ELT teacher education flipped classroom: An analysis of task challenge and student teachers' views and expectations

Hatice Karaaslan [a] *[*] (iD), Hatice Çelebi [b] (iD)

[a] *Ankara Yıldırım Beyazıt University, Ayvalı Mah. Gazze Cad. No:7, Ankara, Turkey*
[b] *Acıbadem University, Kayışdağı Cad. No:32, İstanbul, Turkey*

Abstract

In this study, we explore the interplay between task complexity, task conditions and task difficulty introduced by Robinson (2001) in flipped classroom instruction at tertiary level through the data we collected from undergraduate English Language Teaching (ELT) department students studying at an English-medium state university. For the participants, we adopted the flipped classroom principle that content attainment largely takes place outside the classroom and application attainment inside the classroom. Following this principle, we systematically designed and sequenced tasks according to their complexity, difficulty, and conditions, and assigned them to the ELT student teachers over an academic semester. Data regarding flipped classroom principle in relation to task design features were collected from the students through a focus group meeting and self-report questionnaires. This paper aimed to link ELT student teachers' opinions of task difficulty, complexity and conditions in flipped classroom principle about content and application attainment and discuss the implications of the findings for ELT teacher education.

1. Introduction

While it was reported that in 2013 in the ERIC database, a search of the phrase "flipped classroom" resulted in eight articles (Abeysekera & Dawson, 2014), in 2017 as we write this paper, July 4th, the same phrase results in 337 number of hits. Despite increasing attention to and popularity of flipped classroom approach, in literature, critical evaluations concerning it have been raised (Herreid & Schiller, 2016; Abeysekera & Dawson, 2014; Hamdan, McKnight, McKnight, & Arfstrom, 2013; Bishop, 2013). The main feature of flipped classroom is described as the case when "the information-transmission component of a traditional face-to-face lecture", to be referred as "traditional lecture" thereafter in the article, "is moved out of class time" (Abeysekera & Dawson, 2014, p. 1). Some other characterizations are listed such as focus on development of critical thinking skills; doing activities described as homework in-class for active engagement through extensive interaction with peers and

---

teachers during class time for knowledge construction (Bergmann & Sams, 2012; Missildine, Fountain, Summers, & Gosselin, 2013); class time being dedicated to whole-class brainstorming, group-based hands-on assignments, peer reviews, feedback exchange (Davies, Dean, & Ball, 2013); learner tasks planned to be heavily learner-driven preparation using the advancements to contact both with the teacher and the peer students taking more of a control and ownership of their own learning (Flumerfelt & Green, 2013). Andrews, Leonard, Colgrove, and Kalinowski (2011) maintain that especially at higher-level education, students experience learning difficulties due to traditional lecturing style that does not leave room for active learning. They highlight that active learning happens when "an instructor stops lecturing and students work on a question or task designed to help them understand a concept" (Andrews et al., 2011, p. 394). One approach that advocates active learning through technological devices and the Internet is flipped classroom achieved through coverage of content before class time (i.e. instructional videos, recorded lectures and online digital instructional materials) so that instructors allocate more class time to complex problem solving, deeper conceptual coverage and peer interaction (Strayer, 2012; Tucker, 2012; Gajjar, 2013; Sarawagi, 2013; Lee & Wallace, 2017).

If they intend to adopt the principles of flipped learning in their teaching practice, teachers are advised to consider the following four pillars of flipping while designing their lessons: flexible environment, learning culture, intentional content, and professional educator as described by the Flipped Learning Network (FLN) (2014). Within this framework, a flexible environment is one that allows flexibility in arranging learning spaces and time as well as in grouping students, building study groups, and assessing performance. Such flexibility paves the way for emergence of a new learning culture where students can benefit from deeper discussions and richer learning opportunities. They are active participants in knowledge construction and learning becomes personally meaningful. In order to attain conceptual understanding as well as procedural fluency, learners are exposed to differentiated intentional content tailored around their needs and expectations. Finally, the professional educators providing flipped instruction need to monitor their students' learning process closely, providing feedback when necessary, and reflect on their practice (FLN, 2014).

1.1. Flipped classroom and ELT teacher education

The theoretical underpinning of active learning advocated by flipped classroom is closely linked to constructivist learning theory. Jensen, Kummer, and Godoy (2015) take us back to the roots of constructivist, inquiry-based learning model (Heiss, Obourn, & Hoffman, 1950; Bybee, 1993; Lawson, 2002) and point out two phases under this model: content attainment phase and content application phase (p. 2). Through the first phase, students gain a conceptual understanding and through the second phase, they apply and/or evaluate what they have conceptually learned. When content attainment goes out of the class time and content application replaces it in class time, it appears that, in flipped classroom, the role of the teacher shifts. In flipped classroom, students are responsible for content attainment before class time and teachers are responsible for content application. Moving on with these concepts, Jensen et al. (2015) summarize studies from literature and maintain that studies testing the effectiveness of flipped models have not gone beyond case studies, from practitioners own classrooms, or comparative studies looking into flipped classroom compared to traditional classroom. They continue that the data concerning the differential effect of flipped classroom is limited given "so many potential causative mechanisms are being changed between treatments (e.g., shifting to active learning, including additional technology, using additional teaching materials, implementing peer instruction) that it is difficult, if not possible, to disaggregate them" (Jensen et al., 2015, p. 2). The focus of this study is yet to bring another mechanism, the task, that functions together with flipped classroom and examine the role task complexity, conditions and difficulty play in flipped classroom

learning, through a case study conducted at an English Language Teaching (ELT) department, with senior students, in an ELT course.

Flipped learning is not new; it has been practiced and referred to in different ways for 200 years (Sams & Bergmann, 2013). Recent technological changes have made it more nuanced and more emphasis is being given to learning process (Thoms, 2013) and Bloom's Taxonomy (1956) than videos or content and distribution of videos (Bormann, 2014). The question that lies ahead for educators is to examine whether flipped classroom actually works to increase student achievement towards acquiring 21st century skills and knowledge demands better than any other teaching approach. Abeysekera & Dawson (2014) point out that despite its popularity, there is little research concerning flipped classroom with systematic designs. They propose six testable propositions and call for research "that evaluates the effectiveness of particular components of the flipped classroom approach: pre-class activities, post-class activities, self-paced video lectures versus face-to-face lectures, and so forth" (Abeysekera& Dawson, 2014, p. 11). They further propose different types of investigations; one of which is "small-scale, localized interventions, including experimental studies: what is the efficacy of the flipped classroom approach in this discipline, this classroom, with these students?" (Abeysekera& Dawson, 2014, p. 11). In this study, we address this call as a gap in research and examine task component in a small-scale, localized intervention in an ELT department.

In a previous study, Çelebi et al (2016), investigated the role of corpus use in creating lexico-grammatical awareness, which falls under teacher expertise subject-matter domain, through flipped applications in an ELT department with student teachers. It was observed that the student teachers demonstrated awareness and gained more proficiency with respect to the target domain, examining language use in context. As early as the first year in their four-year teacher education program, they proposed opinions on effective teaching and learning of various lexico-grammatical units. In relation to material presentation through the flipped classroom, student teachers gave relatively high scores to the items on the scale evaluating motivation, course structure, student-instructor interaction, student-student interaction, and student-content interaction. Similarly, research into flipped instruction suggests that it contributes to the teaching-learning process in terms of allowing personalized learning (e.g., learner needs, pace), variety, and student preparation in advance, increasing student participation and interaction, and overcoming time-related limitations (Basal, 2015; Çelebi et al., 2016; Shannon-Chastain & Fell-Kurban, 2016; Lee & Wallace, 2017).

Egbert, Herman and Lee (2015) maintain that teacher learning requires student teachers to deal with different ways of teaching to be able to relate more and be flexible with different teaching contexts. They are advised to learn through applied literacies instead of transmission of disciplinary facts, which Darling-Hammond (2008) emphasizes by her proposition that teachers need to relate their learning of subject-matter into their lives to be able to teach their students through providing accessible and interesting teaching experiences. Loughran and Berry (2005) underline that modeled active teaching practices provide student teachers with more meaningful hands-on experience and teaching practices. This paper links ELT student teachers' opinions of task difficulty, complexity and conditions in flipped classroom principle about content and application attainment and discusses the implications of the findings for ELT teacher education.

1.2. The notion of task

The notion of task is an essential, but an under examined and neglected component of flipped classroom. In simplest terms, if we take the basic content of instruction as "lecture" and the practice and applying concepts as "homework" (Wallace, Walker, Braseby, & Sweet, 2014, p. 254), the practice and application of concepts, the homework part, or rather tasks, become crucial. In the case of

flipped classroom, regardless that educators act on the premise that students have covered and acquired the content outside the class, they need to design learning experiences for content application through meaningful, well-thought and well planned tasks, which is the crux of experiential learning paradigm which informs the applications of flipped instruction. One approach to sequencing tasks in flipped classroom is to utilize Bloom's Taxonomy. When it is applied to flipped classroom, the following is the case: Know is the focus for content attainment and higher-level cognitive processing; Evaluate and Create become the focus of content application, class time (Honeycutt, 2013). However, Wallace et al. (2014) caution us against simplifying Bloom's Taxonomy for flipped:

This compartmentalization of Bloom's taxonomy into lower- and higher-order, and then allocating each to either inside or outside of class, is clearly an oversimplification. Plenty of lectures, when engaged skillfully, can incorporate higher levels of thinking in how students stay involved. Conversely, simply telling students to work on something in class is no guarantee that they will be successfully applying course concepts. These in- and out-of-class "zones" of learning are best suited for particular types of activities that are aimed at particular levels of thinking (p. 258)

They emphasize that while inherently individualistic, information-transformation learning activities may lead to higher level thinking skills, guided practice with peers may not necessarily work towards facilitating higher-level skills. This emphasis they put on design and sequencing of learning activities is echoed in their proposition that the role flipped classroom gives to teachers is that of a cognitive coach, whose approach is "less content-driven and more focused on designing and facilitating activities that engage students in the process of thinking like an expert" (Wallace et al, 2014, p. 260).

Instruction, which is based on the notion of task, rejects the idea that learning happens independently of application. It bases learning on by doing or "experiential learning" (Dewey, 1933). According to Dewey, the principal element to learning is the activities built to engage learners to develop linkages between instructional content and learning environment and the environment that learning would be put to use beyond the learning environment. These tenets of experiential learning has grown into diverse learning theories as cited in Norris (2009, p. 579), "including cognitive psychological learning theories (e.g. Sternberg, 2003), apprenticeship and socialization frameworks (e.g., Lave & Wenger, 1991), and others (e.g., Kolb, 1984)." In these models "task" has become a common ground for implementations of experiential learning across disciplines such as medicine, (e.g., Virjo, Holmberg-Mattila, &Mattila, 2001), environmental studies (e.g., Wright, 2000), social work (e.g., Reid, 1997), and language education (e. g. Samuda and Bygate, 2008).

Once task is acknowledged as an essential component of flipped classroom, questions such as how teachers design and deliver tasks, and in what sequence, in order to sustain learner effort both for content and application attainment; whether complexity, difficulty of tasks and the conditions along which tasks are assigned influence content attainment, application and student performance arise. Research specifically looking into second language education shows that tasks have been categorized depending on task dynamics such as whether tasks are planned (Ellis, 2005); whether there is one fixed solution (closed) versus no fixed solution (open) (Long, 1989); whether the tasks demand references to current time and a shared context (Here-and-Now) or to distant time and unfamiliar context (There-and-Then) (Robinson, 1995). Robinson (2001) emphasizes that task differentials in terms of complexity, conditions and difficulty are important since they are "fixed and invariant features of the tasks" and a systematic use of them "will help explain within learner variance" (p. 30). He (2001, p. 30) outlines task complexity, conditions and difficulty in the figure below:
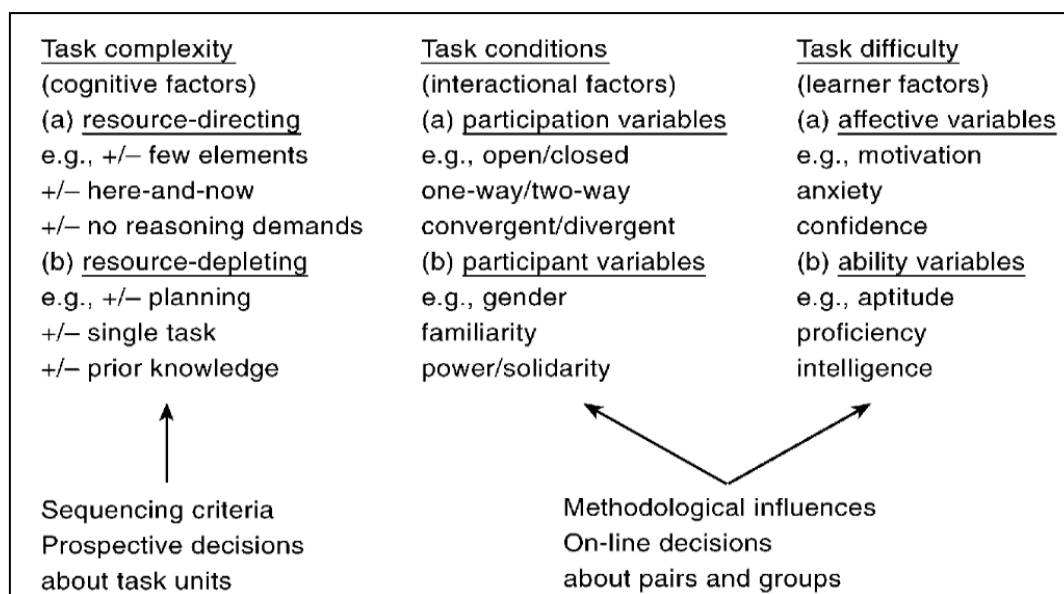
```
Task complexity              Task conditions              Task difficulty
(cognitive factors)          (interactional factors)      (learner factors)
(a) resource-directing       (a) participation variables  (a) affective variables
e.g., +/– few elements       e.g., open/closed            e.g., motivation
+/– here-and-now             one-way/two-way              anxiety
+/– no reasoning demands     convergent/divergent         confidence
(b) resource-depleting       (b) participant variables    (b) ability variables
e.g., +/– planning           e.g., gender                 e.g., aptitude
+/– single task              familiarity                  proficiency
+/– prior knowledge          power/solidarity             intelligence



Sequencing criteria          Methodological influences
Prospective decisions        On-line decisions
about task units             about pairs and groups
```

Figure 1. Task complexity, condition and difficulty (adapted from Robinson, 2001, p.30)

In this figure, it is illustrated that task differentials and conditions can be manipulated for optimum level task complexity and difficulty depending on pedagogical purposes. For example, an increased number of steps of tasks (+/-few elements) as opposed to one-step tasks (+/- few elements) on resource-directing dimension and allowing time for planning (+/- planning time) as opposed to not allowing time (+/- planning time) on resource-depleting dimension suggests a higher level of difficulty when differentials on both dimensions are (-). Moving along with this proposition and outline, in this study, task complexity, conditions and difficulty have been manipulated in an English Language Testing and Evaluation course with senior students of an ELT department in an English-medium state university in order to examine if task design influences student teachers' perceptions of flipped classroom principle that content attainment is to be achieved outside the classroom and content application to be carried out in the classroom. We predicted that as the tasks get more complex, difficult with more challenging interactional factors, both compartments of flipped classroom, content attainment and application attainment will be affected and student teachers' perceptions will positively or negatively shape around these challenges.

In this current study, two research questions were investigated:

1. What are student teachers' views and expectations regarding flipped classroom principle in relation to task design features (task complexity, task conditions, and task difficulty)?

1.1 What have they thought about them while doing (views)?

1.2 What do they expect to find in their future applications (expectations)?

2. What pedagogical implications can be drawn from student teachers' views and expectations regarding flipped classroom principle about content and application attainment?

2. Method

2.1. Context of the study and participants

The study was carried out at the Faculty of Education, Department of English Language Teaching at a state university in Ankara, Turkey. Although English is not an official language of the country, there is a high level of demand for graduates of ELT programs in the country. According to the reports

published in 2013 by The Turkish Student Selection and Placement Center (ÖSYM), the only organization that administers the National Placement Tests in Turkey, it was stated that; for instance, at the tertiary level, there were 164 universities in Turkey with more than 75% of their programs instructed in English, and 18.5% of all bachelor degree programs were carried out in English (Arık & Arık, 2014). In the European Credit Transfer and Accumulation System (ECTS), student teachers of ELT complete 240 ECTS in 4 academic years. Higher Education Council (YÖK) outlines the curriculum of ELT programs across the country to be implemented approving a certain level of flexibility for distinction among universities. In general, the first two years of the ELT programs include theoretical, subject-matter related courses aiming to build linguistic awareness of English language and teaching such as introduction to linguistics, language acquisition, language teaching methods and assessment. The last two years include courses targeting to develop student teacher expertise through practice. In this study, 29 senior year students, 23 female and 6 male with an age-range of 22-24, taking English Language Testing and Evaluation course participated.

2.2. Instruments

Data regarding flipped classroom principle in relation to task design features were collected from the students through a focus group meeting and self-report questionnaires. The focus group meeting in which the participants were asked to offer their opinions on the assigned tasks took place at week nine, halfway through the semester. The questionnaire was administered at the end of the 16-week semester. In total, 41 minutes of spoken data from the focus group meeting was collected. There was a student moderator who signaled the number of the task the members would speak about each time and they talked about each task for about 5 minutes. The entire meeting was recorded on a cell phone. The questionnaire administered was modified from Tang and Chaw (2013), including seven constructs and 37 items in total. The seven constructs were learning flexibility, online learning, online study management, technology use, classroom learning, online interaction and flipped learning. All the items of individual constructs were kept the same except the last part which originally focused on blended learning. These three items were modified to refer to flipped learning instead of blended learning.

2.3. Data collection procedures

Prior to the study, all the ethical procedures of the study were completed and the study was approved by the Ethics Committee of Ankara Yıldırım Beyazıt University. The participants contributed to the study with their informed-consent. Our participants, a group of senior students in an ELT department, were trained on language testing and assessment as part of an English Language Testing and Evaluation course. The students were provided with theoretical background, practical examples, and hands-on experience over an academic semester. Prior to their classes each week, students viewed videos and/or slides and completed the assigned readings, and then participated in in-class discussions as well as completing the weekly tasks. The tasks that varied in terms of complexity, conditions and difficulty level (cognitive and interactional factors as well as learner factors), student groupings and interactional patterns (individual, pair, or group), and setting (in-class, in-lab or online) included the following:

(1) Writing a reflective paragraph summarizing the main points in the video/assigned reading, and presenting their own opinions on it; individual work; online;

(2) Providing written feedback on peer work using the two scales introduced and evaluating the scales in terms of efficiency; individual work; in-class (lab);

(3) Applying the principles of language assessment on the assessment scenarios provided for analysis and evaluation; individual or pair work; online;

(4) Critically analyzing actual tests using the guidelines provided to identify the underlying principles employed, detect the problematic points, and offer alternatives; individual, pair or group work; online;

(5) Developing actual tests using the guidelines provided; individual, pair or group work; both in-class and online;

(6) Revising and editing peer work (tests); individual or pair work; in-class;

(7) Participating in a writing scoring standardization session, reading and scoring sample student paragraphs using the scoring rubric provided; individual and group work; in-class;

(8) Participating in a speaking scoring standardization session, listening to and scoring sample student spoken performance using the scoring rubric provided; individual work; in-class (lab).

Upon the completion of the nine-week period, students were required to respond to the focus group discussion questions with reference to each task described above: What are your views and expectations regarding flipped classroom principle in relation to task design features (task complexity, task conditions, and task difficulty)? What have you thought about them while doing (views)? What do you expect to find in their future applications (expectations)?

In order to arrive at pedagogical implications, the transcribed spoken data from the focus group meeting on student teachers' views and expectations regarding flipped classroom principle were subjected to further content analysis by the researchers with reference to content and application attainment.

Subsequently, a student self-report questionnaire on flipped learning (a slightly modified version of Tang and Chaw's (2013) questionnaire) was administered at the end of the 16-week semester in order to collect data on student teachers' views towards flipped instruction as opposed to face-to-face classroom learning.

2.4. Data analysis

In order to answer the first research question on the student teachers' views and expectations regarding flipped classroom principle in relation to task design features (task complexity, task conditions, and task difficulty), qualitative data were collected through a focus group meeting. The entire meeting was recorded on a cell phone, and 41 minutes of spoken data was collected and transcribed by the researchers. 25 students participated in the focus group meeting. The content analysis (Creswell, 2012) was carried out as follows: (1) organize the data, (2) explore and code the data, (3) construct descriptions and themes, (4) identify the qualitative findings, (5) interpret the findings, and (6) validate the accuracy of the findings. During the data analysis, the transcribed spoken data were read individually and grouped based on the points students considered, including task complexity, task conditions, and task difficulty. At the same time, the researchers analyzed and generated questions in order to determine further common themes or patterns. Both researchers compared and discussed the content analysis to finalize the themes. The explanations were aligned with these findings. The findings were presented (shown below) without comment to demonstrate the actual data, and then the results were interpreted.

Quantitative data were collected through a 37-item five-point Likert-type questionnaire to answer the first research question as to the students' views on flipped learning. 21 students completed the questionnaire. The data gathered through the questionnaire were analyzed using descriptive statistics (means and standard deviations).

In order to answer the second research question regarding the pedagogical implications that can be drawn from student teachers' views and expectations regarding flipped classroom principle about

content and application attainment, qualitative data collected through the focus group meeting were subjected to a secondary content analysis by the researchers. During the data analysis, the transcribed spoken data were read individually and grouped based on the content vs application attainment aspects of flipped instruction. The data were categorized based on the similarities and differences in views, and explanations were aligned with these findings.

3. Results

3.1. Student teachers' views on flipped tasks with reference to their experiences and their expectations for future applications

The students were asked to reflect on their experiences and expectations in relation to the tasks designed and applied for the purposes of this study in a focus group discussion session. They had positive views about the tasks in most respects though they offered some ideas for improvement in some.

In Task 1, the students were asked to write a reflective paragraph summarizing the main points in the video and assigned reading that focused on formative and summative assessment and presenting their own opinions on them individually via website postings. With respect to task complexity, it had five elements which included reading, listening, watching, summarizing, and reflecting and it required prior knowledge of weekly coursework (video, lecture, and reading). All the students found the cognitive demand manageable. In terms of task conditions that manipulate interactional factors, they reported the task allowed peer-interaction and open-ended responses and thus it was creative and effective. They also found it quite familiar in that posting their responses on the website was like posting a comment on facebook. With respect to task difficulty which incorporated learner factors that relate to affective and ability variables, they thought the task led to higher levels of motivation and better learning of the content as it was realistic and relevant to future work, required retrieval of information from multiple sources, and required flexibility (being online) and discipline at the same time making the students take the responsibility of their learning.

As of Task 2, in which the students were required to provide written feedback on each other's reflective paragraphs using the two scales introduced and evaluating the scales in terms of efficiency in the computer lab, they thought the cognitive load was much less in terms of task complexity as they were already familiar with the content and only needed to reconsider their feelings and reactions through alternative perspectives. They reported Task 1 and 2 had similar task conditions, both allowing peer-interaction ad open-ended responses, which was an advantage; however, in terms of task difficulty, they believed affective factors partially interfered with the ability factors in that they had difficulty scoring peer work and avoided being critical not to offend each other. Despite that, they admitted the experience, upon receiving instructor's feedback on their Task 1 and 2 performance, helped them gain this critical perspective for later work and was thus motivating.

In order to meet the requirements of Task 3, the students applied the principles of assessment on the assessment scenarios provided for analysis and reflection. This was an online task and they were allowed to work individually or in pairs. They found this task quite complex and demanding as it required a thorough understanding of the material and proper application of it. Some of the scenarios provided for analysis were not detailed and had room for assumptions, which would allow further speculation and in-depth discussion; however, this was seen as a drawback of the task and they preferred everything to be specified in detail. In terms of task conditions, they found it creative, practical, and realistic and thought it encouraged shifting perspectives; however, some found it really tough due to these qualities. With respect to task difficulty, they believed, being applicable in future

work situations (applying terms in a life-like situation, useful for consolidation, using theoretical information for practice, and no room for memorization), the task was quite motivating. Finally, some still thought it would have been much more beneficial if the task had been conducted as an in-class activity during which students could have received further guidance from their instructor.

In Task 4, the students critically analyzed actual reading and listening tests using the guidelines provided to identify the underlying principles employed, detect the problematic points, and offer alternatives. This was an online task and they were allowed to work individually, in pairs or groups. They found the reasoning demands quite high in that the task required critical examination and analysis of tests and identification of problems as well as offering solutions. Besides, they had to rely on prior knowledge and transfer that information in novel ways. With respect to task conditions and difficulty, some liked the challenging, creative, open-ended, and life-like nature of the task and the great freedom they enjoyed working online as well as interacting with others in class. They said they received important information that could be transferred and used in other courses such as material adaptation and design or school experience. Some others, however, found it very challenging as they said they did not feel knowledgeable or experienced enough to complete the task. These students thought it would have been better if the task had been designed as a solely in-class activity and the part that required them to offer alternatives/solutions had been removed.

In Task 5, the students developed an actual reading test using the guidelines provided. This was both an in-class and online activity and they were allowed to work individually, in pairs or groups. In terms of task complexity, the task had three elements including writing test specifications, writing or adapting a text, and writing test questions though each consisted of further steps such as doing readability and vocabulary profile analysis on the texts, or providing detailed information for the separate sections in the test specifications. They had to rely on prior knowledge from lectures and readings on test writing as well as critically analyzing texts during lab sessions using the recommended web tools. Therefore, most of the students found some parts of the task, adapting a reading text and writing test specifications, very challenging and offered a complete exclusion of these parts for future applications. However, they were all well aware that these were essential aspects of developing tests and they admitted such steps had to be included. With respect to task conditions and difficulty, though most had difficulty completing the task, they found the lab environment and online platforms appealing; besides, they said they benefitted significantly from their experiences doing text analyses and using and becoming familiar with related web tools and found this task very much relevant to their future work. Therefore, they were quite motivated.

In Task 6, the students revised and edited each other's reading tests and this task was completed in the classroom. In terms of complexity, the task included three main elements of analyzing the test in depth, providing a complete moderation by giving feedback and offering alternatives, and they had to rely on previous knowledge on how to write reading tests and conduct text analysis using web tools. They found it quite challenging, and some said they had a really difficult time offering alternatives for the problematic parts, despite their previous experience in Task 4 and 5. With respect to task conditions and difficulty, despite the challenge it posed, they said the task was creative, encouraging, and appropriate for their future work and thus led to higher levels of motivation and confidence. They added that they learnt more about different text and question types and gained further experience into providing appropriate feedback to test-writers, all of which contributed to their confidence-building process.

In Tasks 7 and 8, the students participated in writing and speaking scoring standardization sessions, and scored performance using the scoring rubric provided. With respect to complexity, the tasks had three main elements of attending standardization sessions to work on samples, scoring the spoken and written student records provided, and using rubrics in scoring. They had to rely on previous

knowledge on how to score student performance (written and spoken), and understand and use rubrics appropriately. They found the tasks challenging considering the cognitive demand involved but liked such hands-on experience. In terms of task conditions, they thought the tasks were creative and they were not confusing in any way as all the necessary details were specified during the standardization sessions and rubrics were clear. As of task difficulty, they found the tasks appropriate for internship and future work, authentic, and enjoyable despite the challenges, which led to higher levels of motivation and confidence.

3.2. Student teachers' views on flipped learning and related constructs

The means and standard deviations of the students' ratings in response to the questionnaire items reflecting their attitudes towards flipped learning and related constructors that include learning flexibility, online learning, online study management, technology use, classroom learning, and online interaction are reported in Table 1, below. Prior to the analysis for descriptive statistics in SPSS, the negative items (5,11,12,13,28 respectively) were reverse coded. Nevertheless, the items regarding students' attitudes towards classroom learning were kept as they were because it was aimed to compare the ratings for classroom learning versus online or flipped learning independently of each other.

Table 1. Means and standard deviations of student teachers' attitudes towards flipped learning

	T*	M	SD	N	%
Learning Flexibility	20	17.2	3.27	21	86
Online Learning	40	25.4	8.79	21	64
Online Study Management	30	17.9	7.44	21	60
Technology Use	20	15.9	4.28	21	80
Classroom Learning	25	19.5	4.81	21	78
Online Interaction	35	25	8.14	21	71
Flipped Learning	15	11	2.83	21	73

T* represents the highest mean score possible; e.g., for the learning flexibility construct, it is 5 (the highest point on the scale) *4 (number of items in this subscale) = 20.

In order to be able to compare student ratings across constructs with varying numbers of items, the relevant percentages (the "%" column in Table 1, above) were calculated using the totals (the highest mean possible in each case, the "T" column above) and the actual mean values (the "M" column above) of the data collected. Based on these percentages, it was revealed that the students generally had positive attitudes towards learning flexibility, online learning, online study management, technology use, online interaction and flipped applications, with percentages varying between 60 and 86. On the other hand, they had positive attitudes towards classroom learning with a percentage of 78 as well.

The learning flexibility construct, which included items such as "I like to study at my own pace," received the highest ratings from the students (86%). The constructs of technology use and classroom learning received the second highest set of ratings (80% and 78%, respectively) while the flipped learning and online interaction constructs received relatively lower ratings, standing at 73% and 71% respectively. The lowest set of ratings were given for the constructs of online learning and online study management, with 64% and 60% respectively.

In order to determine the specific items that needed further attention in the flipped class design in later applications, we needed to examine the students' ratings for each item under each construct more closely. Therefore, the individual means of all items in the survey were computed and compared against the mean of the construct each belongs to. Accordingly, five items that were below their respective subscale means were selected and enlisted as high-priority items to be dealt with in later flipped applications, with the exception of the items in the classroom learning construct. As for the items regarding classroom learning, the items that were above their respective subscale mean were selected as they worked based on the assumption that classroom learning is more preferable over online or flipped learning (see Table 2 below).

Table 2. The items on the attitude questionnaire that received lowest ratings

Subscale	Item	Item Mean	Subscale Mean
Online Learning	I would like to have my classes online rather than in the classroom.	2.62	3.17
	I am more likely to miss assignment due dates in an online learning environment.	Reversed: 2.81 Original: 3.19	
	I organize my time better when studying online.	2.86	
Online Study Management	Online learning encourages me to make plans.	2.86	2.98
	Online learning makes me more responsible for my studies.	2.90	
Classroom Learning*	I find learning through collaboration with others face-to-face is more effective.	3.95	3.90

Classroom learning*: The highest ratings were considered for the items in the classroom learning construct.

As Table 2 above indicates, the students reported a preference towards face-to-face classroom learning over online learning, and suffered from online study management issues such as planning studies, managing time and meeting deadlines, and taking the responsibility of their own learning to some extent.

3.3. Pedagogical implications of the flipped learning principle as to content vs application attainment

In the final stage, to collect specific data regarding content vs application attainment aspects of the flipped instruction and draw their implications for pedagogical purposes, the transcribed spoken data from the focus group meeting on student teachers' views and expectations regarding flipped classroom principle were subjected to further content analysis by the researchers and emerging themes were presented. As it is indicated in the summary table (Table 3) below, overall, the data revealed that for the completion of all tasks, the content was attained through both online (outside the classroom) and in-class/lab activities, while for application attainment, activities were designed to be completed either solely in-class/lab or both in-class/lab and online, with the exception of Task 1 in which the application stage was completed outside the classroom (website posting).

In Task 1, writing a reflective paragraph summarizing the main points in the video/assigned reading and presenting their own opinions on it, the content was partially attained through in-class discussions based on the reading and online video-viewing activities prior to the class. The students were then required to summarize and comment on the main points in the content individually through

online postings on the class website, which forms the application attainment aspect of this task. As a result, while the content was presented both in-class and online, the application was carried out on the online platform solely. Due to task aspects such as being familiar, realistic, relevant and allowing peer-interaction, open-ended responses, and a manageable amount of cognitive load, it seemed to achieve the objectives of understanding what formative and summative assessments are and applying that information in new contexts.

Table 3. Content vs application attainment practices observed

Task	Content Attainment	Application Attainment
1: Understanding what formative and summative assessments are and writing a reflective summary on them	Online In-class	Online
2: Understanding the qualities of analytic vs. holistic scoring tools, applying that information in providing written feedback on peer work and evaluating the efficiency of the tools	Online In-class	In-lab
3: Applying the principles of language assessment on the assessment scenarios provided for analysis and evaluation	Online In-class	Online In-class/lab
4: Critically analyzing actual tests using the guidelines provided to identify the underlying principles employed, detect the problematic points, and offer alternatives	Online In-class	Online In-class/lab
5: Developing actual tests using the guidelines provided	Online In-class	Online In-class/lab
6: Revising and editing peer work (tests)	Online In-class/lab	In-class/lab
7: Participating in a writing scoring standardization session, reading and scoring sample student paragraphs using the scoring rubric provided individually and in groups	Online In-class	Online In-class
8: Taking part in a speaking scoring standardization session, listening to and scoring sample student spoken performance using the scoring rubric provided individually	Online In-lab	Online In-lab

In Task 2, providing written peer feedback using the two scales introduced and evaluating the scales in terms of efficiency individually in the lab, the content is attained through both online material available at the coursera website and in-class discussions on the efficiency of the two scales (one analytic, one holistic scoring rubrics). The students were allowed to review the video content again while working on the scales in the lab. Then they were asked to apply that information in the lab: Using the scales to provide feedback on their classmates' reflective paragraphs (Task 1) while at the same time critically evaluating to what extent the scales are efficient. Due to task aspects such as being relevant and allowing peer-interaction, open-ended responses under teacher supervision, it seemed to achieve the objectives of understanding the qualities of analytic vs. holistic scoring tools, applying that information in providing written feedback on peer work and evaluating the efficiency of the tools. However, without the subsequent teacher feedback, it would not have been possible for the students to gain a truly critical perspective considering the affective factors involved such as having difficulty scoring peer work in order not to offend each other by being critical.

In Task 3, applying the principles of language assessment on the assessment scenarios provided for analysis and evaluation, Task 4, critically analyzing actual tests using the guidelines provided to identify the underlying principles employed, detect the problematic points, and offer alternatives, and Task 5, developing actual tests using the guidelines provided, the content was attained through both online materials in the form of videos, slides and readings and in-class discussions on the principles of language assessment and the information regarding test design and preparation. Subsequently, the students were expected to apply testing principles on the scenarios provided, analyze actual reading and listening tests, improve their problematic aspects, and finally write their own reading tests. As the reasoning demands in these three tasks were quite high, they were allowed to work in groups/pairs or individually and both in the lab and online, whereby they could receive teacher guidance as well. However, due to time limitations, they had to work on most aspects on their own outside the class hours. Therefore, some students found the application aspects of the tasks quite complex and challenging, felt incompetent, and thought it could have been designed as an entirely in-class activity rather than a partially online task to be scored, although they agreed these were important skills each need to develop sufficiently prior to graduation to be able to fulfill the requirements of their profession. Some other students, though, stated that they had important gains and even developed skills they could transfer across courses (such as materials development and school experience). On the whole, it turned out that this set of tasks could have produced better results and better attainment of the objectives targeted if more in-class time had been reserved for the application aspect.

In Task 6, revising and editing peer work (tests) individually or in pairs in the classroom, the content was attained both online and in-class/lab focusing on reading test design including the steps of writing test specifications, writing or adapting a text, doing readability and vocabulary profile analysis, and writing questions. Especially the lab session allowed specific hands-on experience under teacher supervision. The actual revising and editing part, the application itself, was conducted in the classroom again in the presence of teacher as the guide and facilitator. In this part, the students were required to analyze the test, provide moderation with feedback and offer alternatives (when necessary), individually or in pairs depending on their preference. They seemed to accomplish the objectives entailed through the tough but encouraging task and built further confidence due to their additional experience into test design (as they had previous practice on test design in Tasks 4 and 5).

In Task 7, participating in a writing scoring standardization session, reading and scoring sample student paragraphs using the scoring rubric provided individually and in groups in the classroom, and Task 8, taking part in a speaking scoring standardization session, listening to and scoring sample student spoken performance using the scoring rubric provided individually in the classroom, the content was attained through both online and in-class activities where the students had the opportunity to build the essential background as to how to score written/spoken student performance working on scored samples. Following that, at the application stage, in the classroom/lab, they were required to actually score student performance attending standardization sessions and using relevant rubrics, simulating the real life scoring context. Despite being an authentic, creative and challenging hands-on task with a high cognitive load, these well-structured tasks seemed to achieve the objectives of scoring student performance (both written and spoken) by attending prior standardization sessions and using rubrics appropriately. The process was facilitated with the presence of teacher providing continuous guidance and clarification.

In sum, as Table 3 above indicates, the instruction employed flipping in complementary ways, allocating room for content as well as application attainment both inside and outside of class.

4. Discussion and Conclusion

In this study, we explored student teachers' views and expectations regarding flipped classroom principle in relation to task design features. We further investigated content vs application attainment aspects of flipped instruction through manipulation of eight tasks in terms of their complexity, conditions and difficulty (Robinson, 2001). We aimed at examining if task design influences student teachers' perceptions of flipped instruction that is based on the principle that content attainment is to be achieved outside the classroom and application to be carried out in the classroom. Our context was an English Language Testing and Evaluation course and the participants were the senior students of an ELT department in an English-medium state university. Data were collected from the participating students through a focus group meeting and a self-report questionnaire.

The results of the focus group meeting in which the students were asked to reflect on their experiences and expectations in relation to the tasks designed and applied for the purposes of this study showed that they had positive views about the tasks in most respects though they offered some ideas for improvement in some. With respect to content, they commented that the tasks in general were realistic and applicable, not just theoretical. They thought although at times overwhelming, tasks required a lot of hands-on, experiential work unlike theoretical exams, which they rated as a positive flipped application. As the tasks got more complex requiring higher cognitive functioning with more prior knowledge and higher reasoning demands as in the case of writing actual tests, learner factors such as anxiety and aptitude came into play and relatively less competent students proposed some updates on the content concerning task scope, sequencing, points allocated, or the number of task elements, and they were actually put into practice online (see the online decisions feature in the task design model in Robinson (2001), Figure 1 reproduced below again for convenience), and especially over the weeks subsequent to the focus group discussion session. As predicted and confirmed by Robinson (2001, 2003) as part of his task design model and the underlying Cognition Hypothesis of task-based learning, which proposes a complexity-based sequencing of tasks from simpler to more complicated ones, a systematic use of task features with various combinations and modifications could explain both within and between learner variance, guiding both the learner and the instructor about the actions to be taken, sometimes for the individual students and sometimes for the entire group.
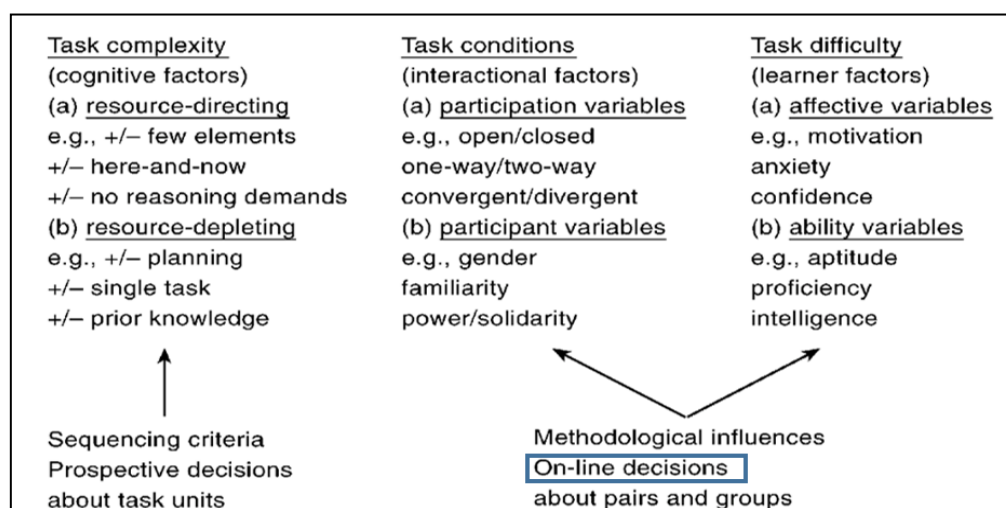

Task complexity (cognitive factors)	Task conditions (interactional factors)	Task difficulty (learner factors)
(a) resource-directing	(a) participation variables	(a) affective variables
e.g., +/− few elements	e.g., open/closed	e.g., motivation
+/− here-and-now	one-way/two-way	anxiety
+/− no reasoning demands	convergent/divergent	confidence
(b) resource-depleting	(b) participant variables	(b) ability variables
e.g., +/− planning	e.g., gender	e.g., aptitude
+/− single task	familiarity	proficiency
+/− prior knowledge	power/solidarity	intelligence

Sequencing criteria
Prospective decisions
about task units

Methodological influences
On-line decisions
about pairs and groups

Figure 1 (rep). Task complexity, condition and difficulty (adapted from Robinson, 2001, p.30)

The views regarding the medium of the course and the tasks, as to whether it was in-class/lab or online, were mainly positive. They commented that the suggested selections of medium for task

completion were appropriate although with majority of the tasks they were provided with both options. For instance, in Task 7, when the students were not able to complete scoring sample written work, they were allowed to work on them further outside the class hours. Again, in terms of student groupings which included individual, pair or group work, students thought that being offered various grouping options for majority of the tasks gave a positive feeling. The idea of providing various grouping options with almost all tasks was mainly due to the variations in students' preferences which seemed to stem from certain interactional factors or learner factors as illustrated in the task design model (Figure 1 above). For instance, some relatively more competent students preferred working individually on tasks as they could safely rely on their prior knowledge and experience for the completion of the tasks or just due to practical reasons such as having limited time to interact with a partner or group members outside the class hours. Similarly, the deadlines regarding task submissions were extended occasionally with a consideration of students' changing workload depending on other coursework including internships, research or data collection duties or exam preparation (since these were senior students preparing for graduate studies or job applications, they were under the pressure of preparing for nationwide aptitude and English language tests or international English language tests). Consequently, online decisions as to mainly task conditions or interactional factors such as task medium, student groupings or deadlines had to be made, which further facilitated the learning process.

Another emerging theme in the focus group meeting concerned student involvement. They expressed a positive feeling in terms of student involvement as they felt they were shaping the course; their views were taken into consideration; there was discipline but the teacher was not "strict" as to course components and allowed flexibility depending on emerging needs and expectations. These were quite competent senior students, well aware of what they already know and what they need to practice further; nevertheless, they varied especially with respect to learner factors, regarding both affective and ability variables as exemplified in the task design model (Figure 1 above). The students who seemed to feel less confident about their testing knowledge and skills were quite anxious and less motivated to welcome the challenge, novelty or creativity required in relatively more real-life, complex and open-ended tasks, and therefore suggested a revision on such tasks turning them into instructor-guided classroom activities. Again with such tasks, online decisions were made as to task scope or components, reasoning demands, points allocated or student groupings in an effort to overcome the issues regarding affective variables and optimize the learning outcomes for the entire group. Most of the students seemed to benefit from this collective learning environment in which tasks were not completely cancelled but exposed to further chunking or filtering not at the expense of complexity. As advised through research into teacher learning and teacher education (Egbert, Herman & Lee, 2015; Darling-Hammond, 2008; Loughran & Berry, 2005), student teachers need hands-on practice on applied literacies and modeled active teaching practices in order to be equipped with the tools and strategies that will help them relate more and be flexible in various teaching environments.

To summarize, student teachers were mainly positive about the task features manipulated in terms of task complexity, task conditions, and task difficulty and designed with a consideration of variations in sequencing, grouping, time demands and familiarity of content and online decisions. In this study, the online decisions in Robinson's model (2001) were perceived in a more flexible manner to include the updates of any kind (not confined to student groupings) made based on the emerging needs (of not only the learners but also the instructors), which, coupled with occasional manipulations on prospective decisions as to task units, facilitated the entire process allowing efficient task use and helped to achieve active learning (Andrews et al., 2011). The incorporation of flipped features in course media, allocating outside the class or online hours as well as in-class/lab time for instruction, further contributed to the accomplishment of such active learning with more class time for complex

problem solving, deeper conceptual coverage and peer interaction (Strayer, 2012; Tucker, 2012; Gajjar, 2013; Sarawagi, 2013; Lee & Wallace, 2017).

Further data on the flipped features employed in this study were collected using a self-report questionnaire. The results of the self-report questionnaire on flipped learning revealed that the students generally had positive attitudes regarding the constructs that contribute to flipped learning such as learning flexibility, online learning, online study management, technology use, online interaction and flipped applications. On the other hand, they still showed a preference for classroom learning over flipped learning, though the percentages reflected a marginal difference. They also reported having some difficulties in certain aspects of flipped or online instruction such as planning studies, managing time, meeting deadlines, and taking the responsibility of their own learning.

Considering these findings, one could argue that this preference for classroom learning and the hardships experienced during this transition from a solely face-to-face to flipped or blended instruction might have stemmed from students' old study habits as well as their expectations regarding the traditional role of instructor as the single source and transmitter of information. With the advances in technology and availability of information anytime and anywhere, students are now quite digitally-skilled, acquainted with online resources and learning platforms, and enjoy the flexibility and variety offered; however, this does not entail that they know how to learn through digital technologies and most need help. When instruction is confined to traditional classroom activities, the overall layout and the content to be covered are quite straightforward and the entire process is mostly planned and conducted by and in the presence of the instructor. However, when classroom instruction composes only one third of the overall course and the rest requires self-directed learning outside the classroom through discovery and experimenting, not necessarily under the supervision of an instructor, even quite competent students have difficulty planning their studies, staying focused, and keeping up with deadlines. Therefore, as it has been reported in previous studies (Abeysekera & Dawson, 2014; Basal, 2015; Çelebi et al., 2016; Shannon-Chastain & Fell-Kurban, 2016; Lee & Wallace, 2017), in order for flipped instruction to contribute to the teaching-learning process, certain issues need to be dealt with efficiently, such as allowing personalized learning (e.g., learner needs, pace), variety, and student preparation in advance, increasing student participation and interaction, and overcoming time-related limitations.

In this study, within the flipped framework (FLN, 2014), a flexible environment was achieved by allowing flexibility in arranging learning spaces and time as well as in grouping students, building study groups, and assessing performance. Nevertheless, there was still limited time for building a new learning culture where students could benefit from deeper discussions and richer learning opportunities due to the overall course load. Considering the nature of this testing and evaluation course in which this study was conducted, it was quite a challenge to allocate adequate time and energy for all of the course objectives (see Appendix A for the course outline developed based on the requirements of the curriculum outline specified by the Higher Education Council, YÖK, in Turkey), and as it was observed in student teachers' reflections on flipped instruction with reference to the tasks assigned in the previous section, they needed more time for consolidation, elaboration and opportunity for student-student and student-instructor interaction to perform well on the tasks.

Similar results were obtained in another study on the summative evaluation of the same course in an English-medium state university in the Turkish context in which the students complained about how quickly they had to cover some important topics due to time limitations and suggested that the ELT curriculum should include at least two but preferably more testing courses for the better accomplishment of course objectives (Hatipoğlu, 2010). This suggestion has been further confirmed in this study and one implication of this study for teacher educators would be to offer a series of testing and evaluation courses. Student teachers would first focus on an understanding of the basic terms and

concepts related to language testing and evaluation and engage in various processes and practices related to assessment of language proficiency, and then take part in studies on developing a variety of testing instruments for various groups of language learners and running the statistical analysis of testing data. Consequently, such curricular revisions would likely increase active student participation in knowledge construction and learning would become personally meaningful as individual learners would be offered differentiated intentional content tailored around their needs and expectations and there would be adequate time for feedback and reflection, which are among the main advantages the flipped instruction offers. Accordingly, this inherently loaded scope of the language testing and evaluation course might have functioned as a confounding limiting factor in this study where flipped instruction was employed. Future research into the integration of task design features within a flipped framework might focus on teacher education courses with a narrower scope so that other course features, such as scope and sequence, do not interfere with the application of the specific intervention under focus (Jensen et al., 2015).

Another topic of investigation in this study concerned the potential pedagogical implications that could be drawn from student teachers' views and expectations regarding flipped classroom principle about content vs application attainment through a secondary content analysis of the focus group meeting data adopting this perspective. We explored whether task design influenced student teachers' perceptions of flipped instruction that is roughly based on the principle that content attainment is to be achieved outside the classroom and application to be carried out in the classroom. The results showed that for the completion of all eight tasks, the content was attained through both online (outside the classroom) and in-class/lab activities, while for application attainment, activities were designed to be completed either solely in-class/lab or both in-class/lab and online, with the exception of Task 1 in which the application stage was completed outside the classroom (website posting). As we predicted, as the tasks got more complex, difficult with more challenging interactional factors or learner factors, both compartments of flipped classroom, content attainment and application attainment seemed to be affected and students' perceptions shaped around these challenges.

With respect to task features, as the tasks got more complex due to higher reasoning demands, most students, especially the less competent ones, needed higher teacher assistance and peer support, both in-class/lab and online/outside the class, for task completion or confirmation. They even requested such open-ended and unfamiliar tasks to be transformed into in-class activities not to be scored. Therefore, the idea of offering more class time for higher-order application tasks in the flipped design (Honeycutt, 2013) and providing variety in student groupings worked well.

However, considering Wallace et al.'s (2014) warning as to the compartmentalization of tasks and skills into lower- and higher-order, allocation of each either outside (content attainment) or inside of class (application attainment) was a proposition which barely reflected the reality in our case, and in fact the boundaries between content and application were quite fuzzy. As the results indicated, both took place in both settings, and not necessarily due to task complexity. Wallace et al.'s (2014) prediction was further confirmed. It was surprising that content attainment also required some class time each week even with these senior students, though at a smaller amount compared to application attainment. This was because students occasionally had difficulty planning their studies and keeping up with the overall pace of the course due to some of the limitations stated above (course scope, other coursework, internships, exam preparation) or some learner factors (confidence, anxiety, proficiency), and thus required brief content revisions prior to application under the supervision of the instructors.

Both findings as to the fuzzy boundaries between content and application attainment and students' content revision needs are indeed indicative of an important implication for the later applications of flipped approach and teacher education in general. Wallace et al. (2014) emphasize the role of instructors adopting the flipped approach as a cognitive coach who designs engaging learning

experiences for students encouraging them to think like experts; nevertheless, our study indicates that it is quite a challenge to free instructors and students from their traditional roles. Flipped instruction requires self-directed learning outside the classroom through discovery and experimenting, not necessarily under the supervision of an instructor, but even quite competent students have difficulty staying focused and keeping up with the coursework. Therefore, without prior training on autonomy-building and self-regulated learning that needs to be integrated into the overall teacher education curriculum, flipped instruction seems to provide less of a service, especially in the case of more complex application tasks, unfortunately defeating the whole purpose. High-quality flipped or blended instruction needs to be personalized (serving individual needs), mastery-based (allowing students move on and receive credit only after they have mastered a core concept), and based on high expectations for all students and student ownership/agency (empowering students with the necessary skills, information, and tools to manage their own learning). Besides, considering the specific context of this study, an English Language Testing and Evaluation course with the senior students of an ELT department in an English-medium state university, certain curricular issues need to be revisited as to the number and scope of compulsory and elective courses to be taught in the fourth year as these might have a significant influence on the design of individual courses and their requirements.

Finally, this study was conducted at a single institution in one country but its findings are relevant to a wider audience receiving a variety of undergraduate courses and training to become teachers. In addition, the results of the current study could guide all the parties involved in the decision-making processes of any teaching-learning setting including students, instructors, curriculum developers and course designers as well as the professionals and administrators in such institutions.

References

Abeysekera, L., & Dawson, P. (2014): Motivation and cognitive load in the flipped classroom: definition, rationale and a call for research, *Higher Education Research & Development*, DOI: 10.1080/07294360.2014.934336

Andrews, T., Leonard, M., Colgrove, C., &Kalinowski, S. (2011). Active learning not associ- ated with student learning in a random sample of college biology courses. Life Sciences Education, *10(4)*, 394–405.

Arık, B.T., & Arık, E. (2014). The role and status of English in Turkish higher education. *English Today, 30*, 5-10.

Basal, A. (2015). The implementation of a flipped classroom in foreign language teaching. *Turkish Online Journal of Distance Education-TOJDE, 16*(4)*, 28–27.

Bergman, J., & Sams, A. (2012). *Flip your classroom: Reach every student in every class every day*. Eugene, OR: International Society for Technology in Education.

Bishop, J. L. (2013). *A controlled study of the flipped classroom with numerical methods for engineers*. Utah State University, Unpublished PhD Dissertation.

Bloom, B. S. (1956). *Taxonomy of Educational Objectives*, Handbook I: The Cognitive Domain. New York: David McKay Co Inc.

Bormann, J. (2014). *Affordances of flipped learning and its effects on student engagement and achievement*. University of Northern Iowa, Unpublished PhD Dissertation.

Bybee, R. (1993). *An Instructional Model for Science Education: Developing Biological Literacy*. Colorado Springs, CO: Biological Sciences Curriculum Studies.

Creswell, J.W. (2012). *Educational Research: Planning, Conducting and Evaluating Quantitative and Qualitative Research, 4th Ed*. Upper Saddle River, NJ: Merrill.

Çelebi, H., Karaaslan, H., &Demir-Vegter, S. (2016). Corpus use in enhancing lexico-grammatical awareness through flipped applications. *Journal of Language and Linguistic Studies, 12(2)*.

Darling-Hammond, L. (2008). Introduction: Teaching and learning for understanding. *Powerful Learning: What We Know About Teaching for Understanding* (pp. 1- 9). San Francisco, CA: Jossey-Bass.

Davies, R. S., Dean, D. L., & Ball, N. (2013). Flipping the classroom and instructional technology integration in a college-level information systems spreadsheet course. *Educational Technology Research and Development*, *61(4)*, 563-580.

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. Boston: Henry Holt.

Egbert, J., Herman, D., & Lee, H. (2015). Flipped instruction in English language teacher education: A design-based study in a complex, open-ended learning environment. *Teaching English as a Second or Foreign Language, 19(2)*, 1-23.

Ellis, R. (2005). Planning and task-based performance: Theory and research. *Planning and Task Performance in a Second Language*, *11*, 3-34.

Flipped Learning Network (FLN). (2014) The Four Pillars of F-L-I-P™.

Flumerfelt, S., & Green, G. (2013). Using lean in the flipped classroom for at risk students. *Educational Technology and Society*, *16*(1), 356-366.

Gajjar, N. B. (2013). The role of technology in 21st century education. *International Journal for Research in Education*, 2(2), 23–25.

Hamdan, N., McKnight, P., McKnight, K., & Arfstrom, K. M. (2013). *The flipped learning model: A white paper based on the literature review titled A Review of Flipped Learning*.

Hatipoğlu, Ç. (2010). Summative evaluation of an undergraduate 'English Language Testing and Evaluation' course by future English language teachers. *English Language Teacher Education and Development (ELTED)*, *13*, 1-12.

Heiss, E. D., Obourn, E. S., & Hoffman, C. W. (1950). *Modern Science Teaching*. New York: Macmillan.

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

Honeycutt, B. (2013). Looking for "flappable" moments in your class. *Faculty Focus.* Retrieved from www.facultyfocus.com/articles/instructional-design/looking-for-flippable-moments-in-your-class/

Jensen, J.L., Kummer, T.A., and Godoy, P.D. (2015). Improvements from a flipped classroom may simply be the fruits of active learning. *CBE Life Sciences Education, 14*, 1-12.

Kolb, D. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice-Hall.

Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.

Lawson, A. E. (2002). *Science Teaching and Development of Thinking*. Belmont, CA: Wadsworth/Thompson Learning.

Lee, G. & Wallace, A. (2017). Flipped learning in the English as a foreign language classroom: Outcomes and perceptions. *TESOL Q.* doi:10.1002/tesq.372

Long, M. H. (1989). Task, group, and task-group interactions. *University of Hawaii Working Papers in ESL,* 8, 1-25.

Loughran, J., & Berry, A., (2005). Modelling by teacher educators. *Teaching and Teacher Education. 21 (2)*, 193-203.

Missildine, K., Fountain, R., Summers, L., & Gosselin, K. (2013). Flipping the classroom to improve student performance and satisfaction. *Journal of Nursing Education*, *52*(10), 597-599.

Norris, J. M. (2009). 30 Task-Based Teaching and Testing. *The Handbook of Language Teaching*, 578.

Reid, W. J. (1997). Research on task-centered practice. *Social Work Research,* 21(3), 132–137.

Robinson, P. (1995). Task complexity and second language narrative discourse. *Language Learning*, *45*(1), 99-140.

Robinson, P. (2001). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. *Cognition and Second Language Instruction*, 287-318.

Robinson, P. (Spring 2003). The cognition hypothesis, task design, and adult task-based language learning. *Second Language Studies*, *21*(2), 45-105.

Sams, A., & Bergmann, J. (2013). Flip your students' learning. *Educational Leadership, 70(6)*.

Samuda, V. & Bygate, M. (2008). *Tasks in second language learning*. New York: Palgrave Macmillan.

Sarawagi N (2013). Flipping an introductory programming course: yes you can! *Journal of Computing Sciences in Colleges, 28(6),* 186-188.

Shannon-Chastain, J. & Fell-Kurban, C. (2016). Just one more hit: Student engagement with pre-class videos in the evolution of an English for Academic Purposes course from traditional to flipped. *The Global e-Learning Journal, 5(2).*

Sternberg, R. J. (2003). What is an "expert student?". *Educational Researcher* 32, 8, 5–9.

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

Thoms, C. L. (2013, June 12-14). Maximizing the blended learning curriculum by using the "flipped classroom" approach in the workplace. Paper presented at *The International Conference on E-Learning in the Workplace*, New York, NY.

Tucker, B. (2012). The flipped classroom. Education Next, 12(1), 82. Retrieved from http://educationnext.org/ the-flipped-classroom/

Virjo, I., Holmberg-Mattila, D., & Mattila, K. (2001). Task-based learning (TBL) in under- graduate medical education. *Medical Teacher, 23(1)*, 55–58.

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

Wallace, M. L., Walker, J. D., Braseby, A. M., & Sweet, M. S. (2014). "Now, what happens during class?" Using team-based learning to optimize the role of expertise within the flipped classroom. *Journal on Excellence in College Teaching, 25*(3&4), 253-273.

Wright, T. (2000, December). No more pencils. No more books? Arguing for the use of experiential learning in post-secondary environmental studies classrooms. *Electronic Green Journal*, n.p.

Appendix A. The course outline

English Language Testing & Evaluation
Course Description

This course aims to present senior student teachers an understanding of how to test and evaluate English language learners at different ages and proficiency levels. The course intends to help student-teachers become familiar with various test types, test preparation techniques for the purpose of measuring the language skills (listening, speaking, reading, writing) and language content (grammar, vocabulary, pronunciation), the practice of preparing various exam questions, the analysis of the test results for learner assessment, and the evaluation of the effectiveness of the tests used.

Course Objective/Aim	Learning Outcome
At the end of this course, students will…	By the end of the course, students will be able to…
become familiar with the basic terms and concepts related to language testing and evaluation	- use basic terms and concepts related to language testing appropriately where/when necessary - express successfully their knowledge related to the role of tests within the curriculum design for language teaching - discuss the importance of test selection according to the profile of the learners and the teaching context
engage in various processes and practices related to assessment of language proficiency successfully	- select tests according to the profile of the learners and the teaching context - use different techniques for adapting language test - write, implement and evaluate a variety of testing instruments for a specific group of language learners
become aware of the statistical analysis of testing data	- perform statistical analysis on the testing data - evaluate tests and test results/items for the improvement of teaching and testing
become aware of a variety of testing instruments for a specific group of language learners	- design, implement and evaluate a variety of testing instruments for a specific group of language learners - discuss the advantages and disadvantages of using published and class teacher-written tests - evaluate the effectiveness of various language tests

Materials

Primary Sources

Brown, H. D. (2004). *Language assessment: Principles and classroom practices.* New York: Pearson Education, Inc.

Council of Europe (2001). *Common European framework of reference for languages: Learning, teaching, assessment.* https://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf

Hughes, A. (2003). *Testing for language teachers.* United Kingdom: Cambridge University Press.

Secondary Sources

Bachman, L.F., & Palmer, A.S. (1996*). Language testing in practice.* Oxford: Oxford University Press.

Brindley, G. (1998). Describing language development? Rating scales and SLA. In L.F. Bachman & A.D Cohen (Eds.), *Interfaces between second language acquisition and language testing research.* Cambridge: Cambridge University Press, pp. 112-140.

Brown, J.D. & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly, 32(4)*, 653-675.

Fulcher, G. *Language Testing Resources* available at http://languagetesting.info

Hamp-Lyons, L. (1997). Washback, impact and validity: ethical concerns. *Language Testing, 14*, 295-303.

Kirschner, M., Wexler, C., & Spector-Cohen, E. (1992). Avoiding obstacles to student comprehension of test questions. *TESOL Quarterly, 26,* 537-556.

Lynch, B. (1997) In search of the ethical test. *Language Testing, 14*, 315-327.

McNamara, T. (2014). *Language testing.* Oxford: Oxford University Press.

Evaluation	
Pre-While- Post Class Tasks (x 9)(Application, Reflection, Review or Coursera* Tasks)	45 %
Midterm	25%
Final Project (Test Specifications & Test Design)	30 %

COURSE OUTLINE		
	Content	**Assigned reading**

Week 1	Teaching and Testing
	  Testing- Assessment- Evaluation
	  Role of Testing in Curriculum Design
	  Classroom Assessment-Large Scale Assessment
		Hughes, 1 & 2

Week 2	Test Types and Testing
	  Formative-Summative; Process-Product;
	  Proficiency- Achievement-Placement-Diagnostic;
	  Norm-based/Criterion-based;
	  High Stakes/Low Stakes;
	  Formal/Informal; Direct/Indirect Assessment
		Hughes, 3
		Brown 1 & 3

Week 3	Principles of Language Assessment:
	  Validity; Reliability; Impact; Backwash/Washback;
	  Authenticity; Practicality; Interactiveness
		Hughes, 4- 5-6
		Brown 2
		Bachman & Palmer, 2

Week 4	Principles of Language Assessment contd.	
Week 5	Stages of Test Development Test Specifications Standardized Tests Common European Framework (CEF)	Hughes, 7 Council of Europe, CEF

Week 6	Common Test Techniques
	  Multiple Choice, True/False, Fill-in-the-blanks,
	  Short Answer, Close Test
		Hughes, 8 & 14
		Brown, 4&5

Week 7	Midterm	

Week 8	Testing Writing
	    Feedback & Error Codes
	    Rater reliability
	    Rubric design & use
		Hughes, 9
		Brown, 9
		Brindley article

Week 9	Testing Speaking Rubric design & use	Hughes, 10 Brown, 7
Week 10	Testing Reading	Hughes, 11 Brown, 8 Kirschner et al. article
Week 11	Testing Listening	Hughes, 12 Brown, 6
Week 12	Testing Grammar and Vocabulary	Hughes, 13
Week 13	Test Administration & Test Ethics Basic Statistical Analysis of Test Data Computer-Based Testing Item Banking	Hughes, 16, Appendix A & Appendix B Hamp-Lyons article Lynch article
Week 14	Alternative Assessment Authentic Assessment Dynamic Assessment	Hughes, 15 Brown, 10 Brown & Hudson article

*You are required to enroll the course "Foundations of Teaching for Learning 6: Introduction to Student Assessment" available online at https://www.coursera.org/learn/learning-assessment and study the weekly materials there (videos and reading assignments). This free course lasts 6 weeks.

İDE öğretmen eğitiminde ters-yüz sınıf uygulamalarına dair bir inceleme: Ödev tasarımı ve öğrenci değerlendirmeleri

Öz

Bu çalışmada, eğitim dili İngilizce olan bir devlet üniversitesinde İngiliz Dili Eğitimi (İDE) Bölümü lisans öğrencilerinden toplanan veriler aracılığıyla ters-yüz sınıf uygulamalarına yer verilen bir derste Robinson (2001) tarafından geliştirilen model kullanarak ödev karmaşıklığı, koşulları ve zorluğu gibi ödev özellikleri arasındaki etkileşim incelenmiştir. Bu bağlamda, ders içeriğinin ve ödevlerin planlanması ve geliştirilmesinde bilgi ediniminin genel itibariyle sınıf dışı faaliyetler, uygulama ediniminin de büyük oranda sınıf içi etkinlikler üzerinden yürütüleceği ilkesi benimsenmiştir. Ödevler karmaşıklık, koşullar ve zorluk özellikleri bakımından sistematik olarak tasarlanmış, sıralanmış ve bir akademik dönem boyunca uygulanmıştır. Uygulamalara yönelik veriler, odak grup toplantısı ve öz-değerlendirme anketleri aracılığıyla katılımcı öğrencilerden toplanmıştır. Bu çalışma ile İDE öğrencilerinin ödev karmaşıklığı, koşulları ve zorluğu gibi ödev özelliklerine ilişkin düşünceleri ile ters-yüz sınıf uygulamalarındaki bilgi ve uygulama edinimi ilkesi arasında bağlantı kurmak ve bulguları İDE öğretmen eğitimi alanına olası katkıları noktasında tartışmak ve değerlendirmek hedeflenmiştir.

Anahtar sözcükler: Ödev; karmaşıklık; zorluk; katılım; ters-yüz sınıf; İDE; öğretmen eğitimi

AUTHOR BIODATA

Hatice Karaaslan is a language instructor at Ankara Yıldırım Beyazıt University, with a PhD in Cognitive Science from METU. Her interests include corpus linguistics, argumentation and critical thinking, advising in language learning and blended learning. [hkaraaslan@ybu.edu.tr]

Hatice Çelebi is a language instructor at Acıbadem University, with a PhD in ELT from METU. Her research interests are (im)politeness, corpus linguistics, collaborative learning and learner training. [Hatice.Celebi@acibadem.edu.tr]