



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

문학석사 학위논문

**The Washback Effect of  
the CSAT-English on Students'  
Learning: Comparison among  
Different Grade Groups**

대학수학능력시험 영어 영역이 학생들의 학습에

미치는 환류효과: 학년 그룹 간의 비교

2020년 2월

서울대학교 대학원

영어영문학과 어학 전공

박승해

**The Washback Effect of  
the CSAT-English on Students'  
Learning: Comparison among  
Different Grade Groups**

지도 교수 이 용 원

이 논문을 문학석사 학위논문으로 제출함

2020년 2월

서울대학교 대학원

영어영문학과 어학 전공

박 승 해

박승해의 문학석사 학위논문을 인준함

2020년 2월

위 원 장           손  창  용           (인)

부위원장           유  은  정           (인)

위      원           이  용  원           (인)

## Abstract

# The Washback Effect of the CSAT-English on Students' Learning: Comparison among Different Grade Groups

Park, Seunghae

Department of English Language and Literature

Seoul National University

Washback is gaining more and more attention not only from language assessment researchers but also from policy makers and educators. The term *washback* refers to the influence of language testing on teaching and learning, and the research into washback investigates the relationship between testing, teaching, and learning (Alderson & Wall, 1993). Washback can be conceptualized as an important component of *consequential validity* (or a consequential aspect of *construct validity*) that has been recently embraced as an important test validation (or evaluation) criterion in the measurement community (Messick, 1989).

The current study attempts to examine the washback effect of the English section of the College Scholastic Ability Test (CSAT-English) on students' learning of English in the South Korean educational context. More specifically, the present study focuses on the impact of two different types of factors that can create or moderate the washback effect of the CSAT-English on learning and teaching: (1) some characteristics of the CSAT (i.e. the required linkage between the CSAT and Educational Broadcasting System (EBS) test prep materials, the newly-introduced criterion-referenced nature of assessment) and (2) the chronological distance between the present time point and the test-takers' CSAT test date represented by the students' school year (or grade). The second type of factor (i.e. time-related) in particular is included here in this study in order to investigate the potential relationship, if there is any, between the students' school grade and the washback intensity of the CSAT-English. It is worthwhile to mention here that the first type of factors can be seen not only as external testing policies (i.e. the EBS-CSAT linkage policy and criterion-referenced assessment) but also as important parts of test characteristics at the same time.

From the causal analysis perspective, the EBS-CSAT linkage and criterion-referenced assessment can be viewed as independent variables in this study, while the students' school grade can be considered to be a moderating variable. On the other hand, the dependent variables represent three different learning practices that can potentially be impacted by the CSAT-English: (1) learning attitudes, (2) learning contents, and (3) learning strategies. Student survey questions were created to

address issues related with these independent, moderating, and dependent variables. For this study, surveys were conducted on high school students in different grades, and an unstructured group interview was conducted with teachers who are currently teaching English in high school.

First, the result of the student survey showed that all three main categories in the students' questionnaire represented a significant difference in the washback intensity between the grades. However, the learning strategies were the only section in which the degree of the washback became stronger as the student's grade went up. In the case of the learning attitudes and contents, on the other hand, a partially or entirely inverse relationship between the students' school year and the intensity of the washback was revealed.

Second, the result of the teacher survey indicated that the differences in the degree of the washback between the different grade groups can be regarded as the consequences of the teachers' behavior that varies across the students' school year. The students' learning contents and learning strategies, in particular, were highly affected by the teaching materials used and the way teachers teach students.

Two important implications can be derived from the findings of the current study. First, it is important to investigate the stakeholders or factors that are related to the learners when conducting the washback study on the learners. A supplementary research conducted on the teachers found that students' learning practices are highly affected by teaching contents and methods. Likewise, the washback on the learners and their learning is not explained just by the

characteristics of the learners themselves but is also highly related to other stakeholders who are in close contact with them, such as teachers.

Another implication of the results of the study is the need for comprehensive study on various learning practices. The current study found a partially or entirely inverse relationship between the students' grade and the intensity of the washback in some of the learning practices, such as the learning attitudes and learning contents. The result of the study indicates that the nature of the washback on the learners could be inaccurately defined if one's research focuses on a single aspect of learning. Hence, the washback studies on the learners ought to consider various kinds of learning practices for the precise understanding of the relationship between the learners and the influence of a test on their learning.

**Keywords:** washback, washback intensity, CSAT, learning, teaching

**Student Number:** 2018-20495

# Table of Contents

<b>CHAPTER I. INTRODUCTION .....</b>	<b>1</b>
1.1 Research Background .....	1
1.2 Research Questions .....	5
1.3 Organization of the Thesis .....	6
<b>CHAPTER II. REVIEW OF LITERATURE .....</b>	<b>8</b>
2.1 Washback .....	8
2.1.1 Washback on Learning .....	11
2.2 Korean Educational Context .....	14
2.3 CSAT-English .....	17
2.3.1 EBS-CSAT Linkage Policy .....	24
2.3.2 Criterion-Referenced Assessment .....	26
<b>CHAPTER III. METHODOLOGY .....</b>	<b>33</b>
3.1 Participants .....	33
3.1.1 Students .....	33
3.1.2 Teachers .....	36
3.2 Instruments .....	37
3.2.1 Students' Questionnaire .....	37
3.2.2 Teachers' Questionnaire .....	39
3.3 Data Collection and Analysis .....	39
3.3.1 Students Data Collection and Analysis .....	39
3.3.2 Teachers Data Collection and Analysis .....	41

<b>CHAPTER IV. RESULTS</b> .....	<b>42</b>
4.1 The Impact of the EBS-CSAT Linkage Policy and CRA on the CSAT-English .....	42
4.2 Descriptive Statistics .....	47
4.3 One-Way Analysis of Variance (ANOVA).....	54
4.4 Students' Responses to the Open-Ended Questions .....	57
4.5 Results of the Teachers' Questionnaire .....	62
<b>CHAPTER V. DISCUSSION</b> .....	<b>65</b>
5.1 The Impact of the EBS-CSAT Linkage Policy and CRA on the Nature/characteristics of the CSAT-English .....	65
5.2 The Impact of the Nature/characteristics of the CSAT-English on Students' Learning.....	66
5.3 The Impact of the Students' School Year on the Washback Effect of the CSAT-English .....	68
5.4 Teachers' Influence on the Washback Effect of the CSAT-English.	74
<b>CHAPTER VI. CONCLUSION</b> .....	<b>78</b>
6.1 Summary and Implication .....	78
6.2 Limitations and Suggestions.....	80
<b>REFERENCES</b> .....	<b>83</b>
<b>APPENDICES</b> .....	<b>92</b>
국문 초록 .....	103

## **List of Tables**

Table 1 Outline of the CSAT-English .....	18
Table 2 Comparison between norm- and criterion-referenced assessment .....	27
Table 3 Summary of the types of college admissions students are preparing for ...	34
Table 4 Summary of the reasons for receiving private English education .....	35
Table 5 Summary of the students' background information .....	36
Table 6 Reliability estimates of the categories in the students' questionnaire .....	38
Table 7 The EBS linkage to the CSAT-English over the past five years .....	44
Table 8 Proportion of the test-takers in each stanine over the past five years .....	45
Table 9 Comparison of the means among the different grade groups .....	50
Table 10 Descriptive statistics of the categories in the students' questionnaire .....	51
Table 11 ANOVA omnibus test results of the means among different grade groups .....	55
Table 12 Responses to Item 31 in the students' questionnaire .....	58
Table 13 Responses to Item 32 in the students' questionnaire .....	60
Table 14 Responses to Items 4 through 6 in the teachers' questionnaire .....	63

## **List of Figures**

Figure 1 A washback model of students' learning .....	12
Figure 2 A washback model of the CSAT-English and its national policies .....	22
Figure 3 Trends of the means for Items 1 through 9 .....	47
Figure 4 Trends of the means for Items 10 through 18 .....	48
Figure 5 Trends of the means for Items 19 through 30 .....	49
Figure 6 Trends of the means for the learning attitudes and its subcategories.....	52
Figure 7 Trends of the means for the learning contents and its subcategories.....	53
Figure 8 Trends of the means for the learning strategies and its subcategories .....	54

# Chapter I

## Introduction

### 1.1 Research Background

In language education, washback refers to the influence of language testing on teaching and learning, and the research into washback investigates the relationship between testing, teaching, and learning (Alderson & Wall, 1993). The importance of investigating the relationship between testing, teaching, and learning has been discussed in relation to the notion of consequential aspect of construct validity<sup>1</sup> (Messick, 1995), in which the validity of a test is determined not only by how well it measures what it is intended to measure but also by its impact on the educational setting or society as a whole.

To put it another way, the term ‘washback’ was introduced as an important component of consequential validity, and washback studies have been conducted to determine the validity of a particular test by empirically investigating the intended or unintended, positive or negative influence of testing on teaching and learning. However, many previous works of washback have generally focused on examining the relationship between testing and teaching, and it was not until recently that

---

<sup>1</sup> Construct validity refers to the degree to which a test measures the construct of interest in a study. More specifically, it addresses how well empirically-collected test data support the theory of the construct, hypothesized relationships among different components of the test, and the intended test structure.

washback studies on the students and their learning began to receive significant attention from researchers.

According to Cheng et al. (2015), the washback studies on the learners are significant as they receive the most direct influence of testing, and the ultimate goal of testing is to promote their learning. Currently, a number of studies have been conducted to investigate the relationship between testing and learning, and many of those research studies have enriched and deepen our understanding of the nature of the washback on the learners. A case in point is the work of Shih (2007), in which he proposed three groups of factors that can affect the washback of a test on the learners: extrinsic, intrinsic, and test factors. An important point to note in the study is that these factors can either impact each other or interact among themselves. That is, while each of the factors can directly affect the washback of a test on the learners, one factor can impact another factor that influences the washback, and there could also be interaction among these factors. For instance, the national policy on the test, which is an extrinsic factor, can have impact on such test factors as the nature of the tested skills and test format which, in turn, results in a change in the teachers' teaching practices and hence student learning. A relevant example is the work of Cheng (1999), in which she investigated the influence of the new Hong Kong Certificate of Education Examination (HKCEE) on the teachers' attitudes and behaviors. The Hong Kong Examination Authority introduced a new policy on the HKCEE which led to a change in the nature of the skills tested and the format of the test. The change in the test factors, in turn, modified the teachers' teaching practices and hence student learning.

While a considerable amount of research has been conducted in Hong Kong educational setting, only a few studies have been carried out on language tests in Korea. Kwon (2015), in particular, pointed out the lack of studies on the washback effect of the English section of the College Scholastic Ability Test (CSAT-English), a Korean university entrance examination, and encouraged more research on the impact of the national policies exerted on the nature and characteristics of the CSAT-English. Currently, there are some major studies that have investigated the washback effect of the Educational Broadcasting System (EBS)-CSAT linkage policy and the newly added criterion-referenced assessment (CRA) nature of the CSAT-English on student learning. Many of these studies, in particular, have compared the effects of the policies on learning between the different grade groups, whose number of days remaining before their scheduled CSAT date is different. Park (2017), for example, conducted a survey on first-, second-, and third-grade students to compare the influence of the EBS-CSAT linkage policy on their perceptions and learning materials. The study found that the EBS-CSAT prep books are more actively used in the third-grade classes than the ones in other grades, but that the third-grade students were not interested in learning English through the EBS materials as the words and sentence structures used in the EBS passages were too difficult.

Although Park's research has revealed the effects of the EBS-CSAT linkage policy on learning (or more precisely the washback effect on the students' learning materials and perceptions of a particular set of characteristics of the CSAT that were introduced by the linkage policy), the study had limitations in terms of the data analysis. That is, the study only provided simple descriptive statistics, such as the

number and percentage of students who are currently learning English through EBS prep books, and comparisons of the effect of the EBS-CSAT linkage policy among the three grade groups were made based on the results of these statistics. Although the study has found some differences in the use and perceptions of the EBS-CSAT prep books between the different grades, it is still not clear whether these differences are statistically significant. Further statistical analysis, such as the t-test and ANOVA, could probably have been done to determine the statistical significance of the differences in the washback among the different grade groups.

Jeong's (2018) research, on the other hand, was one of the very few studies that appraised the statistical significance of the differences in the effect of the CRA between the two different grades through cross tabulation. In the study, she surveyed first- and second-grade students to compare their perceptions of the CRA in the CSAT-English, and the differences between the two groups were indicated in their thoughts about the changes in the academic burden for studying the CSAT subjects other than the CSAT-English. While most studies, including the work of Park (2017), ended with presenting simple comparisons of the effect of the CSAT-related policies on learning between different samples, Jeong went further to appraise whether the differences found between the different grade groups were statistically significant, making the results of the study more reliable and accurate.

Despite the increased accuracy and reliability of the study, however, the work of Jeong indicated some limitations in terms of its methodologies. First, the study failed to survey the third-grade students as they had less than a month left for the upcoming CSAT at the time the survey was conducted. This was a practical problem

that could not be easily solved. However, if the survey had been conducted a little earlier so that third-grade students could take part in the study, the results of the study could have provided deeper insights into the relationship between the impact of the national policy on the test and the student learning. Second, the work focused solely on the students' perceptions of the policy, while the impact of the CRA on other learning practices, such as learning attitudes, learning contents, and learning strategies, remained untouched. Since many studies have already examined the students' perceptions of the CSAT-English as well as the national policies adopted for the test, additional research on their perceptions can be regarded as simple duplication of the previous studies, and hence, research on the effects of the CSAT-English on other various learning practices is needed.

Given the limitations of these previous studies, the current study aims to investigate the washback effect on student learning of the CSAT-English that is characterized by the EBS-CSAT linkage policy and CRA and compare the effects among the three grade groups through statistical methods. The study is expected to deepen our understanding of the nature of the CSAT-English and its impact on various aspects of student learning and provide some future directions for the test.

## **1.2 Research Questions**

The current study intends to quantitatively measure and compare the washback effect on student learning of the CSAT-English that is shaped by the EBS-CSAT linkage policy and the CRA among different grade groups. The comparison of the washback effect between the grades has recently gained considerable importance in the

washback studies on the learners as it is crucial to understanding how students' learning practices are impacted as their time (or number of days) left before their targeted test date changes. As seen in the above section, quite a few studies (Jeong, 2018; Park, 2017) have been conducted to compare this impact of the test and its policies on learning between different grades. However, some major limitations of these studies have been identified which include the lack of statistical computation to appraise the statistical significance in the differences in the washback between the grades (Park, 2017) and the lack of investigation into the impact on other various learning practices (Jeong, 2018). With this as a backdrop, the current study intends to give an answer to each of the following three research questions:

1. How do EBS-CSAT linkage policy and CRA influence the nature of the CSAT-English?
2. Does the CSAT-English that is characterized by the EBS-CSAT linkage policy and the CRA influence student learning?
3. Does students' school year influence the washback effect of the CSAT-English that is characterized by the EBS-CSAT linkage policy and the CRA on student learning?

### **1.3 Organization of the Thesis**

The organization of the present study is as follows. Chapter 2 reviews the theoretical background and previous empirical studies of the washback, washback on learning, Korean educational contexts, and the nature of the CSAT-English in regard to the

EBS-CSAT linkage policy and criterion-referenced assessment. Chapter 3 describes the methods and procedures for data collection and analysis. Chapter 4 presents the results of the statistical analyses. Chapter 5 discusses the major findings in relation to the research questions posed for the current study. Lastly, Chapter 6 concludes the thesis study with a summary of the findings, implications of the current study, and limitations and suggestions for future research.

# Chapter II

## Review of Literature

This chapter provides a general overview of the theoretical backgrounds for washback, previous empirical investigation of the washback effects, and the context of the current study. More specifically, the chapter presents definitions and theoretical frameworks of washback; reviews previous washback studies conducted home and abroad, and finally analyzes the CSAT-English characterized by the EBS-CSAT linkage policy and criterion-referenced assessment.

### 2.1 Washback

In language assessment, washback refers to the influence of language testing on teaching and learning (Alderson & Wall, 1993). While it is relatively a new educational concept, similar studies investigating the relationship among testing, teaching, and learning have long been underway in the field of general education. The notion of measurement-driven instruction, which refers to the use of the high-stakes test to direct teachers and their teaching, gained prominence from studies in the 1980s and early 1990s (Bracey, 1987; Popham, 1987). Popham (1987), in particular, was one of the strongest proponents of the measurement-driven approach that he insisted on using the high-stakes test to adjust the school curriculum to fit in with the test. Also, as a similar concept to the measurement-driven instruction, many researchers supported the notion of curriculum alignment in which curriculums are influenced by the test and hence align with the contents of the test (Shepard, 1991).

The importance of examining the relationship between testing and teaching/learning has been described in relation to the concept of consequential aspect of construct validity that was first introduced in the early work of Messick (1989). According to Messick (1995), test validity is not just determined by “the value implications of score interpretation (p.6)” but rather by “the actual and potential consequences of test use (p.6),” such as the influence of a particular test on individuals, school settings, the educational system, or society as a whole. That is, testing is not a remote event in which the validity of a particular test is determined solely by appraising how well it measured what it intended to measure but is highly connected to the social setting that is under the influence of testing.

In this regard, the concept of washback was introduced as a significant constituent of the consequential aspect of construct validity, and the washback research determines the validity of a test by examining the intended or unintended, positive or negative impact of testing on teaching and learning.

In addition, washback differs from previous notions, such as measurement-driven instruction and curriculum alignment, in that teaching and learning are considered as having the equal status in terms of the relationship between testing, teaching, and learning. That is, while previous concepts regarded the test influence on the learning as the consequence of the test impact on teaching, learning in the washback theory is directly affected by testing. Indeed, the fifteen washback hypotheses provided by Alderson and Wall (1993) show that teaching and learning have the same importance in terms of the number of hypotheses that are directly related to them:

- (1) A test will influence teaching.
- (2) A test will influence learning.
- (3) A test will influence what teachers teach.
- (4) A test will influence how teachers teach.
- (5) A test will influence what learners learn.
- (6) A test will influence how learners learn.
- (7) A test will influence the rate and sequence of teaching.
- (8) A test will influence the rate and sequence of learning.
- (9) A test will influence the degree and depth of teaching.
- (10) A test will influence the degree and depth of learning.
- (11) A test will influence attitudes to the content, method, etc. of teaching and learning.
- (12) Tests that have important consequences will have washback.
- (13) Tests that do not have important consequences will have no washback.
- (14) Tests will have washback on all learners and teachers.
- (15) Tests will have washback effects for some learners and some teachers, but not for others.

(Alderson and Wall, 1993, p.120-121)

There are five statements each that describe the washback on teaching (washback hypotheses 1, 3, 4, 7, 9) and learning (washback hypotheses 2, 5, 6, 8,

10), and these hypotheses have played a substantial role in laying out the directions of future washback studies.

Many washback studies that have been conducted based on these hypotheses have investigated either the influence of the public high-stakes tests, such as the College Scholastic Ability Test (CSAT) in Korea (Kwon, Lee, & Shin, 2017) and the Hong Kong Certificate of Education Examination (HKCEE) in Hong Kong (Cheng, 1997, 1999; Cheng, Andrews, & Yu, 2011) on a particular educational system in which the test is being held or the impact of major English-language tests, such as the Test of English as a Foreign Language (TOEFL) (Alderson & Hamp-Lyons, 1996) and the International English Language Testing System (IELTS) (Green, 2006) in worldwide educational settings.

However, despite the substantial number of washback studies that were carried out after the work of Alderson and Wall (1993), many studies have generally focused on the relationship between testing and teaching (Cheng, 1997, 1999; Watanabe, 2004), and it was not until recently that washback research on the students and their learning began to receive significant attention from researchers.

### **2.1.1 Washback on Learning**

According to Cheng et al. (2015), the washback studies on the learners and their learning are important as students receive the most direct influence of testing, and the final goal of assessment is to encourage their learning. In fact, the importance of investigating the washback effect on the learners has been widely emphasized since the 1990s. Bailey (1996), for instance, assigned a special status to the washback on

the learners by separating the test impact on the learners (i.e. washback to the learners) from other possible types of test consequences, such as the influence of testing on the teachers, materials, writers, curriculum designers, and researchers (i.e. washback on the program).

For more than three decades, knowledge and insights from the washback studies on the learners have been accumulated, and some of these studies have significantly contributed to the understanding of the nature of the washback on the learners. A case in point is the work of Shih (2007) in which he proposed some factors that can affect the washback of the test on students' learning and psychology in the washback model of students' learning. The model is represented in Figure 1.

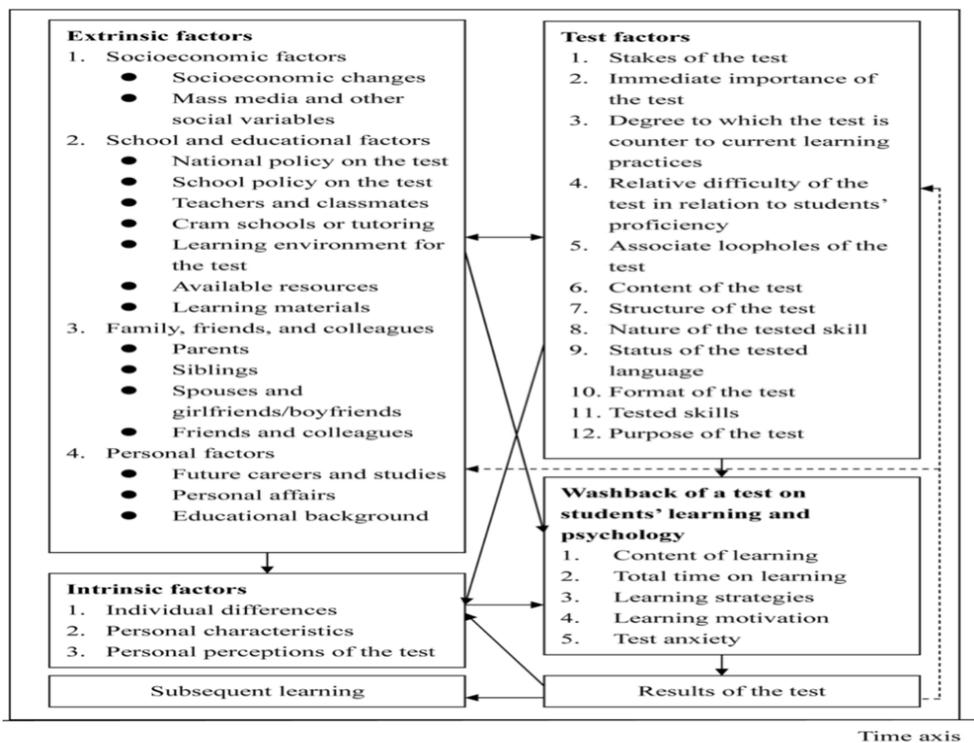


Figure 1 A washback model of students' learning

(Shih, 2007, p.151)

As shown in Figure 1, Shih categorized the factors into three groups: test, extrinsic, and intrinsic factors. First, test factors consist of 12 components that describe the overall nature of the test, such as the stakes of the test, immediate significance of the test, test content, test structure, etc. Second, extrinsic factors are comprised of four sub-factors: socioeconomic factors, school and educational factors, personal factors, and family, friends, and colleagues. Finally, intrinsic factors include individual differences, personal characteristics, and personal perceptions of the test.

One of the important points to note in the model is that these three factors can be interconnected with each other. The extrinsic factors, for instance, can have influence on the intrinsic factors. A case in point is the work of Cheng et al. (2011) that investigated the influence of the parents on the students' perceptions of the school-based assessment (SBA) in the HKCEE. The study found that the relationship between the parents and their children is very strong that the parents' views are highly reflected in their children's perceptions of the SBA and suggested that the educational phenomenon in which the washback on the learners are affected by other stakeholders requires "further investigation within other contexts worldwide (p.240)."

Extrinsic factors can also interact with the test factors. The national policy on the test, for instance, can impact, or trigger changes in, the test factors, such as the nature of the tested skills and format of the test which can, in turn, lead to change in the teachers' teaching practices and hence student learning. The work of Cheng (1999) was one of the studies that investigated the washback effect that was caused by a series of this processes, although the model was not introduced at the time her

study was conducted. Here, she examined the washback effect of the new HKCEE on the teaching practices of the teachers. Under the control of the Hong Kong Examination Authority, the HKCEE underwent a major change in the skills tested in the test and its format, and in response to the changes in the test, teachers also made considerable changes in their teaching contents, materials, and class activities.

The work of Shih shows that washback research on the learners is very much complicated that it requires the consideration of the relationship between these factors and the washback of a test on student learning and also the relationship among the factors. While a number of studies have been carried out in the Hong Kong educational setting, there are only a few studies that have been conducted in Korea with the consideration of this complex structure of the washback on the learners in mind. Kwon (2015), in particular, pointed out the lack of studies on the washback effect of the English section of the CSAT (CSAT-English), a Korean university entrance examination, and suggested that the washback effects of the CSAT-English on learning and the impact of the national policies exerted on various characteristics of the test should be investigated more actively and thoroughly. In the following sections, general discussions of the Korean educational setting and the important characteristics of the English section of the CSAT are provided.

## **2.2 Korean Educational Context**

According to the World Bank (2019a), South Korea posted a \$1.619 trillion nominal gross domestic product (GDP) in 2018, ranking 12<sup>th</sup> among 205 countries worldwide, with the per capita gross national income (GNI) standing at \$30,600. As a result,

South Korea has become the seventh member of the so-called the '30-50 club' with a per capita income of more than \$30,000 and a population larger than 50 million. This placed South Korea as one of the most advanced countries in the world. Before Korea entered the club, there were only six countries that met the criteria of being the member of the 30-50 club: The United States, Japan, Germany, the United Kingdom, France, and Italy.

South Korea's economy has been growing at a high rate under the government's management and support of the export-led model. Given the limitations of poor natural resources and its small domestic market compared to other six members of the 30-50 club, it was a truly natural and inevitable choice for the Korean government to rely on international trade. According to a statistical report from the World Bank (2019b), South Korea's trade-to-GDP ratio, which is defined by the ratio of exports and imports to GDP, has reached 83% in 2018, showing a very high economical dependency on trade among the Organization for Economic Co-operation and Development (OECD) countries.

The export-driven economic growth policy generally led to an increase in the demand for language experts to deal with international businesses (Ingram, 1986). English, in particular, has a special status in Korean society, and the benefits of learning English are greater than acquiring other languages as it is the most widely spoken language in the world (Song, 2011). The English language learning market in Korea has indeed grown exponentially during the past few decades. According to a report from Statistics Korea (2019), the market size of private education in 2018 was \$16.177 billion, and of the total expenditure, private English education

accounted for about 30% (\$4.85 billion), which was the largest portion among the whole subjects that were examined. As a result of many Koreans' interest in learning English, Korea earned quite a high mark from institutions that measure the English proficiency of countries worldwide, especially non-English-speaking countries. According to the EF English Proficiency Index's (EPI) report (2018), South Korea has a moderate proficiency EF EPI score of 56.27 which was ranked 6<sup>th</sup> out of 21 countries in Asia.

Despite the achievements, however, there are some crucial limitations that have long existed in the Korean English education system. Most of all, instead of promoting active learning and individual thinking to students, the cramming method of teaching has been prevalent in both public and private educations. According to Park (2009), most of the English learning processes in Korea are carried out by following exactly the same as the instructions given by the teacher, dictating the contents written on the blackboard without skipping a single word, and memorizing them. By reducing the students into passive patients in the classroom, students are blocked from expressing their own individual thoughts and establishing their active position through involvement in the class.

Another problem of the current English education system is that it is excessively focused on listening and reading, while there are not enough courses for speaking and writing. The Ministry of Education's (MOE) (2015) investigation of the educational environment in primary and secondary schools revealed that the number of English conversation instructors decreased from 6,480 in 2013 to 2,837 in 2015. Also, the number of native English teachers declined from 7,790 to 3,265

during the same period. In terms of class hours, the percentage of secondary schools that assigned English conversation courses for at least one hour a week marked 71.5% in 2015, which is 8%p drop from 2013.

These problems, however, remain unsolved as most Koreans are learning English to get desired scores on English tests, rather than to use it as a communication tool with people speaking English. As most Koreans are learning English to the test, the English curriculums in regular schools and private educations are set to satisfy their needs. According to Cheng (1997), this examination-led educational system is strengthened with the use of the high-stakes test. For students, in particular, the CSAT is a very high-stakes public examination that its scores are one of the most crucial factors in determining which university to go to, and the English section of the test has considerable impact on their English learning practices.

### **2.3 CSAT-English**

The English section of the CSAT (CSAT-English) has considerable significance in English education in Korea. The test is designed to measure test-takers' academic achievement based on achievement standards set by the National English Curriculum and the levels of English proficiency required for higher education, and the scores of the test have a major impact on determining the university the test-takers can enter. Since its inception in 1993, the CSAT-English has gone through many changes in the number of items, the type of questions, passage length, testing time, and maximum raw score. The formal name of the test even changed from Foreign Language to English in 2013. A list of major features of the current test is as follows.

**Table 1 Outline of the CSAT-English**

<b>Category</b>	<b>CSAT-English</b>
Test contents	Topics and passages based on English I, II
Testing time	13:10-14:20 (70min)
Test structure	45 five-choice multiple-choice items (17 Listening, 18 Reading)
Score range	0-100 (2 or 3 points for each item)
Score report	Stanine score only

Despite the numerous changes that have been made on the test, the stakes or importance of the test on the test-takers has remained identical, and the CSAT-English is still believed to have a strong influence on student English learning. Although not enough studies have been conducted to investigate the impact of the CSAT-English on the learners (Kwon, 2015), there were actually some research studies that played an important role in understanding the relationship between the CSAT-English and student learning in early 2000s.

Hwang (2003), for instance, conducted a survey of 120 third-grade students in high schools to examine the washback of the CSAT-English on their learning practices. Since it was one of the earliest studies that examined the influence of the CSAT-English, the study focused on revealing the actual existence of the washback effect of the test and the nature of the washback, whether it was positive or negative. The results of the study showed that there was a negative impact of the CSAT-English on both teaching and learning, and that the participants of the study had negative attitudes towards the test. While Hwang's research surveyed a group of

students who were in the same grade, the work of O (2001) observed second- and third-grade classes and also interviewed some of the third-grade students from the observed classes to see if there were any significant differences in the washback effect of the CSAT-English between the two grades. The study found that learning areas, such as learning materials, contents, and class activities, were more CSAT-oriented in the third-grade classes than the second-grade ones.

While the existence of the washback of the CSAT-English has been verified through previous studies, including Hwang's, the kind of research that is intended to compare the washback effect between the different grade groups is still being carried out frequently. As with O's research, these studies aimed to investigate the possible influence, on the washback intensity of the test, of the chronological distance between the time point when the study was undertaken and the students' expected CSAT test date, which can be represented by the students' school year or grade. However, unlike the work of O in which the qualitative research methods, such as in-depth interview and class observation, were used, recent studies in 2010s were generally conducted based on the quantitative analyses and methods of research, such as survey questionnaires, to make comparisons of the washback between the different grade groups. According to Queirós et al. (2017), the use of qualitative research methods allows deeper understanding of a given problem, but these studies are not usually concerned with the quantitative representations of the data that are often carried out through statistical analysis. Although the results of the class observation and interviews in O's research indicated that there was a difference in

the degree of the washback among the three grade groups, it was difficult to identify how big these differences were.

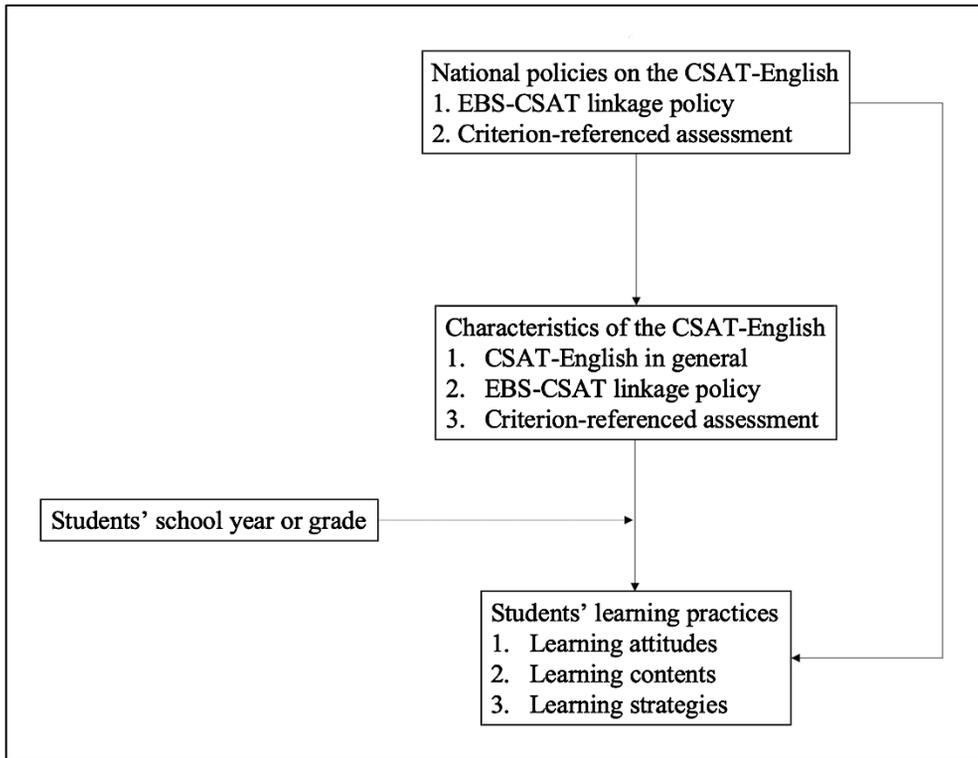
Another change that has appeared in recent studies is that the studies are not only concerned with the washback effect of the CSAT-English itself but also the impact of the national policies exerted on the test. The situation behind the change was the excessive rise in the total spending on private English education with the increase in the difficulty level of the test. Indeed, as the English section of the CSAT became very difficult over the past few years, the students who received private English education were more advantageous in getting a good grade than those who did not. Consequently, people's demand for private English cram schools have increased so dramatically over the years that annual expense for private English education in 2012 reached \$ 5.48 billion according to Statistics Korea (2013), accounting for the largest portion of the total private education expenditure.

In response to the excessive increase in the difficulty of the test and total spending on private English education, the Korean government came up with various solutions to lower the cost of private English education and to reduce the achievement gap between students who took private English courses and those who did not. Among them, the introduction of the Educational Broadcasting System (EBS)-CSAT linkage policy in 2010 and the criterion-referenced assessment (CRA) in 2017 were the two major measures for tackling these problems.

According to the washback model of students' learning (Shih, 2007), these national policies of the test, which belongs to the extrinsic factor, can also have direct impact on the learners or have multidimensional impact on student learning by

interacting with other factors, such as the test factors. Over the past few years, quite a few studies have been conducted to investigate these influence of the EBS-CSAT linkage policy and the adoption of the CRA approach to the CSAT-English on student learning.

Given the recent trends of the washback studies on the learners in which the influence of the national testing policies on the characteristics of the test and on the learners and the impact of the students' academic year on the washback intensity of the test are the main concerns, the current study aims to examine the washback effects of the CSAT-English that are directly and indirectly caused by the EBS-CSAT linkage policy and the adoption of the CRA for the CSAT-English on student learning and further compare them among the three different grade groups through appropriate statistical methods. In order to clarify the relationship between the variables, a tentative washback model, which was created based on Shih's (2007) washback model of students' learning, is provided in Figure 2.



**Figure 2 A washback model of the CSAT-English and its national policies**

As seen in Figure 2, two national policies on the CSAT-English (i.e. EBS-CSAT linkage policy and CRA) can either have direct impact on student learning or have indirect influence by triggering changes in the test characteristics of the CSAT-English, which can, in turn, impact student learning. While the former seems to be the main topic of research for the field of language policy, the latter, which mediates the change in the nature of the test, is the area of main focus in the current washback study. With this as backdrop, the national policies on the CSAT-English are assumed to have influence on the test by modifying its characteristics, which, in turn, generates washback effect on the learners. Here, new characteristics are added to the test that are highly related to the policies that affect it. For instance, with the

introduction of the EBS-CSAT linkage policy in the CSAT-English, 70% of the test contents came up from the EBS-CSAT prep books, while the implementation of the CRA has changed the grading system of the test. These changes in the test consequently could affect student learning and cause washback effect. Likewise, causal relationships are found in Figure 2 between the national policies on the CSAT-English and the characteristics of the test, and between the nature of the CSAT-English and student learning, which is referred to as washback. While national policies on the CSAT-English and students' learning practices function as an independent and dependent variable respectively, the characteristics (or nature) of the CSAT-English can either be an independent or a dependent variable depending on which variable it matches with. That is, when defining the relationship between the policies on the test and the characteristics of the test, it is regarded as a dependent variable that is influenced by the test policies, while it is an independent variable when matched with the students' learning practices. Meanwhile, the students' school year or grade functions as a moderating variable that modifies the strength or magnitude of the washback effect of the CSAT-English that is shaped by the two major policies on student learning. The current study attempts to investigate not only the impact of the above-mentioned, two national CSAT related policies on the characteristics of the CSAT-English and the overall washback effect of the CSAT-English on students' learning practices but also the influence of the students' school year (or grade) on the washback intensity of the test.

### **2.3.1 EBS-CSAT Linkage Policy**

Under the government's plan to reduce private education expenditures and to normalize public education, the MOE declared that, beginning from 2010 CSAT, 70% of the test contents in all six study subject areas would be based on the EBS-CSAT prep books reviewed by the Korea Institute for Curriculum and Evaluation (KICE). In the first year of policy implementation, however, students' academic and financial burdens rather increased as they had to buy and study more than 30 EBS textbooks in total to keep up with all six study areas in the CSAT, including 11 CSAT English prep books. To minimize their academic and financial stress, the MOE gradually reduced the number of the EBS-CSAT prep books that were recommended to be reviewed by students. As of 2019, there were less than 20 EBS textbooks that a student had to study, and among them, the number of textbooks for the CSAT-English decreased to 4.

Currently, the contents from the designated test prep books are reused to create a large portion of the CSAT-English items by revising the EBS passages, changing the question formats, and referring to the tables, figures, topics, and concepts mentioned in the book. Beginning from the 2021 CSAT, however, when the first-grade students in high schools surveyed in this study will take the test, the percentage of the test contents to be generated based on the EBS materials are expected to decrease from 70% to 50%.

The EBS-CSAT linkage policy has been active for approximately 10 years since 2010, but annual statistics showed that the results were quite the opposite of what was expected by the Korean government. As to the private English education,

for instance, reports from Statistics Korea (2011, 2012, 2013, 2014, 2015, 2016, 2017) indicate that the monthly expenditures on private English education per high school student have been increasing steadily since 2010. In particular, the cost jumped approximately 7% from \$60.98 to \$65.22 in 2016, which saw the largest year-over-year increase. On the other hand, the percentage of high school students who were receiving private English education showed a downward curve from 2010 to 2015, indicating some promising results of the EBS-CSAT linkage policy. However, the rate began to rise again in 2016, with 30.3% of high school students taking at least one course from private English institutions.

In addition to these statistical reports, a number of research studies have been conducted on the EBS-CSAT linkage policy. Song (2015), for instance, examined the number and percentage of third grade students in high schools who were using the EBS textbooks to prepare for the CSAT-English and how much they were satisfied with the materials. The results of the study showed that most of the students were using the EBS materials to prepare for the CSAT, and they were overall satisfied with the materials. Also, the EBS policy helped to reduce their academic burden for studying English but did not increase their interest in learning English. Her work has enhanced our understanding of the EBS-CSAT linkage policy by providing up-to-date statistics about the number of students using the EBS prep books to prepare for the CSAT-English as well as the overall perceptions towards the policy. Furthermore, Park (2017) conducted a study similar to Song's (2015), but her work not only surveyed the third graders but also first and second grade students in their high schools. Thereby, the study was able to provide the results of analysis

from each grade on their actual use and perceptions of the EBS-CSAT prep books and make numerical comparisons between the different grade groups. The results showed that EBS textbooks were more actively used in the third-grade classrooms than in other lower grades' classes, while the third-grade classes were quite similar without significant differences.

Both studies have revealed how the EBS materials are being used in the actual English classes and how students think about them. However, the results were limited to providing simple descriptive statistics, such as the number and percentage of students who chose to use the EBS textbooks to prepare for the CSAT-English. In particular, the work of Park (2017) made comparisons between the different grade groups based on the frequency analysis, but the results would have been more accurate if additional statistical analyses, such as the t-test or analysis of variance (ANOVA), had been conducted to determine whether those differences were statistically significant.

### **2.3.2 Criterion-Referenced Assessment**

As a countermeasure to the recent increase in the monthly cost of private English education per high school student and the percentage of high school students who were receiving private English education, the Korean government decided to adopt a criterion-referenced assessment (CRA) approach for the CSAT-English from the 2017 administration. The test had been norm-referenced for more than 20 years on the basis of the stanine system.

**Table 2 Comparison between norm- and criterion-referenced assessment**

	<b>Norm-Referenced Assessment</b>	<b>Criterion-Referenced Assessment</b>
Score interpretation	Relative interpretation of scores	Absolute interpretation of scores
Purpose of evaluation	To rank students or test-takers according to their test scores or proficiencies	To measure students' or test-takers' academic achievement against the National English Curriculum
Test items	A certain number of items are made complicated to secure the test's discriminating power	Test items being generated based on well-defined and specific criteria
English learning	Intense competition among test-takers to get desired scores or grades.	Learn English to improve their communication skills
English classes	English classes focusing on solving CSAT-English practice questions rather than improving students' English communication skills	English classes focusing on improving students' English communication skills rather than solving CSAT-English practice questions

A brief comparison of major characteristics between the norm-referenced and criterion-referenced assessment is presented in Table 2. One of the important features of the CRA is its absolute interpretation of test scores based on cut-off scores, as opposed to the relative score interpretation of NRA (norm-referenced assessment) (Shin, 2018). The MOE is still maintaining the method of scaling, and putting, students' test scores on a nine-point standard scale (i.e., the stanine scale) in the new grading system of the CSAT-English. However, the way of converting the test scores into the stanine scores has changed. In the norm-referenced grading system, the stanine scores were assigned based on the percentile rank of the students' test scores, but with the introduction of the absolute scale system, the raw scores of the test were

directly converted to the stanine scores on the basis of preset cut-off scores for each stanine. If one's CSAT-English score is 90 or above out of 100, then his/her report card will be given a stanine score of 1 regardless of the percentile rank of the raw score.

Despite the easiness of getting the desired scores or grades, however, the data collected by the Statistics Korea (2017, 2018, 2019) have shown quite different results from what was expected by the Korean government and the KICE. That is, the monthly expenditures on private English education per high school student were consistent with \$65.22 over the two years in 2016 and 2017, but the following year, the cost jumped approximately 13% from \$65.22 to \$73.64, which saw the largest year-over-year increase. Also, the percentage of high school students who received private English education fell 0.1%p to 30.2% in 2017, compared to the previous year, but in 2018, the rate increased by 2.3%p at 32.5%.

As done on the EBS-CSAT linkage policy, many studies have been conducted on the CRA policy for the CSAT-English. First, Lee (2018) conducted interviews with 11 third grade students in high school to investigate their perceptions of the CRA features in the CSAT-English and the impact of the CRA characteristics on their private English education, English learning methods, and English classes in their school. The study focused on the negative impacts of the CRA features of the CSAT-English on the students' learning practices. Her study found that the introduction of the CRA in the CSAT-English brought about the "balloon effect." That is, the new grading system decreased the importance of the CSAT-English and

thereby, many students shifted their focus to other CSAT subjects of non-CRA, such as Korean and Mathematics.

Indeed, the balloon effect of the CRA in the CST-English is one of the major concerns among the stakeholders of the CSAT. Recent data from Statistics Korea (2017, 2018, 2019) showed that monthly expenditures on private Korean and Mathematics education per high school student in 2018 increased by 52% to \$28.90 and 13% to \$97.44, respectively, in comparison with those of 2016. Also, the percentage of high school students who received private Korean education hit 16 % with an increase of 4.4%p, and the percentage of high school students who received private Mathematics education was 38%, an increase of 1.9%p.

Second, Lee's (2018) study also found that the influence of the CRA-related features of the CSAT-English on the students' learning practices was very weak and that their learning contents and learning methods did not change significantly from those of the CSAT-English based on the previous grading system. The positive washback of the CRA, which was anticipated by the government, was that the new system would normalize public education and reduce the test-driven educational setting in the classrooms. In contrary to the expectation, however, the result of the study showed that the effect of switching to the CRA approach for the CSAT-English might not lead to the actual gain in learning outcomes or changes in learning contents and strategies of the students.

In addition to the work of Lee (2018), which was conducted by employing the qualitative research method, there were quite a few studies that attempted to quantify the influence of the CRA on the students' learning practices through survey

questionnaires. Jang (2017), for instance, surveyed 450 second-grade students in high school to examine their perceptions of the CRA features of the CSAT-English that had been newly introduced to it with the adoption of the new grading system. The result showed that most of the students agreed with the intended purposes and rationales of the CRA in the CSAT-English and believed that the new system would lessen the competition for getting good scores and academic burden for studying English. Although the work of Jang (2017) contributed to the understanding of the effectiveness of the new evaluation system in the CSAT-English, the survey questionnaire consisted of only 8 items in total that it seemed quite difficult to measure the students' perception of the policy in detail. Also, the study did not go further to examine the influence of the CRA on the students' learning practices, such as their learning attitudes, learning contents, and learning strategies.

Similar to the work of Jang (2017), Jeong (2018) investigated the students' perceptions of the CRA in the CSAT-English through a survey, but the participants of the study included not only the second-grade students but also the first graders. Therefore, the study was able to make comparisons between the two grade groups in light of the competition for acquiring good scores, importance of the CSAT-English, academic burden for studying English, and academic burden for learning other CSAT subjects. Here, cross tabulation was used to determine whether the differences that occurred between the two grade groups were statistically significant. The results showed that most of the participants in both first and second grades thought that the perceived importance of the CSAT-English did not change with the introduction of the CRA, but the competition for getting desired scores on the CSAT-English and

academic burden for learning English seemed to have been reduced. On the other hand, the percentage of students who said the academic burden for studying other CSAT subjects would increase due to the new system was higher for the second-grade students than the first graders. Also, the second-grade groups were more negative than the freshmen on the statement that the English classes would become more communicative with the implementation of the CRA in the CSAT-English.

Jeong's (2018) research is one of the few studies that attempted to determine whether the differences in the washback intensity of the CRA in the CSAT-English would be statistically significant between different grade groups. The study used the cross tabulation, which is one of the most commonly used methods to determine the statistical significance, along with the t-test and ANOVA. The work of Jeong (2018), however, could not include the third-grade students as the participants in the study since the survey was conducted on October when the number of days left for the upcoming CSAT was less than a month. Also, as with other previous studies, the influence of the CRA on other various learning practices, such as the learning attitudes, learning contents, and learning strategies, were not investigated in the study.

In summary, the main purpose of introducing the EBS-CSAT linkage policy and the CRA in the CSAT-English has been the reduction of the private English expenditures and normalization of the public educations. However, major changes in the high-stakes test can also trigger other various impacts on the teachers, students, classrooms, schools, and educational system (Wall, 1997). There were quite a few studies that examined the influence of the EBS-CSAT linkage policy and CRA in

the CSAT-English, but some of these research had their main focus on the students' perceptions of the policies and did not go further to explore in detail the washback effect of these changes on other learning practices, such as their learning attitudes, contents, and methods. Also, some other limitations, such as the lack of statistical significance testing, to determine the statistical significance in the differences in the washback effect between different grade groups (Park, 2017) and the lack of study on the third-grade students (Jeong, 2018) were found in the studies that compared the washback effect between the different grade groups.

# Chapter III

## Methodology

The following chapter describes the methods used for data collection and analysis. More specifically, it presents background information of the students and teachers who took part in the current study, a detailed description of the survey questionnaires that are given to the participants, and the procedures to collect and analyze the data.

### 3.1 Participants

#### 3.1.1 Students

A total of 334 high school students participated in the current study. Of the 334, 111 (33.2%) were first-grade students, 113 (33.8%) were second graders, and remaining 110 (32.9%) were in the third grade. With regard to the location of the school, 72 (21.6%) students were from schools located in the capital city of South Korea (Seoul), 81 (24.3%) from local metropolitan cities (Busan = 16 (4.8%), Daegu = 16 (4.8%), Incheon = 17 (5.1%), Gwangju = 15 (4.5%), Daejeon = 10 (3.0%), and Ulsan = 7 (2.1%)), 178 (53.3%) from local provinces (Gyeonggi = 92 (27.5%), Gangwon = 6 (1.8%), Chungcheong = 23 (6.9%), Jeolla = 21 (6.3%), and Gyeongsang = 36 (10.8%)), and 3 (0.9%) from other regions (Sejong = 1 (0.3%) and Jeju = 2 (0.6%)).

While the students' schools were found to be distributed across the country, all of the study participants were students attending non-special, general high schools. Students attending special-purposed high schools (e.g., foreign language, science high schools) were excluded from the current study as they were highly likely to be

in the test-driven educational system regardless of their grade in the school and most of their learning practices were shown to be already test-oriented (Sohn & Ju, 2010).

Students were able to give multiple answers to the questions about the type of college admissions they were intensively preparing for and about the reason for receiving private English education. First, students' responses for the former question is summarized in Table 3. It indicates that more than half of the students plan to enter universities through school records procedure of screening (53.8%), and it was followed by CSAT screening (28.2%), performance screening (6.3%), and essay screening (4.6%).

**Table 3 Summary of the types of college admissions students are preparing for**

Response	Frequency		% of Cases
	N	%	
CSAT screening	124	28.2	37.1
School records screening	236	53.8	70.7
Essay screening	20	4.6	6.0
Performance screening	28	6.3	8.4
Others	31	7.1	9.3
Total	439	100.0	131.4

Next, descriptive statistics of the students' reason for receiving private English education are presented in Table 4. It shows that 135 (33.1%) students were not taking private English courses at all. On the other hand, those who were receiving private English education were mostly preparing for either school English tests or

CSAT-English, or both. There were also a few students who were preparing for the English-language tests such TOEIC, TOEFL, and TEPS.

**Table 4 Summary of the reasons for receiving private English education**

Response	Frequency		% of Cases
	N	%	
Not receiving	135	33.1	40.4
To prepare for the CSAT-English	113	27.7	33.8
To prepare for the school English tests	139	34.1	41.6
To prepare for the English-language tests such as TOEIC, TOEFL and TEPS	6	1.5	1.8
To improve practical English	15	3.7	4.5
Total	408	100.0	122.2

Other specifics of the students' background information are summarized in Table 5: The mean values of the most recent grade/score of the PCSAT-English (1 = Stanine of 1, 2 = Stanine of 2, 3 = Stanine of 3, 4 = Stanine of 4, and 5 = Stanine of 5 or below), hours of studying English per day (1 = Less than 1 Hour, 2 = 1 to 2 Hours, 3 = 2 to 3 Hours, 4 = 3 to 4 Hours, and 5 = Over 4 Hours), and self-assessed English levels rated on a five-point Likert scale (1 = Very Poor, 2 = Poor, 3 = Average, 4 = Good, 5 = Very Good).

**Table 5 Summary of the students' background information**

Measure	Grade			Total
	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>	
The PCSAT-English grade/score	3.50	3.61	3.59	3.57
Hours of studying English per day	1.73	1.70	2.00	1.81
Self-assessed proficiency of English				
Listening	3.40	3.42	3.62	3.48
Reading	3.30	3.33	3.20	3.28
Speaking	2.78	2.81	2.59	2.73
Writing	2.70	2.80	2.63	2.71

### 3.1.2 Teachers

A total of 11 high school teachers teaching English at a high school in Gyeongsang province took part in the teachers' survey (three males, eight females). As with the students' questionnaire, the teachers also had to enter some personal information before moving on to the main survey questions.

First, of the eleven participants, nine (81.8%) teachers had more than four years of teaching experience, while the other two (18.2%) worked as English teachers for three to four years. Also, a total of nine (81.8%) teachers had the experience of teaching all three different grades of students, whereas the remaining two (18.2%) have taught first and second grade students only. Currently, three (27.2%) teachers were teaching the first-grade students, and the number of teachers teaching second- and third-year students in high school was four (36.4%) each.

## **3.2 Instruments**

### **3.2.1 Students' Questionnaire**

The students' questionnaire consists of three main sections with 32 questions in total, all of which were created to examine the washback effect of the CSAT-English on students' learning. Students' learning was assessed in three major areas, including learning attitudes, learning contents, and learning strategies (see Appendix A). These three areas of learning correspond to the three sections of the questionnaire. Each question in each of these sections is designed to be answered on a five-point Likert scale of agreement. The scales are labeled as 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, and 5 = Strongly Agree. After finishing all the sections, two open-ended questions are given to the students so that they can write down additional opinions or experience about the possible effects of the CSAT-English and its policies on their learning attitudes, contents of learning, and strategies use.

Also, the questions of five-point Likert scale in each main section are again sub-divided into three subsections. The first group of questions are generated in order to investigate the washback of the CSAT-English in general (hereafter CSAT-English), while the second and third sections include statements about the changes in the learning practices due to the introduction of the EBS-CSAT linkage policy (hereafter EBS) and the CRA (hereafter CRA), respectively. Items 1 to 9, for instance, are a group of questions that ask about the washback on the students' learning attitudes. Of the nine items, the first three statements (Items 1, 2, 3) are related to the impact of the CSAT-English in general, and the following three items (Items 4, 5, 6) are about the effects of the CSAT-English attributable to the EBS-CSAT linkage policy, and the last three questions (Items 7, 8, 9) are about the effects of the CSAT-

English due to the adoption of the CRA. In this manner, the first and second sections comprise nine items in total, while the third section consists of twelve questions.

Reliability estimates for each of the categories in the students' questionnaire are presented in Table 6. Reliability analysis was conducted to examine how reliable the students' section and subsection scores from the questionnaire would be and how consistent the students' scores were across different items.

**Table 6 Reliability estimates of the categories in the students' questionnaire**

<b>Category</b>	<b>Item</b>	<b>Cronbach's Alpha</b>
Learning attitudes	1-9	.933
CSAT-English	1-3	.881
EBS	4-6	.886
CRA	7-9	.857
Learning contents	10-18	.913
CSAT-English	10-12	.813
EBS	13-15	.761
CRA	16-18	.925
Learning strategies	19-30	.957
CSAT-English	19-22	.861
EBS	23-26	.894
CRA	27-30	.888

As shown in the table, Cronbach's alpha for all three main categories were above .9, consistently marking high values of reliability. The results indicate high internal consistencies of the questionnaire as a whole and among the items inside the category. Cronbach's alpha for each subcategory was also computed in the table.

Since the number of items included in each section is relatively small compared to the ones in each main category, the overall reliability estimates for the subcategories are also lower than the ones calculated to represent the main categories. However, except for the subcategory made up of items about the washback of the EBS-CSAT linkage policy on the students' contents of learning (Items 13, 14, 15), all other values are all above .8.

### **3.2.2 Teachers' Questionnaire**

While open-ended questions in the students' questionnaire functioned as incidental items where students can give additional opinions or comments about the five-point Likert scale questions, the main items in the teachers' questionnaire were all open-ended type of questions (see Appendix B). With six open-ended questions in total, the first three items (Items 1, 2, 3) aimed to investigate the washback effect of the CSAT-English that was shaped by its major policies on the teachers' attitudes, teaching contents, and methods of teaching, whereas the last three items (Items 4, 5, 6) asked whether there were any differences in the teaching attitudes, contents, and methods depending on the students' grade in high school.

## **3.3 Data Collection and Analysis**

### **3.3.1 Students Data Collection and Analysis**

The survey data of the students were collected for four days between August 9, 2019 to August 12, 2019 through online. The survey took approximately 10 to 15 minutes on average for each participant. Since participants could not jump to the next item

without responding to the previous question, there was no missing cases in the data collected.

The IBM SPSS Statistics 25 was used to analyze the students' survey data. First, means of the results for each item were calculated by participants' grade. Second, in order to perform an analysis of variance (ANOVA) with the main and sub-categories, new variables were created by computing the mean value of the summated scale of the items that are grouped into a category. A one-way ANOVA with the significant value of  $p < .05$  was used to compare the means of three samples: students in grades 1, 2, and 3. Since t-test can only cover up to two samples, the survey result was analyzed by employing the ANOVA method of mean comparisons, which can provide statistical test for more than two samples. Significance value for each category was calculated to compare the mean scores among the three samples. If the result shows meaningful differences in the means with the significance value of less than .05, it can be concluded that the washback intensity is affected by the learners' grade in high school. Also, a post-hoc test was conducted to determine the exact points in which the differences in the washback intensity have occurred.

Lastly, the results of the two open-ended questions were analyzed with the qualitative method of data analysis and interpretation. Contents that were not covered in the five-point Likert scale questions but frequently came up as additional opinions or experience of the students were mainly focused. In this manner, it will complement the missing parts of the Likert items.

### **3.3.2 Teachers Data Collection and Analysis**

For the teachers' questionnaire, an offline survey was conducted September 10, 2019 through September 25, 2019 with 11 high school English teachers. It took 15 to 20 minutes on average per participant, and the collected data were analyzed in the same way as the ones from the open-ended questions in the students' questionnaire.

Whereas the open-ended items in the student survey were created to complement the five-point Likert scale items by seeking additional information that is not captured by the items, the same type of open-ended questions in the teachers' questionnaire were the core items that measured the washback of the CSAT-English and its major feature changes caused by the government policy on different teaching practices and the relationship between the washback on teachers' teaching attitudes, contents, and strategies and the grade level of their students. The analysis was mainly focused on the responses that were commonly shared or agreed among the participants.

# Chapter IV

## Results

The current chapter reports the results of analyses of the data collected from the students' and teachers' questionnaires. First, it presents the impact of the EBS-CSAT linkage policy and CRA on the CSAT-English. Second, the analysis results from the students' questionnaire are presented, which include descriptive (i.e., score distribution) and inferential statistics (i.e., ANOVA and post-hoc test results). Also, the results of qualitative data analysis on the open-ended questions in both questionnaires are summarized.

### **4.1 The Impact of the EBS-CSAT Linkage Policy and CRA on the CSAT-English**

Before moving on to the analyses of the results of the students' and teachers' questionnaires, the impact of the EBS-CSAT linkage policy and CRA on the CSAT-English is first described and presented in this section to investigate how the nature or characteristics of the test has changed with the introduction of the two major CSAT-related policies by the government.

Regarding the impact of the EBS-CSAT linkage policy on the nature or characteristics of the CSAT-English, the EBS policy has changed the way the test contents are generated. That is, according to the newly-introduced EBS-CSAT linkage policy, at least 70% of the total CSAT-English contents or items must be created based on those of the EBS-CSAT prep books that are designated (or approved) by the KICE and MOE. Therefore, out of 45 items in total, 32 or more

had to be linked with the EBS materials, and currently, as of 2019, the contents from the designated test prep books were reused to create a large portion of the CSAT-English items by revising the EBS passages, changing the question formats, and referring to tables, figures, topics, and concepts mentioned in the book.

The types of CSAT-English questions that were generated based on the EBS prep books over the last five years are represented in Table 7. The table is based on the EBS' annual reports of the CSAT contents or items that are made by referring to the EBS materials (EBS, 2015, 2016, 2017, 2018, 2019). As seen in the table, the number of items that were based on the EBS-CSAT prep books over the past five years were either 32 or 33 out of 45 questions, reaching above 70%. More specifically, as to the listening section, 15 out of 17 questions referred to the EBS materials in all five years. This marks approximately 88% of the total listening items, exceeding the basic 70% standard. On the contrary, the number of reading items that were created by revising the EBS prep books were either 17 or 18, and the percentages were all less than 70%. The percentage of reading items being less than 70% is acceptable, however, as the EBS-CSAT linkage policy states that 70% of all CSAT-English items, not each of the sections, have to refer to the EBS prep books. Also, as most of the items in the listening section were linked to the EBS materials with the percentage being approximately 88%, the overall percentage were able to reach over 70% even though the reading section marked 60 or 64% only.

**Table 7 The EBS linkage to the CSAT-English over the past five years**

Item Type	Year				
	2015	2016	2017	2018	2019
	Item				
Listening section	1-15	1-15	1-15	1-15	1-15
Reading section					
Purpose	18	18	18	18	18
Feeling	19	19	19	19	19
Argument, point of view, title, topic	20-23	20-23	20-23	20, 22-24	20, 22-24
Implication				21	21
Graph	24	24	24	25	25
True/False	25-27	25-27	25-27	26-28	26-28
Grammar	28	28	28	29	29
Vocabulary	29	29	29	30	30
Reference	30	30	30		
Fill-in-the-blank	31, 32	31	31	31, 32	31, 32
Irrelevant sentence	35	35	35	35	35
Paragraph ordering	37	36	36	36	36
Sentence insertion	39	38	38	38	38
Summarization		40			
Listening	15/17 (88%)	15/17 (88%)	15/17 (88%)	15/17 (88%)	15/17 (88%)
Reading	18/28 (64%)	18/28 (64%)	17/28 (60%)	18/28 (64%)	18/28 (64%)
Total	33/45 (73%)	33/45 (73%)	32/45 (71%)	33/45 (73%)	33/45 (73%)

(EBS, 2015-2019)

Furthermore, the types and format of reading questions that were created by referring to the designated EBS prep books were quite fixed during the past five

years, with only small changes made across the years. When the CRA was introduced, this additional policy had some impact on the way and pattern the contents of EBS prep books were linked to those of the CSAT-English items. After the CRA implementation, reference items were no longer linked to the EBS materials, whereas the implication questions were newly added (the second year of implementation).

**Table 8 Proportion of the test-takers in each stanine over the past five years**

Stanine	Year				
	2015	2016	2017	2018	2019
	% (Cumulative %)				
1	4.62 (4.62)	4.42 (4.42)	10.03 (10.03)	5.3 (5.3)	7.43 (7.43)
2	7.68 (12.3)	6.87 (11.29)	19.65 (29.68)	14.3 (19.6)	16.25 (23.68)
3	11.59 (23.89)	12.31 (23.6)	25.43 (55.11)	18.5 (38.1)	21.88 (45.56)
4	17.27 (41.16)	18.02 (41.62)	17.97 (73.08)	20.9 (59.0)	18.48 (64.04)
5	19.36 (60.52)	18.68 (60.3)	10.49 (83.57)	16.5 (75.5)	12.27 (76.31)
6	17.42 (77.94)	17.09 (77.39)	6.74 (90.31)	10.7 (86.2)	9.21 (85.52)
7	11.51 (89.45)	11.98 (89.37)	4.74 (95.05)	7.4 (93.6)	7.37 (92.89)
8	6.76 (96.21)	6.69 (96.06)	3.49 (98.54)	4.6 (98.2)	5.24 (98.13)
9	3.79 (100.0)	3.94 (100.0)	1.46 (100.0)	1.7 (99.9)	1.87 (100.0)

(KICE, 2015-2019)

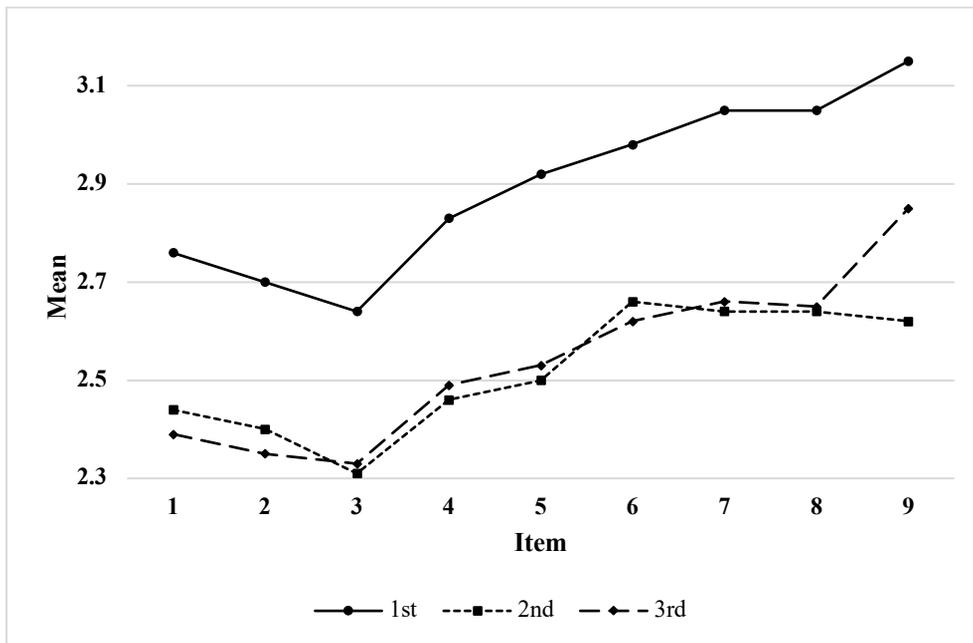
In respect to the CRA, it has changed the grading and score reporting system of the CSAT-English, and many stakeholders of the CSAT-English believed that it would become easier to get the desired scores or grades with the implementation of

the CRA since cutoff scores for each of the stanines became lower, compared to the ones that used to be used in the norm-referenced system. The overall summary of the percentage of the test-takers belonging to each stanine is presented in Table 8, and it is based on the KICE's annual report of the CSAT results from 2015 to 2019 (KICE, 2015, 2016, 2017, 2018, 2019). According to Table 8, the percentage of the test-takers who got the stanine of either 1, 2, or 3 has indeed increased sharply in 2017, marking 10.03%, 19.65%, and 25.43%, respectively. Before the introduction of the CRA, the cumulative percentage of obtaining the top two grades was generally between 11 to 12%, but in the first year of implementation, the percentage of stanine of 1 alone has reached the level close to that of the two stanines combined, and the cumulative percentage of the test-takers who received one of the top three grades in 2017 (55.11%) exceeded those of stanine 4 in 2015 (41.16%) and 2016 (41.62%). Also, in the 2018 and 2019 CSAT-English, the cumulative percentage of students with top three grades were 38.1% and 45.56%, respectively, which are still higher than original 23%.

On the other hand, the percentage of test-takers who received below the stanine score of 4 was significantly lower in 2017 to 2019 than 2015 and 2016, when the CRA was not yet implemented. A case in point is the percentage of the test-takers who got the lowest stanine from the test. Before CRA, the proportion of the test-takers receiving the stanine of 9 was between 3 and 4 percent, but after the introduction of the CRA, the rate reduced to less than 2 percent.

## 4.2 Descriptive Statistics

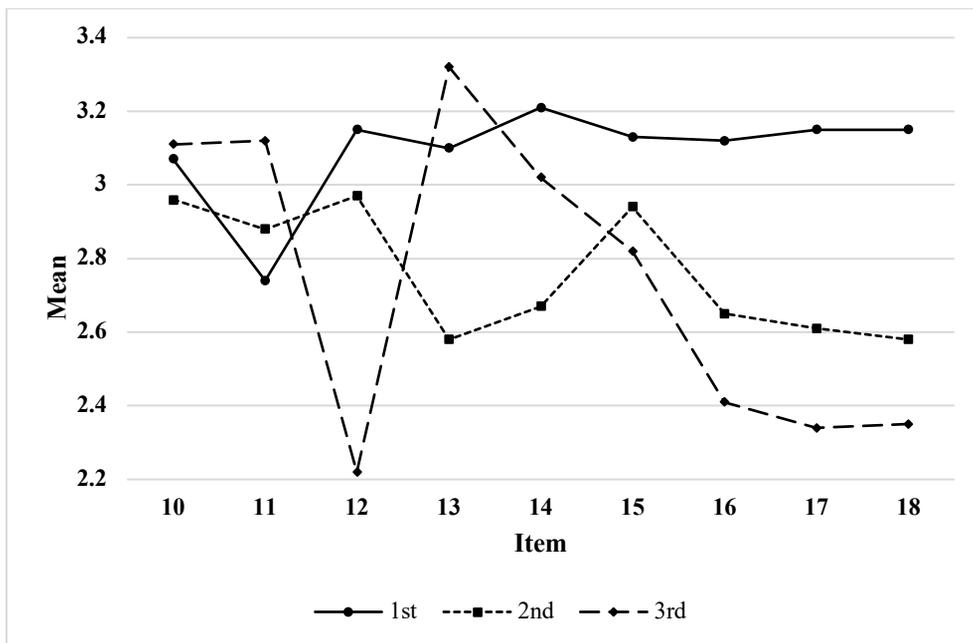
In this section, the statistical results of the five-point Likert scale items in the students' questionnaire are presented. More specifically, the means for each grade sample in each item are provided through several graphs. First, the trends of the means for Items 1 to 9 are displayed in Figure 3. As seen in the graph, the means for the first-grade student group marks the highest for all the items in the category of learning attitudes, while second- and third-year students alternates the second highest means through the items. Also, overall averages increase as the item number increases, especially in the first and third graders.



**Figure 3 Trends of the means for Items 1 through 9**

Second, the trends of means for Items 10 through 18 are graphically represented in Figure 4. Unlike the mean tendencies shown in Figure 3, it is much more complicated to detect noticeable commonalities between the students' grades

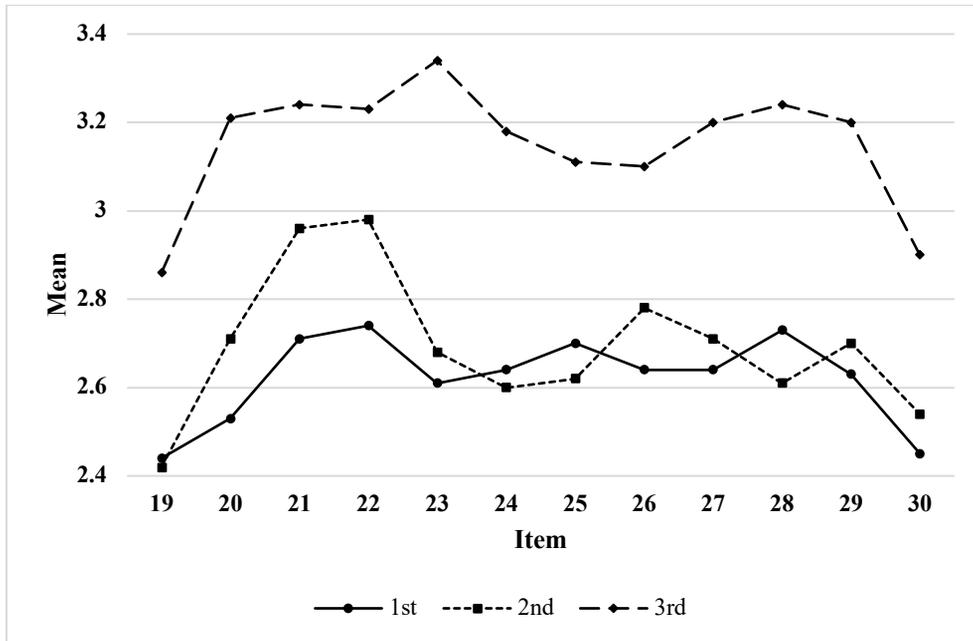
of the trends of the means for items that belong to the category of learning contents, and each grade indicates its own unique shape; First-graders show quite a constant trends of the means throughout the items, except for Item 11; Second-year students have their means fall across the board, except for a temporary rise in the averages for Items 14 and 15, and third-grade group also indicates declining trends of the means after a dramatic change in Item 12.



**Figure 4 Trends of the means for Items 10 through 18**

Lastly, according to Figure 5, all of the twelve items in the category of learning strategies show a pattern in which the means are the highest for the third-grade students. Also, the grade mean scores for the items in this figure have the patterns that are symmetric left and right around Item 24 or 25. In respect to the first-graders, for instance, the items that come before Item 25 indicate a pattern in which the means rise until Item 22, decrease between Item 22 and 23, and increase until

Item 25. The remaining items form a symmetry with the items before Item 25 by having the opposite trends of the means for Items 25 through 30. The results for the second- and third-grade students do not have as complete symmetry as the one found in the first graders but still show quite similar patterns between left and right items.



**Figure 5 Trends of the means for Items 19 through 30**

The results of the means in each item can be classified into several patterns in Table 9. First, all of the nine items in the category of learning attitudes show an overall pattern in which the first-grade student group attains the highest mean score. Despite the first grader group being on top across items, a close comparison of item means reveals two different sub-patterns in terms of the order of the other two grade groups. For six of the nine items, the second highest mean score is observed for the third-grade student group followed by the second grader group. For the remaining

three items (1, 2, 6), the order was reversed, and the mean score decreases as the grade level increases.

Second, for the category of learning contents, the largest number of items (5 out of 9) show a pattern where the means score is the highest for the first-grade group and the lowest for the third-grade students. More specifically, the mean score increases as the grade level increases in this pattern. Interestingly, the remaining four items belong to three different patterns: one for the  $1^{st} > 3^{rd} > 2^{nd}$  pattern, two for the  $3^{rd} > 1^{st} > 2^{nd}$  pattern, and one for  $3^{rd} > 1^{st} > 2^{nd}$  pattern. The learning contents is the only category where four different patterns are observed.

**Table 9 Comparison of the means among the different grade groups**

Category	Item	Comparison of Means
Learning attitudes	1, 2, 6	$1^{st} > 2^{nd} > 3^{rd}$
	3, 4, 5, 7, 8, 9	$1^{st} > 3^{rd} > 2^{nd}$
Learning contents	12, 15, 16, 17, 18	$1^{st} > 2^{nd} > 3^{rd}$
	14	$1^{st} > 3^{rd} > 2^{nd}$
	10, 13	$3^{rd} > 1^{st} > 2^{nd}$
	11	$3^{rd} > 2^{nd} > 1^{st}$
Learning strategies	19, 24, 25, 28	$3^{rd} > 1^{st} > 2^{nd}$
	20, 21, 22, 23, 26, 27, 29, 30	$3^{rd} > 2^{nd} > 1^{st}$

Lastly, most of the items (8 out of 12) in the category of learning strategies displayed a pattern where the average is the highest for the third-grade students and the lowest for the first-grade group. In this typical pattern for the third category, the mean score increases as the grade level increases. For the remaining four items, only

the order of the first- and second-grade groups was reversed, with the third-grade group being on top (a just slight variation from the dominant pattern for this category).

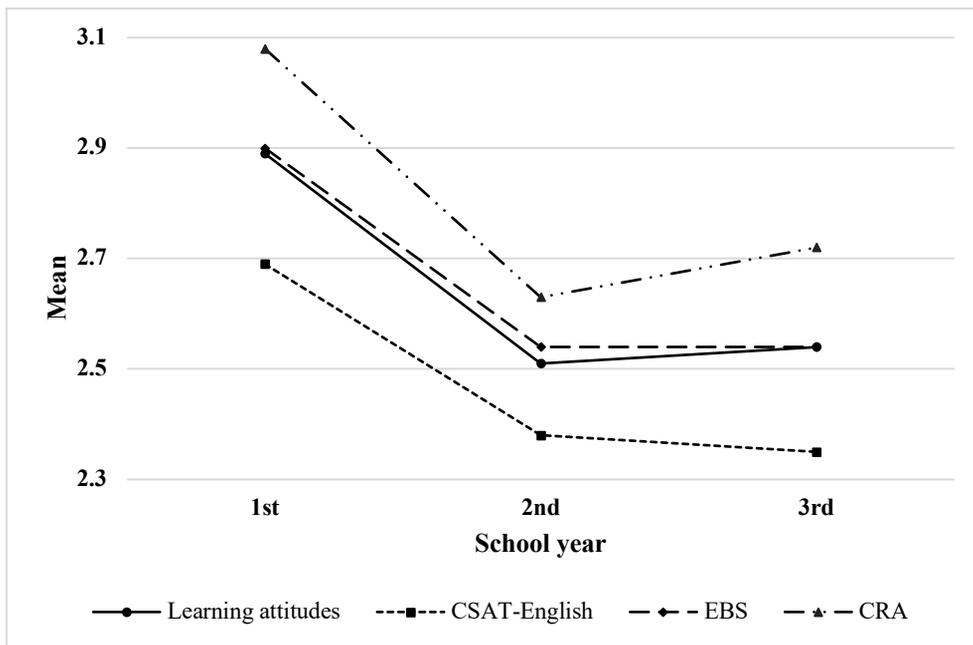
Table 10 presents the averaged grade means for the three main categories and three subcategories under each of the main categories. (Here, the averaged grade mean refers to the mean score averaged across all of items belonging to a particular category (or subcategory) for a particular grade.).

**Table 10 Descriptive statistics of the categories in the students' questionnaire**

Category	Grade		
	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>
	Mean		
Learning attitudes	2.89	2.51	2.54
CSAT-English	2.69	2.38	2.35
EBS	2.90	2.54	2.54
CRA	3.08	2.63	2.72
Learning contents	3.09	2.76	2.74
CSAT-English	2.98	2.94	2.81
EBS	3.14	2.72	3.05
CRA	3.14	2.61	2.36
Learning strategies	2.62	2.69	3.14
CSAT-English	2.60	2.76	3.13
EBS	2.64	2.67	3.16
CRA	2.61	2.63	3.13

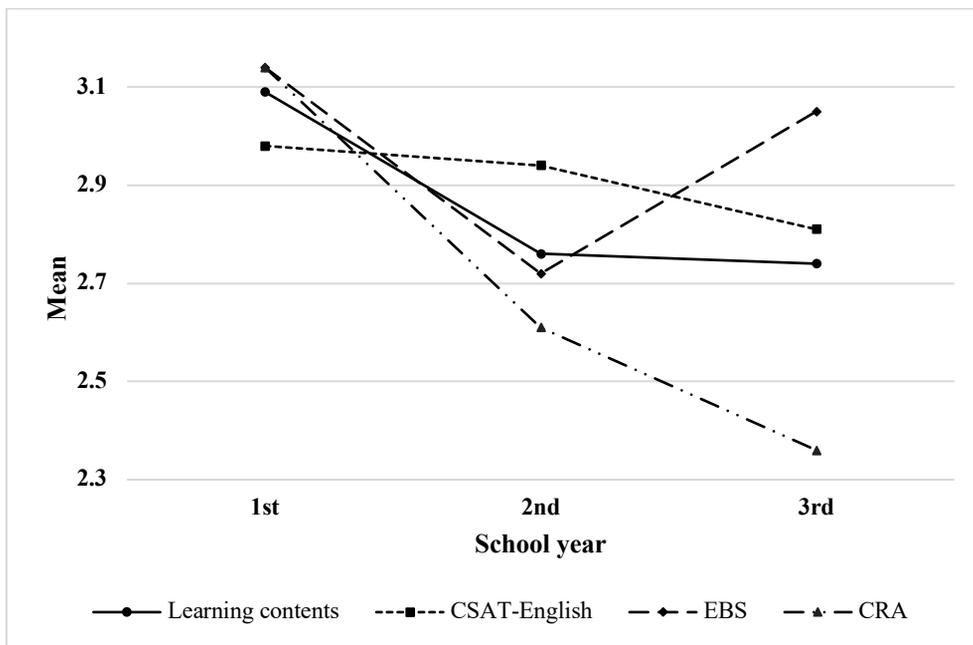
First, the averaged grade mean for the category of learning attitudes shows the 1<sup>st</sup> > 3<sup>rd</sup> > 2<sup>nd</sup> pattern in which the mean score is the highest for the first graders and the lowest for the second graders. However, the three subcategories under this category show three different patterns of 1<sup>st</sup> > 2<sup>nd</sup> > 3<sup>rd</sup>, 1<sup>st</sup> > 2<sup>nd</sup> = 3<sup>rd</sup>, and 1<sup>st</sup> > 3<sup>rd</sup> > 2<sup>nd</sup> patterns, in all of which the first-grade group had the highest mean score.

These patterns are also graphically represented in Figure 6. All these taken together, it can be said that the first-grade students in general had more positive attitudes than the second and third graders toward various features of the CSAT-English, some of which might have been either introduced by the testing policies or existed even before the introduction of such policies.



**Figure 6 Trends of the means for the learning attitudes and its subcategories**

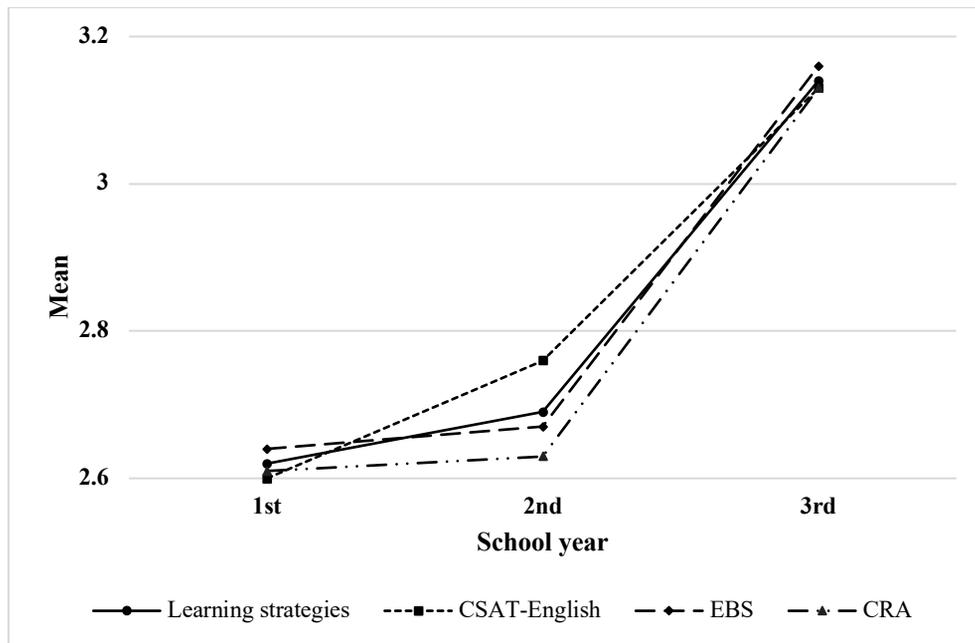
Second, the grade mean scores of the second category (learning contents) and its subcategories had very interesting patterns. As shown in Table 10 and Figure 7, the main category and two of its sub-categories (CSAT-English, CRA) had the  $1^{st} > 2^{nd} > 3^{rd}$  where the mean score decreases as the grade level increases. However, one subcategory (EBS linkage) under this category shows  $1^{st} > 3^{rd} > 2^{nd}$  pattern where the first graders remain to be the highest mean group while the second graders turns out to be the lowest mean group.



**Figure 7 Trends of the means for the learning contents and its subcategories**

Third, according to Table 10 and Figure 8, the grade mean scores for the third main category (learning strategies) and its three subcategories all exhibit the same pattern of  $3^{rd} > 2^{nd} > 1^{st}$ . In this pattern, the mean score is the highest for the third-

grade students and the lowest for the first-grade group. In other words, the mean score rises as the grade level increases.



**Figure 8 Trends of the means for the learning strategies and its subcategories**

### 4.3 One-way Analysis of Variance (ANOVA)

A one-way analysis of variance (ANOVA) was used to compare the mean scores among three different grade groups and to test the statistical significance of mean score differences among the grade groups, which were previously presented in Table 10.

Table 11 shows the F value and significance level (or p-value) for each of the categories. According to the table, the differences in the means were significant with the significance level of all three main categories and 8 out of 9 subcategories being less than .05, which is considered to be the cutoff rate for statistical significance.

Only the category of the CSAT-English in the learning contents did not indicate significant difference in the mean values of three different samples,  $F(2,231) = .916$ ,  $p = .401$ .

**Table 11 ANOVA omnibus test results of the means among different grade groups**

Category	F(2, 331)	Sig.
Learning attitudes	6.453	.002
CSAT-English	4.510	.012
EBS	5.180	.006
CRA	6.254	.002
Learning contents	5.742	.004
CSAT-English	.916	.401
EBS	6.391	.002
CRA	17.160	.000
Learning strategies	11.360	.000
CSAT-English	9.987	.000
EBS	10.412	.000
CRA	11.209	.000

*Note:* \* Significant at .05 level

Although most of the differences in the mean scores among the different grade groups were demonstrated to be statistically significant when the ANOVA was conducted on the data, the learning strategies and its subcategories were the only categories in which the mean scores increased in accordance with the students' grade.

On the other hand, the average values for learning attitudes, learning contents, and their subcategories did not become greater as the grade in high school went up and rather decreased in some or all grades.

Also, since ANOVA omnibus test is used only to appraise the overall statistical significance of the differences of the washback intensity between the different grade groups, an additional post-hoc test, a Scheffe test, was conducted to determine the exact points in which the statistical differences between the students' grades appear. More specifically, there are three possible pairs of grades, i.e. first and second graders, first and third graders, and second and third-grade students, in which statistical differences between the grades can occur, and quite different results were found in each of the main and sub-categories.

First, regarding the category of learning attitudes and its subcategories, all four main and sub-categories had the significance value of less than .05 in the relationships between first- and second-grade students and first and third graders, but not between second and third graders. The post-hoc test result indicates that the significance value that marked lower than .05 in the learning attitudes and its three subcategories is attributed to the statistical differences found between first and second graders and first- and third-grade students.

Second, as to the category of learning contents and its three subcategories, the main category had the significant differences in the results between first and second graders and first- and third-grade student groups, and the CRA was the only subcategory that showed the same post-hoc test result with its upper category of learning contents. The subcategory of the EBS had the significant differences in the

washback intensity of the CSAT-English that is characterized by the EBS linkage policy between first- and second-grade students and second and third graders, whereas the remaining category of the CSAT-English did not show statistical significance between any of the grades. Indeed, the CSAT-English was the only subcategory in the learning contents that had the significance value that is higher than .05 ( $p = .401$ ).

Lastly, with respect to the learning strategies as well as its subcategories, all four main and sub-categories indicated significant differences between first- and third-year students and second- and third graders in the washback intensity of the CSAT-English and the impact of the EBS and CRA policies. The mean difference between first- and second-grade student groups was not statistically significant when the post-hoc Scheffe test was conducted on the data.

#### **4.4 Students' Responses to the Open-Ended Questions**

A qualitative data analysis was done to examine the participants' responses to the open-ended questions in the students' questionnaire. In particular, the information that was not elicited in the five-point Likert scale items but was repeated frequently in the students' responses to the open-ended questions was considered especially important. Those sets of topics or themes frequently-appearing in the students' answers across different students were grouped into some meaningful response categories.

First, the participants' responses to the first open-ended question which asked them about their opinions or experience in relation to their perceptions and learning

attitudes were summarized into several categories in Table 12. Of a total of 334 units of participant responses, more than half of the participants (N = 172, 51.5%) said they had nothing to add or comment about the washback of the CSAT-English and the introduction of the EBS-CSAT linkage policy and CRA on their perceptions and learning attitudes. Fifty-six (16.7%) responses were either unclear in meaning as there was not enough information given (e.g. I think I should study harder.) or irrelevant to what the question asked (e.g. It is interesting).

**Table 12 Responses to Item 31 in the students' questionnaire**

Response	Frequency	
	N	%
Incomprehensible or irrelevant	56	16.7
Discriminating power of the CSAT-English	12	3.6
Difficulty of the CSAT-English	28	8.4
Importance of the CSAT-English	17	5.1
Academic and psychological burdens of preparing for the CSAT-English	42	12.6
Others	7	2.1
Nothing to add	172	51.5
Total	334	100

The answers from the remaining 106 participants were categorized into five groups. Among the responses, many students (N = 42, 12.6%) commented on academic and psychological burdens of studying the CSAT-English. In particular, a considerable number of people (N = 34, 10.2%) said that the introduction of the EBS-

CSAT linkage policy and CRA in the CSAT-English had relieved their anxiety and stress of preparing for the test. In contrast, there were people (N = 8, 2.4%) who insisted that the change in the test rather raised their academic and mental stress as the amount of study increased with the use of the EBS-CSAT prep books in the CSAT. Also, they mentioned that the criterion-referenced CSAT-English lost its discriminating power, and consequently students' concern about getting good grades in other subjects, such as the Korean language and Mathematics, increased.

Indeed, 12 (3.6%) students maintained that the discriminating power of the CSAT-English dropped significantly after it adopted the new grading system. Furthermore, 17 (5.1%) people commented on the importance of the CSAT-English, and most of them (N = 16, 4.8%) argued that they became negligent in studying for the criterion-referenced CSAT-English.

There was, however, a mixture of affirmative and negative opinions about the difficulty of the CSAT-English. Of the 28 participants (8.4%), 15 of them (4.5%) said that the current criterion-referenced test is harder than the ones that were administered under the norm-referenced system, while 13 students (3.9%) argued that the difficulty of the CSAT-English did not change compared to the past grading system and rather decreased.

Lastly, the statements that were grouped into 'others' category (N = 7, 2.1%) were about the increase in the students' confidence, motivation, and willingness to study English and the decrease in the expenditures on the private English education after the implementation of the EBS-CSAT linkage and CRA policies for the CSAT-English. Also, three participants (.9%) criticized the current Korean English

education system that does not provide a sufficient number of English-speaking classes or a sufficient amount of time resources for English speaking instruction and activities to students.

**Table 13 Responses to Item 32 in the students' questionnaire**

Response	Frequency	
	N	%
Incomprehensible or irrelevant	51	15.3
Use of the EBS-CSAT prep books and related materials	32	9.6
Use of the previous CSAT and PCSAT papers	2	0.6
Memorizing the words and sentences in the EBS-CSAT prep books	7	2.1
Others	7	2.1
Nothing to add	235	70.3
Total	334	100

Next, Table 13 shows the frequencies of the participants' responses to the second open-ended question, which was about the impact of the CSAT-English and its two major changes on the students' learning contents and learning strategies. As with the first open-ended question, a significant number of participants (N = 235, 70.3%) said they did not have any other opinions or comments to add. Also, there were 51 (15.3%) answers in total that were not related to what the question asked (e.g. competition) or were not very interpretable in terms of their exact meaning or intention (e.g. It got tough.). Forty-eight (14.4%) replies were left after these responses were excluded. However, unlike the previous question, most of the

significant replies written in response to the second item were similar to those asked in the five-point Likert scale items.

Of the 48 statements, most of the participants (N = 32, 9.6%) wrote about the use of the EBS-CSAT prep books and related materials, such as the EBS-modified practice books. Although EBS-related contents had already been mentioned in the five-point Likert scale items (Items 13, 14, 15), many participants reemphasized the importance of mentioning them in their responses to the open-ended question. Moreover, two students (.6%) said they were studying the sample items from the previous CSAT and PCSAT, which were similarly mentioned in their answers to Items 10 and 11.

In terms of the learning strategies, 7 (2.1%) students said they were memorizing the vocabularies and sentences in the EBS-CSAT prep books and thought that this method of rote learning was most effective in getting a desired grade or score on the CSAT-English. This kind of responses were also repetitions of what was covered by the students' responses to some of the Likert-type items (Items 23, 24).

The remaining answers (N = 7, 2.1%) in the "others" category contained the statements about the use of the high school English textbooks to prepare for the CSAT-English, the importance of analyzing the proven tips and skills for solving various types of questions, the effectiveness of one-to-one learning, and the increase in the time spent for studying English grammar.

## **4.5 Results of the Teachers' Questionnaire**

In addition to the students' responses to the open-ended questions in the students' questionnaire, the teachers' replies to the open-ended items in the teachers' questionnaire were analyzed by using a qualitative method of interpretation. Since the teacher survey was conducted to supplement the student survey, the analysis of teachers' data was based on the answers that were commonly raised among the respondents.

As mentioned in the previous chapter, the participants gave answers to the items from two major sections. Regarding the first section of questions, which asked about the washback of the CSAT-English and its related policies on their different teaching practices, many teachers said they felt a sense of loss as their students' interest in studying English had been decreasing over time due to the changes in the high-stakes test. Also, they also claimed that the credit hours in English courses had been reduced in the past few years since the introduction of the revised National Curriculum in 2015.

In terms of the teaching contents, most of the teachers used the EBS-CSAT prep books to prepare students for the CSAT-English. In particular, there were quite a few teachers who were using only the EBS textbooks to teach the CSAT-English. The result demonstrates that simplified and limited types of teaching materials are currently being used in English classes at public high schools in Korea.

As to the teaching methods, on the other hand, task-based, process-based teaching methods became popular in regular English classes after the implementation of the revised National Curriculum (2015) and the introduction of

the CRA in the CSAT-English. The change in the teaching method led to the reduction in the teacher-dominated talk and the increase in various types of activities in the classroom.

For the second section of questions, which dealt with the relationship between the teachers' different teaching practices and the students' grade, many teachers emphasized the change in their teaching contents and methods in accordance with their student's level of grade. They said that their attitudes toward preparing the classes and teaching the students in the classroom were not affected or changed by what grade the students are in. An overall summary of the responses to the questions is presented in Table 14.

**Table 14 Responses to Items 4 through 6 in the teachers' questionnaire**

Category	Grade	
	1 <sup>st</sup> and 2 <sup>nd</sup>	3 <sup>rd</sup>
Teaching attitudes	No difference between grades	
Teaching contents	High school regular textbooks	EBS materials
Teaching methods	Oral activities	Highly CSAT-oriented

Regarding the teaching contents or materials used, regular textbooks for high school were mostly used to teach first- and second-year students, whereas the EBS-CSAT prep books and other CSAT-English practice materials were used for third-grade students, who were about to take the CSAT in the near future.

The teaching methods were also differentiated between the third-year students and other lower grade students. While oral activities to improve students'

communication skills and practical English were often carried out in the classrooms, particularly with first- and second-grade students, the classes for the seniors were highly CSAT-oriented with the EBS-CSAT prep books being the main textbook.

# Chapter V

## Discussion

This chapter discusses the major findings of the current study in relation to the major research questions. More specifically, the following research questions were posed for this study: (1) how EBS-CSAT linkage policy and CRA influence the nature or characteristics of the CSAT-English, (2) how student learning is influenced not only by the initial characteristics of the CSAT-English but also by some other features shaped up and added due to the related policy changes and (3) whether students' school year also influence (or moderate) the washback effect of the CSAT-English, particularly the intensity of the effect. In addition, the teacher's influence on, and perception, of the washback impact were addressed. Intriguing results were obtained in relation to these research questions and issue. The chapter provides a brief summary of the major findings along with discussions on some key issues.

### **5.1 The Impact of the EBS-CSAT Linkage Policy and CRA on the Nature/characteristics of the CSAT-English**

Regarding the first research question, the study found that EBS-CSAT linkage policy and CRA have triggered many changes to the CSAT-English. First, the EBS policy requires that at least 70% of the test contents should be generated based on the EBS-CSAT prep books that are designated by the KICE and MOE. In particular, the types and percentage of the test questions that should be linked to the EBS materials have mostly been fixed, and there have been only small changes over the past five years.

More specifically, 15 out of 17 listening items and 17 or 18 out of 28 reading items were consistently created or prepared from the EBS materials, and the remaining 12 or 13 questions were generated without referring to the EBS textbooks. Given that the listening section is generally not difficult for students with an intermediate or high level of English proficiency, the difficulty and discriminating power of the test are mostly determined by the unfamiliar reading items that do not come from the EBS prep books. In particular, as fill-in-the-blank questions in the CSAT-English are regarded as one of the most difficult and complicated items among the test-takers, two or three fill-in-the-blank items that are made without referring to the EBS contents determine the difficulty and discriminating power of the test. Second, with the introduction of the CRA in 2017, it became easier for the test-takers to get higher grades or stanines in the CSAT-English. According to Table 8 in the previous chapter, the percentage of test-takers who got one of the top three grades increased steeply compared to the ones shown in the norm-referenced system, whereas the proportion of test-takers receiving stanines lower than 4 have decreased considerably. It is quite difficult to insist that the difficulty level of the test became easier based on these results, but it is possible to conclude that the discriminating power of the test became weaker as many students are receiving high grades more readily.

## **5.2 The Impact of the Nature/characteristics of the CSAT-English on Students' Learning**

In respect to the second research question, the study found that the CSAT-English, which had undergone crucial changes in its contents, grading methods, and score reporting system since the introduction and implementation of the new policies, had

a considerable impact on student learning practices, especially their learning attitudes, contents, and strategies. With the EBS linkage policy in place, for instance, 70% of the CSAT-English contents were created by referring to the EBS-CSAT prep books, and a student survey has confirmed that many students were preparing for the CSAT-English by using the EBS textbooks. The result matches with the works of Song (2015) and Park (2017), in which they investigated the number and percentage of students using the EBS materials to prepare for the upcoming CSAT-English. Both studies showed that third-grade students were actively studying English with the EBS prep books, and Park (2017), in particular, found that third-grade students were more frequently using the EBS contents than the lower grade students. In addition, the CRA changed the grading and score reporting system of the CSAT-English, and some individual students, whose academic burden or pressure decreased with the implementation of this new policy, spent more time improving their English communication skills rather than solving as many CSAT-English practice questions as possible. It matches with the results of Jang's (2017) and Jeong's (2018) studies, where they commonly found that students' competition for getting good scores and academic burden for studying English were reduced with the new grading system. The work of Lee (2018), on the other hand, produced a different result, with the influence of the CRA on the students' learning practices being very small that their learning contents and methods did not show significant differences compared to the previous norm-referenced system.

Despite the change in the students' learning attitudes and behaviors, however, it is still unclear whether these changes are solely due to the influence of the nature

or characteristics of the test that is affected by its policies. In other words, these changes in student learning can also be regarded as the direct influence of the national policies exerted on the learning, without the influence being mediated through the change in the test. The current study, however, did not intensively investigate the direct influence of the test policies on student learning since washback studies on learners, including the current one, generally focus on the relationship between testing and learning, and thereby requires the inclusion of the influence of the test on learning. Although washback studies on the learners have their main concerns on the influence of testing on learning, consideration of the impact of the test policies on the learners and their learning will enrich our understanding of the detailed mechanism that causes change in the students' learning practices.

### **5.3 The Impact of the Students' School Year on the Washback Effect of the CSAT-English**

As to the last research question, each of the learning practices that was investigated in this study indicated different results. First, regarding the learning attitudes, the degree of the washback, which was numerically displayed via mean values showed a significant difference ( $p = .002$ ) among the students in different grades. The average, however, did not grow higher as the students advance to the upper grade, with the mean of the first-grade students being the highest. This tendency of the learning attitudes is explained by the patterns of its subcategories and items' mean values. In particular, the trend of the main category appeared to follow the most frequent pattern among its items. Of the nine items included in the learning attitudes,

a total of six items indicated the same trend with that of the main category, while other three items had the highest average in the first-grade group and the lowest in the third-grade students. Since more than 60% of the items showed the highest mean value for the first-grade sample and the lowest rate for the second-grade students, it may be natural to see the superordinate (main) category also show the same pattern of change across the grade groups in the average score.

The result of the current study matches with that of the work of Hwang (2003), with the student participants in both studies having negative attitudes towards the CSAT-English. However, while Hwang included third-grade students only, the current took it one step further by investigating not only the attitudes of the third-grade students but also the ones of the first- and second-grade students and comparing among these three different grade groups. Consequently, the result showed that negative attitudes towards the CSAT-English were dominant in all three grade groups, with the means being less than 3.0.

One point to note from the finding is that, although approximately 16 years has passed since Hwang's work was published in 2003, and although several major policies have been introduced to promote positive washback of the test during that period, students' learning attitudes did not change and still remained negative. Indeed, except for the first-grade students' mean value in the CRA category, all other mean scores (ratings) in every grade marked lower than 3.0, denying the positive effect of the CSAT-English on their learning attitudes.

Second, as to the learning contents, the washback of the CSAT-English and its major changes on the students' learning contents showed significant level of

difference ( $p = .012$ ) among the three grade groups of students. Also, the average consistently decreased as the grade increased from year one to year three. As with the first issue, the trend for the learning contents is also due to the pattern that is most frequently seen among its items. Five out of nine items had the highest average for the first-grade group and the lowest for the third-grade sample, indicating the same inverse relationship between the students' grade and the washback intensity of the CSAT-English as the main category. Also, of the three subcategories in the learning contents, the CSAT-English and the CRA presented the same trends as their upper category, whereas the EBS category had the highest average for the first-grade students and the second highest for the third-grade students.

The results of the learning contents can also be discussed in relation to the previous studies. Among the three items (Items 13, 14, 15) in the EBS category, for instance, Item 13 asked whether the students' time spent for studying the EBS-CSAT prep books have increased with the introduction of the EBS-CSAT linkage policy, and the result showed that the mean was the highest for the third-grade group and the lowest for the second-grade students. The result is quite similar to the one obtained in Park's (2017) study, where the EBS materials were more actively used in the third-grade classes than the first- and second-grade ones. In particular, most of the third-grade students (82.4%) in her study were using the EBS-CSAT prep books at the time the survey was conducted, while only 27.5% of the second-grade students were studying EBS textbooks to prepare for the CSAT, which was the lowest among the three different grade groups. The results from both studies showed that third-grade students who were about to take the CSAT in the very near future

tended to focus on the learning contents or materials that were directly related to the test, which were the EBS-CSAT prep books. Unlike the EBS materials being used most actively in the third-grade groups, however, the results negate the positive relationship between the students' grade and EBS use since first-grade students indicated higher levels of effect than the second-grade students in both studies. Further investigation is required to explain the phenomenon where second-grade students use the EBS materials the least.

Furthermore, while EBS-CSAT prep books were most frequently used among the third-grade students, the result for Item 15 indicated a complete opposite tendency from that of Item 13. That is, the time spent for taking the EBS-CSAT courses has marked the lowest for the third, meaning an entirely inverse direction from the students' grade. Following the results, the study has found that third-grade students generally prepare for the CSAT-English by studying the EBS textbooks but are not likely to take courses that cover EBS contents.

An important point to note from the results of the two categories is that quite many items indicated higher means in the first- and second-grade groups than the third-grade students. The results show that the washback effect of a particular test, whether it is positive or negative, does not always become stronger as the students' school year or grade increases. The national policies established newly for the test can have numerous impact on the washback effect of the test and can lead to various changes in the students' learning practices, but the intensity of the effect can be different for different grade groups of students. As mentioned above, for instance, the EBS-CSAT linkage policy in the CSAT-English caused students to prepare for

the test through the EBS materials, but the results from the related items showed that the EBS-CSAT prep books were most frequently used among the third-grade students and least frequently in the first-grade group, while the EBS-CSAT courses were less frequently taken as the students' school year went up. The current study attempted to describe this specific phenomenon through the students' behavior to reduce their learning contents or materials and focus on the materials that were highly essential for preparing for the test as the number of days left for the upcoming CSAT decreased. However, further research is required to reveal a whole array of relevant factors that can contribute to possible differences in the intensity of the impact, in addition to the students' school year or grade.

While the categories in the learning attitudes and contents showed partially or entirely negative correlations between the students' grade and the washback effect of the test, the statistical results for the last issue, which was about the students' learning strategies, were quite opposite from the prior ones. First, the comparison of the mean values through a one-way ANOVA showed that there is significant difference ( $p = .000$ ) among the group of students in different grades. Also, the intensity of the washback became consistently stronger as the students' grade went up. The average value of the third-grade students was the highest among the grade groups, whereas mean of the first-grade students marked the lowest. All the items and subcategories in the learning strategies also followed the same trends of the mean scores across the groups as their main category, indicating the positive relationship between the students' grade and the intensity of the washback of the test.

The results of the CSAT-English and EBS categories showed that students' learning strategies became more test-driven as their school year went up, and these coincide with the ones represented in O's (2001) research, where she found that students' learning practices were more CSAT-oriented in the third-grade classes than the second-grade ones. In addition to reconfirming the results found in O's research, the current study also enriched and deepened our understanding of the tendencies indicated in the students' learning strategies by expanding the scope of participants to the first-grade students.

The CRA category also showed higher mean values for higher students' grades, but the results ought to be interpreted differently. That is, as for the items in the CRA subcategory concerning whether students have reduced their time spent for solving CSAT-English practice questions, memorizing, and studying difficult CSAT-English passages, it was found that third-grade students indicated the highest mean, meaning that time spent for those behaviors have decreased. Given the results in the CSAT-English and EBS categories, however, it is difficult to conclude that students' tendencies to be test-driven have weakened with the introduction of the CRA, but it is possible to believe that the negative washback of the CSAT-English that have been criticized among the stakeholders for a long time have been reduced and students seemed to change their way of learning with the new grading system, especially the higher grade students. The result of the current study is quite different from the work of Lee (2018), who found that the CRA did not have positive effects on students' learning methods, and since it has been only three years since the CRA

was first introduced, additional research are required to determine the effectiveness of the new system precisely.

In short, all three main categories in the students' questionnaire revealed significant differences in the magnitude of the washback between the grades. Above all, the second category of "learning strategies" and its subcategories were the only sections where the degree of the washback became stronger as the student's grade went up. Nevertheless, the learning attitudes, contents, and their subcategories showed very different pattern, which were partially or entirely inverse relationships between the students' grade and intensity of the washback.

#### **5.4 Teachers' Influence on the Washback Effect of the CSAT-English**

The impact of the students' grade on the washback effect of the CSAT-English that is characterized by its major polices can also be discussed in relation to the results of the teacher survey, especially the second section of the questions, which asked about the relationship between the students' grade and teachers' attitudes and behaviors. The results for the second section of the teachers' questionnaire showed that there was no difference in the teachers' attitudes toward teaching students with different grades, and most of the noticeable changes occurred in their teaching contents and methods.

First, regarding the teaching contents, most of the teachers were using high school English textbooks in the first- and second-grade students, whereas the EBS-CSAT prep books and other CSAT-English related materials were used as the main

contents in the classrooms with third-grade students. Much of this teachers' behavior of using different types of textbooks and materials depending on their students' grade seems to have been reflected in the students' learning contents. For instance, the CSAT-English in general (Items 10, 11, 12) in the learning contents category indicated a downward tendency with the mean values of the first-grade students being the highest and the third-grade students being the lowest. Of the three items in this category, Item 12, which asked about the use of the high school English textbooks to prepare for the CSAT-English played a crucial role in shaping the trend of the category. That is, the tendency of the category changed completely with the third-grade group indicating the greatest average and the first-grade sample having the lowest rate when the means were calculated without Item 12, whereas the trends remained the same as the original one when each of the other two items were deleted.

Here, the trend of Item 12 that marked the highest mean value in the first-grade students and the lowest for the third-grade students can be seen to be substantially influenced by the teachers' behavior of using different teaching materials in accordance with the students' grade. That is, the teachers' tendencies to use the regular English textbooks in the classes with first- and second-grade students and other textbooks and materials in the courses for third-grade students have affected the contents and materials that students use to prepare for the CSAT-English, and the effects are strong enough to determine the tendency of the category.

Second, as to the teaching methods, numerous kinds of activities were carried out in the lower grade classrooms to improve their English communication skills. However, as the grade went up, the English courses became highly CSAT-oriented

with the teacher dominating most of the in-class conversations. The class was more routinized than the ones in the lower grades, and the main purpose of the class was to improve the students' CSAT-English scores.

The difference in the teachers' teaching methods in accordance with the students' grade can also be considered to have affected the students' learning strategies. That is, the mean values of the CSAT-English in general and EBS-CSAT linkage policy in the learning strategies being the highest in the third-grade students and lowest in the first-grade students can be attributed to the change in the teachers' teaching method which becomes more CSAT-driven as the students' grade goes up. On the other hand, it seems difficult to relate the result of the CRA to the teachers' teaching practices, although it had the highest mean for the third-grade students and the lowest for the first-grade students, because the results indicated the reduction of the negative washback of the CSAT-English on the students, especially the higher-grade students by decreasing the time spent for solving practice questions, memorizing passages, and studying difficult English passages.

Furthermore, the change in the teaching methods depending on the students' level of grade can also describe the trend of the CRA in the learning contents. The CRA in the learning contents had a downward tendency with the mean value being the greatest for the first-grade group and the lowest for the third-grade sample. The items in this category (Items 16, 17, 18) asked whether the time spent for learning practical English has increased with the introduction of the CRA in the CSAT-English. The answers from the first-grade students were the most positive, and as the grade increased, the number of negative responses also increased.

This result can also be seen as the influence of the varying style of teaching in accordance with the grade level of the students. Since numerous types of activities that aim to improve the students' practical English are often used in the lower grade students' classes, these students are more likely to encounter materials or sources that are appropriate for learning practical English used in daily life. On the other hand, as the English classes for the seniors are much more test-oriented, they would not have time to learn practical English in the regular courses.

Taken together, the differences in the degree of the washback between the different grade groups that have been revealed in the student survey can be considered to be the consequences of the teachers' behavior that varies with the students' grade. The students' learning contents and learning strategies, in particular, were highly affected by the teaching materials used and the way teachers teach students. These results can be regarded as a representative example to the washback model of students' learning (Shih, 2007), in which he proposed that teachers, who are one of the extrinsic factors, can have significant influence on students' learning practices.

# Chapter VI

## Conclusion

### 6.1 Summary and Implications

The primary goals of this study were to investigate whether the initial and recently added characteristics of the CSAT-English have an impact on students' learning practices and to examine whether the washback intensity of the CSAT-English on student learning is influenced (or moderated) by the students' grade. To put it another way, the core issue was to see whether the CSAT-English has differential impact on different grade groups of students.

The results showed that all three learning practices that were investigated in this study had significant differences in the washback intensity among the different grade groups. However, the learning strategies were the only main category whose degree of the washback became stronger as the students' grade went up. The washback intensity on the learning attitudes and learning contents became partially or entirely weaker with the grade increase.

These relationships between the students' grade and washback intensity on different learning practices were discussed in relation to the teachers' varying behaviors by students' grade. Some parts of the students' learning contents and learning strategies were affected by the teaching contents and methods that teachers employ.

One major implication that is derived from the findings is the need for comprehensive study on various learning practices. Previous studies that examined the washback effect of either the EBS-CSAT linkage policy or CRA in the CSAT-English generally focused on the students' perceptions of the policy, and many studies concluded that the impact of the changes in the test became stronger as the school year progressed. However, the current study found that some of the learning practices, such as the learning attitudes and learning contents, showed a partially or entirely inverse relationship between the students' grade and intensity of the washback. The result of the study indicates that the nature of the washback on the learners could be wrongly defined if one's research focuses on a single aspect of learning. Hence, the washback studies on the learners ought to consider various kinds of learning practices for the precise understanding of the relationship between the learners and the influence of a test on their learning.

Another implication of the study is the importance of investigating the stakeholders or factors that are related to the learners when conducting the washback study on the learners. The current study, which aimed to examine the relationship between the students' grade and the washback intensity of the CSAT-English and its major policies, included a supplementary research on the teachers by measuring the washback of the CSAT-English on the teachers and the relationship between the students' grade and teachers' different teaching practices. Here, the study found that many of the students' learning practices could be highly affected by what contents the teachers use and how they teach them. Likewise, the washback on the learners and their learning is not just explained by the characteristics of the learners

themselves but is highly related to the other stakeholders who are in close contact with them, such as teachers. According to a washback model of students' learning (Shih, 2007), the washback study on the learner is very much complicated that there are a number of other factors to consider. The current study has again confirmed that research on the participants that have significant influence on the students and their learning is as much important as investigating the characteristics and features of the students.

## **6.2 Limitations and Suggestions**

Despite the major findings and implication in this study, there are some limitations in terms of the methodology and analysis of data collected for the study. First of all, the student questionnaire did not clearly distinguish among the washback effect of the CSAT-English in general, the EBS-CSAT linkage policy, and the washback of the CRA. For instance, students may have been confused when responding to the items that asked about their learning attitudes. In the learning attitude section, the students had to choose whether their motivation for, and confidence and interest in, learning English have increased in the process of preparing for the upcoming CSAT-English and with the introduction of the EBS-CSAT linkage policy and CRA in the CSAT-English. However, it could have been difficult for the participants to determine whether it was the EBS-CSAT linkage policy or the CRA that led to the change in their learning attitudes. Also, since the current CSAT-English is a test in which the EBS-CSAT linkage policy and CRA are already applied, the first three items which asked about the change in the students' learning attitudes during the

preparation of the CSAT-English can be seen to include all the remaining items in the learning attitude section. Hence, the questions may have seemed to be overlapping to some participants. The study could have produced more accurate and reliable results if the items among the categories had conceptualized, stated, and separated in a more clear-cut way.

Second, the study did not survey the teachers who were teaching the student participants at the time when the student data were collected. That is, 11 English teachers that participated in the teacher survey were all recruited from the same high school in Gyeongsang province, and none of their students participated in the study. The result of the study showed that teachers had considerable impact on the students' learning contents and learning strategies, but the mismatch of the school between the teachers and students could undermine the reliability of the result. The study would have been more accurate and reliable if the school between the teachers and students had been identical.

Lastly, for a practical reason, the study could not include the students who took classes in high schools before the EBS-CSAT linkage policy, and all the student participants that took part in the current study had never experienced the high school classes before the policy was implemented. The EBS-CSAT linkage policy has been active for 10 years, and in the year when it was first introduced in the CSAT-English, the current high school students were attending either primary schools or kindergarten. The inclusion of the students that can compare the educational setting before and after the EBS-CSAT linkage policy would have enriched the

understanding of the washback of the policy, but the same limitation is likely to be found in the future studies.

In conclusion, the future studies will help understand the nature of the washback on the learners if researchers review the limitations of the current study carefully and find ways to overcome them in their research design. In particular, in the current situation where multiple policies are applied simultaneously to a single test, it is essential to fully analyze the influence generated from each of these policies and the impacts caused by the interactions between them. Taking the limitations and suggestion of this study into consideration, the washback studies will have promising prospects of further research, leading to a more in-depth understanding of the relationship between testing, teaching, and learning.

## References

- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing*, 13(3), 280-297.
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115-129.
- Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, 13(3), 257-279.
- Bracey, G. W. (1987). Measurement-driven instruction: Catchy phrase, dangerous practice. *Phi Delta Kappa*, 68, 683-686.
- Cheng, L. (1997). How does washback influence teaching? Implications for Hong Kong. *Language and Education*, 11(1), 38-54.
- Cheng, L. (1999). Changing assessment: Washback on teacher perceptions and actions. *Teacher and Teacher Education*, 15(3), 253-271.
- Cheng, L., Andrews, S., & Yu, Y. (2011). Impact and consequences of school-based assessment (SBA): Students' and parents' views of SBA in Hong Kong. *Language Testing*, 28(2), 221-249.
- Cheng, L., Sun, Y., & Ma, J. (2015). Review of washback research literature within Kane's argument-based validation framework. *Language Teaching*, 48(4), 436-470.

- EBS. (2015). *2016hag-nyeon-do dae-hag-su-hag-neung-lyeog-si-heom yeon-gye-nae-yeog-bun-seog (A report of the EBS linkage to the 2015 CSAT)*. Retrieved from <https://www.ebsi.co.kr/ebs/ent/enta/retrieveEntAnlyStrdDataVw.ebs>
- EBS. (2016). *2017hag-nyeon-do dae-hag-su-hag-neung-lyeog-si-heom yeon-gye-nae-yeog-bun-seog (A report of the EBS linkage to the 2016 CSAT)*. Retrieved from <https://www.ebsi.co.kr/ebs/ent/enta/retrieveEntAnlyStrdDataVw.ebs>
- EBS. (2017). *2018hag-nyeon-do dae-hag-su-hag-neung-lyeog-si-heom yeon-gye-nae-yeog-bun-seog (A report of the EBS linkage to the 2017 CSAT)*. Retrieved from <https://www.ebsi.co.kr/ebs/ent/enta/retrieveEntAnlyStrdDataVw.ebs>
- EBS. (2018). *2019hag-nyeon-do dae-hag-su-hag-neung-lyeog-si-heom yeon-gye-nae-yeog-bun-seog (A report of the EBS linkage to the 2018 CSAT)*. Retrieved from <https://www.ebsi.co.kr/ebs/ent/enta/retrieveEntAnlyStrdDataVw.ebs>
- EBS. (2019). *2020hag-nyeon-do dae-hag-su-hag-neung-lyeog-si-heom yeon-gye-nae-yeog-bun-seog (A report of the EBS linkage to the 2019 CSAT)*. Retrieved from <https://www.ebsi.co.kr/ebs/ent/enta/retrieveEntAnlyStrdDataVw.ebs>
- EF EPI. (2018). *EF EPI* (8<sup>th</sup> ed.). EF Education First.
- Green, A. (2006). Washback to the learner: Learner and teacher perspectives on IELTS preparation course expectations and outcomes. *Assessing Writing*, 11(2), 113-134.

- Hwang, H. J. (2003). *The impact of high-stakes exams on teachers and students: A washback study of the university entrance exam at the secondary school level in South Korea* (Unpublished master's thesis). McGill University, Canada.
- Ingram, D. E. (1986). *Language and the export economy*. Paper presented to the 6th Biennial National Languages Conference of the Australian Federation of Modern Language Teachers Association, Adelaide, September.
- Jang, H. S. (2017). *A study on the perception of English education due to the change of English CSAT to the criterion-referenced system* (Unpublished master's thesis). Korea National University of Education, South Korea.
- Jeong, H. N. (2018). *A study of high school students' and parents' perception on criterion-referenced evaluation of the CSAT in the English section* (Unpublished master's thesis). Sogang University, South Korea.
- KICE. (2015). *2016hag-nyeon-do dae-hag-su-hag-neung-lyeog-si-heom chae-jeom-gyeol-gwa bo-do-ja-lyo (A report of the 2015 CSAT scoring results)*. Retrieved from <http://www.suneung.re.kr/boardCnts/view.do?boardID=1500230&boardSeq=5003182&lev=0&m=0302&searchType=S&statusYN=W&page=1&s=suneung>
- KICE. (2016). *2017hag-nyeon-do dae-hag-su-hag-neung-lyeog-si-heom chae-jeom-gyeol-gwa bo-do-ja-lyo (A report of the 2016 CSAT scoring results)*. Retrieved from <http://www.suneung.re.kr/boardCnts/view.do?boardID=1500230&boardSeq>

=5011367&lev=0&m=0302&searchType=S&statusYN=W&page=1&s=sun  
eung

KICE. (2017). *2018hag-nyeon-do dae-hag-su-hag-neung-lyeog-si-heom chae-jeom-  
gyeol-gwa bo-do-ja-lyo (A report of the 2017 CSAT scoring results)*. Retrieved  
from

[http://www.suneung.re.kr/boardCnts/view.do?boardID=1500230&boardSeq  
=5021737&lev=0&m=0302&searchType=S&statusYN=W&page=1&s=sun](http://www.suneung.re.kr/boardCnts/view.do?boardID=1500230&boardSeq=5021737&lev=0&m=0302&searchType=S&statusYN=W&page=1&s=sun)  
eung

KICE. (2018). *2019hag-nyeon-do dae-hag-su-hag-neung-lyeog-si-heom chae-jeom-  
gyeol-gwa bo-do-ja-lyo (A report of the 2018 CSAT scoring results)*. Retrieved  
from

[http://www.suneung.re.kr/boardCnts/view.do?boardID=1500230&boardSeq  
=5033661&lev=0&m=0302&searchType=S&statusYN=W&page=1&s=sun](http://www.suneung.re.kr/boardCnts/view.do?boardID=1500230&boardSeq=5033661&lev=0&m=0302&searchType=S&statusYN=W&page=1&s=sun)  
eung

KICE. (2019). *2020hag-nyeon-do dae-hag-su-hag-neung-lyeog-si-heom chae-jeom-  
gyeol-gwa bo-do-ja-lyo (A report of the 2019 CSAT scoring results)*. Retrieved  
from

[http://www.suneung.re.kr/boardCnts/view.do?boardID=1500230&boardSeq  
=5048274&lev=0&m=0302&searchType=null&statusYN=W&page=1&s=s](http://www.suneung.re.kr/boardCnts/view.do?boardID=1500230&boardSeq=5048274&lev=0&m=0302&searchType=null&statusYN=W&page=1&s=s)  
uneung

Kwon, O. R. (2015). A history of policies regarding the English section of Korea's  
college scholastic ability test. *English Teaching*, 70(5), 3-34.

- Kwon, S. K., Lee, M. B., & Shin, D. K. (2017). Educational assessment in the Republic of Korea: Lights and shadows of high-stake exam-based education system. *Assessment in Education: Principles, Policy, and Practice*, 24(1), 60-77.
- Lee, Y. J. (2018). *Test-takers' perception on the criterion-referenced evaluation of the CSAT English section* (Unpublished master's thesis). Kyunghee University, South Korea.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (pp. 13-103). New York: Macmillan.
- Messick, S. (1995). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice*, 14(4), 5-8.
- MOE. (2015). *2015 nyeon cho Jeong go yeong-eo-gyo-yug hwan-gyeong bun-seog mich yeong-eo mal-ha-gi gyo-yug gang-hwa bang-an (2015 report of the primary and secondary schools' English education settings and the way to enhance English-speaking classes)*. Retrieved from [http://www.prism.go.kr/homepage/entire/retrieveEntireDetail.do;jsessionid=578D662BC9A841BA4EF2F130E7787929.node02?cond\\_research\\_name=&cond\\_research\\_start\\_date=&cond\\_research\\_end\\_date=&research\\_id=1342000-201600026&pageIndex=538&leftMenuLevel=160](http://www.prism.go.kr/homepage/entire/retrieveEntireDetail.do;jsessionid=578D662BC9A841BA4EF2F130E7787929.node02?cond_research_name=&cond_research_start_date=&cond_research_end_date=&research_id=1342000-201600026&pageIndex=538&leftMenuLevel=160)

- O, K. M. (2001). *Washback on high school classrooms of the English tests within Korean university entrance exams* (Unpublished doctoral dissertation). Columbia University, USA.
- Park, J. K. (2009). 'English fever' in South Korea: Its history and symptoms. *English today*, 25(1), 50-57.
- Park, J. E. (2017). *High school students' and teachers' actual application and perceptions of EBS English reference materials for the preparation of the college scholastic ability test* (Unpublished master's thesis). Hankuk University of Foreign Studies, South Korea.
- Popham, W. J. (1987). The merits of measurement-driven instruction. *Phi Delta Kappa*, 68, 679-682.
- Queirós, A., Faria, D., & Almeida, F. (2017). Strengths and limitations qualitative and quantitative research methods. *European Journal of Education Studies*, 3(9), 369-387.
- Shepard, L. A. (1991). Psychometricians' beliefs about learning. *Educational Researcher*, 20(6), 2-16.
- Shih, C. M. (2007). A new washback model of students' learning. *Canadian Modern Language Review*, 64(1), 135-161.
- Shin, S. K. (2018). The characteristics of absolute grading of the College Scholastic Ability Test English section. *English Teaching*, 73(3), 95-113.

- Sohn, S. Y., & Ju, Y. H. (2010). Conjoint analysis for recruiting high quality students for college education. *Expert Systems with Applications*, 37(5), 3777-3783.
- Song, J. J. (2011). English as an official language in South Korea: Global English or social malady? *Language Problems and Language Planning*, 35(1), 35-55.
- Song, J. H. (2015). *A study on the students' and teachers' perceptions of EBS English materials for CSAT* (Unpublished master's thesis). Inha University, South Korea.
- Statistics Korea. (2011). *Private education expenditures survey 2010*. Retrieved from <http://kostat.go.kr/portal/eng/pressReleases/11/2/index.board?bmode=read&bSeq=&aSeq=273260&pageNo=1&rowNum=10&navCount=10&currPg=&searchInfo=&sTarget=title&sTxt=>
- Statistics Korea. (2012). *Private education expenditures survey 2011*. Retrieved from <http://kostat.go.kr/portal/eng/pressReleases/11/2/index.board?bmode=read&bSeq=&aSeq=273259&pageNo=1&rowNum=10&navCount=10&currPg=&searchInfo=&sTarget=title&sTxt=>
- Statistics Korea. (2013). *Private education expenditures survey 2012*. Retrieved from <http://kostat.go.kr/portal/eng/pressReleases/11/2/index.board?bmode=read&bSeq=&aSeq=286462&pageNo=1&rowNum=10&navCount=10&currPg=&searchInfo=&sTarget=title&sTxt=>

Statistics Korea. (2014). *Private education expenditures survey 2013*. Retrieved from  
<http://kostat.go.kr/portal/eng/pressReleases/11/2/index.board?bmode=read&bSeq=&aSeq=312788&pageNo=1&rowNum=10&navCount=10&currPg=&searchInfo=&sTarget=title&sTxt=>

Statistics Korea. (2015). *Private education expenditures survey 2014*. Retrieved from  
<http://kostat.go.kr/portal/eng/pressReleases/11/2/index.board?bmode=read&bSeq=&aSeq=334971&pageNo=1&rowNum=10&navCount=10&currPg=&searchInfo=&sTarget=title&sTxt=>

Statistics Korea. (2016). *Private education expenditures survey 2015*. Retrieved from  
<http://kostat.go.kr/portal/eng/pressReleases/11/2/index.board?bmode=read&bSeq=&aSeq=352520&pageNo=1&rowNum=10&navCount=10&currPg=&searchInfo=&sTarget=title&sTxt=>

Statistics Korea. (2017). *Private education expenditures survey 2016*. Retrieved from  
<http://kostat.go.kr/portal/eng/pressReleases/11/2/index.board?bmode=read&bSeq=&aSeq=359882&pageNo=1&rowNum=10&navCount=10&currPg=&searchInfo=&sTarget=title&sTxt=>

Statistics Korea. (2018). *Private education expenditures survey 2017*. Retrieved from <http://kostat.go.kr/portal/eng/pressReleases/11/2/index.board?bmode=read&bSeq=&aSeq=367176&pageNo=1&rowNum=10&navCount=10&currPg=&searchInfo=&sTarget=title&sTxt=>

Statistics Korea. (2019). *Private education expenditures survey 2018*. Retrieved from <http://kostat.go.kr/portal/eng/pressReleases/11/2/index.board?bmode=read&bSeq=&aSeq=374184&pageNo=1&rowNum=10&navCount=10&currPg=&searchInfo=&sTarget=title&sTxt=>

Wall, D. (1997). Impact and washback in language testing. In C. Clapham & D. Corson (Eds.), *Encyclopedia of language and education* (pp. 291–302). Dordrecht: Kluwer Academic.

Watanabe. Y. (2004). Teacher factors mediating washback. In L. Cheng & Y. Watanabe with A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 129–146). New York: Lawrence Erlbaum Associates, Inc.

World Bank, World Development Indicators. (2019a). *Gross domestic product 2018*. Retrieved from <https://databank.worldbank.org/data/download/GDP.pdf>

World Bank, World Development Indicators. (2019b). *Trade (% of GDP)*. Retrieved from <https://data.worldbank.org/indicator/NE.TRD.GNFS.ZS>

# Appendices

## Appendix A: Students' Questionnaire

### 1) Students' Background

1. 귀하의 성별은 어떻게 되십니까?

(What is your gender?)

- ① 남자 ② 여자

2. 귀하는 현재 몇 학년에 재학 중입니까?

(What is your grade in high school?)

- ① 1 학년 ② 2 학년 ③ 3 학년

3. 귀하가 재학 중인 고등학교는 어느 지역에 속해 있습니까??

(Where is your high school located?)

- ① 서울특별시 ② 광주광역시 ③ 대구광역시 ④ 대전광역시  
⑤ 부산광역시 ⑥ 울산광역시 ⑦ 인천광역시 ⑧ 세종특별자치시  
⑨ 강원도 ⑩ 경기도 ⑪ 경상도 ⑫ 전라도 ⑬ 충청도 ⑭ 제주도

4. 귀하가 가장 최근에 본 국가 모의고사 영어 등급은 무엇입니까?

(What is your most recent national PCSAT-English grade?)

- ① 1 등급 ② 2 등급 ③ 3 등급 ④ 4 등급 ⑤ 5 등급 이하

5. 학교 교육과 사교육을 통한 영어 학습 시간 외 자율적 영어 학습 시간은 하루 평균 몇 시간입니까?

(On average, how many hours a day do you spend for studying English except for regular classes given by the school or private education institution?)

- ① 1 시간 미만 ② 1 시간 이상 2 시간 미만 ③ 2 시간 이상 3 시간 미만  
④ 3 시간 이상 4 시간 미만 ⑤ 4 시간 이상

6. 귀하가 주력하고 있는 대학입학전형은 무엇입니까? (복수 선택 가능)

(What type of college admission are you intensively preparing for?)

- ① 정시 ② 학생부종합전형 ③ 논술고사 ④ 특별전형 ⑤ 특기자전형  
⑥ 기타

7. 영어 사교육(학원, 과외, 인터넷 강의 등)을 받고 있다면, 그 이유는 무엇입니까? (복수 선택 가능)

(If you are receiving private English education such as hagwon classes, private tutoring, and internet lectures, why is that?)

- ① 영어 사교육을 받고 있지 않다 ② 수능 영어 대비 ③ 내신 영어 대비  
④ 공인영어시험(토익, 토플 텡스 등) 대비 ⑤ 실용영어 향상 ⑥ 기타

8. 귀하가 생각한 귀하의 영역별 영어 실력은 무엇입니까?

(How good is your English listening/speaking/writing/reading?)

8-1) 듣기 ① 매우 못한다 ② 못한다 ③ 보통이다 ④ 잘한다 ⑤ 매우 잘한다

8-2) 말하기 ① 매우 못한다 ② 못한다 ③ 보통이다 ④ 잘한다 ⑤ 매우 잘한다

8-3) 쓰기 ① 매우 못한다 ② 못한다 ③ 보통이다 ④ 잘한다 ⑤ 매우 잘한다

8-4) 읽기 ① 매우 못한다 ② 못한다 ③ 보통이다 ④ 잘한다 ⑤ 매우 잘한다

## 2) Main items

### 2-1) Learning attitudes

1. 수능 영어 준비를 하면서 영어 학습에 대한 의욕이 증가했다.

(As I prepared for the CSAT-English, my motivation for English increased.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

2. 수능 영어 준비를 하면서 영어에 대한 자신감이 증가했다.

(As I prepared for the CSAT-English, my confidence in English increased.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

3. 수능 영어 준비를 하면서 영어에 대한 흥미가 증가했다.

(As I prepared for the CSAT-English, my interest in English increased.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

4. EBS-수능 연계로 인해 영어 학습에 대한 의욕이 증가했다.

(The EBS-CSAT linkage policy increased the motivation for English learning.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

5. EBS-수능 연계로 인해 영어 학습에 대한 자신감이 증가했다.

(The EBS-CSAT linkage policy increased the confidence in English learning.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

6. EBS-수능 연계로 인해 영어 학습에 대한 부담이 줄었다.

(The EBS-CSAT linkage policy reduced the academic burden of studying English.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

7. 수능 영어 절대평가로 인해 영어 학습에 대한 의욕이 증가했다.

(The CRA increased the motivation for English learning.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

8. 수능 영어 절대평가로 인해 영어 학습에 대한 자신감이 증가했다.

(The CRA increased the confidence in English learning.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

9. 수능 영어 절대평가로 인해 영어 학습에 대한 부담이 줄었다.

(The CRA reduced the academic burden of studying English.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

## 2-2) Learning contents

10. 수능 영어 준비를 위해 수능 영어 기출문제를 학습하고 있다.

(To prepare for the CSAT-English, I am studying sample items from the previous CSAT-English.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

11. 수능 영어 준비를 위해 수능 영어 사설 모의고사를 학습하고 있다.

(To prepare for the CSAT-English, I am studying the PCSAT-English made by private educational institutions.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

12. 수능 영어 준비를 위해 고등학교 영어 교과서를 학습하고 있다.

(To prepare for the CSAT-English, I am studying high school English textbooks.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

13. EBS-수능 연계로 인해 EBS 연계 교재 학습 시간이 증가했다.

(The EBS-CSAT linkage policy increased the time spent for studying the EBS-CSAT prep books.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

14. EBS-수능 연계로 인해 EBS 변형 문제 학습 시간이 증가했다.

(The EBS-CSAT linkage policy increased the time spent for studying the EBS-modified practice questions.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

15. EBS-수능 연계로 인해 EBS 수능 영어 강의 학습 시간이 증가했다.

(The EBS-CSAT linkage policy increased the time spent for taking the EBS-CSAT English courses.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

16. 수능 영어 절대평가로 인해 영어 교과서와 EBS 연계 교재 외 영어 뉴스, 영미 영화 및 드라마, 팝송 등을 활용한 영어 듣기 시간이 증가했다.

(The CRA in the CSAT-English increased the time spent for studying English listening by using English news, American or British movies and dramas, or pop songs.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

17. 수능 영어 절대평가로 인해 영어 교과서와 EBS 연계 교재 외 영자신문, 영어잡지, 영어소설 등을 활용한 영어 읽기 시간이 증가했다.

(The CRA in the CSAT-English increased the time spent for studying English reading by using English newspaper, magazines, and novels.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

18. 수능 영어 절대평가로 인해 시사 이슈와 특정 주제에 대한 영어 말하기와 쓰기 시간이 증가했다.

(The CRA in the CSAT-English increased the time spent for studying English speaking and writing about current issues and topics.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

### 2-3) Learning strategies

19. 모의고사와 수능 일정에 맞춰 생활 패턴을 조절하고 있다.

(I am adjusting my daily life pattern in accordance with the number of days left for the upcoming PCSAT and CSAT.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

20. 모의고사와 수능 일정에 맞춰 영어 학습 계획을 세우고 있다.

(I am making my English study plan in accordance with the number of days left for the upcoming PCSAT and CSAT.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

21. 수능 영어 준비를 위해 유형별 문제풀이법을 분석하면서 학습하고 있다.

(To prepare for the CSAT-English, I am studying English by analyzing the proven tips for answering various types of questions in the CSAT-English.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

22. 수능 영어 준비를 위해 자신이 취약한 부분을 집중적으로 보완하면서 학습하고 있다.

(To prepare for the CSAT-English, I am studying English by intensively strengthening my weaknesses.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

23. EBS-수능 연계로 인해 EBS 연계 교재 단어를 외우면서 학습하는 시간이 증가했다.

(The EBS-CSAT linkage policy increased the time spent for memorizing the words in the EBS-CSAT prep books.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

24. EBS-수능 연계로 인해 EBS 연계 교재의 지문 내용을 외우면서 학습하는 시간이 증가했다.

(The EBS-CSAT linkage policy increased the time spent for memorizing the passages of the EBS-CSAT prep books.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

25. EBS-수능 연계로 인해 EBS 연계 교재의 지문 내용을 요약정리하면서 학습하는 시간이 증가했다.

(The EBS-CSAT linkage policy increased the time spent for summarizing the passages of the EBS-CSAT prep books.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

26. EBS-수능 연계로 인해 EBS 연계 교재에 나오는 지문의 한글 번역문을 활용하면서 학습하는 시간이 증가했다.

(The EBS-CSAT linkage policy increased the time spent for studying English by using the Korean translations of the passages in EBS-CSAT prep books.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

27. 수능 영어 절대평가로 인해 주변 사람들과 의사소통을 하면서 학습하는 시간이 증가했다.

(The CRA in the CSAT-English increased the time spent for communicating with other people.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

28. 수능 영어 절대평가로 인해 문제풀이 중심의 학습을 하는 시간이 줄었다.

(The CRA in the CSAT-English reduced the time spent for solving CSAT-English practice questions.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

29. 수능 영어 절대평가로 인해 암기 중심의 학습을 하는 시간이 줄었다.

(The CRA in the CSAT-English reduced the time spent for memorization-oriented learning.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

30. 수능 영어 절대평가로 인해 이해하기 어려운 난해한 영어 지문을 학습하는 시간이 줄었다.

(The CRA in the CSAT-English reduced the time spent for studying CSAT-English practice passages that are complicated and hard to understand.)

① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 매우 그렇다

#### 2-4) Open-ended questions

31. EBS-수능 연계와 수능 영어 절대평가로 인해 달라진 학습 태도에 대해 덧붙일 내용이 있다면 최대한 자세히 서술해 주시기 바랍니다.

(Please explain in detail if you have anything to add about the learning attitudes with the introduction of the EBS-CSAT linkage policy and the criterion-referenced assessment.)

32. EBS-수능 연계와 수능 영어 절대평가로 인해 달라진 학습 교재와 시험 대비 전략에 대해 덧붙일 내용이 있다면 최대한 자세히 서술해 주시기 바랍니다.

(Please explain in detail if you have anything to add about the learning contents and learning strategies for the CSAT-English changed with the introduction of the EBS-CSAT linkage policy and the criterion-referenced assessment.)

## Appendix B: Teachers' Questionnaire

### 1) Teachers' background

1. 귀하의 성별은 어떻게 되십니까?

(What is your gender?)

① 남자 ② 여자

2. 귀하가 영어 선생님이로 근무한 기간은 총 얼마입니까?

(How long have you worked as an English teacher?)

① 1년 미만 ② 1년 이상 2년 미만 ③ 2년 이상 3년 미만

④ 3년 이상 4년 미만 ⑤ 4년 이상

3. 귀하가 지금까지 가르친 경험이 있는 학년은 무엇입니까? (복수 선택 가능)

(What grade have you taught so far?)

① 1학년 ② 2학년 ③ 3학년

4. 귀하가 현재 가르치고 있는 학생들은 몇 학년입니까? (복수 선택 가능)

(What grade are your current students in?)

① 1학년 ② 2학년 ③ 3학년

## 2) Main items

1. 수능 영어 영역에 대한 교육부와 평가원의 정책(EBS 수능 연계, 수능 영어 절대평가 등)으로 인해 수업 준비 태도 혹은 수능 영어를 가르치는 태도에 달라진 점이 있다면 최대한 자세히 서술해 주시기 바랍니다.

(Please explain in detail if there are any differences in your attitude toward preparing for class or teaching CSAT-English due to the policies of the Ministry of Education and the KICE on the CSAT-English such as the introduction of the EBS-CSAT linkage policy and criterion-referenced assessment.)

2. 수능 영어 영역에 대한 교육부와 평가원의 정책으로 인해 수능 영어를 가르치는 데 사용하는 교재에 달라진 점이 있다면 최대한 자세히 서술해 주시기 바랍니다.

(Please explain in detail if there are any differences in the textbooks or materials used to teach CSAT-English due to the policies of the Ministry of Education and the KICE on the CSAT-English.)

3. 수능 영어 영역에 대한 교육부와 평가원의 정책으로 인해 수능 영어를 가르치는 방식에 달라진 점이 있다면 최대한 자세히 서술해 주시기 바랍니다.

(Please explain in detail if there are any differences in the way you teach CSAT-English due to the policies of the Ministry of Education and the KICE on the CSAT-English.)

4. 수업 학생의 학년에 따라 수업 준비 태도 혹은 수능 영어를 가르치는 태도에 달라진 점이 있다면 최대한 자세히 서술해 주시기 바랍니다.

(Please explain in detail if there are any differences in your attitude toward preparing for class or teaching CSAT-English in accordance with the grade your students are in.)

5. 수업 학생의 학년에 따라 수업 교재 혹은 수능 영어를 가르치는 데 사용하는 교재에 달라진 점이 있다면 최대한 자세히 서술해 주시기 바랍니다.

(Please explain in detail if there are any differences in the textbooks or materials used to teach CSAT-English in accordance with the grade your students are in.)

6. 수업 학생의 학년에 따라 수업 방식 혹은 수능 영어를 가르치는 방식에 달라진 점이 있다면 최대한 자세히 서술해 주시기 바랍니다.

(Please explain in detail if there are any differences in the way you teach CSAT-English in accordance with the grade your students are in.)

## 국문 초록

### 대학수학능력시험 영어 영역이 학생들의 학습에

#### 미치는 환류효과: 학년 그룹 간의 비교

본 연구는 학생의 학년에 따라 달라지는 수능 영어, EBS 수능 연계 정책 및 수능 영어 절대평가 정책의 환류효과에 대해 살펴보고자 하였으며, 이를 위해 고등학교 1, 2, 3 학년 학생과 고등학교 영어 교사 각각에 대한 설문조사를 실시하였다. 설문 조사의 분석 결과는 다음과 같다.

우선, 학생의 학년에 따른 환류효과의 차이와 관련하여, 학생들의 학습 태도, 학습 자료 및 시험 대비 전략 모두에서 통계적으로 유의미한 환류효과의 차이가 나타났으며, 이를 통해 학년에 따른 환류효과의 정도 차이가 존재한다는 것을 알 수 있다. 그러나 학년에 따른 평균의 변화 양상을 살펴보면, 시험 대비 전략만이 유일하게 학년의 상승에 따라 환류효과의 강도가 세지는 것으로 나타나고, 나머지 두 항목은 학년의 상승에 따라 부분적으로 혹은 전체적으로 환류효과의 강도가 약해지는 것으로 나타났다.

다음으로, 고등학교 영어 교사를 대상으로 진행한 설문 조사를 통해 학생의 학년에 따라 달라지는 교사의 교습의 차이를 중점적으로 살펴보고자 하였다. 그 결과, 많은 교사들이 수업 태도는 모든 학년에 걸쳐 동일하지만, 사용하는 교재와 이를 가르치는 방식은 학생의 학년에 따라 달라진다고

답변하였다. 특히, 수업 교재와 관련해서, 대부분의 교사들이 1, 2 학년 수업에서는 정규 교과서를 주로 사용하는 데 반해, 수능 시험이 임박한 3 학년 학생 수업에서는 EBS 연계 교재를 주로 다룬다고 하였다. 또, 수업 방식과 관련해서, 저학년 수업에서는 학생들의 의사소통능력을 향상시키기 위한 활동이 다수 진행되는 한편, 3 학년 수업은 수능 중심으로 이루어지거나 학생들이 주로 준비하고 있는 대학 입시 제도에 부합하는 활동들이 주로 진행되었다.

이상에서 살펴본 연구 결과를 바탕으로 본 연구의 시사점은 다음과 같다. 우선, 학습 측면별로 서로 다른 변화 양상을 보인 것을 토대로, 앞으로의 환류효과에 대한 연구는 학생들의 다양한 학습 측면을 종합적으로 고려하여 진행되어야 한다. 특히, 특정 학습 측면에서 나타난 특성을 학생들의 학습 전체에 일반화하는 것을 경계해야 한다.

다음으로, 학년에 따른 교사의 교습에 차이가 나타나고 이로 인해 학생들의 학습 교재 및 학습 방식에서 학년에 따른 차이가 발생한 점을 바탕으로, 학생에 대한 환류효과를 분석하는 연구에서 학생 외 이들의 학습에 영향을 주는 이해관계자 혹은 주변인에 대한 분석 또한 이루어져야 한다. 학생들의 학습이 이들 스스로의 의지만으로 결정되기보다는 일정 부분 주변 이해관계자에 의해 영향을 받기 때문이다.

**주요어:** 환류효과, 환류효과 강도, 대학수학능력시험, 학습, 교습

**학번:** 2018-20495