance between a useful model with specific and practical guidelines and a wide-reaching model with general and flexible guidelines is crucial, but challenging. In the process of internal validation in this study, the experts closer to being practitioners preferred the former approach whereas the experts who were closer to being theorists prefer the latter. These preferences reflect the theoretical and practical roles of ID models, that is, to promote understandings of ID realities and to guide ID performance respectively (Branch and Kopcha 2014; Davies 1996; Gustafson and Branch 2002; Jung and Rha 1989; Lee and Jang 2014; Rubinstein 1975; Seels and Glasgow 1998). Since an ID model for designing FL is not yet available in the research literature, the final FL design model developed in this study was intended for general application within the higher education context, while clearly specifying the meanings of the component steps so that the steps can cognitively guide designers to make intelligent design decisions.

This study may be limited in that the FL design model was developed from a single case, suggesting the need for confirmation in more cases. Also, since the case was an algebra course, the model may include some features relevant to courses in natural sciences but not to disciplines such as liberal arts, social sciences, or art and music, which may require different design processes. Further, even though the FL design model in this study underwent internal validation from experts from diverse disciplines like engineering, physics, mathematics education, and educational technology, actual implementation of the model in courses in these fields might reveal different aspects of FL design.

Another potential limitation of this study that suggests future research topics regards the broad definitions used for outcomes of FL. Many FL studies have found meaningful improvements in students' satisfaction or attitude (e.g. Bland 2006; Kellogg 2009; Talbert 2012). However, other studies have reported insignificant or marginal increases in academic performance when comparing FL to traditional methods (e.g. Kellogg 2009; Papadopoulos and Roman 2010). In those cases, the academic achievement of students in FL courses was measured using the same instruments as in traditional courses. By contrast, in this study the outcomes of the FL course were broadly defined as epistemology change, quality of reflection, and satisfaction. Observation throughout the semester revealed other areas in which FL has positive outcomes, including study skills, presentation skills, collaborative or communicative skills, and inquisitive attitudes toward learning, some of which are mentioned in previous studies (Bishop and Verleger 2013; Mason et al. 2013; Talbert 2012). Thus, balancing qualitative and quantitative assessment, and content knowledge and general competencies, research on more comprehensive approaches to evaluating the learning impacts of FL is recommended for the future.

The interview data from students taking the course pointed to changes in their epistemology and classroom culture. Thus, a potential impact of the course designed using the FL design model may be innovation in learning from a broader institutional context, which may be due to the transformative potential of design by envisioning different types of learning experiences (Collins et al. 2004). Design has been described as systematic planning for future innovation (Rowland 2008; Simon 1969; Smith and Boling 2009), an

approach that diverges from the widely-accepted view of design as a solution to a problem. Increasingly, design is seen as providing "a big picture lens of the problem" in a larger context rather than "a direct path" (Daly et al. 2012, p. 200) or an immediate intervention to a problem at hand. Some scholars have even argued that design is a matter of "shaping the world into desired states" (Molenda and Boling 2008, p. 121) or "how things ought to be" (Simon 1969, p. 133), thus "making the world a better place" (Laurillard 2012, p. 225). Design models, then, can inform, guide, and lead to successful educational innovations (Smith and Boling 2009). Although proposed innovative potential of FL design models is somewhat speculative, they nevertheless may serve as a bridge that approaches FL from the individual course level as well as from the level of institutional support and policy. It is hoped that the FL design model developed in this study would contribute at both level.

References

Allen, I. E., Seaman, J., & Garrett, R. (2007). *Blending in: The extent and promise of blended education in the United States*. Newburyport, MA: Sloan Consortium.

Arum, R., Cho, E., Kim, J., & Roska, J. (2012). *Documenting uncertain times: post- graduate transitions of the academically adrift cohort*. New York: Social Science Research Council.

Arum, R., & Roska, J. (2011). *Academically adrift: Limited learning on college campuses*. Chicago: University of Chicago Press.

Baker, J. W. (2000). *The classroom flipped: Using web course management tools to become the guide on the side*. Paper presented at the 11th International Conference on College Teaching and Learning, Jacksonville, FL.

Barlas, Y. (1994). *Model validation in system dynamics*. Paper presented at the International System Dynamics Conference.

Bergmann, J., & Sams, A. (2012). *Flip your classroom*. Eugene, OR: International Society for Technology in Education

Bergmann, J., & Sams, A. (2013). Flipping for mastery. *Educational Leadership, 71*(4), 24–29.

Bishop, J. L., & Verleger, M. A. (2013). T*he flipped classroom: A survey of the research*. Paper presented at 120th ASEE Annual Conference & Exposition, Atlanta, GA.

Bland, L. (2006). *Applying flip/inverted classroom model in electrical engineering to establish life-long learning*. Paper presented at 113th ASEE Annual Conference & Exposition, Chicago, IL.

Branch, R. M., & Kopcha, T. J. (2014). Instructional design models. In M. J. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (4th ed., pp. 77–87). New York: Springer.

Bush, M. D. (2013). Educational technology points of inflection: what MOOCs, flipped classrooms, and OLPC teach us about individualization of learning. *Educational Technology, 53*(6), 60–63.

Carlson, M. (1999). The mathematical behavior of six successful mathematics graduate students: influences leading to mathematical success. *Educational Studies in Mathematics, 40*(3), 237–258.

Collins, B., de Boer, W., & van der Veen, J. (2001). Building on learner contributions: a web-supported pedagogic strategy. *Educational Media International, 38*(4), 229–240.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: theoretical and methodological issues. *Journal of the Learning Sciences, 13*(1), 15–42.

Covil, D., Patel, B. A., & Gill, D. S. (2013). Flipping the classroom to support learning: an overview of flipped class from science, engineering and product design. *School Science Review, 95*(350), 73–80.

Daft, R. L., & Lengel, R. H. (1984). Information richness: A new approach to managerial behavior and organization design. In B. M. Staw & L. L. Cummings (Eds.), *Research in organizational behavior* (Vol. 6, pp. 191–233). Greenwich: JAI Press.

Daly, S. R., Adams, R. S., & Bodner, G. M. (2012). What does it mean to design? A qualitative investigation of design professionals' experiences. *Journal of Engineering Education, 101*(2), 187–219.

Davies, I. K. (1996). Educational technology: Archetypes, paradigms and models. In D. P. Ely & T. Plomp (Eds.), *Classic writings on instructional technology* (pp. 15–30). Englewood. CO: Libraries Unlimited Inc.

Davis, L. (1992). Instrument review: getting the most from your panel of experts. *Applied Nursing Research, 5,* 194–197.

Dennis, A., & Valacich, J. (1999). *Rethinking media richness: Towards a theory of media synchronicity.* Paper presented at the 32th Hawaii International Conference of Systems Sciences, Maui, HI.

Dick, W. (1997). A model for the systematic design of instruction. In R. D. Tennyson, E. Schott, N. M. Seel, & S. Dijkstras (Eds.), *Instructional design: international perspectives* (pp. 361–369). Mahwah: Lawrence Erlbaum Associates.

Downes, S. (2003). Design and reusability of learning objects in an academic context: a new economy of education? *Journal of the United States Distance Learning Association, 17*(1), 3–22.

Gannod, G. C., Burge, J. E., & Helmick, M. T. (2008). *Using the inverted classroom to teach software engineering.* Paper presented at the 30th International Conference on Software Engineering, Leipzig, Germany.

Garrison, D. R., & Kanuka, H. (2004). Blended learning: uncovering its transformative potential in higher education. *Internet and Higher Education, 7*(2), 95–105.

Garrison, D. R., & Vaughan, N. D. (2008). *Blended learning in higher education: Framework, principles, and guidelines.* San Francisco: John Wiley & Sons.

Gedik, N., Kiraz, E., & Ozden, M. Y. (2013). Design of a blended learning environment: considerations and implementation issues. *Australasian Journal of Educational Technology, 29*(1), 1–19.

Ginns, P., & Ellis, R. (2007). Quality in blended learning: exploring the relationships between on-line and face-to-face teaching and learning. *The Internet and Higher Education, 10*(1), 53–64.

Goodwin, B., & Miller, K. (2013). Evidence on flipped classrooms is still coming in. *Educational Leadership, 70*(6), 78–80.

Graham, C. R. (2006). Blended learning systems: Definition, current trends, and future directions. In C. J. Bonk & C. R. Graham (Eds.), *Handbook of blended learning: global perspectives, local designs* (pp. 3–21). San Francisco: Pfeiffer Publishing.

Graham, C. R., Woodfield, W., & Harrison, J. B. (2013). A framework for institutional adoption and implementation of blended learning in higher education. *Internet and Higher Education, 18*(3), 4–14.

Gustafson, K., & Branch, R. M. (2002). *Survey of instructional development models* (4th ed.). Syracuse: Eric Clearinghouse on Information & Technology.

Halverson, L. R., Graham, C. R., Spring, K. J., Drysdale, J. S., & Henrie, C. R. (2014). A thematic analysis of the most highly cited scholarship in the first decade of blended learning research. *The Internet and Higher Education, 20*(1), 20–34.

Hamdan, N., McKnight, P., McKnight, K. & Arfstrom, K.M. (2013). *A review of flipped learning.* http://flippedlearning.org/review.Retrieved 18 Feb 2014

Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* New York: Routledge.

Herreid, C. F., & Schiller, N. A. (2013). Case studies and the flipped classroom. *Journal of College Science Teaching, 42*(5), 62–66.

Higgins, K. M. (1997). The effect of year-long instruction in mathematical problem solving on middle-school student's attitude, beliefs, and abilities. *Journal of Experimental Education, 66*(1), 5–28.

Hofer, B. K. (2004). Epistemological understanding as a meta-cognitive process: thinking aloud during online searching. *Educational Psychologist, 39*(1), 43–55.

Hofmann, J. (2002). Blended learning case study. In A. Rossett (Ed.), *The ASTD e-learning handbook: Best practices, strategies, and case studies for an emerging field* (pp. 516–519). New York: McGraw-Hill.

Hollingshead, A. B. (1996). The rank-order effect in group decision making. *Organizational Behavior and Human Decision Processes, 68*(3), 181–193.

Huon, G., Spehar, B., Adam, P., & Rifkin, W. (2007). Resource use and academic performance among first year psychology students. *Higher Education, 53*(1), 1–27.

Jung, I., & Rha, I. (1989). *Instructional design theories.* Seoul: Educational Science.

Kellogg, S. (2009). *Developing online materials to facilitate and inverted classroom approach.* Paper presented at the 39th ASEE/IEEE Frontiers in Education Conference, San Antonio, TX.

Kember, D., McNaught, C., Chong, F. C. Y., Lam, P., & Cheng, K. F. (2010). Understanding the ways in which design features of educational websites impact upon student learning outcomes in blended learning environments. *Computers & Education, 55*(3), 1183–1192.

Kerres, M., & de Witt, C. (2003). A didactical framework for the design of blended learning arrangements. *Journal of Educational Media, 28*(2–3), 101–113.

Khan, S. (2012). *The one world schoolhouse: Education re-imagined*. London: Hodder & Stoughton Ltd.

Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: a gateway to creating an inclusive learning environment. *The Journal of Economic Education, 31*(1), 30–43.

Laurillard, D. (2012). *Teaching as a design science: building pedagogical patterns for learning and technology*. New York: Routledge.

Lee, J., & Jang, S. (2014). A methodological framework for instructional design model development: critical dimensions and synthesized procedures. *Educational Technology Research and Development, 62*(6), 743–765.

López-Pérez, M. V., Pérez-López, M. C., Rodríguez-Ariza, L., & Argente-Linares, E. (2013). The influence of the use of technology on student outcomes in a blended learning context. *Educational Technology Research and Development, 61*(4), 625–638.

Lou, Y., Bernard, R. M., & Abrami, P. C. (2006). Media and pedagogy in undergraduate distance education: a theory-based meta-analysis of empirical literature. *Educational Technology Research and Development, 54*(2), 141–176.

Lynn, M. (1986). Determination and quantification of content validity. *Nursing Research, 35,* 382–385.

Mason, G., Shuman, T. R., & Cook, K. E. (2013). Comparing the effectiveness of an inverted classroom to a traditional classroom in an upper-division engineering course. *IEEE Transactions on Education, 56*(4), 430–435.

McLaughlin, J. E., Roth, M. T., Glatt, D. M., Gharkholonarehe, N., Davidson, C. A., Griffin, L. M., et al. (2014). The flipped classroom: a course redesign to foster learning and engagement in a health professions school. *Academic Medicine, 89*(2), 236–243.

Mezirow, J. (Ed.). (1990a). *Fostering critical reflection in adulthood: A guide to transformative and emancipatory learning*. San Francisco: Jossey-Bass.

Mezirow, J. (1990b). *Fostering critical reflection in adulthood: a guide to transformative and emancipatory learning*. San Francisco: Jossey-Bass.

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

Milman, N. B. (2013). The flipped classroom strategy: what is it and how can it best be used? *Distance Learning, 9*(3), 85–87.

Molenda, M., & Boling, E. (2008). Creating. Januszewski, A., & Molenda, M. (Eds.). Educational technology: A definition with commentary. New York, NY: Routledge.

Muis, K. R. (2008). Epistemic profiles and self-regulated learning: examining relations in the context of mathematics problem solving. *Contemporary Educational Psychology, 32*(2), 177–208.

Orton-Johnson, K. (2009). 'I've stuck to the path I'm afraid': exploring student non-use of blended learning. *British Journal of Educational Technology, 40*(5), 837–847.

Overmyer, J. (2012). Flipped classrooms 101. *Principal, 9–10,* 46–47.

Papadopoulos, C., & Roman, S.A. (2010). *Implementing and inverted classroom model in engineering statics: Initial results*. Paper presented at the 117th ASEE Annual Conference & Exposition, Louisville, KY.

Rasmussen, C., Kwon, O. N., Allen, K., Marrongelle, K., & Burtch, M. (2006). Capitalizing on advances in mathematics and K-12 mathematics education in undergraduate mathematics: an inquiry-oriented approach to differential equations. *Asia Pacific Education Review, 7*(1), 85–93.

Reigeluth, C. M. (1992). Elaborating the elaboration theory. *Educational Technology Research and Development, 40*(3), 80–86.

Rice, R. (1992). Task analysability, use of new media and effectiveness—A multisite exploration of media richness. *Organization Science, 3*(3), 475–500.

Richey, R. C. (1986). *The theoretical and conceptual bases of instructional design*. New York: Kogan Page.

Richey, R. C. (2005). *Validating instruction design and development models*. In: Spector J. M. and Wiley D. A. (Eds.), Innovations in instructional technology: Essays in honor of M. David Merrill (pp. 171–185.) Mahwah, NJ: Lawrence Erlbaum Associates.

Richey, R. C., & Klein, J. D. (2007). *Design and development research*. Mahwah: Lawrence Erlbaum.

Roehl, A., Reddy, S. L., & Shannon, G. J. (2013). The flipped classroom: an opportunity to engage millennial students through active learning strategies. *Journal of Family & Consumer Sciences, 105*(2), 44–49.

Rowland, T. (2008). The purpose, design and use of examples in the teaching of elementary mathematics. *Educational Studies in Mathematics, 69*(2), 149–163.

Rubinstein, M. (1975). *Patterns of problem solving*. New York: Prentice Hall.

Seels, B. A., & Glasgow, Z. (1998). *Making instructional design decisions* (2nd ed.). Upper Saddle River: Prentice Hall.

Shim, H. (2013). *A commentary on the flipped classroom instructional model*. Paper presented at the 30th Symposium on Mathematics Education. Seoul, Korea.

Simon, H. A. (1969). *The sciences of the artificial*. Cambridge: MIT Press.

Singh, H. (2003). Building effective blended learning programs. *Educational Technology, 43*(6), 51–54.

Slomanson, W. R. (2014). Blended learning: a flipped classroom experiment. *Journal of Legal Education, 64,* 1–10.

Smith, K. M., & Boling, E. (2009). What do we make of design? Design as a concept in educational technology. *Educational Technology, 49*(4), 3–17.

Smith, C. M., & McDonald, K. (2013). The flipped classroom for professional development (Part II): making podcasts and videos. *Journal of Continuing Education in Nursing, 44*(11), 486–487.

Stannard, R. (2012). The flipped classroom or the connected classroom? Looks at the benefits of putting more input online. *Modern English Teacher, 21*(1), 35–37.

Strayer, J. (2012). How learning in an inverted classroom influences cooperation, innovation and task orientation. *Learning Environments Research, 15*(2), 171–193.

Talbert, R. (2012). *Learning matlab in the inverted classroom*. Paper presented at 119th ASEE Annual Conference and Exposition, San Antonio, TX.

Tracey, M.W., (2001). *The construction and validation of an instructional systems design model incorporating multiple intelligences. Unpublished doctoral dissertation*. Detroit: Wayne State University.

Dr. Jihyun Lee is an Assistant Professor with the School of Dentistry at the Seoul National University. She received her Master Degree with Technology in Education Program from Harvard Graduate School of Education, and her Ph.D. in Educational Technology from Seoul National University. Her research interests include Dental/Medical Education, Model Development Methodology, Media Theories, Flipped Learning, MOOCs, and Technology Integration for Higher-order Thinking.

Dr. Cheolil Lim is a Professor in the School of Education, Seoul National University of Korea. He is currently the Director of Education Research Institute, and Chair of the department of education, and worked as the Director of Center for Teaching and Learning and Associate dean of Education Affairs of SNU. He served as Vice President of the Korean Society for Educational Technology and President of the Korean Society for Learning and Performance. His recent research interests include Instructional Systems Design Model, e-Learning Design, Instructional Design Strategies for Creativity and Self-regulated Learning, and Flipped Learning Models.

Hyeonsu Kim finished his master program in the Department of Education at the Seoul National University in Korea. His research interests include Game-Based Learning, Distance Learning with Videoconferencing, and Technology-based Learning Environment Design.