Education and Science

Original Article

Vol 46 (2021) No 208 157-188

Alignment of the Intended, Enacted, Received and Assessed Curriculum in EFL Pre-Service Measurement and Evaluation Education

Serpil Tekir [1]

Abstract

Recent research points out a problem regarding the pre-service assessment education given by English as a Foreign Language (EFL) programs in the national context. One factor contributing to the problem can be the poor alignment among program components as numerous teacher education researchers assert that alignment is a necessary condition for learning experiences and practices of pre-service teachers. Accordingly, in this study, the author examined the alignment among the components of the pre-service assessment course curriculum; that is, the intended, enacted, assessed, and received curricula. The study adopted a mixed-methods case study approach using both qualitative and quantitative data collection tools, and to examine the alignment among the curricula, an adaptation of Surveys of the Enacted Curriculum (SEC) alignment methodology was used. Data for the intended and assessed curricula were collected through curricular and assessment documents and analyzed through content analysis, while the data regarding the enacted and received curricula were collected through surveys with teacher educators and pre-service teachers and analyzed through descriptive statistics. In both cases, the data were transferred to matrices for further analysis. To calculate the alignment among each pair of curricula, Porter's alignment index formula was used. Results indicated that the pre-service assessment course curriculum has moderate to high alignment indices ranging from 0, 44 to 0, 78. The study found a strong alignment between the intended and enacted curricula and the assessed and received curricula but a moderate alignment between each pair of the intended and assessed; enacted and assessed; enacted and received and intended and received curricula. The results of this study serve to identify the points of misalignment within the pre-service assessment curriculum. The study concludes with implications for improving the alignment in teacher preparation and suggestions for future research on pre-service assessment education.

Keywords

Assessment education
Assessed curriculum
Curriculum alignment
Enacted curriculum
Pre-service teacher education
Intended curriculum
Received curriculum

Article Info

Received: 05.09.2020
Accepted: 04.22.2021
Online Published: 05.18.2021

DOI: 10.15390/EB.2021.9750

[1] Middle East Technical University, School of Foreign Languages, Turkey, stekir@metu.edu.tr

Introduction

Assessment literacy is a core professional requirement for teachers and it has been the main focus of teacher education over the last two decades (DeLuca, LaPointe-McEwan, & Luhanga, 2016). Assessment plays a central role within the current education context as it might have significant effects on students' academic achievement and teachers' professional development. By implementing proper assessment techniques and grading practices, teachers can evaluate and advance their instruction, increase student motivation for learning, make valid judgments about student achievement and enhance their success (Dinther, Dochy, & Segers, 2015; Mellati, Khademi, & Shirzadeh, 2015; Mertler, 2003).

Given the critical role of assessment literacy, supporting teacher assessment knowledge has become a systemic priority. To illustrate, at the national level, the Turkish Ministry of National Education describes assessment and evaluation as a required teacher competency in the *General Competencies for Teaching Profession* document (Ministry of National Education [MoNE], 2017). To help teacher candidates maintain the required assessment literacy levels and the necessary assessment competencies, the Higher Education Council (HEC) requires teacher preparatory programs in Turkey to offer explicit assessment courses to teacher candidates (Higher Education Council, 2016). Accordingly, teacher education programs offer must courses to improve the assessment literacy and competency of pre-service teachers.

Despite such explicit assessment courses given at EFL teacher education programs, research suggests that beginning EFL teachers still feel insecure for assessing student learning and they have low assessment literacy levels (Hatipoğlu, 2015; Karaman & Şahin, 2014; Mertler, 2003; Öz & Atay, 2017). Most teachers were observed to be mainly untrained to effectively integrate assessment with their instruction, especially beginning teachers reported to be incompetent in this area (Mertler, 2003). In-service teachers in several contexts report feeling ill-prepared to assess student learning and claim that their lack of preparation is due mainly to inadequate pre-service training in educational measurement (Karaman & Şahin, 2014). Likewise, Mede and Atay (2017) found that EFL teachers have quite low assessment literacy levels and need further training in almost all areas of testing and assessment.

It is quite interesting that although EFL teacher education programs in Turkey offer explicit courses to equip teacher candidates with the necessary assessment literacy, beginning EFL teachers do not have the expected assessment literacy levels and do not feel competent enough to assess student learning. Previous research has identified curriculum misalignment as a factor contributing to the low assessment literacy levels of teachers (DeLuca & Klinger, 2010; Graham, 2005; Stiggins, 2005). While research on assessment has examined the assessment literacy and competency levels of new teachers and teacher candidates and pointed out a problem in this area, to date, there has yet to be a study that examines the alignment of the pre-service assessment education, which may be a factor leading to teachers' low level of assessment literacy. This study addresses this gap in the literature by analyzing the alignment of the curriculum of an explicit assessment course offered by an EFL teacher education program.

Curriculum Alignment

The word alignment was defined by Webb (2007) as the degree "to which learner expectations and assessments are in agreement and serve in conjunction with one another to guide the system towards students learning of what they are expected to know and do" (p. 1). Roach, Niebling, and Kurz (2008) defined alignment "as the extent to which curricular expectations and assessments are in

agreement and work together to provide guidance to educators' efforts to facilitate students' progress toward desired academic outcomes" (p. 1). According to Squire (2012), alignment is making sure that assessments and standards coverage are addressed in the instructional process.

Curriculum alignment is rooted in the belief that learning may be enhanced when learners encounter consistent ideas across learning experiences (Bransford, Brown, & Cocking, 2000; Bruner, 1990). Biggs (1999) supported the relation between alignment and achievement by claiming on the condition that all the components of the education system are aligned, it is not possible for a learner not to learn content. Cohen (1987) expanded this idea by asserting that alignment between the assessment and teaching objectives brings about an increase in student learning as much as two standard deviations. In a similar way, several other research findings indicated that alignment predicts student achievement (EdSource, 2006; Kercheval, 2001; McFadden, 2009; Murphy, 2007; Schuenemann, Jones, & Brown, 2011; Zavadsky, 2006).

As literature reveals, the improved achievement is a promise made by alignment, and to achieve that, it is necessary to examine the match among the curriculum elements. However, measuring the match among these elements is not easy. According to Porter (2002, 2004), the challenge is that the content of each element, that is, curriculum, instruction, and assessment, should be translated into a common language framework that enables us to make comparisons and judgments of the match among their contents. For that purpose, Porter (2004) differentiated between the content of three types of curricula: the intended, enacted, and assessed curricula. In this description, the intended curriculum defines the instructional content that needs to be covered in the classroom, the enacted curriculum emphasizes the content that students were given a chance to learn, and the assessed curriculum reveals the content students are assessed. The alignment among them shows the level of agreement among the contents of these curricula and the extent to which they work in cooperation with each other to improve student learning, that is, the learned or received curriculum.

The SEC Curriculum Alignment Model

For measuring alignment, Webb and Porter developed two major methodologies and designed tools: the Webb model and the Surveys of Enacted Curriculum (SEC) model. The Webb Alignment Model is a framework to measure the alignment between the content standards and the assessments, while the SEC model measures the alignment among the standards, assessments, and instruction through a common content matrix (Bhola, Impara, & Buckendahl, 2003; Council of Chief State School Officers [CCSSO], 2006; Roach et al., 2008). Since it can measure the alignment at the enacted curriculum level, the SEC model was used for the purposes of this study. In addition, the SEC model offers two advantages over the Webb model: 1) it uses a common language framework that helps the researcher make subsequent comparisons and judgments of agreement between different types of curricula and 2) it provides the alignment index as a single figure within the range of 0 (no alignment) to 1 (perfect alignment), which makes alignment judgment easier (Kurz, Elliott, Wehby, & Smithson, 2010).

The model also allows translating the content of each curriculum into a specific content matrix across two dimensions. The first dimension represents specific content topics. The second dimension represents the respective expectations for learner performance, that is, the expected cognitive demands. Like Bloom's (1949) taxonomy, the SEC uses five categories of cognitive demands to define the cognitive tasks required, such as 1) recall, 2) procedural skills, 3) application, 4) analysis and 5) synthesis (Smithson, 2007).

Content translations of the intended and assessed curricula are typically made by several subject matter experts. First, they are trained in the SEC coding principles. Then, they code assessed curriculum (i.e., test items) and the written curriculum (i.e., standards or objectives in course syllabi, course materials etc.) using the subject-specific content language and cognitive demands (Smithson, 2007). After that, the codes are transferred to the matrix that represents the content translation of the specific curriculum. Information on the content of the enacted curriculum is gathered from the teachers themselves. After being trained in the SEC coding rules, teachers report their instructional content, topic by topic, through the subject-specific content language. To do that, they give a rating to show the emphasis their teaching gives on each topic in terms of the level of coverage and the degree of emphasis (Blank, 2002).

Content data for each curriculum type are then reduced to cell-by-cell proportions with their sum across all rows and columns equaling 1.00. That means each unit analyzed is a proportion of the whole (Roach et al., 2008). That is, in any content matrix, the sum of all ratings should be 1.00. To calculate the proportional values (PV), the rating in each cell is divided by the total number of ratings in the matrix. This is a way of measuring relative emphasis of the content category at a particular cognitive demand level as compared to the total number of ratings on the matrix (DeLuca & Bellara, 2013). To illustrate, the total of the values in the written curriculum matrix is 43. For an individual cell having a value of 3, the proportional value is 3/43= 0, 06. In this way, data are reduced to cell-by-cell proportions with their sum across all rows and columns equaling 1. This calculation applies to any calculation of the SEC content matrix. As the content matrices for each of the intended, enacted, assessed and received curricula consist of 25 cells (5 columns × 5 topics) with the sum of all ratings across cells equaling 1, they can be paired and subjected to quantitative analyses to measure the alignment between them.

Research on Curriculum Alignment

Most of the research on alignment has examined the relationship between alignment and student achievement. For example, Schmidt et al. (2001) studied state standards (written curriculum) of maths and science over forty countries, the test items in TIMMS and test scores of students from these countries and found that alignment between the content of the written curriculum and the tested curriculum improves student achievement on tests. That is, the amount of coverage of topics in the textbook (written curriculum) determines how well students do on the TIMSS test (Schmidt et al. 2001). In a similar way, Porter, Kirst, Osthoff, Smithson, and Schneider (1994) studied the alignment of the taught curriculum and the written curriculum of the first-year high school mathematics courses and student achievement in these courses. The findings of their study demonstrated that when the taught curriculum is aligned to the written curriculum, it leads to a positive and significant increase in student achievement (Porter et al., 1994). In addition, in their study, Schmidt et al. (2001) studied the eight grade US students' maths scores in TIMMS and the instructional time. The findings of their study revealed that the alignment between the taught and the tested curriculum contribute to student achievement scores. Typically, spending about one week more on a topic would lead to a 3 to 24 % increase in students' achievement scores. Thus, it seems likely that additional instruction on key topics will bring about better learning and higher scores for students (Schmidt et al., 2001). Cohen (1987) reported similar findings. In his experimental design study, the researcher explored the relation between instruction and curriculum embedded testing and found that when instruction and assessment are aligned, both low-aptitude and high-aptitude students make gains. Interestingly, he suggested that low-aptitude students make greater gains than high-aptitude students.

In one of the recent research on curriculum alignment, Kurz and his colleagues investigated the alignment among the intended, enacted, and assessed curricula in an eighth-grade mathematics course in general and special education via the Surveys of the Enacted Curriculum (SEC). At the end of the study, they reported that alignment between the intended, planned, and enacted curricula for general and special education was relatively low (.12 out of 1.00). They also announced there is a significant correlation between the alignment between the enacted and assessed curricula and student achievement (Kurz et al., 2010). In the same vein, Seitz (2017) explored the curriculum alignment among the intended, enacted, and assessed curricula in the ninth-grade mathematics along with two domains: content and cognitive processes. The findings of the study suggested high alignment among the intended, enacted, and assessed curricula in terms of content but low alignment among the intended, enacted, and assessed curricula in terms of cognitive processes. Research supports that alignment among curriculum types brings about student success; however, it is still not known much how these curriculum types communicate with each other. The limited literature makes the following points concerning the communication among different curricula.

Available research has found a varying pattern of relations among curriculum types. For example, Glatthorn (2000) proposed a stronger communication between the enacted curriculum and the assessed curriculum, as shown in Figure 1. That was explained to be a result of accountability efforts. As Gooding (1994) claimed, these efforts led teachers to be more concerned about the performance of their students on tests and spend most of their teaching time on improving students' test competence and on practicing sample questions from the district, state, or national tests.

Despite teachers' considerable efforts in class, research findings suggested that the match between the enacted and the learned curricula is not very strong. In other words, students do not always learn what their teachers teach in class. There are several reasons for that, such as poor motivation, limited cognitive abilities or short attention span of students, and the inability of teachers' monitoring student learning or making the curriculum meaningful and challenging for them (Glatthorn, Carr, & Harris, 2001).

Glatthorn et al. (2001) also claimed that the written curriculum has just a partial influence on both the taught and the assessed curricula. To justify the relatively weaker communication among the written and the taught curricula, the researchers explained that most teachers, especially the experienced ones, check the curriculum document at the beginning of the year and do not refer to it much for the rest of the year as they consider some other factors in deciding what and how to teach in class. These factors are learner interest, their previous experience as regards what would and would not work in class, as well as what is more likely to be tested in national tests to plan their instruction.

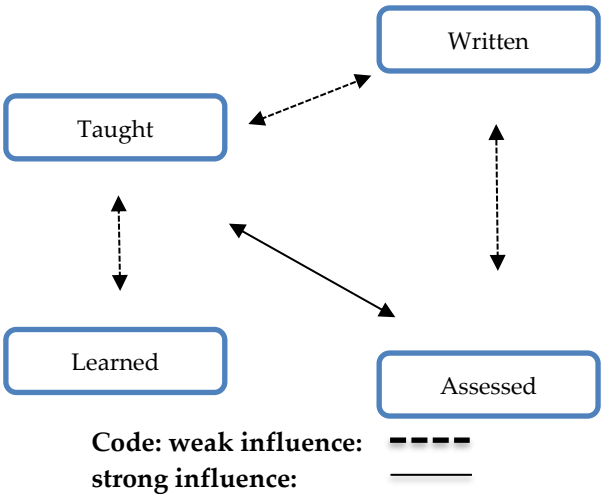

Figure 1. Relations among curriculum types (Adapted from Glatthorn, 2000).

Previous research on curriculum alignment has presented support for the alignment of the intended, enacted, and assessed curricula for effective learning (Elliott, Braden, & White, 2001; Webb, 1997a, 1997b). Similarly, research on teacher education claimed that if teacher candidates are not provided with coherent learning experiences, they may suffer from low-level literacy and competency (Darling-Hammond, 2006; Howey & Zimpher, 1989; Russell & Mcpherson, 2001). While the available research on pre-service assessment education has reported poor assessment competence and performance of EFL teachers in the national context (Hatipoğlu, 2015; Karaman & Şahin, 2014; Mede & Atay, 2017; Mertler, 2003; Öz & Atay, 2017), it has yet to be examined if the low literacy level of teachers is because of a potential misalignment among the components of the pre-service assessment course curriculum; that is, if a mismatch among the intended, enacted and assessed curricula of the assessment course is responsible for the poor assessment literacy of EFL teachers. In addition, most of the studies conducted on alignment so far have investigated the alignment between standards and assessments (e.g., Resnick, Rothman, Slattery, & Vranek, 2003; Webb, 1997a, 1997b; Webb, Herman, & Webb, 2006). The very limited number of alignment studies on the enacted curriculum has focused primarily on the instructional content delivered in the K-12 context. The enacted curriculum in teacher education and its alignment to other curricula remains relatively unexamined. Thus, the current research will contribute to curriculum alignment literature as it aims to investigate the enacted curriculum in a teacher education context, which is an area that has yet to receive research attention and to the teacher education literature as it intends to find the gaps in pre-service assessment education and lead to improvements in teacher preparation to better meet the needs of teacher candidates.

For these purposes, the author intended to explore the alignment among the intended, enacted, received and assessed curricula of the pre-service assessment course offered by an EFL teacher education program in central Turkey. Additionally, finding the pattern of communication among these curriculum types is another motivation of the current study. In this respect, the study was guided by the following research questions:

1. What is the degree of alignment among the intended, enacted, received and assessed curricula of the EFL pre-service assessment course?
2. How do the intended, enacted, received and assessed curricula of the EFL pre-service assessment course communicate with each other?

Method

Research Design

The study adopted a mixed-methods case study design, which is a sort of mixed-methods research using both quantitative and qualitative data collection, results, and integration in order to yield in-depth evidence for a case (Creswell & Plano Clark, 2018). Case study and mixed-methods research are not distinct units instead, the line between the two is absorbent and flexible, which permits each to either support or lead in a research effort (Carolan, Forbat, & Smith, 2016).

Consistent with the definitions in the literature, the current study adopted a mixed-methods case study design as the individual elements and their relations within a particular curriculum are explored in depth through both qualitative and quantitative data collection instruments. As Rowley (2002) claimed, case study research can be based on any mix of quantitative and qualitative approaches and tools such as surveys, interviews, document analysis and observation. For the purposes of this study, surveys and document analysis were used as data collection tools. Data regarding the content of the enacted and received curricula were gathered from the teacher educators and the pre-service teachers via surveys. For the intended and assessed curricula, curricular documents (e.g., course syllabi, course materials and readings) and assessment tools (e.g., midterm and final exams) were subjected to content analysis to determine the content coverage and cognitive demand levels.

Context and Participants

To make the research more accurate and pragmatic, the researcher aimed to evaluate the alignment of a pre-service assessment curriculum in a specific teacher education program rather than looking for the shared and persistent features of EFL teacher education curricula to make generalizations. The study sought to figure out the alignment of the program in its own implementation, context and complexity (Stake, 1995). Additionally, the notion of alignment is a distinctive attribute of individual programs. Each program is unique with its distinct profile of stakeholders, so it is likely that each program has a different alignment pattern. Therefore, investigating the alignment of all or a group of programs in the country and making a general decision about the alignment of the EFL programs in Turkey will not be reasonable or scientific. Thus, to be able to make a thorough and precise conclusion about the alignment of the pre-service assessment curriculum, a sample research site was selected.

The researcher selected to study the alignment of the English Language Testing and Evaluation course offered by an EFL teacher education program in central Turkey. The specific teacher education institution was preferred to be the research site due to its convenience and potential to provide rich and detailed insights required for the investigation. As Pickart (2018) suggests selecting a site that will provide rich data from the onset is critical for researchers. They are recommended to make sure they will be able to use multiple data collection techniques, access to artifacts and people having significant information about the phenomena. To get detailed data, researchers need more than gaining permission to enter the research site. They also need to build up confidence and relationship with all of the stakeholders (Pickart, 2018). As the current research requires accessing some confidential documents like assessment instruments, the researcher selected the specific research site that is a convenient source of relevant data.

The pre-service teachers who were taking the undergraduate English Language Testing and Evaluation course offered by the case program in the 2017-2018 academic year participated in the study through surveys. Surveys were given in the last session of the course. Of the 90 pre-service teachers taking the course, 57 (51.3%) were present in the classrooms. All of the 57 pre-service teachers present in the class attended the survey. Forty-two of the attendees were female and fifteen were male.

The teacher educators giving the English Language Testing and Evaluation course at the case program in the same academic year constituted the other participant group. Three teacher educators

were giving the course in three sections. As Table 1 shows, one of the teacher educators was a professor, and the others were associate professors. One of them was male, while the others were female. Their experience in EFL teacher education ranged from 15 to 30 years. All of the teacher educators received their Ph.D. degrees from state universities in Turkey. Two of them had majored in English Language Teaching (ELT) in their Ph.D. degrees while the other had her degree in English Linguistics. The researcher contacted the course instructors at the beginning of the term to inform them about the research and the contribution expected from them. Getting their consent and agreement to involve in the study from the outset, the researcher started off the research process.

Table 1. Profiles of Teacher Educators

Participants	Age	Gender	Title	Teaching experience	Ph.D. Majors
Educator 1	62	Male	Professor	30	ELT
Educator 2	43	Female	Assoc. Prof	15	Linguistics
Educator 3	45	Female	Assoc. Prof	18	ELT

Data Sources

As Table 2 indicates, data in this study were received from two main sources: the course documents and surveys. Data regarding the intended and assessed curricula were collected through the curricular and assessment documents, while the data for the enacted and received curricula were collected through closed surveys from teacher educators and pre-service teachers.

Table 2. Data Collection Instruments

Intended	Enacted	Received	Assessed
Curricular documents	Teacher educator survey	Pre-service teacher survey	Assessment instruments

Curricular documents. The course syllabus and course materials (i.e., textbooks, readings and the printed course notes) provided the data for the intended curriculum. To inform the teacher educators about the research and data collection process, the researcher talked to each one in person. After getting their initial consent for participating in the study, the researcher asked for a copy of their most current course syllabus, the textbooks, course readings and course notes. All of the teacher educators provided digital copies of the documents through e-mail in two weeks. The syllabus was prepared by the three educators collaboratively. Thus, the course content and objectives were the same in all of the three sections. They were following two main coursebooks and six articles on different aspects of assessment in ELT were chosen by the educators to be shared with pre-service teachers during the term. The teacher educators were using slides to present the topic of each week in class. These slides were very similar to each other, but there were minor differences in content and design. Briefly, the syllabi, two main coursebooks, six outside readings and the slides of the instructors constituted the curricular documents analyzed for the intended curriculum of the assessment course.

Assessment tools. As the course requirement, pre-service teachers were expected to do a presentation on one of the topics in the syllabus and show active participation in classroom activities besides taking two midterms and a final exam. To define the assessment content of the course, the researcher decided to include the official written student assessment instruments (i.e., the midterms and final exams). To collect the assessment tools at the end of the term, after each instrument was implemented in class, the researcher contacted the course instructors in person and requested a copy of the midterms and the final exams they used. As the assessment tools are confidential documents, the researcher provided the teacher educators with a letter of confidentiality explaining that the tools would be used only for research purposes and not be shared with anyone except the experts, who would review and code the documents. The researcher collected copies of the midterms and the final exam at the end of the term. The exams were designed by the three course instructors collaboratively. However,

the way they were administered varied in some aspects. For example, the final exam was project work requiring pre-service teachers to design an assessment instrument for a particular situation. While two of the educators permitted online submission, the other teacher educator required pre-service teachers to submit the assignment as a printed document. Also, the second midterm, which was a take-home exam, was administered as an individual work by the teacher educator one and two; it was designed as a pair work assessment instrument by the third teacher educator. Briefly, the two midterms and the final exam used for the assessment of pre-service teachers' performance in the course were collected to be analyzed to define the content of the assessed curriculum.

Surveys with teacher educators. Data regarding the content of the enacted curriculum was collected from the teacher educators through a survey. Teacher educators described the coverage of their teaching content topic by topic based on the subject-specific content language. They rated their instructional content on two dimensions: the level of coverage for each content topic and the degree of emphasis for each category of cognitive demands (Blank, 2002).

The participating teacher educators completed the SEC content survey at the end of the term. The Surveys of Enacted Curriculum are practical and consistent data collection tools to obtain reliable data on the participating teachers' instructional practices and the content they teach in classrooms (CCSSO, 2006). The survey contained the content language formed by the two experts at the end of their initial coding of the curricular documents. The content language featured 5 broad content areas including multiple subtopics: ELT measurement methods and principles (6 topics), evaluating language skills and test development (8 topics), alternative assessment measurement (5 topics), analysis of the test results for learner assessment (4 topics), evaluation of the effectiveness of the tests used (6 topics). The teacher educators gave emphasis ratings on the level of coverage and categories of cognitive demands for each content topic. In other words, teacher educators were asked to indicate if they allowed students to learn the content in the learning expectations at the cognitive level stated in the intended curriculum. Four-point Likert scale was used for the instructional coverage: *"0" for "None, not covered"*, *"1" for "Slight Coverage (less than one class/lesson)"*, *"2" for "Moderate Coverage (one to five classes/lessons)"*, *and "3" for "Sustained Coverage (more than five classes/lessons"*. The surveys were e-mailed to each teacher educator and they returned the surveys in a week.

Surveys with pre-service teachers. The pre-service teachers taking the English Language Testing and Evaluation course were given a survey in which they were asked to indicate their self-perceived learning gains of the topics covered in the course across their cognitive demand levels. In other words, pre-service teachers were asked to indicate whether or not they learned the content in the learning expectations at the cognitive level stated in the intended curriculum. The list of content topics and the cognitive demand levels, which were formed by the initial coding of the curricular documents by two experts, were provided to the pre-service teachers and they were asked to indicate their competence level by ticking the appropriate point of the four-point Likert scale. The points were "*0*" for *"Not competent"*, "*1*" for "*Slightly competent*", "*2*" for "*Competent"*, and "*3*" for "*Very competent"*.

Data Analysis

The data collected for the intended and assessed curricula via documents were analyzed through content analysis, while the data collected regarding the enacted and received curricula via surveys were analyzed through descriptive statistics. In both cases, the data were transferred to matrices for further analysis.

Curricular documents. First of all, the intended objectives, the expected outcomes and the course content in the curricular documents were analyzed through content analysis. The course syllabus, coursebooks, readings, slides were coded by the two experts to determine a list of intended learning outcomes. Two experts, both of whom have a background in English language testing and assessment

and a Ph.D. degree in educational sciences, were trained on how to use the coding conventions for alignment methodology and how to use the subject-specific content language to code the curricular documents. For the trustworthiness of the content analysis process, intercoder reliability was used. The two experts worked together and coded the sample curricular documents. After talking about what to look for in documents, they established a preliminary code list. Later, experts independently examined the documents. After finishing the coding process, they came together to compare their codes and proposed themes. The correlation coefficient was found to be .916, which indicates good reliability between the reviewers. By coding all documents together and reconciling the differences in coding of individual items, the experts created the final set of tables for analysis.

For purposes of alignment analysis among the curriculum types, first, the experts reviewed the course documents such as the course syllabi, course materials and readings to determine the content area of the intended curriculum. Initially, 29 codes were recognized in total. Later, they were thematically categorized into five content themes. At the end of this process, the content of the intended curriculum was determined topic by topic. The subtopics were grouped under general topics using a subject-specific content language. The content language featured five broad content areas: *1) ELT measurement methods and principles; 2) Evaluating language skills and test development; 3) Alternative assessment; 4) Analysis of test results for learner assessment; 5) Evaluation of the effectiveness of the tests used.*

After determining the content areas, the experts reviewed all the curricular documents again and coded them for the emphasis they put on each topic at each cognitive demand level this time. In the coding system, they used, *"0" for "no emphasis", "1" for "slight emphasis"* (less than 25% of objectives/content spent on topic), *"2" for "moderate emphasis"* (25-33% of objectives/content spent on topic) and *"3" for "sustained emphasis"* (more than 33% of objectives/content spent on-topic). At the end of this process, the ratings were transferred to the matrix for the intended curriculum to be compared with other types of curricula and to make an alignment judgment. They were shown on the value (V) column of the matrices.

Assessed curriculum. The assessment instruments of the course, collected by the researcher, were analyzed by the experts. They initially divided the tasks into steps and each step was coded for the emphasis they put on each topic at each cognitive demand level by the content experts. In the coding system, *"0"* is used for *"no emphasis"*, *"1"* for *"slight emphasis*" (less than 25% of assessment content spent on the topic), "2" for "*moderate emphasis"* (25-33% of assessment content spent on the topic) and *"3"* is used for *"sustained emphasis"* (more than 33% of assessment content spent on the topic). The final codes were sent to the teacher educators for the credibility of the analysis of the assessment tasks. In other words, through the member checking method, the credibility of the study was supported as suggested by Lincoln and Guba (1985). To clarify, the codes of the assessment tasks were sent to the participant teacher educators and they were asked to review them for accuracy to eliminate the possibility of misinterpretations. All of the three teacher educators responded to this request, and they approved the coding of the tasks in the assessment tools. After getting their approval, the emerged ratings were transferred to the value (V) column on the matrices for the assessed curriculum.

Survey data for the enacted and received curricula. The quantitative data gathered through pre-service and teacher educator surveys were examined using descriptive statistics. In order to realize the degree of responses, the means for each item were calculated through the SPSS 18 program. Then, the mean scores from each survey were transferred to the value (V) column on the matrices designed for each of the received and enacted curricula for further analysis.

Alignment analysis

The data transferred to the value (V) column of the alignment matrices were transformed into proportional values (PV) (DeLuca & Bellara, 2013). In this step, the data counts in each cell were divided

by the total number of ratings in the data set in the matrix. This is a way of measuring the relative emphasis of the content category at a particular cognitive demand level as compared to the total number of ratings on the matrix.

Matrices formed for each curriculum type were paired up and the alignment between them was measured through quantitative analyses (Kurz et al., 2010). At this point, six tables were formed for each of the curriculum pairings: intended and enacted curriculum, intended and assessed curriculum, intended and received curriculum, enacted and assessed curriculum, enacted and received curriculum and assessed and received curriculum. The values (V) and proportional values (PV) were presented in the matrices to make the comparison easier. These tables are given in the results section.

In the next step, a cell-by-cell comparison was made between the matrices of each pair using Ms. Excel because the total of absolute discrepancies between data sets ($\sum|$ x-y$|$) was needed to calculate Porter's alignment index. For example, the absolute discrepancy between the proportional values in the first cell (i.e., Content 1 at recall level) in the intended curriculum column (x) and the enacted curriculum column (y), was calculated to be $|0\text{-}0{,}08| = 0{,}08$ and transferred to the absolute discrepancy column (x-y) on the matrix. After getting the difference between the proportional values of each corresponding cell, the sum of these differences was calculated for each matrix (Fulmer, 2011). For example, the sum of the absolute discrepancies in the matrix for the intended and the enacted curricula pair was calculated to be 0,08 + 0,2 +,0,11 + 0,12 + 0,16 = 0, 67. In the final step, the alignment between the data sets of each pair was calculated using Porter's alignment index formula below (Porter, 2002).

Porter's Alignment Index = $1.0 - \frac{\sum|x-y|}{2}$

Alignment Index for the Intended and Enacted Curricula = $1.0 - \frac{0{,}67}{2} = 0, 67$

The same formula was used to calculate the alignment indices for the other curricula pairings and the results are reported in the findings section.

Limitation of the Study

Although the study reported important findings, it has some limitations as well. First of all, the study was limited to only the case teacher education program where the research was conducted. The findings of the study concerning the alignment among different types of curricula are not generalizable beyond the program that participated in this study. It is very likely that other teacher education programs will have different intended outcomes, classroom applications, assessments and, as a result, different learner gains. Therefore, the alignment indices among their curricula may be different from the ones found in this study. However, the methodology of the study can be repeated in similar contexts. The conceptual and methodological framework of the study will provide a model for following research on curriculum alignment. In addition, the study is also limited to one course offered by the teacher education program, namely the English Language Testing and Evaluation course. Therefore, to have a general perception of the program alignment, further studies evaluating each course in the program, as well as the practicum component, is going to be carried out by the researcher.

Another limitation of the study is the concerns about the quality of survey data. Teacher educators might have failed to report their instructional coverage accurately; however, available literature suggests that teacher surveys on classroom teaching have a strong correlation with classroom observations and teacher logs (Mullens & Gayler, 1999; Shavelson, Webb, & Burstein, 1986; Smithson & Porter, 1994). Also, Porter and his colleagues included detailed data for the reliability and validity of the SEC measures in their report (Porter et al., 1994).

Research Ethics

Ethical approval for this study was obtained from METU Human Subjects Ethics Committee with the approval number 28620816/158-300. Through e-mails, the participating teacher educators were informed about the aims and objectives of the study as well as the documents the researcher would like to receive. For that purpose, the researcher sent the teacher educators a written signed contract stating her agreement that the data and course documents obtained from teacher educators would remain protected from disclosure and from unauthorized persons and would be handled and stored properly. After explaining the purpose of the study and the confidentiality issues, the researcher asked for teacher educators' and pre-service teachers' voluntary participation in the research and received their informed consent to participate in the research.

Results

Degree of Alignment among the Intended, Enacted, Received and Assessed Curricula

Table 3, which shows the alignment indices among the intended, enacted, received and assessed curricula, suggests that the pre-service English Language Testing and Evaluation course curriculum has moderate to high alignment indices for each group of six alignment matrix pairings ranging from 0,44 to 0,78. Alignment decision was made based on Porter and Smithson's (2000, 2001, 2002) explanation that an alignment index of 1,00 would suggest perfect alignment across both dimensions, whereas an alignment index of 0,00 would indicate that there is no agreement on topic coverage and cognitive demand between the two curricula.

Table 3. Alignment Indices for Six Alignment Matrix Pairings

AI 1	**AI 2**	**AI 3**	**AI 4**	**AI 5**	**AI 6**
0,67	0,48	0,45	0,44	0,78	0,44

AI 1: Intended-Enacted, AI 2: Intended-Assessed, AI 3: Intended-Received,
AI 4: Enacted-Received, AI 5: Assessed-Received, AI 6: Enacted-Assessed

The first alignment index for the analysis (AI 1), which refers to the alignment between the intended curriculum and the enacted curriculum, was found to be 0, 67. That suggests that 67% of what teacher educators' intended to teach in their written curriculum matches what they actually taught in the classroom. The intended curriculum maintained representation across categorical concurrence themes ranging from 13% to 23%, which is similar to the enacted curriculum, but with a less balanced distribution with a range of %16 to %24 (See Table 4). Regarding the cognitive demand levels, there is again a parallel distribution of the content across the cognitive levels in both data sets with minor differences. The most significant mismatch between the intended and the enacted curriculum is that the "procedural skills" demand level was represented with a higher figure (26%) in the intended curriculum than it was in the enacted curriculum (16%).

Table 4. Alignment Matrix for the Intended and Enacted Curriculum

	Recall					Procedural skills					Application					Analysis					Synthesis					Total			
Content	**Intended**		**Enacted**		**Absolute difference**	**Intended**		**Enacted**		**Absolute difference**	**Intended**		**Enacted**		**Absolute difference**	**Intended**		**Enacted**		**Absolute difference**	**Intended**		**Enacted**		**Absolute difference**	**Intended**		**Enacted**	
	V	**PV (x)**	**V**	**PV (y)**	**|x-y|**	**V**	**PV (x)**	**V**	**PV (y)**	**|x-y|**	**V**	**PV (x)**	**V**	**PV (y)**	**|x-y|**	**V**	**PV (x)**	**V**	**PV (y)**	**|x-y|**	**V**	**PV (x)**	**V**	**PV (y)**	**|x-y|**	**V**	**PV (x)**	**V**	**PV (y)**
1	0	0	3	,08	,08	3	,06	3	,08	,02	1	,02	0	0	,02	0	0	0	0	0	3	,06	0	0	,06	7	**,13**	**6**	**,16**
2	0	0	0	0	0	3	,06	3	,08	,02	3	,06	3	,08	,01	3	,06	0	0	,06	2	,04	0	0	,04	11	**,21**	**6**	**,16**
3	0	0	0	0	0	3	,06	0	0	,06	3	,06	2	,05	,02	3	,06	3	,08	,02	3	,06	3	,08	,02	12	**,23**	**8**	**,21**
4	0	0	0	0	0	2	,04	0	0	,04	3	,06	3	,08	,02	3	,06	3	,08	,02	3	,06	3	,08	,02	11	**,21**	**9**	**,24**
5	0	0	0	0	0	3	,06	0	0	,06	3	,06	3	,08	,04	3	,06	3	,08	,02	3	,06	3	,08	,02	2	**,22**	**9**	**,24**
Total	0	0	3	**,08**	,08	14	**,26**	6	**,16**	**,2**	13	**,25**	11	**,29**	,11	12	**,23**	9	**,24**	**,12**	14	**,26**	9	**,24**	**,16**	43	**1**	38	**1**

1: ELT measurement methods and principles; 2: Evaluating language skills and test development; 3: Alternative assessment; 4: Analysis of test results for learner assessment; 5: Evaluation of the effectiveness of the tests used

V: Value, PV: Proportional Value

The second alignment analysis (AI 2) showing the alignment between the intended and the assessed curricula was found to be 0,48. The figure is slightly less than 0,50, so it shows a moderate alignment between what teacher educators planned to teach and what they tested through assessment tools. There are minor differences between the two curriculum types in terms of categorical concurrence. Still, the main gaps between them lay in the emphasis they put on the cognitive demand levels (See Table 5). In both curriculum types, each of the content categories had adequate coverage with a range of 17% to 23%, except for "ELT measurement methods and principles" being represented less in the intended curriculum (13%) than it was in the assessed curriculum (19%). When it comes to the complexity of cognitive demands, there are surprisingly weak associations between the two curricula. The lower cognitive levels like "recall" and "procedural skills" were highly addressed in the assessed curriculum with a percentage of 36 and 42, respectively, while the higher cognitive levels like "application" and "analysis" were less addressed with a representation of 17% and 6%. It is also interesting that the highest complexity level (i.e., synthesis) received no representation at all in the assessed curriculum. In the intended curriculum, however, each of the cognitive demand levels was adequately represented with a percentage range of 23% to 26% except for the lowest demand level, "recall", which was not represented at all.

Table 5. Alignment Matrix for the Intended and Assessed Curriculum

Content	Recall					Procedural skills					Application					Analysis					Synthesis					Total			
	Intended		Assessed		Absolute difference	Intended		Assessed		Absolute difference	Intended		Assessed		Absolute difference	Intended		Assessed		Absolute difference	Intended		Assessed		Absolute difference	Intended		Assessed	
	V	PV (x)	V	PV (y)	|x-y|	V	PV (x)	V	PV (y)	|x-y|	V	PV (x)	V	PV (y)	|x-y|	V	PV (x)	V	PV (y)	|x-y|	V	PV (x)	V	PV (y)	|x-y|	V	PV	V	PV
1	0	0	3	,08	,08	3	,06	3	,08	,02	1	,02	1	,03	,01	0	0	0	0	0	3	,06	0	0	,06	7	**,13**	**7**	**,19**
2	0	0	3	,08	,08	3	,06	3	,08	,02	3	,06	2	,06	0	3	,06	0	0	,06	2	,04	0	0	,04	11	**,21**	**8**	**,22**
3	0	0	2	,06	,06	3	,06	3	,08	,02	3	,06	2	,06	0	3	,06	0	0	,06	3	,06	0	0	,06	12	**,23**	**7**	**,19**
4	0	0	3	,08	,08	2	,04	3	,08	,04	3	,06	1	,03	,03	3	,06	1	,03	,03	3	,06	0	0	,06	11	**,21**	**8**	**,22**
5	0	0	2	,06	,06	3	,06	3	,08	,02	3	,06	0	0	,06	3	,06	1	,03	,03	3	,06	0	0	,06	2	**,22**	**6**	**,17**
Total	0	0	13	**,36**	**,36**	14	**,26**	15	**,42**	**,12**	13	**,25**	6	**,17**	**,1**	12	**,23**	2	**,06**	**,18**	14	**,26**	0	**0**	**,28**	43	**1**	36	**1**

1: ELT measurement methods and principles; 2: Evaluating language skills and test development; 3: Alternative assessment; 4: Analysis of test results for learner assessment; 5: Evaluation of the effectiveness of the tests used

V: Value, PV: Proportional Value

The third alignment analysis (AI 3) represents the content match between teachers' intended curriculum and the students' received curriculum. The alignment of this pair is moderate, with an index of 0.45, which indicates that 45% of what teacher educators intended was learned by the pre-service teachers. The matrix tables of both curricula suggested that in terms of a measure of matching topics in the intended and received curricula there are discrepancies. Specifically, the received curriculum maintained a higher emphasis on "ELT measurement methods and principles" (24%) but a lower emphasis on "evaluation of the effectiveness of the tests used" (12%) than the intended curriculum, in which each categorical content theme received 13% and 22% respectively (See Table 6).

There are also mismatches in the measure of relative emphasis of cognitive demands. For instance, in the written curriculum, there was no content at "recall" level, while in the received curriculum, almost half of the codes were at "recall" level (44%). Participants reported having half of their gains at this cognitive level, which their instructors did not intend in their curricular documents at all. At the "procedural skills" cognitive level, there was a similarly strong emphasis both in the written curriculum (26%) and in the received curriculum (28%). Interestingly enough, about the higher demand levels like "application" (16%), "analysis" (4%) and "synthesis" (8%), the received curriculum involved far less content than the intended curriculum, which addressed all of these higher cognitive levels with over 20% representation.

Table 6. Alignment Matrix for the Intended and Received Curriculum

Content	Recall					Procedural skills					Application					Analysis					Synthesis					Total			
	Intended		Received		Absolute difference	Intended		Received		Absolute difference	Intended		Received		Absolute difference	Intended		Received		Absolute difference	Intended		Received		Absolute difference	Intended		Received	
	V	PV (x)	V	PV (y)	|x-y|	V	PV (x)	V	PV (y)	|x-y|	V	PV (x)	V	PV (y)	|x-y|	V	PV (x)	V	PV (y)	|x-y|	V	PV (x)	V	PV (y)	|x-y|	V	PV (x)	V	PV (y)
1	0	0	3	,12	,12	3	,06	2	,08	,02	1	,02	0	0	,02	0	0	0	0	0	3	,06	1	,04	,02	7	**,13**	6	**,24**
2	0	0	2	,08	,08	3	,06	1	,04	,02	3	,06	1	,04	,02	3	,06	0	0	,06	2	,04	0	0	,04	11	**,21**	4	**,16**
3	0	0	2	08	,08	3	,06	2	,08	,02	3	,06	1	,04	,02	3	,06	0	0	,06	3	,06	1	,04	,02	12	**,23**	6	**,24**
4	0	0	2	,08	,08	2	,04	2	,08	,04	3	,06	1	,04	,02	3	,06	1	,04	,02	3	,06	0	0	,06	11	**,21**	6	**,24**
5	0	0	2	,08	,08	3	,06	0	0	,06	3	,06	1	,04	,02	3	,06	0	0	,06	3	,06	0	0	,06	2	**,22**	3	**,12**
Total	0	0	11	**,44**	**,44**	14	**,26**	7	**,28**	**,16**	13	**,25**	4	**,16**	**,1**	12	**,23**	1	**,04**	**,2**	14	**,26**	2	**,8**	**,2**	43	**1**	25	**1**

1: ELT measurement methods and principles; 2: Evaluating language skills and test development; 3: Alternative assessment; 4: Analysis of test results for learner assessment; 5: Evaluation of the effectiveness of the tests used

V: Value, PV: Proportional Value

The fourth alignment analysis (AI 4), which is for the content alignment between the enacted curriculum and received curriculum was found to be 0,45. Looking at the categorical concurrence of the topics, the two curricula had a similar representation of the content categories except for the "ELT measurement methods and principles" and "evaluation of the effectiveness of the tests used". In the enacted curriculum, "evaluation of the effectiveness of the tests used" theme had far less representation (16%) than it had in the received curriculum (24%) (See Table 7). However, the ELT measurement methods and principles theme was more addressed (23%) in the enacted curriculum than it was in the received curriculum (12%). The differences in the cognitive complexities of the content might also lead to the relatively low level of alignment. The cognitive demand level, "recall", maintained very limited representation in the enacted curriculum (8%), but it was measured to have a relatively high representation (44%) in the received curriculum. In a similar way, the "procedural skills" cognitive level had more emphasis in the received curriculum (28%) than it had in the enacted curriculum (16%). Interesting enough, the other cognitive demands, especially the highest two demands (i.e., analysis and synthesis), received very little emphasis in the received curriculum of pre-service teachers (4% and 8% repectively) while they had quite a high relative emphasis in the teacher educators' enacted curriculum (both had 24%). This pointed out a difference between the cognitive complexity of what teachers taught in class and what pre-service teachers learned.

Table 7. Alignment Matrix for the Enacted and Received Curriculum

Content	Recall					Procedural skills					Application					Analysis					Synthesis					Total			
	Enacted		Received		Absolute difference	Enacted		Received		Absolute difference	Enacted		Received		Absolute difference	Enacted		Received		Absolute difference	Enacted		Received		Absolute difference	Enacted		Received	
	V	PV (x)	V	PV (y)	|x-y|	V	PV (x)	V	PV (y)	|x-y|	V	PV (x)	V	PV (y)	|x-y|	V	PV (x)	V	PV (y)	|x-y|	V	PV (x)	V	PV (y)	|x-y|	V	PV	V	PV
1	3	,08	3	,12	,04	3	,08	2	,08	0	0	0	0	0	0	0	0	0	0	0	0	0	1	,04	,04	**6**	**,16**	6	**,24**
2	0	0	2	,08	,08	3	,08	1	,04	,04	3	,08	1	,04	,04	0	0	0	0	0	0	0	0	0	0	**6**	**,16**	4	**,16**
3	0	0	2	,08	,08	0	0	2	,08	,08	2	,05	1	,04	,01	3	,08	0	0	,08	3	,08	1	,04	,04	**8**	**,21**	6	**,24**
4	0	0	2	,08	,08	0	0	2	,08	,08	3	,08	1	,04	,04	3	,08	1	,04	,04	3	,08	0	0	,08	**9**	**,23**	6	**,24**
5	0	0	2	,08	,08	0	0	0	0	0	3	,08	1	,04	,04	3	,08	0	0	,08	3	,08	0	0	,08	**9**	**,23**	3	**,12**
Total	3	**,08**	11	**,44**	**,36**	6	**,16**	7	**,28**	**,2**	11	**,29**	4	**,16**	**,13**	9	**,24**	1	**,04**	**,2**	9	**,24**	2	**,08**	**,24**	38	**1**	25	**1**

1: ELT measurement methods and principles; 2: Evaluating language skills and test development; 3: Alternative assessment; 4: Analysis of test results for learner assessment; 5: Evaluation of the effectiveness of the tests used

V: Value, PV: Proportional Value

The fifth alignment analysis (AI 5), which was pertained to the alignment between teacher educators' assessed curriculum and student teachers' received curriculum was 0,78. The result shows that both curricula put the relatively balanced emphasis on all content categories, so there are no significant differences except for the less emphasis that "evaluating language skills and test development" theme got in the received curriculum (16%) than it did in the assessed curriculum (22%) (See Table 8). There are also some minor discrepancies between the received curriculum and the tested curriculum on the relative emphasis across the cognitive levels. In the assessed curriculum, there is slightly more representation of the content at higher cognitive complexity than there is in the received curriculum. To illustrate, "procedural skills" has a description of 42% in the assessed curriculum, while the same cognitive level had quite a high but still relatively lower representation in the received curriculum (28%). Similarly, in terms of "application" and "analysis", the assessed curriculum had slightly more content (18%, 6% respectively) than the received curriculum (16%, 4% respectively). However, at the "synthesis" demand level, there was no content in the assessed curriculum, while the same cognitive level was represented in 8% of the content in the received curriculum.

Table 8. Alignment Matrix for the Assessed and Received Curriculum

Content	Recall					Procedural skills					Application					Analysis					Synthesis					Total			
	Assessed		Received		Absolute difference	Assessed		Received		Absolute difference	Assessed		Received		Absolute difference	Assessed		Received		Absolute difference	Assessed		Received		Absolute difference	Assessed		Received	
	V	PV (x)	V	PV (y)	|x-y|	V	PV (x)	V	PV (y)	|x-y|	V	PV (x)	V	PV (y)	|x-y|	V	PV (x)	V	PV (y)	|x-y|	V	PV (x)	V	PV (y)	|x-y|	V	PV (x)	V	PV (y)
1	3	,08	3	,12	,04	3	,08	2	,08	0	1	,03	0	0	,03	0	0	0	0	0	0	0	1	,04	,04	**7**	**,19**	**6**	**,24**
2	3	,08	2	,08	0	3	,08	1	,04	,04	2	,06	1	,04	,02	0	0	0	0	0	0	0	0	0	0	**8**	**,22**	**4**	**,16**
3	2	,06	2	,08	,02	3	,08	2	,08	0	2	,06	1	,04	,02	0	0	0	0	0	0	0	1	,04	,04	**7**	**,19**	**6**	**,24**
4	3	,08	2	,08	0	3	,08	2	,08	0	1	,03	1	,04	,01	1	,03	1	,04	,01	0	0	0	0	0	**8**	**,22**	**6**	**,24**
5	2	,06	2	,08	,02	3	,08	0	0	,08	0	0	1	,04	,04	1	,03	0	0	,03	0	0	0	0	0	**6**	**,17**	**3**	**,12**
Total	13	**,36**	11	**,44**	**,08**	15	**,42**	7	**,28**	**,12**	6	**,17**	4	**,16**	**,12**	2	**,06**	1	**,04**	**,04**	0	**0**	2	**,08**	,08	36	**1**	**25**	**1**

1: ELT measurement methods and principles; 2: Evaluating language skills and test development; 3: Alternative assessment; 4: Analysis of test results for learner assessment; 5: Evaluation of the effectiveness of the tests used

V: Value, PV: Proportional Value

The sixth alignment analysis (AI 6) is for the content alignment between teachers' enacted curriculum and teachers' assessed curriculum. The alignment for this pair was found to be 0,44. That indicates a moderate alignment between what teacher educators did in class and what they tested in the exams. That is because of the gap both in the matching topics and in the relative emphasis of their cognitive demands. Most of the content topics have more representation in the enacted curriculum than they have in the assessed curriculum except for "ELT measurement methods and principles" and "evaluating language skills and test development" (See Table 9). These topics maintained less representation in the enacted curriculum, both with 16%. They are slightly more represented in the assessed curriculum with 19 and 22 percentages, respectively. In terms of the cognitive demand levels, content at lower levels of demand was much more emphasized in the assessed curriculum than it was in the enacted curriculum. The "recall" (36%) and "procedural skills" (42%) were highly represented in the assessed curriculum, whereas in the enacted curriculum, they were represented with 8 and 16 percentages, respectively. However, the higher cognitive demand levels like "application", "analysis" and "synthesis" were strongly addressed in the enacted curriculum with over 20% representation. In the assessed curriculum, on the other hand, different from the "application" level, which had quite a good representation (17%), the "analysis" demand level was minimally represented with 6%, and the "synthesis" level was not represented at all.

Table 9. Alignment Matrix for the Enacted and Assessed Curriculum

	Recall					**Procedural skills**					**Application**					**Analysis**					**Synthesis**					**Total**			
Content	**Enacted**		**Assessed**		**Absolute difference**	**Enacted**		**Assessed**		**Absolute difference**	**Enacted**		**Assessed**		**Absolute difference**	**Enacted**		**Assessed**		**Absolute difference**	**Enacted**		**Assessed**		**Absolute difference**	**Enacted**		**Assessed**	
	V	**PV (x)**	**V**	**PV (y)**	**|x-y|**	**V**	**PV (x)**	**V**	**PV (y)**	**|x-y|**	**V**	**PV (x)**	**V**	**PV (y)**	**|x-y|**	**V**	**PV (x)**	**V**	**PV (y)**	**|x-y|**	**V**	**PV (x)**	**V**	**PV (y)**	**|x-y|**	**V**	**PV (x)**	**V**	**PV (y)**
1	3	,08	3	,08	0	3	,08	3	,08	0	0	0	1	,03	,03	0	0	0	0	0	0	0	0	0	0	**6**	**,16**	7	**,19**
2	0	0	3	,08	,08	3	,08	3	,08	0	3	,08	2	,06	,02	0	0	0	0	0	0	0	0	0	0	**6**	**,16**	8	**,22**
3	0	0	2	,06	,06	0	0	3	,08	,08	2	,05	2	,06	,01	3	,08	0	0	,08	3	,08	0	0	,08	**8**	**,21**	7	**,19**
4	0	0	3	,08	,08	0	0	3	,08	,08	3	,08	1	,03	,05	3	,08	1	,03	,05	3	,08	0	0	,08	**9**	**,24**	8	**,22**
5	0	0	2	,06	,06	0	0	3	,08	,08	3	,08	0	0	,08	3	,08	1	,03	,05	3	,08	0	0	,08	**9**	**,24**	6	**,17**
Total	3	**,08**	13	**,36**	**,28**	6	**,16**	15	**,42**	**,24**	11	**,289**	6	**,17**	**,19**	9	**,24**	2	**,06**	**,18**	9	**,24**	0	**0**	**,24**	38	**1**	36	**1**

1: ELT measurement methods and principles; 2: Evaluating language skills and test development; 3: Alternative assessment; 4: Analysis of test results for learner assessment; 5: Evaluation of the effectiveness of the tests used

V: Value, PV: Proportional Value

Relation among the Intended, Enacted, Received and Assessed Curricula

Another important finding of the study is about the relation between the different curricula of the same course. The alignment indices of each pair of curricula indicate the pattern of communication among curriculum types, which is illustrated in Figure 2. As the figure suggests, the intended and the enacted curriculum has a strong relationship (AI 1>0.50) while the intended and the assessed curricula and the intended and the received curricula have moderate communications (AI 2 and AI 3< 0, 50). The enacted curriculum has a moderate relation with both the assessed and the received curricula (AI 6 and AI 4< 0, 50). The assessed curriculum has a moderate relation with both the intended and enacted curricula (AI 2 and AI 6< 0, 50). Learners' received curriculum has stronger communication with the assessed curriculum (AI 5> 0, 50) than it has with the intended or enacted curricula (AI 3 and AI 4 < 0, 50)

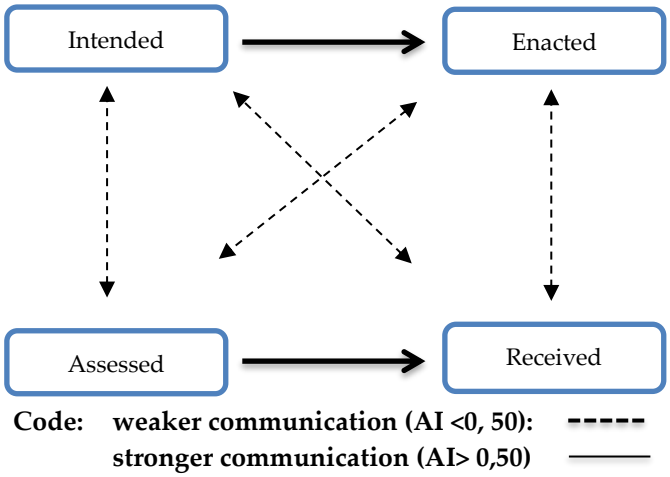

Figure 2. Relations among curriculum types.

Discussion

The alignment indices for each pair of curricula provided us with very useful information not only about the extent of alignment among the intended, enacted, received and assessed curricula of the pre-service assessment course but also about the relation among each curriculum type. The findings of the study indicate that the alignment between the teacher's intended and enacted curriculum (0,67) is higher than the alignment between the intended and assessed curricula (0,48) and the alignment between the enacted and assessed curricula (0,44). The analysis for the alignment between teachers' planned and enacted curricula provided information about how much teachers enacted their own planned instructional content. The alignment index suggests that teacher educators taught 67% of the content in their intended plan, which indicates a high alignment. The finding of the study contradicts the findings of Kurz and his colleagues. They investigated the alignment among the intended, enacted and assessed curricula of the eighth-grade mathematics course in general and special education via the Surveys of the Enacted Curriculum and found a low alignment between the intended and the enacted curricula. They stated that half of the participating teachers implemented less than half of their planned curriculum (Kurz et al., 2010).

The findings of the study also noted that teacher educators tested about half (48%) of the content in their written curriculum and they tested less than half (44%) of what they actually taught in class. The relatively lower level of alignment between the written and the tested curricula and the one between the taught and the tested curricula of the pre-service assessment education can be the reason for the low assessment literacy level of teacher candidates, as Schmidt et al. (2001) proposed the alignment between the written and the tested curricula and the match between the taught and the tested curricula predict student achievement gains.

The alignment indices demonstrated that the alignment between teachers' assessed curriculum and learners' received curriculum is quite high (0, 78), while the alignment between the teachers' intended and learners' received curriculum (0, 45) and the teachers' enacted and the learners' received curriculum (0, 44) is lower. The relatively lower alignment between these pairs of curricula is not mainly because of the differences in the content themes but mostly due to the differences in the cognitive levels they incorporate. In teacher educators' written and taught curricula, the content was mainly represented at relatively higher cognitive levels while teacher candidates reported learning the content at mostly lower cognitive demands. In other words, in their written curriculum, teacher educators are more likely to set their learning objectives and design classroom-teaching activities at relatively higher cognitive demand levels while they tend to design tasks in the assessment tools, most of which measured pre-service teachers' lower cognitive level abilities. This finding of the study is in line with previous research. Studying the coherence among the intended, enacted, and assessed curricula in the ninth-grade mathematics, Seitz (2017) found that the curricula are aligned as regards content but poorly aligned in terms of cognitive processes. Likewise, Fuhrman (2001) came up with similar findings suggesting that items in tests are more likely to assess lower-level cognitive processes such as remember and understand instead of higher-level cognitive processes such as evaluate and create.

It is rational to see changes in relative emphasis in teachers' planned, enacted and assessed curriculum and learners' received curriculum; however, it could become a concern if a particular curriculum constantly over-represent or under-represent specific cognitive demands. That appears to be the situation of the curricula examined in this study. To illustrate, the intended curriculum over-represented higher cognitive levels, while not representing the "recall" level at all; the enacted curriculum over-represented "application", "analysis" and "synthesis" cognitive levels, while under-represented "recall" level; the assessed curricula over-represented "recall" and "procedural skills" cognitive levels and under-represented "analysis" and "synthesis" levels. This pattern is worrying as it suggests shifts in teacher educators' emphasis in planning, teaching and testing and, even worse, a tendency towards lower-order thinking skills in pre-service teachers' learning. The over-emphasis on lower-order cognitive skills in the intended and assessed curricula can be affecting the pre-service teachers' learning. Thus, to improve their performance, simply covering topics at "recall" or "procedural skills" levels may not be enough. The content should also be taught at higher-order cognitive levels. Likewise, they should be tested at the same levels on tests. In other words, in order to improve learners' higher-order skills as well as lower-order skills, the course content should not only be taught but also tested at higher cognitive levels, not only at lower levels.

The over-representation of the lower level cognitive demands in the received curriculum may suggest poor learning of the pre-service teachers and this can be because of the confusion they probably had at the end of the mismatches between the cognitive demands represented in the course syllabus, instructional activities in and the assessment tools of the same course. However, the researcher wants to point out that to be well prepared for their future careers, teacher candidates need higher-order cognitive skills. In other words, they should be able to apply and implement what they learn about foreign language assessment. This calls for learning objectives, teaching and learning activities and performance tasks aiming for higher cognitive levels. Yet, pre-service assessment course curricula appear to need some revisions to meet this expectation. Teacher educators' intended and assessed curricula should not only emphasize lower cognitive levels like "recall" and "procedural skills", but they should also adequately emphasize the higher cognitive levels like "analysis" and "synthesis". It is certainly desirable to increase the number of learning objectives and performance tasks at higher cognitive levels. That requires teacher educators to revise and improve their syllabus and assessment instruments.

Another important finding of the study is about the relation among the types of curricula of pre-service assessment education. The alignment indices suggest a strong alignment among the intended and enacted curricula and the assessed and received curricula but a moderate alignment between each pair of the intended and assessed; enacted and assessed; enacted and received and intended and received curricula. The strong communication between the intended and the enacted curricula suggested by this study contradicts previous research of Glatthorn et al. (2001). In their study, the researchers claimed that the intended curriculum has a partial impact on the enacted curriculum. They further explained that most teachers, especially the experienced ones, consider other classroom-related factors while deciding on what to teach. They just check the curriculum document at the beginning of the year and then forget about it. Therefore, the intended curriculum does not have a strong effect on what teachers do in class. Similarly, in an official document of UNICEF, it was stated to be very natural to observe differences between the intended and enacted curricula, both of which belong to the same teacher since when the intended curriculum is not adequate, teachers figure out strategies to enable their students to achieve certain performance standards (UNICEF, 2000).

Different from Glatthorn et al.' study (2001), the researcher found a strong relationship between the intended and the enacted curricula in this study. This can be because the research was conducted at a teacher education program, that is, at a higher education institution. The course instructors at the program prepared their intended or written curriculum without any pressure from an external authority. Especially in the K-12 context, the district and school administrators can use the intended curriculum as a tool to control what is enacted in class, so teachers in K-12 may prefer to base their teaching on their experience and intuition rather than the directives of an external body. However, the intended curriculum that was investigated in this research was designed by the course instructors as the research site was a higher education institution. That may explain the differences between the findings. As a result, the findings of the study adds to the literature with the proposal that when the written curriculum is not enforced on teachers but instead designed by them, there is a stronger alignment between the written curriculum and the enacted curriculum.

Another finding of the study is that the intended curriculum has a weaker influence on the assessed curriculum. The intended curriculum includes the content that the teacher educators expect the learners to achieve at the end of the course. Thus, it is very likely that teacher educators evaluate their intended learning objectives in the exams they prepare for. The assessed curriculum should be a good representative of the intended curriculum, so the content of what is assessed should match well with what is intended in the written curriculum. Indeed, the weaker alignment found in this study is not because of the differences in content coverage but because of the differences in the cognitive demand levels. This can be because while the learning expectations were being written, their cognitive demand levels might not be worded properly. Instead, teacher educators might emphasize the content categories to be covered in the course syllabus, assuming that the outcomes they intend in the course were all in their minds. This may explain why the intended curriculum addressed the content topics at mostly "recall" level and did not adequately reflect the other cognitive demands that emerged in the assessed curriculum.

The study also indicates a weaker control of the enacted curriculum on the assessed curriculum. In regular classrooms, teachers are inclined to assess what they actually do in class. Thus, the content that they allocate the most class hours is expected to be tested on exams. However, the study found that it is not the case in the pre-service assessment course. This can be again because of the characteristics of the research context. As it was a higher education institution, in particular, a teacher education program, it is the course instructors who design the assessment tools. The course instructors, that is, the teacher educators, might have different goals in the assessment tasks they used. As in many other teacher education programs, it is quite possible that they aimed to enhance pre-service teachers' oral presentation skills to prepare them for their future professional roles by assigning them to do presentations in class as a part of course assessment. This may explain the incongruence between what teacher educators do in class and what they test.

The study also found that the assessed curriculum has a stronger influence on the received curriculum than the intended and the enacted curricula. This stronger relation between the assessed and the received curricula matches with previous research. In his study, Glatthorn (2000) claimed a strong alignment between the two curricula and added that the assessed content on exams is more likely to be remembered to be acquired by learners. In addition, according to Glatthorn, students take only the things that they are held responsible for seriously. Thus, whatever objectives the teacher declares or whatever the teacher highlights in class, students tend to only recall the content that is tested on exams.

Regarding the weak influence of the enacted curriculum on the received curriculum, Marsh and Willis (2003) claimed that it is difficult to know exactly how students understand the formal curriculum and how their understanding is integrated and associated with both previous in-class learning and with outclass learning, which can be attained through media, or some other sources. Because of having been exposed to different experiences, socio-political effects and analytical attitudes, they tend to have a different understanding of the same lesson. Likewise, Glatthorn et al. (2001) mentioned a gap between the enacted and the received curricula in their study and named several factors such as the poor motivation, limited cognitive abilities or short attention span of students and inability of teachers' in monitoring student learning or making the curriculum meaningful and challenging for their learners.

In brief, the study found a strong alignment between the intended and enacted curricula and the assessed and received curricula but a moderate alignment between each pair of the intended and assessed; enacted and assessed; enacted and received and intended and received curricula. The gaps between them, however, are not on the content congruence but mainly on the cognitive demand levels of the content. That indicates that teacher educators need to attend to cognitive demands more while planning the learning expectations, teaching enactments and assessment tasks. However, they should expect and maintain objectives, classroom activities and assessment tasks not only of low demands but also of high cognitive demands, which will lead to better learning and performance of pre-service teachers. What is needed is a balance in the cognitive processes, where both lower and higher-level cognitive skills are taught and assessed appropriately.

Conclusion

Previous research on curriculum alignment has suggested that providing teacher candidates with consistent messages through aligned instruction and assessment to the intended goals will lead to improved achievement in teacher preparation (Elliott et al., 2001; Webb, 1997a, 1997b). On the contrary, if teacher candidates are not provided with aligned learning experiences, they are likely to have low literacy and competency levels (Darling-Hammond, 2006; Howey & Zimpher, 1989; Russell & Mcpherson, 2001). In the national context, several researchers pointed out a problem regarding the assessment competence and performance of EFL teachers (Hatipoğlu, 2015; Karaman & Şahin, 2014; Mede & Atay, 2017; Mertler, 2003; Öz & Atay, 2017). However, there has been no study investigating if the low assessment literacy levels of EFL teachers is because of a curriculum alignment problem. In this respect, the motive of this study was to investigate the alignment among the intended, enacted, received and assessed curricula of a pre-service assessment course offered by an EFL teacher education program in central Turkey. In this study, the researcher attempted to emphasize the importance of a well-aligned teacher education curriculum in which all components such as the learning objectives, teaching and learning activities and assessment are in accord and, ultimately, support pre-service teachers' learning.

The study adopted a mixed-methods case study approach using both qualitative and quantitative data collection tools and used an adaptation of Surveys of the Enacted Curriculum (SEC) alignment methodology (Porter & Smithson, 2000, 2001, 2002) to examine the alignment among the curricula. Content translations of the intended and assessed curricula were made by subject matter experts who coded the written curriculum documents and assessment instruments in terms of the content coverage and cognitive demand levels of each topic. Information on the content of the enacted and received curricula was gathered by the participating teacher educators and pre-service teachers themselves. That is, teacher educators reported on the content coverage of their teaching via an enacted

curriculum survey, and the pre-service teachers reported their level of learning of the listed topics at the end of the course through a survey instrument. At the end of this process, all the emerged codes were transferred to alignment matrices. The data on the matrices were processed for proportional quantification. In this way, the data counts were transformed into proportional values (DeLuca & Bellara, 2013). Matrices created in this way were consequently paired and subjected to quantitative analyses using Porter's alignment index formula in order to measure the alignment of each pair of curricula (2002).

The alignment indices between the intended and enacted and the assessed and received curricula for the English Language Testing and Evaluation course were found to be relatively high. However, the alignment between the intended and assessed, enacted and assessed, intended and received and enacted and received curricula was found to be moderate. The gaps, however, are not on the content coverage but mainly on the cognitive demand levels of the content. That indicates that successful pre-service teacher learning and attainment can be more consistently achieved if the course expectations, instruction in the classroom, and assessment align with each other not only in terms of content coverage but also in terms of cognitive levels. The findings of the study require teacher educators to attend to cognitive demands while planning their written curriculum and assessment instruments and maintain them in course conduct as well. Including more objectives having high cognitive demands besides the existing objectives with low demands and planning teaching activities and assessment instruments, facilitating higher-order thinking related to assessment knowledge will lead to better learning of pre-service teachers.

To sum up, the areas of misalignment that the study identified, especially regarding the most critical type of alignment between the enacted and the received curricula, require additional research that will examine the reasons for the gap. In this study, the pre-service assessment education was only investigated by analyzing the curriculum of a must assessment course; however, future research may also investigate the other pre-service courses as well as the practicum experiences through which teacher candidates may learn about assessment.

References

Bhola, D. S., Impara, J. C., & Buckendahl, C. W. (2003). Aligning tests with states' content standards: Methods and issues. *Educational Measurement: Issues and Practices, 22*(3), 21-29.

Biggs, J. B. (1999). *Teaching for quality learning at university*. Buckingham: Open University Press.

Blank, R. K. (2002). Using surveys of the enacted curriculum to advance evaluation of instruction in relation to standards. *Peabody Journal of Education, 77*(4), 86-121.

Bloom, B. S. (1949, November). *A taxonomy of educational objectives.* Paper presented at the Meeting of Examiners, Monticello, IL.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school.* Washington DC: National Academy Press.

Bruner, J. (1990). *Acts of meaning.* Cambridge, MA: Harvard University Press.

Carolan, C. M., Forbat, L., & Smith, A. (2016). Developing the DESCARTE model: The design of case study in health care. *Qualitative Health Research, 26*(5), 626-639. doi:10.1177/1049732315602488

Cohen, S. A. (1987). Instructional alignment: Searching for a magic bullet. *Educational Researcher, 16*(8), 16-20.

Council of Chief State School Officers. (2006). *Alignment analysis*. Retrieved from http://www.ccsso.org/Projects/Alignment_Analysis/

Creswell, J. W., & Plano Clark, V. (2018). *Designing and conducting mixed methods research* (3rd ed.). Thousand Oaks, CA: SAGE.

Darling-Hammond, L. (2006). Constructing 21st-century teacher education. *Journal of Teacher Education, 57*(3), 1-15.

DeLuca, C., & Bellara, A. (2013). The current state of assessment education: Aligning policy, standards, and teacher education. *Journal of Teacher Education, 64*(4), 356-372.

DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice, 17*(4), 419-438.

DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016). Approaches to classroom assessment inventory. A new instrument to support teacher assessment literacy. *Educational Assessment, 21*(4), 248-266. doi:10.1080/10627197.2016.1236677

Dinther, M. V., Dochy, F., & Segers, M. (2015). The contribution of assessment experiences to student teachers' self-efficacy in competence-based education. *Teaching and Teacher Education, 49*, 45-55. doi:10.1016/j.tate.2015.02.013

EdSource. (2006). *Similar students, different results (SSDR): Why do some schools do better?.* Retrieved from http://www.edsource.org/pub_abs_simstu05.cfm

Elliott, S. N., Braden, J. P., & White, J. (2001). *Assessing one and all: Educational accountability for students with disabilities.* Arlington, VA: Council for Exceptional Children.

Fuhrman, S. (Ed.). (2001). *From the capitol to the classroom: Standards-based reform in the States: One-hundredth yearbook of the National Society for the Study of Education.* Chicago, IL: University of Chicago Press.

Fulmer, G. W. (2011). Estimating critical values for the strength of alignment among curriculum, assessments, and instruction. *Journal of Educational and Behavioural Statistics, 36*(3), 381-402.

Glatthorn, A. A. (2000). *The principal as curriculum leader: shaping what is enacted and assessed* (2nd ed.). Thousand Oaks, Calif: Corwin Press.

Glatthorn, A. A., Carr, J. F., & Harris, D. E. (2001). *Planning and organizing for curriculum renewal.* Retrieved from http://www.ascd.org/publications/ curriculum-handbook/398.aspx

Gooding, A. (1994). Recognition of effective group discussion. In G. Bell, B. Wright, N. Leeson, & J. Geake (Eds.), *Challenges in mathematics education: Constraints on construction* (pp. 295-305). Lismore, NSW: MERGA.

Graham, P. (2005). Classroom-based assessment: Changing knowledge and practice through pre-service teacher education. *Teaching and Teacher Education, 21*(6), 607-612.

Hatipoğlu, Ç. (2015). English language testing and evaluation (ELTE) training in Turkey: Expectations and needs of pre-service English language teachers. *ELT Research Journal, 4*(2), 111-128. Retrieved from https://dergipark.org.tr/tr/pub/eltrj/issue/28780/308006

Higher Education Council. (2016). *Egitim fakültesi öğretmen yetiştirme lisans programları.* Ankara: Yüksek Öğretim Kurulu Başkanlığı.

Howey, K. R., & Zimpher, N. L. (1989). Pre-service teacher educators' role in programs for beginning teachers. *The Elementary School Journal, 89*(4), 450-470.

Karaman, P., & Şahin, Ç. (2014). Öğretmen adaylarının ölçme değerlendirme okuryazarlıklarının belirlenmesi. *Ahi Evran University Journal of Kırsehir Faculty of Education, 15*(2), 175-189.

Kercheval, A. (2001). *A case study of key effective practices in Ohio's improved school districts.* Bloomington, IN: Indiana Center for Evaluation. Retrieved from http://www.schooleffects.com/gen/corp_generated_bin/documents/basic_module/Indiana_Effective_Practices.pdf

Kurz, A., Elliott, S., Wehby, J., & Smithson, J. (2010). Alignment of the intended, planned and enacted curriculum in general and special education and its relation to student achievement. *Journal of Special Education, 44*(3), 131-145.

Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry.* Newbury Park, CA: Sage.

Marsh, C. J., & Willis, G. (2003). *Curriculum: Alternative approaches, ongoing issues* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.

McFadden, L. (2009). District learning tied to student learning. *Phi Delta Kappan, 90*(8), 545-553.

Mede, E., & Atay, D. (2017). English language teachers' assessment literacy: The Turkish context. *Dil Dergisi, 168*(1), 1-5.

Mellati, M., Khademi, M., & Shirzad, A. (2015). The relationships among sources of teacher pedagogical beliefs, teaching experiences, and student outcomes. *International Journal of Applied Linguistics and English Literature, 4*(2), 177-184. doi:10.7575/aiac.ijalel.v.4n.2p.177

Mertler, C. A. (2003). *Pre-service versus in-service teachers' assessment literacy: Does classroom experience make a difference?*. Paper presented at the Annual Meeting of the Mid-Western Educational Research Association, Columbus.

Ministry of National Education. (2017). *General competencies for teaching profession*. Retrieved from http://oygm.meb.gov.tr.

Mullens, J. E., & Gayler, K. (1999). *Measuring classroom instructional processes: Using survey and case study field test results to improve item construction* (NES 1990-08). Washington, DC: National Center for Education Statistics.

Murphy, J. (2007). Restructuring through learning-focused leadership. In H. J. Walberg (Ed.), *Handbook on restructuring and substantial school improvement* (pp. 71-83). Charlotte, NC: Information Age Publishing. Retrieved from https://www.adi.org/pla/docs/Restructuring_Handbook.pdf

Öz, S., & Atay, D. (2017). Turkish EFL instructors' in-class language assessment literacy: Perceptions and practices. *ELT Research Journal, 6*(1), 25-44.

Pickart, A. J. (2018). *Research methods in information.* London: Facet Publishing.

Porter, A. C. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher, 31*(7), 3-14.

Porter, A. C. (2004). Curriculum assessment. In J. Green, G. Camilli, & P. Elmore (Eds.), *Complementary methods for Research in Education.* Washington, DC: American Educational Research Association.

Porter, A. C., & Smithson, J. L. (2000, April). *Alignment of state testing programs NAEP and reports of teacher practice in mathematics and science in Grades 4 and 8.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Porter, A. C., & Smithson, J. L. (2001). Are content standards being implemented in the classroom? A methodology and some tentative answers. In S. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the states. One-hundredth yearbook of the national society for the study of education* (pp. 60-80). Chicago: University of Chicago Press.

Porter, A. C., & Smithson, J. L. (2002, April). *Alignment of assessments, standards, and instruction using curriculum indicator data.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Porter, A. C., Kirst, M. W., Osthoff, E., Smithson, J. L., & Schneider, S. (1994). *Reform of high school mathematics and science and opportunity to learn*. Consortium for Policy Research in Education Policy Briefs. New Brunswick, NJ: Rutgers University, Consortium for Policy Research in Education.

Resnick, L. B., Rothman, R., Slattery, J. B., & Vranek, J. L. (2003). Benchmarking and alignment of standards and testing. *Educational Assessment, 9*(1), 1-27.

Roach, A. T., Niebling, B. C., & Kurz, A. (2008). Evaluating the alignment among curriculum, instruction, and assessments: Implications and applications for research and practice. *Psychology in the Schools, 45*(2), 158-176.

Rowley, J. (2002) Using case studies in research. *Management Research News, 25*(1), 16-27. doi:10.1108/01409170210782990

Russell, T., & Mcpherson, S. (2001, May). *Indicators of success in teacher education: A review and analysis of recent research*. Paper presented at the Pan-Canadian Education Research Agenda Symposium Teacher Education/Educator Training: Current Trends and Future Directions. Laval University, Quebec City, Canada.

Schmidt, W. H., McKnight, C. C., Houang, R. T., Wang, H. C., Wiley, D. E., Cogan, L. S., … Wolfe, R. G. (2001). *Why schools matter: A cross-national comparison of curriculum and learning.* San Francisco: Jossey-Bass.

Schuenemann, D., Jones, D., & Brown, M. (2011). The impact of a curriculum model on the mathematics and science achievement of economically disadvantaged students. *National Forum of Educational Administration and Supervision Journal, 29*(1), 63-87.

Seitz, P. (2017). Curriculum alignment among the ınтended, enacted and assessed curricula for grade 9 mathematics. *Journal of the Canadian Association for Curriculum Studies (JCACS), 15*(1), 72-94.

Shavelson, R. J., Webb, N. M., & Burstein, L. (1986). Measurement of teaching. In M. Wittrock (Ed.), *Handbook of research on teachin*g (3rd ed., pp. 55-91). Washington, DC: American Educational Research Association.

Smithson, J. L. (2007). *Report of alignment analyses of state n content standards and assessment instruments* (Unpublished manuscript). University of Wisconsin-Madison, Madison.

Smithson, J. L., & Porter, A. C. (1994). *Measuring classroom practice: Lessons learned from efforts to describe the enacted curriculum. The Reform Up Close study* (CPRE Research Rep. Series No. 31). Madison: University of Wisconsin-Madison, Consortium for Policy Research in Education.

Squires, D. A. (2012). Curriculum alignment research suggests that alignment can improve student achievement. *The Clearing House, 85*(4), 129-135.

Stake, R. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.

Stiggins, R. (2005). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice, 18*, 23-27.

UNICEF. (2000). *Curriculum report card*. New York: Working Paper Series Education Section Programme Division.

Webb, N., Herman, J., & Webb, N. L. (2006). *Alignment of mathematics state-level standards and assessments: The role of reviewer agreement* (CSE Report 685). Los Angeles: University of California, National Center for Research on Evaluations, Standards, and Student Testing.

Webb, N. L. (1997a). *Criteria for alignment of expectations and assessments in mathematics and science education* (Research Monograph No. 6). Madison: University of Wisconsin, National Institute for Science Education.

Webb, N. L. (1997b). *Determining alignment of expectations and assessments in mathematics and science education* (NISE Brief, Vol. 1, No. 2). Madison: University of Wisconsin, National Institute for Science Education.

Webb, N. L. (2007). Issues related to judging the alignment of curriculum standards and assessments. *Applied Measurement in Education, 20*(1), 7-25.

Zavadsky, H. (2006). How NLCB drives success in urban schools. *Educational Leadership, 64*(3), 69-73.