RELATIONSHIP BETWEEN IB DP ENGLISH SCORES,
ENGLISH PROFICIENCY EXAM SCORES AND STUDENT
SELECTION EXAMINATION SCORES:
A CORRELATIONAL STUDY

A MASTER'S THESIS

BY

ERAY BİÇER

THE PROGRAM OF CURRICULUM AND INSTRUCTION
BİLKENT UNIVERSITY
ANKARA

JANUARY 2015

Relationship between IB DP English Scores, English Proficiency Exam Scores and
Student Selection Examination Scores: A Correlational Study

The Graduate School of Education

of

Bilkent University

by

Eray Biçer

In Partial Fulfilment of the Requirements for the Degree of

Master of Arts

in

The Program of Curriculum and Instruction
Bilkent University
Ankara

January, 2015

BİLKENT UNIVERSITY

Graduate School of Education

Relationship between IB DP English Scores, English Proficiency Exam Scores and
Student Selection Examination Scores: A Correlational Study

SUPERVISEE: ERAY BİÇER

January, 2015

I certify that I have read this thesis and have found that it is fully adequate, in scope
and in quality, as a thesis for the degree of Master of Arts in Curriculum and
Instruction.

Asst. Prof. Dr. İlker Kalender

I certify that I have read this thesis and have found that it is fully adequate, in scope
and in quality, as a thesis for the degree of Master of Arts in Curriculum and
Instruction.

Prof. Dr. Alipaşa Ayas

I certify that I have read this thesis and have found that it is fully adequate, in scope
and in quality, as a thesis for the degree of Master of Arts in Curriculum and
Instruction.

Asst. Prof. Dr. Deniz Ortaçtepe

Approval of the Graduate School of Education

Prof. Dr. Margaret K. Sands

ABSTRACT

RELATIONSHIP BETWEEN IB DP ENGLISH SCORES, ENGLISH PROFICIENCY EXAM SCORES AND STUDENT SELECTION EXAMINATION SCORES: A CORRELATIONAL STUDY

Eray Biçer

M.A., Program of Curriculum and Instruction
Supervisor: Asst. Prof. Dr. İlker Kalender

January 2015

English language teaching has become one of the most inseparable parts of Turkish National Education Curriculum. With the emergence of international programs such as International Baccalaureate, International General Certificate of Secondary Education and Advanced Placement English language teaching and language proficiency gained new perspectives in Turkey and around the world. While the English curriculum in Turkish national system focuses on receptive and productive skills; reading, listening, writing and speaking, the curriculum of international programs focus more on critical thinking skills in the target language. The question of how well a student who received IB Diploma Programme English education can perform in a standardized English proficiency test which focuses on critical thinking skills rather than four skills remains unanswered in Turkish context.

This study aimed to explore the relationship between IB DP English scores, a standardized English test scores (COPE) and student Selection Examination (OSS) scores of Bilkent Erzurum Laboratory School graduates. This study was completed

by using the IB DP English, COPE and SSE scores of 119 students as instruments. The data were analysed with correlational analysis approach. Statistically significant relationships were discovered between COPE exam scores and IB predicted grades.

Keywords: International Baccalaureate, Diploma Program, Language Proficiency, English education, COPE exam, Student Selection Examination.

ÖZET

ULUSLARARASI BAKALORYA DİPLOMA PROGRAMI İNGİLİZCE
PUANLARI İNGİLİZCE YETERLİLİK SINAVI PUANLARI VE ÖSS PUANLARI
ARASINDAKİ İLİŞKİ: BİR KORELASYON ÇALIŞMASI

Eray BİÇER

Yüksek Lisans, Eğitim Programları ve Öğretim
Tez Yöneticisi: Yrd. Doç. Dr. İlker Kalender

Ocak 2015

İngilizce eğitimi ulusal Türk eğitim sisteminin ayrılmaz bir parçası olmuştur.
Uluslararası Bakalorya, Uluslararası Ortaöğretim Genel Sertifikası ve İleri Düzey
Yerleştirme Programı gibi uluslararası programların ortaya çıkmasıyla İngilizce
eğitimi Türkiye'de ve dünyada yeni bakış açıları kazanmıştır. Türk eğitim
sistemindeki İngilizce eğitimi daha çok okuma, dinleme, yazma ve konuşma gibi
kavrama ve üretici beceriler üzerine yoğunlaşırken uluslararası programların
müfredatları daha çok hedef dilde eleştirel düşünebilme becerilerine yoğunlaşmıştır.
Hedef dilde eleştirel düşünebilme daha çok üzerine yoğunlaşmış olan Uluslararası
Bakalorya Diploma Programı İngilizce eğitimi almış bir öğrencinin yazma, konuşma,
dinleme ve okuma becerilerine yönelik İngilizce yeterlilik sınavında Türkiye
bağlamında nasıl bir performans göstereceği sorusu cevapsız bırakılmıştır.
Bu çalışmanın amacı Bilkent Erzurum Laboratuvar Okulu mezun öğrencilerinin
Uluslararası Bakalorya İngilizce puanları, İngilizce yeterlilik sınavı (COPE) puanları

ve Öğrenci Seçme Sınavı puanları arasında nasıl bir bağlantı olduğunu saptamaktır. Bu çalışma 119 Bilkent Laboratuvar Okulu mezunu öğrencilerinin Uluslararası Bakalorya Diploma Programı İngilizce puanları, İngilizce yeterlilik sınavı (COPE) puanları ve Öğrenci Seçme Sınavı Puanlarının veri olarak kullanılmasıyla gerçekleşmiştir. Veriler Korelasyon yöntemiyle analiz edilmiştir. Uluslararası Bakalorya Diploma Programı İngilizce tahmini puanlar ve COPE puanları arasında istatiksel olarak güçlü bir ilişki saptanmıştır.

Anahtar Kelimeler: Uluslararası Bakalorya, Diploma Programı, İngilizce yeterliliği, İngilizce eğitimi, Bilkent COPE sınavı, Öğrenci Seçme Sınavı.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BELS……..Bilkent Erzurum Laboratory School

CEFRL…...Common European Framework Reference for Languages

COPE…….Certificate of Proficiency in English

DP………..Diploma Programme

HEE………Higher Education Examination

HL………...High Level

IB…………International Baccalaureate

IELTS…….International English Language Testing System

MYP………Middle Years Programme

PYP……….Primary Years Programme

SL…………Standard Level

SSE……….Student Selection Examination

TOEFL……Test of English as a Foreign Language

UPE……....Undergraduate Placement Examination

CHAPTER 1: INTRODUCTION

Introduction

There are a number of international educational organizations such as International General Certificate of Secondary Education (IGCSE) and Advanced Placement (AP) that offer bilingual education and strive to raise students who can use their second language effectively. International Baccalaureate Diploma Programme (IB DP) (IBO, 2014) is among these international programs which were established on the idea of raising international- minded students. One of the main aims of the IB Programme is to educate students as those who can communicate effectively in more than one language. In IB curriculum there are six groups of subjects and Group 1 consists of language subjects. Entitled under Group 1 subjects in IB curriculum Language A: Literature, Language A: Language, Literature and Literature and Performance courses were developed for the students to use English as academic medium. Language B, Language ab initio and Classical languages which are categorized under Group 2 subjects were designed for students to communicate in a second language effectively (IBO, 2014). In non-English speaking countries, English predominates Group 1 or Group 2 options. Schools that implement IB DP curriculum can offer students two options according to the type of literacy they want to build; they can either offer English as second language or along with their native language students can take English as first language.

Like in many other countries IB DP has become one of the widely used international programs in Turkish private schools. As of 2014, there are 50 (June, 2014) IB schools in Turkey and thirty five of these schools implement IB DP curriculum along

with Turkish Ministry of National Education Curriculum except Bilkent International Laboratory School which has an agreement with the government to implement only international curricula. It is evident from their curricular aims and objectives that IB DP claims to offer an intensive foreign language education. Therefore, it is crucial to find the correlation between the students IB DP English scores and a standardized English proficiency test.

Founded in 1968 International Baccalaureate Organization (IBO) is a non-profit educational foundation (IBO, 2014) that offers programs at primary, middle and high school levels. International Baccalaureate Organization started offering Diploma Programme in 1968 for high school students, Middle Years Programme (MYP) in 1994 for middle school students, and Primary Years Programme (PYP) in 1997 for primary school students. DP is a two-year high school program which is being implemented widely by international schools around the world.

DP curriculum includes six groups of subjects and a central core. Six groups of subjects are; group one- studies in language and literature, group two- language acquisition, group three- individuals and societies, group four- sciences, group five- mathematics, and group six- the arts. The students are expected to choose one course from each category. However, instead of a group six subject they can select one subject from other categories. In order to complete the requirements of the central core, the students are expected to write an Extended Essay and Theory of Knowledge Essay and spend hours to complete Creativity, Action, Service (IBO, 2014).

Group 1 subjects include; Language A- Literature, Language A- Language and Literature, and Literature and Performance (IBO, 2011). Designed for the students who have used the language in academic context, Group 1 subjects aims to "support

future academic study by developing a high social, aesthetic and cultural literacy, as well as effective communication skills" (IBO, 2011, p.4). Group 2 subjects include two modern languages and classical languages course. Language ab initio and Language B course which can be delivered in many different languages are designed as language acquisition courses. Language ab initio and Language B courses aim to enable students to communicate effectively in an environment where the language is spoken and acquire an intercultural understanding (IBO, 2011). The classical languages course is developed to have students understand the language, literature and culture of the classical world (IBO, 2011).

Background

Language proficiency measures how well an individual has excelled in a particular language and how well the individual uses, and communicates in a target language (Stephen, Welan & Joradan, 2004). Language proficiency is measured via two different types of skills; receptive and productive skills. Receptive skills include reading and writing whereas productive skills include writing and speaking. As well as receptive and productive skills language proficiency exams evaluate how well a learner uses the grammar, vocabulary and mechanics of the target language (Vinke & Jochems, 1993).

Since the English Language is considered as the global lingua franca (Stephen et al., 2004) of 21st Century there are many English proficiency exams that test a learner's English proficiency. Language proficiency tests such as TOEFL (Test of English as a Foreign Language), IELTS (International English Language Testing System), and ESOL (English for Speakers of Other Languages) enable a learner to prove and validate his/her proficiency in the target language. According to Doey and Oliver

(2002) TOEFL (Test of English as a Foreign Language) and IELTS (International English Language Testing System) are the most popular English proficiency exams that non-English speaking background students use in order to apply overseas universities and by the help of these two most popular proficiency tests international students are able to provide evidence of proficiency in the English before being admitted.

Test of English as a Foreign Language (TOEFL) is a standardized test that aims to measure fluency and knowledge of the foreign language learner. TOEFL IBT (internet based test) consists of reading, listening, speaking and writing and the exam takes four hours. In reading part, the test takers are expected to answer questions related to four or five different reading passages. The listening part assesses the test takers ability to understand the main ideas in the conversations they hear and answer related questions. In speaking part the test takers are asked to talk about the prompts for twenty minutes. The writing part takes fifty minutes and the test takers are expected to write a composition on one of the given prompts (ETS, 2014). Similar to TOEFL, IELTS "measures the language proficiency of people who want to study or work in environments where English is used as a language of communication'' (UCLES, 2014). In this 2 hours and 55 minutes test the individuals are tested on four areas; listening, speaking, writing and reading.

English Language proficiency has become an important component of Turkish education due to the increasing role of English as global language. English has become an inseparable part of Turkish national education in all levels. Although English education starts with the kinder garden schools in private schools, the students start to receive English education in second grade with three periods a week

in state schools. English education in state high schools shows variety in terms of curriculum and teaching hours. For example, while in Anatolian high schools the students receive 6 periods of English classes a week, the number of English classes descends to four periods a week in regular high schools. Besides state schools which implement the curriculum of Turkish Ministry of Education there some 61 schools in Turkey that implement international curricula. Advanced Placement (AP) is implemented by 11 (June 2014) schools while International Baccalaureate implemented by 50 (June 2014) schools in Turkey. When compared to state schools the intensity of English education is higher in these two international curricula.

Many Turkish universities such as Hacettepe and Bilkent University demand their students to validate their English proficiency. By using the scores of these standardized tests the students are allowed to skip prep year in the university and they can immediately start studying in their departments. Bilkent University demands to see a satisfactory score from one of these English proficiency tests; IELTS, TOEFL, Cambridge English CAE (Certificate in Advanced English), Cambridge English FCE (First Certificate in English), and YDS (Foreign Language Competency Exam). The students who do not have valid English score from one of these English proficiency exams are required to take COPE exam which is designed by Bilkent University School of Language. Bilkent COPE Exam measures students' English proficiency according to the standards of Common European Framework of Reference of Languages: Learning, Teaching, Assessment which is a guideline to measure individuals' foreign language proficiency prepared by Council of Europe. With the help of COPE Exam Bilkent University measures if the students, who want to skip one year of English prep school, are at B2 Level

Problem

The examinations in IB Diploma program are regarded as high-stakes and international since they allow students to enter universities anywhere in the world (Coates, Rosicka & Ball 2007,). Therefore, considering the students' achievement, the schools that apply IB DP need to make the most suitable decision for their students in terms of offering IB DP courses. For instance, an IB DP school can offer English Language B course instead of Language ab initio considering the students' level of English proficiency or the use of English in students' future academic life. Moreover, IB DP schools are allowed to offer two Group 1 courses instead of one group one and one group two course. As well as taking Group 1 course in their native language the students are sometimes required to take another group one course in a foreign language. There are many studies that examine the relationship between language proficiency and academic attainment. For example, in her correlational case study Sert (2006) examined the relationship between academic attainment of Hacettepe medicine and economy students and their English proficiency levels. Similarly, Watt (2009) investigated the correlation between the TOEFL scores and academic success of engineering students in an unspecified American university abroad. However, there is no study that investigates the correlation between IB DP Group 1 or Group 2 options and a standardized English proficiency test in Turkish context.

Purpose

The purpose of the study is to explore; the correlation between (i) IB DP English scores and COPE scores and the correlation between (ii) IB DP English scores and Student Selection Examination (SSE) scores. Specifically, the research aims to find out the most suitable Group 1 and Group 2 courses for the students who are

6

considering taking English proficiency test. IB DP schools offer English courses according to the levels and needs of the students. This study can provide information about the nature of English proficiency tests and IB DP English examinations which is essential for the schools when they are making decisions about which IB DP course to offer.

Research questions

1. Is there any relationship between students' (i) IB DP English scores and COPE scores; (ii) IB DP English scores and SSE scores?

2. What are the factors explaining differences between three IB English options?

3. What are the factors explaining differences between students who are below and above COPE average score set as 75?

Significance

This study provides an insight into the IB DP English curriculum which is implemented in some high schools in Turkey with different options. It also gives information about English proficiency in Turkey and several standardized high-stakes tests such as IELTS, TOEFL and COPE. This study might be useful for researchers who are investigating the effectiveness of IB DP Group 1 and Group 2 courses in terms of delivering English language proficiency. Furthermore, schools that implement IB DP may benefit from this study when making choices about Group 1 or Group 2 courses to offer in their school.

Definition of key terms

International Baccalaureate: International Baccalaureate Organization is a non-profit educational foundation established in 1968 (IBO, 2014). The organization offers programs for students between the ages of 6-19. IB curriculum is implemented in 3612 schools around the world.

Diploma Programme: Offered by International Baccalaureate Diploma Programme is a two-year high school programme for students aged 16 to 19. Throughout this programme, which is widely used in international high schools, the students' works are internally and externally assessed by their teachers and examiners around the world.

Language A: The curriculum of International Baccalaureate Diploma Programme consists of six groups of subjects and the students are required to take one subject from each group. Language A is the common name given to Group 1 subjects. Language A course options include; Language and Literature, Literature and Literature and Performance. An IBDP school can offer any of these courses which are designed for students who have experience in using the language in academic context.

Language B: The curriculum of International Baccalaureate Diploma Programme consists of six groups of subjects and the students are required to take one subject from each group. Language B is the common name of the courses offered as Group 2 subjects. Language B course options include Language B, Language ab initio and Classical Languages. Designed as language acquisition courses Language B courses aim to enable students communicate effectively in another language.

Language Proficiency: The term language proficiency is used to describe an individual's ability to use, and perform in a foreign language. There are many high-stakes language proficiency tests such as TOEFL and IELTS that measures an individuals' language proficiency level. These tests generally examine the level of language proficiency by testing an individuals' ability to write, speak, read, listen and use the grammar and mechanics of the foreign language.

COPE: Offered by Bilkent University COPE is a language proficiency exam. "COPE exam is set for B2 Level of Common European Framework of Reference" (COPE, 2013) which is a framework prepared by Council of Europe for assessing and teaching foreign languages. Unless they provide a satisfactory score from other English proficiency exams such as TOEFL, IELTS and YDS all freshmen students are required to take COPE exam in order to skip an English prep year. The results of COPE exam is only used by Bilkent University.

CHAPTER 2: LITERATURE REVIEW

Introduction

The aim of the study is to explore; the correlation between (i) IB DP English scores and COPE scores and the correlation between (ii) IB DP English scores and Student Selection Examination (SSE) scores. Therefore, this chapter provides brief information about high-stakes exams, language proficiency, language proficiency exams, English education in Turkey, IB DP, Language education in IB DP, and SSE. The chapter ends with brief discussion about the studies that were conducted and the gap in literature.

High-stakes exams

High-stakes test is a test whose outcomes drastically effect or alter the test-takers' life. Passing the high-stakes examination has important benefits for the test-taker while failing conditions cause certain disadvantages (Embse & Hasson, 2012). Tests such as Language proficiency, student selection examinations for higher education, driver's license test, certification examinations, and high school diploma tests can be categorized under high-stakes examination which change the test-takers' life in a positive way.

There are many issues involving the advantages and disadvantages of high-stakes exams. Amrein & Berliner (2002) summarize the advantages of high-stakes testing in schools as; students and teachers know what is important to learn and teach; teachers and students are motivated to teach and learn better and low-achiever students try

harder; high-stakes tests measure the effectiveness of the curricula being implemented in the schools; high-stakes tests are good measure for the students' performance; results of the high-stakes tests can be used to make alteration in instruction and curriculum and finally the results of high-stakes examination provide parents with a parameter to understand their children's performance.

According to the studies that were conducted by Embse & Hasson (2012) and Amrein & Berliner (2002) the disadvantages of high-stakes tests can be summarized as; high-stakes exams limits the creativity of students and teachers and they just needed to focus on examination subjects; students with economic disadvantage cannot get extra help and they study in disadvantaged schools but they are measured with the same examinations that the students with economic advantage studying in better schools. The most prominent disadvantage of the high-stakes examination is that test-takers face with exam anxiety and they cannot demonstrate the performance they do in classes (Embse & Hasson, 2012).

Language proficiency and English proficiency exams

In general sense, language proficiency means an individuals' ability to use a foreign language in various ways such as writing, reading, speaking and listening. Llurda (2000) describes language proficiency as "the skills needed to put language knowledge into practice that is to transform knowledge into language use" (p.91). In terms of reflecting the communicative aspect of language proficiency definition of Stern (1996) shows parallelism to that of Llurda (2000). Stern (1996) defines an individual with language proficiency having the following attributes; the intuitive mastery of the forms of the target language and linguistic cognitive, affective and sociocultural meanings expressed by the language forms; creativity and capacity to

11

use the language with maximum attention to communication and minimum attention to form.

The term proficiency is preferred to competence in recent studies since proficiency also refers to potential ability to learn a second language (North, 2012). In his seminal work Taylor (1988) explains the reason of using competence instead of proficiency as "competence is clearly a state and not a process and has nothing to do with capacity or ability" (p.151).

Studies of researchers such as Bachman (1990), Canale and Swain (1980) brought new dimensions to the language proficiency and the term communicative competence ,which is currently being tested in many language proficiency tests such as TOEFL and IELTS, came into use instead of the broad term proficiency. In general sense, communicative competence refers to having knowledge or capacity to implement and execute language proficiency in appropriate, contextualized communicative language use (Llurda, 2000).

Canale and Swain (1980) suggested three sets of competency areas which produce communicative competence; grammatical competence, sociolinguistic competence, strategic competence and discourse competence. Grammatical competence refers to "knowledge of lexical items and rules of morphology, syntax, sentence grammar semantics and phonology" (Canale & Swain, 1980, p.29). Canale and Swain (1980) describe sociolinguistic competence as "sociocultural rules of use and rules of discourse" (p.29). Canale and Swain (1980) describe strategic competence as "verbal and non-verbal communicative strategies that may be called into action to compensate for breakdowns in communication due to performance variables or to insufficient competence" (p.29). The last competence is discourse competence which

can be described as "mastery of cohesion and coherence in different genres" (Razmjoo, 2011, p.87).

Bachman (1990) extended the communicative competence framework proposed by Canale & Swain (1980) and his framework attempts to "characterize the process by which the various components interact with each other and with the context in which language use occurs" (p.81). the framework proposed by Bachman (1990) includes three components; language competence, strategic competence and psychophysiological competence. Language competence which covers organizational and pragmatic competence is described as "a set of components that are utilized in communication via language" (Bachman, 1990, p.84). Assessment, planning, and execution are categorized as strategic competence which refers to "the mental capacity to implement language competence appropriately in the situation which communication takes place, and involves sociocultural and real world knowledge" (Razmjoo, 2011, p.88). Psychophysiological competence can be described as neurological and psychological processes which result in

Language proficiency exams test an individual's ability to perform in a second language. Second language learners take language proficiency exam in order to validate their second language level. Many of the international standardized, high-stakes language proficiency exams such as IELTS and TOEFL measures an individual's level of reading, writing, speaking, listening and use of language. International language proficiency exams, whose scores are valid around the world, such as TOEFL and IELTS can be categorized as criterion-referenced assessment rather than norm-referenced assessment. That is, the performances of second

language learners are "judged in relation to defined standard and not in relation to" (North, 2012, p.131) other test takers.

There are two major standardized international language proficiency exams; Test of English as a Foreign Language (TOEFL) and International Language Testing system (IELTS). TOEFL is a standardized English language test designed by Educational Testing Service (ETS) which is a non-profit American testing service specialized in educational testing and assessment. Non-native English speaking students take TOEFL exam in order to demonstrate the level of their English proficiency especially when applying to colleges and universities which use English language as a medium (Cho & Bridgeman, 2012). Many English medium universities require a satisfactory TOEFL score, which depends on the level of English proficiency necessary to use in the institution. These universities view TOEFL as "linguistic threshold that enables them to approach academic work in English in a meaningful manner" (Chalhoub-Deville & Deville, 2006, p.520). TOEFL can be taken in two different ways, TOEFL Paper Based Exam and Internet Based Exam (TOEFL, IBT) which is more commonly used than the former. First implemented in 2005, TOEFL IBT consists of reading, listening, speaking and writing and all of the tasks in the test are completed in computers in test centers. Table 1 represents structure of TOEFL Exam;

Table 1 Structure of TOEFL exam (ETS, 2014)

Skill	Score Range	Level
Reading	0-30	High (22–30)
		Intermediate (15–21)
		Low (0–14)
Listening	0-30	High (22–30)
		Intermediate (15–21)
		Low (0–14)
Speaking	0-30	Good (26–30)
		Fair (18–25)
		Limited (10–17)
		Weak (0–9)
Writing	0-30	Good (24–30)
		Fair (17–23)
		Limited (1–16)
TOEFL Score	0-120	

In the reading section the test takers are expected to answer 30 questions related to 4-6 passages. "The Reading section measures the test taker's ability to understand university-level academic texts and passages. In many academic settings around the world, students are expected to read and understand information from textbooks and other academic materials written in English" (TOEFL, 2014, p.8). Listening section consists of six passages, each 3–5 minutes in length and it "measures the test taker's ability to understand spoken English" (TOEFL, 2014, p.12). Speaking section includes two independent and four integrated speaking tasks. In independent tasks the test takers answer the questions according to their opinions and in integrated tasks the test takers are expected to answer the questions related to another reading

and listening passages. "The speaking section measures the test taker's ability to speak effectively in academic settings" (TOEFL, 2014, p.16). Like speaking section, writing section consists of two different tasks; independent and integrated writing questions. For the integrated part the test takers are expected to listen to an academic passage or conversation and summarize the main points in the listening passage. In the independent part the test takers are expected to produce a coherent composition on a given writing prompt. "The Writing section measures a test taker's ability to write in an academic setting" (TOEFL, 2014, p.22).

Similar to TOEFL, IELTS is a standardized high stakes English proficiency test that measures test takers ability to perform in reading, writing, speaking and listening tasks. Prepared by Cambridge English Language Assessment, the British Council and IDP Education "the International English Language Testing System (IELTS) is a test that measures the language proficiency of people who want to study or work in environments where English is used as a language of communication" (ELTS, 2014, p.2).

There are two versions of IELTS; Academic version and General Training Version. Academic Version of IELTS is taken by the individuals who want to study in the universities in English-speaking countries or who want to be employed in the occupations that required high level of English. General Training Version of IELTS is taken by the individuals who want to have non-academic training, or immigrate to English- speaking countries.

Listening section of IELTS which consists of four listening passages and related questions measures the test takers' "ability to understand main ideas and detailed factual information; understand the opinions and attitudes of speakers; to understand

the purpose of an utterance and to follow the development of ideas" (IELTS, 2014).

Reading section consists of three reading passages and forty related questions.

Reading section of IELTS measures different abilities of reading such as "reading for

gist, reading for main ideas, reading for detail, skimming, and understanding logical

argument, recognizing writers' opinions, attitudes and purpose" (IELTS, 2014, p.11).

Writing section of the exam which consists of two questions, general interest

question and academic question, tests the test takers' ability to communicate in

written form of English language. In the speaking section the test-takers are required

to complete three speaking tasks which consist of personal questions, questions

related to particular topic and further questions related to a given topic. This section

tests the test-taker's ability to use spoken English (IELTS, 2013). At the end of the

examination the test-takers are given an overall score from 1-9. Score one describes

the test-taker as non-user and the test-taker who have score 1 in this examination is

assumed to have "no ability to use the language beyond possibly a few isolated

words" (IELTS, 2014, p.12). Score nine describes the test-taker as expert user who

"has fully operational command of the language: appropriate, accurate and fluent

with complete understanding" (IELTS, 2014, p.12).

Certificate of English proficiency exam (COPE)

Certificate of Proficiency Exam (COPE) is a standardized English proficiency exam

prepared by Bilkent University School of English Language to assess the level of

students' English proficiency. Bilkent COPE exam is prepared by using the standards

of Common European Framework of Reference for Languages (CEFRL) which is a

project of Council of Europe aiming to "provide a common basis for the elaboration

of language syllabuses, curriculum guidelines, examinations, textbooks, etc. across

Europe (CEFR, 2001). COPE takes B2 level of CEFRL as basis and measures if the students are at B2 level. CEFRL refers B2 level of language learners as independent learners and describes them as having the following attributes; B2 level of language learners;

> Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialization; can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party; can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and Independent disadvantages of various options (CEFRL,2001, p.33).

Offered three times a year, COPE exam tests the students' skills of reading, listening, writing speaking and using language and grammar. Designed as five-option multiple choice test reading paper of COPE consists of three parts and the duration is one hour twenty minutes. The higest point that the test-taker can get from COPE is 150 while they have to score at east 83,5 in order to pass the exam. Bilkent university offers alternatives to the students that do not wish to take COPE exam. As of 2015, following scores from the related exams are accepted by Bilkent University instead of COPE exam; Certificate in Advanced English (C), First Certiicate in Englsh (B), IELTS (6.5), TOEFL (87), and Foreign Language Examination (YDS, 87).

In part one of the paper there are three short passages and the students are tested on their ability to make inference, determine the tone, purpose and audience of the text, guess meaning from the context and paraphrase (COPE, 2013). Part two of the paper consists of two middle length passages and the test-takers are assessed on their ability to find main ideas, supporting details and specific details and make inferences (COPE, 2013, p.3). Part three of the paper includes one long passage which is in between 1,300 and 1,600 words in length. Part three of the reading paper assesses the

test-taker's ability to find main ideas, supporting ideas and specific information; guess meaning from the context; understand references; and make inference (COPE, 2013).

Listening paper consists of two lectures (passages) and each lecture takes approximately fifteen minutes in length. In total the test-takers are given an hour to complete the paper. The students are allowed to take notes on the paper while listening. After listening for each lecture the students are expected to answer ten short answer questions and five multiple choice questions. Listening paper test the test-takers' "ability to listen for main ideas, supporting ideas, specific information; and ability to infer meaning and identify speaker's tone and purpose" (COPE, 2013, p.4).

Language paper of COPE consists of three parts and the test-takers are given forty minutes to complete the paper. Part one of the paper includes two passages and the test-takers answer ten cloze test questions by filling in the gaps of the passages using one word. Part one of the language paper tests the students "ability to produce grammatical structures according to the meaning required" (COPE, 2013, p.4). In part two of the language paper the students fill in seven gaps by using the correct forms of the words provided in the paper. This part tests the students "ability to recognize words and use them in the correct context" (COPE, 2013, p.4). Part three of the language paper includes a passage with eight gaps and box filled with words. The students are expected to write the correct forms of the words by using prefixes and suffixes. This part assesses the students' ability to "manipulate word forms by adding suffixes and prefixes according to the meaning required; understand the context; focus on spelling and grammar" (COPE, 2013, p.4).

Unlike other papers of the exam which included receptive skills, productive skills of the students are assessed by using well-described criteria. Writing paper of COPE includes two essay questions that students can choose from. The students are expected to write a 350 word essay in an hour. Writing paper assesses the students' "ability to comprehend a given prompt; analyze and reflect on the argument; generate and organize ideas; justify, support and exemplify; and produce coherent, accurate and relevant prose" (COPE, 2013, p.5). At the end of the examination the students are given a score out of ten. An essay that is scored ten is described as having a very good coverage of the topic; very good use of linking devices; accurate use of a wide range appropriate structures; complex sentences and a wide range of vocabulary (COPE, 2013). On the other hand, an essay scored one or two is described as having no relevance to the topic, limited vocabulary, serious language errors and incoherency (COPE, 2013).

COPE speaking exam is carried as a paired format exam which involves two or three students. An interlocutor asks questions and carries on a conversation while an assessor examines the speaking abilities of the students. The speaking exam which lasts 15 minutes consists of two parts. In part one the students are asked tree questions and expected to talk about them for three minutes. In part two of the exam the students are given speaking prompts and they are expected to talk for one minute (COPE, 2013). COPE speaking exam criteria includes the bands entitled as; pronunciation, interaction, fluency, discourse and language. The students who score five in any bands are described as "displaying a very high level of communicative effectiveness for B2 level" whereas the students who score one in any bands are referred as "displaying a very low level of communicative effectiveness and as not clearly at B2 level" (COPE, 2013, p.145).

COPE can be considered to measure communicative competence of the test-takers. Kathleen and Kenji (1996) argue that "the communicativeness of a language proficiency test might be seen as being on a continuum (p.2). Many language proficiency tests such as COPE have some element of communicativeness and only few tests are entirely communicative (Kathleen & Kenji, 1996). In the light of this information it can be discussed that listening, writing and speaking tests are clearly aligned with the premises of communicative competence since these tasks ask the test-takers to put themselves in a certain situation by providing context. These tests measure if the test-takers can express themselves clearly in certain situations.

English education in Turkey

English language teaching has become prominent in 1980s in Turkey due to the increasing role of English as lingua Franca in the rapidly globalizing world. English language education has entered into every level of Turkish national education system after 1980s. English language teaching has gone through many changes in the last three decades (Dinçer, Takkaç, & Akalın, 2010). In the current educational system English language education starts in the second grade with three periods a week. In the middle school students receive 7- 8 periods of English lessons. In high school the number of English lesson periods differs according to the type of high schools which were all turned into Anatolian high schools in 2014. In terms of English education high schools in Turkey are classified as high schools with prep year and high schools without prep year. Anatolian high schools offer three- two hours of English classes a week whereas Anatolian Science high schools and offer twice as many English courses a week. On the other hand high schools with prep classes offer the same

number of English classes a week and they also offer twenty hours of English classes in the prep class.

The number of English classes shows variety in the private schools which gives more importance to English language teaching than the state schools. English education in private schools can even start as early as kinder garden level. Moreover, in Turkey international curricula such as Primary Years Programme, Middle Years Programme, Diploma Programme, International General Certificate of Secondary Education, and Advanced Placement are being implemented by the authorized private schools. In these private schools the number of English classes and the expectations from the students who receive any these international curricula are higher when compared to English education in state schools.

International Baccalaureate and English education

Founded in 1968, International Baccalaureate (IB) is "is a non-profit educational foundation, motivated by its mission, focused on the student (IBO, 2014). There are four programs in IB for the students aged 3-19; these programs are Primary Years Programme (PYP), Middle Years Programme (MYP) and Diploma Programme (DP) and Career-related Certificate (CC). "Each programme includes a curriculum and pedagogy, student assessment appropriate to the age range, professional development for teachers and a process of school authorization and evaluation" (IBO, 2014). Around the world there are 3,930 schools implementing at least one of the IB programs in 147 countries (2014). In Turkey there are 35 (June, 2014) high schools that implement IB DP curriculum. Except Prof. Dr. Mümtaz Turhan Social Sciences High School all of these high schools are private high schools. With the exception of

Bilkent Laboratory International School IB DP schools in Turkey implement national curriculum with DP curriculum (Kondakçı, 2014).

 IB DP curriculum was prepared for the students aged 16-19 and its curriculum includes six groups of subjects that prepare students for the higher education during two years. These groups are Group 1- studies in language and literature, Group 2- language acquisition, Group 3- individuals and societies, Group 4- sciences, Group 5- mathematics, and Group 6- the arts. The schools offer different subjects for each group and the students are required to choose at least one subject from each group. The courses can be taken in High Level (HL) or Standard Level (SL) and the students have to choose at most four high level subjects and three standard level subjects. HL courses require 240 hours of study whereas SL courses require 150 hours of study and the assessment rubrics are different for each level. It can be said that HL students are more harshly graded than SL students. For high level subjects the students receive 240 hours of education on a particular course while they receive 150 hours of education for each standard level lesson.

While offering Group 1 subjects DP schools can choose any of the following subjects; Language A: Literature, Language A: Language and Literature, and Literature and Performance (IBO, 2011). While Literature and Performance is only available at standard level, Language A: Literature and Language A: Language and Literature provide students opportunity to choose either standard level or high level. Group 1 courses are prepared for the students who have substantial experience in using the language in academic context. Group 1 courses aim to develop the following skill in the DP students; "a personal appreciation of language and literature; skills in literary criticism; an understanding of the formal, stylistic and

aesthetic qualities of texts; strong powers of expression, both written and oral; an appreciation of cultural differences in perspective" (IBO, 2013, p.12).

Group 2 subjects include Language ab initio, Language B, and classical languages. While classical languages course is only available in Latin or Classical Greek Language ab initio and Language B can be offered in many languages. Designed as language acquisition courses Language ab initio and Language B aim to have students communicate successfully in the target language and gain intercultural understanding.

IB DP examinations can be categorized as summative achievement tests which determine the achievement of the students, teachers or the courses by previously set assessment objectives (Hughes, 2003). In other words, IB DP courses are designed according to the objectives that students need to achieve till the end of the second year of IB DP. Brown (1996) argues that achievement tests can be used to measure how much learners have learned the topics included in a course and how much they achieve the objectives of a curriculum. In addition to that, Spolsky (1995) points out that achievement an achievement test is a tool for teachers by the help of which they can check the progress of their students. In the light of these arguments, it can be stated that IB DP examinations provide teachers and students a way to evaluate their teaching and learning practices. Taras (2005) defines summative assessments as the tests given at the end of instructional unit to evaluate students' learning by using certain criteria. This definition fits into the structure of IB DP examinations since the students are assessed at the end of the second year by certain sets of criteria.

Student Selection Examination (SSE) in Turkey

One of the high-stakes tests the students need to take in Turkey is SSE. Enrollment into the university education in Turkey is only possible with the Student Selection Examination (ÖSS) which is a standardized test. It is compulsory for every individual to take SSE to apply for the universities after graduating from secondary education. The Center of Selection and Placement of Students in Higher Education Institutions (ÖSYM) prepares SSE papers and the sole aim of the institution for the SSE exam is to select students whose chances of being successful are more than other students taking the exam (ÖSYM, 1982).

SSE has been prepared and administered by ÖSYM since 1979. Between the years of 1980 and 1999 ÖSYM applied two different test to accept students to the higher education; SSE (ÖSS) and Student Placement Test (ÖYS). Individuals who were successful in   SSE exam gained right to attend ÖYS exam which determined the university and the department they could attend. Between the years of 1999-2005 ÖSYM changed the exam structure into a single test with 180 questions including the subjects; mathematics, geometry, physics, chemistry, biology, Turkish language, history, geography and philosophy. ÖSYM made another change into the exam structure in 2006 and they introduced new tests to the examination; SSE 1 and SSE 2. While SSE 1 questions had to be answered by every student irrespective of their department students taking SSE 2 were required to answer questions according to their departments.

In 2010 ÖSYM introduced a new system for enrollment into higher education. Similar to the examination conducted between the years of 1980-1990 ÖSYM introduced two-phased exam system which is still in use as of 2015. Higher

Education Examination (HEE) is taken in April and the students who are successful

in this exam are eligible to take Undergraduate Placement Examination (UPE) which

is conducted in June. Both exams consist of multiple-choice questions. Irrespective

of their departments the students are required to answer 160 questions in HHE from

the all the subjects; mathematics, geometry, physics, chemistry, biology, Turkish

language, history, geography, religious studies and philosophy and the time allocated

for this examination is 160 minutes. Students who score at least 140 are permitted to

get UPE exams. There are five different exams that students can take; UPE 1

(Mathematics – Geometry), UPE 2 (Science), UPE 3 (Literature- Geography), UPE 4

(Social Studies) and UPE 5 (Foreign Language). Students can decide on the exams

they want to take according to the department they want to be enrolled in.

Conclusion

There are many studies conducted about the relationship between language

proficiency and academic achievement, education in IB DP and SSE. Studies that

were conducted by Baybliss & Raymond (2004), İnal, Evin & Saracıoğlu (2005),

Sert (2006), Wait (2009), Kazazoğlu (2013) concluded that there is a strong

relationship between academic achievement and language proficiency levels. In other

words, these studies claim that students with high proficiency levels tend to perform

better in courses other than English.

Studies of Burris and Murphy (2014) and Beaty and Tarc point out that IB DP

prepares students well for the university education. Moreover, Kondakçı (2014)

compared Ministry of National Education English curriculum with the IB DP and AP

English curricula in terms of ideology, intentions, content and assessment by using

Schiro's four curriculum ideologies.

Since, SSE is considered as one of the most important examinations in Turkey and it effects the students life drastically (Ekici, 2005), there are many studies conducted about SSE. Karakaya and Tavşancıl (2008) conducted a research on the predictive validity of the SSE. Yıldırım (2007) investigated the relationship between depression, test anxiety and social support of students the students preparing for SSE. However, there is no study conducted in Turkish context investigating the relationship between the scores of English proficiency exam, IB DP English scores and SSE.

CHAPTER 3: METHOD

Introduction

The purpose of the study is to explore; the correlation between (i) IB DP English scores and COPE scores and the correlation between (ii) IB DP English scores and Student Selection Examination (SSE) scores. Chapter two focused on the terms that are related to aforementioned examinations. This chapter includes brief information about the research design, context, instruments, and participants, how data was collected and analyzed. The chapter ends with the further clarification of the variables and the reasons they were included in this study.

Context

The study was conducted in Bilkent Erzurum Laboratory School with the participation of its graduate students. Established as a government project to promote high-quality education in the Eastern provinces of Turkey, BELS offers an entrance exam whose components have been changed through the years. Generally, the entrance exam includes open ended questions for all major subjects; social studies, Turkish, Math and sciences. The exam also includes a reasoning test. In the second phase of the admission process the students who are successful in the entrance exam are interviewed by the teachers. Students, who are successful in the exam and the interview, are admitted to the school and they start with a prep year going through an intensive English prep program. For the students in the ninth and tenth grades the school implements the curriculum of International General Certificate of Secondary Education (IGCSE). Although it was changed into English as First Language course

in 2013, the graduate students of the school took English as second language course which measured students' ability to write, listen, read and speak in English. The students who are successful in IGCSE exams are offered a place in the school to study eleventh and twelfth grades in which they study IB courses.

Participants

The sample was drawn from the students graduated from BELS. Each graduate group of students of BELS took a different IB DP English option: Language A2 between 2010-2012; Language B between 2011-2013; Language A: Language and Literature between 2012-2014. Each of the mentioned IB DP English course is different in terms of its nature and objectives. In order to be granted with a scholarship in Bilkent University the graduate students had to take minimum acceptable score from Student Selection Examination (SSE). The numbers in Table 2 and Figure 1 present descriptives and histogram of SSE scores respectively. Both SSE and COPE scores were relatively high as shown by the minimum scores. Also lower standard deviations indicated that scores did not show large dispersion. Figure 1 and 2 show histograms for COPE and SSE scores, respectively.

Table 2

Descriptive of participants in COPE and SSE

	SSE	COPE
N	119	119
Mean	381,83	106.16
Median	390.13	106.5
Mode	273,10	98.75
SD	52,22	14.28
Minimum	273,10	79,00
Maximum	517,11	142.25
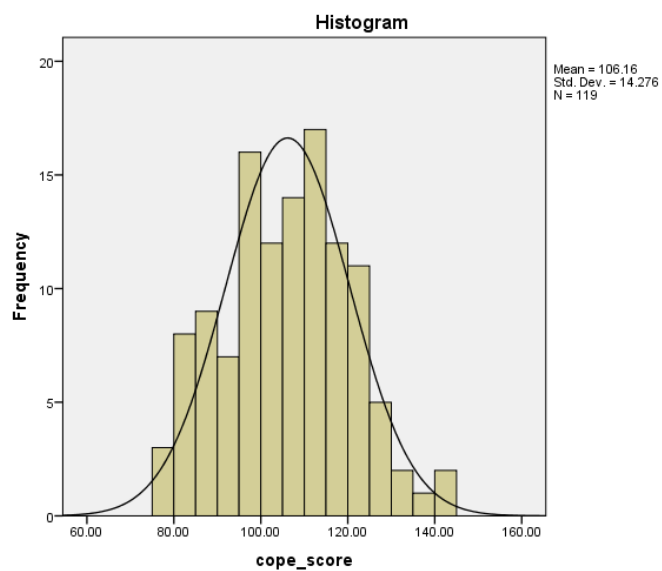
Figure 1. Histogram of COPE scores
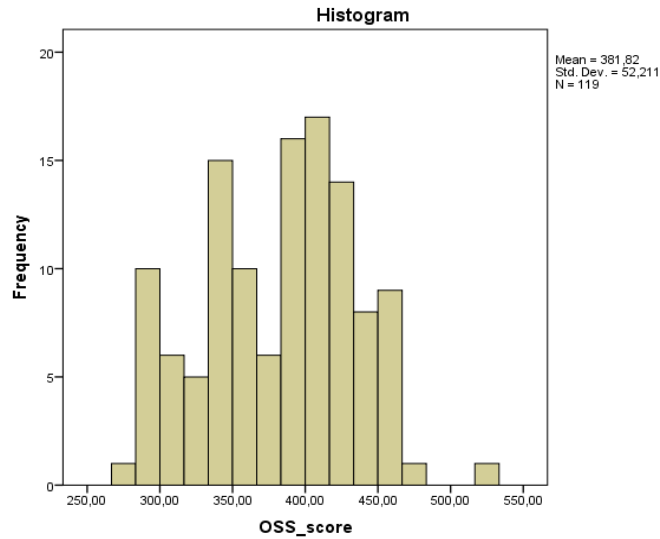
Figure 2. Histogram of SSE scores

Other relevant descriptive information regarding IB DP paths were given in Table 3. The other requirement for students who want to study in Bilkent University is to take COPE exam which measures students' English proficiency and score 83 at least (out of 150). Table 3 shows the number of BELS students who enrolled to Bilkent University after taking COPE exam.

Table 4 presents the number of students with respect to IB DP English option. Most of the students were HL options students (n=79), while there were 40 students at the low level.

Table 3
Descriptive statistics of IB-related and COPE scores

	Predicted	Actual	Scaled
Mean	5.09	5.38	68.15
Median	5	5	67
Mode	5	5	59
SD	1.05	0.748	10.74
Minimum	3	3	41
Maximum	7	7	91
N	119	119	119

Table 4
IB DP descriptives across years

Graduation Year	No of Students	IB DP English Option	No. of HL students	Number of SL students
2012	50	Language A2	11	39
2013	42	Language B	42	0
2014	27	Language A: Lang. and Lit.	26	1

Since the sample size for the present study was limited, discrimination between HL and SL was not used and students were categorized into one of the options (A, A2 and B) throughout the analyses.

Research design

Since the research questions for the present study were mainly related to the relationship seen among several variables (Gall, Gall, & Borg, 2007), correlational research design and discriminant analysis were preferred in order to answer these questions; Is there any relationship between students' university English proficiency

scores and IB DP English Proficiency exam results?; What are the factors explaining differences between three IB English options?; What are the factors explaining differences between students who are below and above COPE average (75)? Correlational research design refers to the studies in which the purpose is to discover relationships between variables through the use of correlational statistics. Discriminant analysis was conducted in order to "determine the correlation between a set of predicted variable that is in the form of categories (Gall, Gall, & Borg, 2007, p.354)

Instruments

The instruments of this study included assessment material in IB DP English options, COPE test and SSE . Each of the assessment material is described below.

Assessment in IB DP language options

Each course in IB DP requires two years of study. During two years, students' competency was measured by external assessments, evaluated by the IB examiners, and internal assessments which were internally assessed by the class teacher externally moderated by the IB examiners. At the end of the second year the students' scores from different assessments are added up and they get an overall IB score out of 100. Each year according to the students' performance IB determines grade boundaries. Students are given a number out of seven according to their scores.

Language A2
BELS students took IB DP English Language A2 between the years of 2011-2012. Language A2 course, which is no longer available, was designed for the students who already have high competency of the target language. Although it was not

entitled as language acquisition course Language A2, included further studies of the language and culture of the target language. Details of the assessment of the course are as follows in table 5;

Table 5
Language A2 assessment components and their weights

External Assessment	Weight
Paper 1: Comparative Commentary	25%
Paper 2: Essay	25%
Written Tasks	20%

Internal Assessment	Weight
Individual Oral	15%
Interactive Oral	15%

Grouped under the category of external assessment, Paper 1 which assess the students' ability to "demonstrate the analytical skills of comparison and stylistic and thematic appreciation by commenting on two texts in a clear and coherent manner" (IBO, 2003, p.30) consists of two pairs of unseen texts. HL students are expected to choose a pair of thematically linked texts and write an essay comparing and contrasting them. SL students are provided with guiding questions while comparing the texts. Paper 2 exam assesses the students' ability to "demonstrate critical thinking by constructing a relevant, organized and well-supported discussion or argument about a given topic or issue" (IBO, 2003, p.33). In paper 2 exam the students are required to write an essay either on the global issues discussed in the course or the literary works they studied in class. For the written tasks, which measures the students' ability to write in arrange of styles and registers, the students are expected to write two essays 1000-1500 words in total. Internal assessment of the course consists of two oral examinations. Individual oral commentary measures the

students' ability to engage in critical examination of a particular text. Individual oral commentary requires students to talk about an extract of the text they studied in class for fifteen minutes whereas further oral commentary, requires students to engage in any oral activity such as group discussion, role play, dramatic presentation or oral presentation for fifteen minutes (see Appendix A) Although the nature of the task remains the same in SL and HL the students are evaluated by different criteria for the oral components of the task.

Language B

2013 graduates of BELS took Language B education for two years before they graduate. Designed as language acquisition course Language B course is designed for the students who have limited knowledge of the target language. Language B course aims to provide students with the opportunities to develop their language skills and the course focuses on productive, receptive and interactive skills. Details of the assessment of the course are as follows in Table 6;

Table 6
Language B assessment components and their weights

Assessment Component	Weighting
External assessment	70%
Paper 1 (1 hour 30 minutes): Receptive skills	25%
Paper 2 (1 hour 30 minutes): Written productive skills	25%
Written assignment: Receptive and written productive skills	20%
Internal assessment	30%
Individual oral (8–10 minutes)	20%
Interactive oral activity	10%

Paper 1 assesses "a range of receptive skills such as understanding overall meaning, scanning texts for particular details, and understanding structural features" (IBO, 2012, p.31). Paper 1 consists of four texts for SL and five texts for HL students. In this reading examination the students are expected to answer questions related to the texts. The questions include gap filling exercises, matching summary sentences, and identifying the key points. Paper 2 assesses students' productive skills such as writing for a variety of purposes and it requires students to write in a specific genre in 250-400 words. As for the written assignments, SL students need to produce 1 task whereas HL students need to produce two tasks between 300-400 words in length. The written assignments are designed "to provide the student with the chance to reflect upon and develop further understanding of a core topic, as well as to develop intertextual receptive and productive skills" (IBO, 2013, p. 33). The students choose topics from the suggested list of topics offered by IB and they first read texts, such as articles or interviews, discussing or representing the issue they have chosen. After reading they are asked to write an essay discussing the issue presented in the article or interview. The essay is submitted with a 100 word rationale in which the students explain their ways of analyzing and reasons for choosing the text. In individual oral activity, which lasts for fifteen minutes, the students are required to talk about a photograph related to the topics they previously discussed in the class. After the students deliver their speech about the photograph, they are expected to engage in discussion with the teacher. Interactive oral activity can be any type of oral activity that Language B teacher decides on. Students complete three further oral activities one of which should be based on a listening part (see Appendix B). While assessing SL and HL students for the internal assessment, a Language B teacher uses different criteria for different level of students.

Language A: Language and literature

BELS students took Language A: Language and Literature course between the years of 2012-2014. Designed as group one course Language A: Language and Literature aims to "develop in students skills of textual analysis and the understanding that texts, both literary and non-literary, can be seen as autonomous yet simultaneously related to culturally determined reading practices" (IBO, 2012, p.5). Students who take two group one courses are provided with a bilingual IBDP diploma. The details of the assessment of Language A: Language and Literature course is illustrated in Table 7;

Table 7
Language A: Language and literature assessment components and their weights

Assessment Component	Weighting
External assessment (4 hours)	70%
Paper 1: Comparative textual analysis (2 hours)	25%
Paper 2: Essay (2 hours)	25%
Written tasks	20%
Internal assessment	30%
Individual oral commentary	15%
Further oral activity	15%

Language A: Language and Literature paper 1 assesses the students' ability to analyze and compare and contrast texts. In paper 1, SL students analyze only one text following the guiding questions whereas HL students compare and contrast two unseen, thematically combined texts. Paper 2 assesses students' understanding of the literary works studied in class. In the exam, the students are given six generic essay questions and they are expected to answer one of the questions considering the

literary works they studied in class. For paper 2 HL students are required to study four texts whereas SL students study three texts. HL students are required to produce four written tasks throughout the course. At least one of the written tasks is based on the topics they discuss in the class and they are expected to produce a formal essay. For the other written tasks the students are expected to produce a creative body of writing such as pastiche, or diary entry considering the literary works they study in class. SL students produce two written tasks one of which should be a formal essay. The written tasks are expected to be between 800-1000 words in length plus a rationale. In the individual oral commentary, which assess students' ability to engage in critical discussion about the text previously read in class, the students are given extracts and two guiding questions related to literary works they studied in class. They are expected to talk about the extract and answer the questions in fifteen minutes. Throughout the course students complete two further oral activities and their oral abilities are assessed. Further oral activities can be in any form of oral activity such as discussion, presentation or speech (see Appendix C).

COPE exam

In order to be granted with a scholarship in Bilkent University each group of Bilkent Erzurum Laboratory School students entered COPE exam which is prepared by the Bilkent School of English Language. Cope Exam is a standardized English proficiency exam that measures test takers' ability to read, listen, speak and use English. Test takers' scores are calculated out of 150 and they are expected to get at least 83,5 in order to start studying in their departments otherwise they are expected to pay the tuition for the prep year in Bilkent University.

The details about the assessment of the test are as follows;

Reading paper

Reading paper consists of multiple-choice questions and the students code their answers in optic forms. Reading paper measures the test takers' ability to guess meaning from the context; find main ideas, supporting details and specific information; and make inferences from the text (COPE, 2013). In each part there is a reading passage whose length shows variety according to the part. In this part four wrong answers cancel out one correct answer. Students are given one hour and twenty minutes to answer thirty-five questions. Parts of the Reading section are given in Table 8;

Table 8
Structure of COPE reading paper

Part	Content
Part one	Three short texts of approximately 250-300 words each
Part two	Two middle length texts of between 600 and 800 words each
Part three	One long text of between 1,300- 1,600 words

Listening paper

Listening paper measures the test takers' ability to listen for main ideas, supporting ideas and specific information; and infer meaning and identify speaker's tone and purpose (COPE, 2013). Listening paper consists of two lectures and related questions. Each lecture is fifteen minutes long and the students are provided with a structured note taking sheet to use while listening. After they listen to the first lecture the students have thirty minutes to check their notes. After listening the second lecture the students are distributed the questions and they have twenty-five minutes to answer thirty questions. For each lecture the students are required to fill in a chart

and answer five multiple-choice questions. Maximum point that the students can get from this section is thirty.

Language paper

Language paper measures the test taker's ability to recognize grammatical structures; recognize words and use them in the correct context; produce grammatical structures according to the meaning required; manipulate world forms by adding suffixes and prefixes (COPE, 2013). Language paper consists of three parts. Students are given forty minutes to answer thirty-five questions in language paper. The maximum point they can get from this section is thirty-five (Table 9).

Table 9
Structure of COPE language paper

Part	Format
Part one	Two cloze tests
	Ten gaps in each text using one word only
Part two	Gap fill
	Completion of total seven gaps by choosing the correct word from the box.
	There are five extra words in the box
Part three	Gap fill
	Completion of a total of eight gaps in a text by making one or two changes in a word given

Writing paper

Writing paper of COPE measures the test taker's ability to comprehend a given prompt; analyze and reflect on the argument; generate and organize ideas; justify, support and exemplify; and produce accurate and relevant prose (COPE, 2013). In

writing paper the students are asked to write a 350 word essay in an hour. The

students are given two prompts which are not longer than five lines. The maximum

point the students can get from this paper is thirty points (see Appendix D).

Speaking paper

In order to assess students' ability to engage in discussions the speaking test is

carried in a paired format. The students have fifteen minutes to complete the task and

the maximum point they can get is twenty (Table 10) (see Appendix D).

Table 10
Structure of COPE speaking test

Part	Format
Part 1	Three personal or general questions
Part 2	The students are given two prompts
	The students choose one prompt and talk about it for one minute

Student Selection Examination

Since 2010 SSE examination has been conducted as two phased multiple-choice

examination in order to select students for higher education. Four incorrect answers

cancel out one correct answer in both parts of the examination. In the first phase of

the examination (HEE) which is taken in April, the students are required to answer

questions from mathematics, geometry, physics, chemistry, biology, Turkish

language, history, geography, religious studies and philosophy in 160 minutes. The

candidates are responsible from the curriculum of ninth grade curriculum and they

are required to answer all questions irrespective of the departments they have chosen

to study in eleventh and twelfth grades.

The second phase of the examination (UPE) is held in June and there are five

different examinations that students can choose from according to the departments

they studied in high school and the departments they want to study in the university.

The details of the examinations are given in tables 11, 12, 13, and 14.

Table 11
 UPE 1 Mathematics and geometry test

Subjects Covered	Number of Questions	Percentage of Questions	Duration
Mathematics	50	% 62,5	75 minutes
Geometry	30	% 37,5	60 minutes
Total	80	% 100	135 minutes

Table 12
 UPE 2 Science test

Subjects Covered	Number of Questions	Percentage of Questions	Duration
Physics	30	%33,3	45 minutes
Chemistry	30	%33,3	45 minutes
Biology	30	%33,3	45 minutes
Total	90	%100	135 minutes

Table 13
UPE 3 Turkish language and literature- geography test

Subjects Covered	Number of Questions	Percentage of Questions	Duration
Turkish Language and Literature	56	% 70	85 minutes
Geography 1	24	% 30	35 minutes
Total	80	% 100	120 minutes

Table 14
UPE 4 Social sciences test

Subjects Covered	Number of Questions	Percentage of Questions	Duration
History	44	% 49	65 minutes
Geography 2	12	% 15	25 minutes
Philosophy and religious studies	32	% 36	45 minutes
Total	90	% 100	120 minutes

At the end of the examinations raw scores of that the students take from HEE and UPE test ranks in the test are calculated. In order to find an overall score of the students their high school grade point average, their schools achievement score in the SSE examinations are added up. Departments in the universities require combinations of the students HEE scores and department- related scores from UPE tests.

Each phase of the examination measures different sets of constructs. In HEE examination the students' levels of knowledge and comprehension are measured through multiple-choice questions. Recalling, matching, defining, calculating and commenting are the constructs that characterize the questions in the UPE test (Koç, Çftçi & Sönmez, 2013). Second phase of the examination, UPE tests, measure the students ability to analyze, infer, make associations, synthesize, compare and contrast, (Koç, Çftçi & Sönmez, 2013) the knowledge they have from various subjects. According to Karakaya and Tavşancıl (2008) SSE provides the stake-holders with a measurement of students' cognitive level since the questions in the examination are clearly aligned with the Bloom's taxonomy of educational objectives which refers to the different levels of thinking skills.

Method of data collection

In the present study, BELS administration provided information related to IB DP. Students' SSE and COPE scores were given by Bilkent University School of English Language of Bilkent University. Then two sets of data were linked using citizen IDs of the students. IDs were only used for linking purposes and no attempt to identify any student was made during the study. Also results were reported in group-level, not individual level. As a result, the following information was collected for each student in the sample: SSE score, COPE score, IB predicted grade which were given by students' English teachers at the end of the final year of IB DP (out of 7), and IB scaled total (out of 100).

Method of data analysis

First, data were checked for missing values and it was seen that data included no missing. After that, several correlational analyses were conducted. First correlation between IB and COPE scores was calculated. Then correlation between IB and COPE scores and between COPE scores and SSE scores were investigated. In all correlational analyses, Pearson Product Moment correlation was employed. Furthermore, these analyses were repeated for years and IB DP options. In all analyses, alpha level was set to 0.05. Then two discriminant analyses were conducted in order to determine the correlation between the variables in the form of categories (Gall, Gall & Borg, 2007). First one was to explain differences between three IB DP language options in terms of the variables Order of Preference, COPE Score, Predicted IB Score, Actual IB Score, Scaled IB Scores, and Rank in SSE. The second discriminant analysis was conducted to classify students into two groups, those below and above the COPE average. Discriminant analysis results were investigated in terms of standardized coefficient, canonical correlations, and classification rates.

<center>**List of variables**</center>

IB related variables

Predicted grade: By closely examining the students' performance in classes during two years and mock exams IB DP teachers give a predicted grade out of seven. Predicted grades are given for two reasons; to find out if there is a discrepancy between the students' performance in the classroom and in the exams and if their exams papers are lost in the shipment the predicted grades are considered as the students' final IB DP grade. Predicted grades are included in this study to find the relationship between the participants performance in IB DP English classes and English proficiency exam.

Scaled total: There are five different assessments in IB DP English options and each assessment has a different weight. Students' scores from each assessment are added up and they get a cumulative score out of 100 at the end of the second year.

Actual grade: IB DP examinations are held in IB schools ever year and students taking the same course all around the world are assessed with the same questions and criteria by IB DP examiners. According to the students' performance IB DP determines a different set of grade boundaries each year. Grade boundaries are designed to find the equivalence of the scaled total grade out of seven. For instance, seventy in scaled total grade might equal to five one year or it might equal to four next year according to the students' performance. Actual grades were used in this study in order to determine the relationship between the participants' performance in IB DP English options in a specific year and their performance in an English proficiency exam.

COPE scores

COPE is a standardized language proficiency exam prepared by Bilkent University for the students who desire to pass English prep year. Participants COPE scores were included in this study so as to determine the relationship between the variables related to different IB DP English options and English proficiency test.

SSE related variables

Rank in SSE: Student selection examination is held every year in order to choose students for higher education. Students rank in SSE show the students' performance in the examination. Rank in SSE was included in this study in order to explore the relationship between students' IB DP English scores and their performance in SSE. Order of preference: After SSE students who get satisfactory scores make an official list of the universities and departments they want to be placed. Order of preference shows the rank of the university and its department the student was placed in the list. This variable is also an indication of students' cognitive level.

CHAPTER 4: RESULTS

The purpose of the study is to explore; the correlation between (i) IB DP English scores and COPE scores and the correlation between (ii) IB DP English scores and Student Selection Examination (SSE) scores. Chapter three included methodology and detailed information about the instruments. In this chapter results of the analyses are presented under the related research questions. First, correlational analyses were presented. Correlations between investigated in the following order: (i) IB Scores and COPE scores, (ii) IB and COPE scores with respect to Options, (iii) COPE scores and SSE scores, and (iv) COPE scores and SSE scores with respect to IB DP English options. Then results of the discriminant analyses were given. First results of a discriminant analyses were given conducted to explain differences between three IB DP language options. Then, results of the second discriminant analysis were given conducted to find out differences between those above and below the average of COPE exam (75 out of 150). At the end of the chapter, brief summary of the results is also given.

Research question 1: Is there any relationship between students' (i) IB DP English scores and COPE scores; (ii) IB DP English scores and SSE scores?

Correlation between IB scores and COPE scores

The correlation between students' predicted IB grades, which are given by their English teachers, and students' actual IB scores was found to be .700 (df=117, p < .001). It is interesting to note that there is a strong correlation between two types of scores. Thus, it can be said that IB teachers can predict students' actual scores prior to the exam. Similarly, correlations were computed between students' COPE scores,

IB predicted grades assigned by the teachers and actual IB grades. The most prominent finding is that the correlation between students' COPE scores and IB predicted grade (r=.462, df=117, p < .001) is stronger than the correlation between students' COPE scores and actual IB grades (r=.379, df=117, p < .001). These results indicated that teachers' predictions related to students' performances during two years are more predictive than students' actual IB grades with respect to COPE exam requirements. In terms of explained variance, teachers' prediction explained 21.34% of the total variance in COPE scores, while actual IB grades did just 14.31%.

Correlation between IB and COPE scores with respect to IB DP English options

Table 15 illustrates the correlations between COPE scores, IB predicted grades and IB actual grades for each group separately. The correlation between COPE scores and IB predicted grades in group B is stronger than (r=.678) group A and A2 which shows that the relationship between COPE scores and IB predicted grade is more significant in Language B option. The weakest correlation (r=0.499) between COPE scores and IB predicted grade belongs to group A. It can be deduced from these results that actual (scaled) grades predict COPE scores better for the option A, teachers' prediction worked well to estimate COPE scores for the options A2 and B.

Table 15
Correlations between COPE scores, IB predicted and actual grades with respect to options

	IB Scores		
Option	Predicted	Actual	Scaled
A	.499[**]	.608[**]	.613[**]
A2	.527[**]	.405[**]	.507[**]
B	.678[**]	.495[**]	.701[**]

*. Correlation is significant at the 0.05 level (2-tailed)

**. Correlation is significant at the 0.01 level (2-tailed)

Correlations between IB, COPE, and SSE rank

Table 16 shows that the correlation between SSE rank, which represents the cognitive development of the students, and COPE score is negative with the correlation coefficient r=-.189. This was expected since ranks in SSE are reverse order. In other words, low ranks correspond to higher cognitive levels. That is, the relationship between SSE rank and COPE score is positive in terms of achievement. Similarly the correlation between SSE rank, IB predicted grade and IB actual grade is positive in terms of achievement although the correlation is found negative with the correlation coefficients r=-.297 and r=-.395. Actual IB scores was the variable which has the highest correlation with SSE ranking and, in turn, highest predicted power. Furthermore, the correlation between predicted and actual (scale) IB scores was found to be .700 (.745). Nevertheless, predicted IB scores seemed to have a weak relationship between ranks in SSE. Taking all these calculations into consideration, it can be put forward that students' cognitive development levels are not strongly correlated with language proficiency.

Table 16
Correlation between students' COPE scores, IB predicted grades and IB actual grades

	Actual	Scaled	COPE	Rank in SSE
Predicted	.700[**]	.745[**]	.462[**]	-.297[**]
Actual		.878[**]	.379[**]	-.395[**]
Scaled			.379[**]	-.366[**]
COPE				-.189[*]

*. Correlation is significant at the 0.05 level (2-tailed)

**. Correlation is significant at the 0.01 level (2-tailed)

When the same relationship was investigated with respect to IB language options, it was observed that ranks in SSE could not be predicted by IB variables (for the option A, see Table 17). However, the relationships within IB-related variables were found to be considerably higher.

Table 17
Correlation between students' COPE scores and SSE scores for option A

	Actual	Scaled	COPE	Rank in SSE
Predicted	.562**	.685**	.678**	-.227
Actual		.838**	.495**	-.285
Scaled			.701**	-.339*
COPE				-.209

*. Correlation is significant at the 0.05 level (2-tailed)
**. Correlation is significant at the 0.01 level (2-tailed)

Similarly, for the options A2 and B, Table 18 and 19 show that the correlation between students' COPE scores and SSE scores is weak which means there is not a strong relationship between COPE scores and SSE rank. The same pattern is observed in these tables as well. Ranks in SSE do not have strong relationships with IB scores but IB scores (actual, scaled and COPE) are inter-related.

Table 18
Correlation between students' COPE scores and SSE scores for option A2

	Actual	Scaled	COPE	Rank in SSE
Predicted	.487*	.575**	.499**	-.146*
Actual		.895**	.608**	-.210*
Scaled			.613**	-.178*
COPE				-.209*

*. Correlation is significant at the 0.05 level (2-tailed)
**. Correlation is significant at the 0.01 level (2-tailed)

Table 19
Correlation between students' COPE scores and SSE scores for option B

	Actual	Scaled	COPE	Rank in SSE
Predicted	.654**	.666**	.527**	-.237*
Actual		.913**	.405**	-.420**
Scaled			.507**	-.352*
COPE				-.203*

*. Correlation is significant at the 0.05 level (2-tailed)

**. Correlation is significant at the 0.01 level (2-tailed)

Research question 2: What are the factors explaining differences between three

IB English options?

Discriminant analysis across groups based on IB options

Results of the cluster analysis are given in Table 20. Results indicated that both of

the functions were overall significant ($p < .001$) (Table 19).

Table 20
Significance of the discriminant functions

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	.221	171.574	12	.000
2	.778	28.439	5	.000

In the cluster analysis, three groups were used (A, A2, and B). Means (SDs) of the

three groups and the whole one were given in the Table 21.

Table 21
Mean and SD of the groups (based on IB DP English options)

Options	Variables	Mean	SD
A	Order of Preference	2.96	1.76
	COPE Score	107.69	16.08
	Predicted IB Score	4.22	0.70
	Actual IB Score	4.67	0.68
	Scaled IB Scores	57.11	6.26
	Rank In SSE	785.14	50.50
A2	Order of Preference	3.58	2.24
	COPE Score	105.18	12.97
	Predicted IB Score	5.10	1.05
	Actual IB Score	5.44	0.67
	Scaled IB Scores	66.04	8.25
	Rank In SSE	587.52	48,09
B	Order of Preference	2.90	1.48
	COPE Score	106.36	14.79
	Predicted IB Score	5.64	0.85
	Actual IB Score	5.76	0.53
	Scaled IB Scores	77.76	6.76
	Rank In SSE	492.02	41.42
Total	Order of Preference	3.20	1.91
	COPE Score	106.16	14.28
	Predicted IB Score	5.09	1.05
	Actual IB Score	5.38	0.75
	Scaled IB Scores	68.15	10.74
	Rank In SSE	598.78	47.26

The first discriminant functions explained 71.7% (R=.847) of the total variance between three groups, while the second did 22.2% (R=.471). Details of the Eigenvalues were given in the Table 22.

Table 22
Eigenvalues groups

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	2.529[a]	89.9	89.9	.847
2	.285[a]	10.1	100.0	.471

Standardized canonical coefficients were used to assess the strength of the relationships between variables and group discrimination. Results showed that in both functions actual and predicted IB scores were the most significant predictors to estimate group membership.

Table 23
Standardized canonical discriminant function coefficients for IB-related variables

	Function	
	1	2
Order of Preference	-0.032	0.402
COPE Score	-0.678	-0.539
Predicted IB Score	0.039	0.540
Actual IB Score	-1.159	1.491
Scaled IB Scores	2.067	-1.065
Rank In SSE	-0.004	-0.066

The relationship between the variables and the discriminant function were evaluated using the structure matrix given in Table 24.

Table 24
Correlations between the predictor variables and discriminant functions

	Function	
	1	2
Order of Preference	0.678	0.287
COPE Score	0.365	0.593
Predicted IB Score	0.341	0.410
Actual IB Score	-0.046	0.293
Scaled IB Scores	-0.135	-0.199
Rank In SSE	-0.007	-0.127[*]

Group means were estimated as -1.91, -0.62, and 0.026 for function 1 and -0.731,

0.577, and -0.217 for function 2. The means can be said well separated and three IB

options were differed considerably. Classification results were also promising. 80.7%

of the original cases were correctly classified. The same ratio the cross-validated

classification was 76.5%. Thus it can be concluded that IB related variables were

significantly differed across IB DP Language options.

Research question 3: What are the factors explaining differences between students who are below and above COPE average score set as 75?

Discriminant analysis across groups based on COPE scores

Only a single discriminant function was defined since there are two groups (below

and above the COPE average). The discriminant function was overall significant (p <

0.001) (Table 25) Canonical correlation Total variance explained by the discriminant

function was 0.458 (R=.458).

Table 25
Significance of the discriminant functions

Test of Function(s)	Wilks' Lambda	Chi-square	Df	Sig.
1	.790	26.866	6	.000

Mean and SD of the variables across 3 IB groups and whole body were given in

Table 26.

Table 26
Significance of the discriminant functions for three IB DP English option groups

Cope discriminant analysis with respect to options	Variables	Mean	SD
A	Actual IB Score	671.70	51.59
	Rank In SSE	3.05	1.54
	Predicted IB Score	4.63	0.82
	Actual IB Score	5.14	0.72
	Scaled IB Scores	64.49	9.47
A2	Order of Preference	531.74	42.2
	Rank In SSE	3.34	2.2
	Predicted IB Score	5.52	1.07
	Actual IB Score	5.6	0.71
	Scaled IB Scores	71.52	10.82
B	Order of Preference	598.782	47.26
	Rank In SSE	3.2	1.91
	Predicted IB Score	5.09	1.05
	Actual IB Score	5.38	0.75
	Scaled IB Scores	68.15	10.74

Standardized coefficients indicated that the most significant variables were predicted

IB scores to predict group membership of above and below COPE scores.

Table 27
Standardized canonical discriminant function coefficients to predict group
membership (below and above 75)

	Function
	1
Rank In SSE	-.060
Order of Preference	.237
Predicted IB Score	.915
Actual IB Score	.049
Scaled IB Scores	.040

In terms of the relationships, predicted IB scores were the variables with the highest

correlation with discriminant function (Table 28).

Table 28
Correlations between the predictor variables and discriminant functions

	Function
	1
Predicted IB Score	.905
Scaled IB Scores	.674
Actual IB Score	.624
Rank In SSE	-.291
Order of Preference	.146

Group means (-0.533 for those below the COPE average and 0.490 for those above)

indicated that using the variables entered into discriminant function two groups of

students could be well separated. Classification results were also confirmed the

predictive power of the discriminant function. 67.2% of the cases were correctly

classified. After the cross-validation, the ratio was found to be 65.5%.

Summary

Results related to the correlational analyses indicated that scores of SSE, an exam to assess cognitive traits, were not correlated with IB scores but COPE scores, an English language proficiency exam, can be predicted by IB-related scores (actual, predicted and scaled). Analyses also showed that the relationships significantly varied across IB options. Discriminant analysis on IB DP groups revealed that actual and scaled IB scores were the strongest predictors explaining differences between three IB groups. The other discriminant analysis on COPE scores indicated that, predicted IB scores could be used to effectively categorize students into two groups, above and below the COPE average.

CHAPTER 5: DISCUSSION

Introduction

Chapter four included the results of the correlational analysis and discriminant analysis. The main results can be summarized as;

(1) The correlation between IB DP English and COPE scores is not strong. On the other hand there is a strong correlation between IB DP English predicted grades and COPE scores.

(2) The correlation between COPE scores and IB DP English predicted grades has the strongest correlation for option B.

(3) There is no statistically significant relationship between SSE scores and English language proficiency scores.

(4) Discriminant analysis of the variables revealed that IB actual and predicted scores were the strongest predictors explaining the differences between three IB DP language options.

(5) Discriminant analysis indicated that predicted grades could be used to categorize students into two groups as students who are above and below COPE average score set as 75.

This chapter includes a discussion of the major findings under the related research questions, implications of this study for further practice and research. At the end of the chapter, limitations of the present study were given.

Discussion of the major findings

Research question 1: Is there any relationship between students' (i) IB DP English scores and COPE scores; (ii) IB DP English scores and SSE scores?

Correlation between IB scores and COPE scores

There is not a strong correlation between IB actual grades and COPE scores. This shows that, the requirements of IB assessment and COPE exam do not correspond significantly. While Group 1 and Group 2 courses in IB strives to develop critical thinking skills as well as language proficiency (IBO, 2013), the COPE exam measures how well an individual performs in English by the help of structured tests (COPE, 2013). COPE exam measures communicative competence of the test-takers. As Llurda (2000) and Stern (1996) discuss language proficiency exams that measure communicative competence focus more on how the test-takers communicate and express themselves rather than how they use forms or structures of the target language. On the other hand, in IB DP English papers the examiners pay close attention to the language and use of certain forms. Moreover, the emphasis on critical thinking skills in IB Group 1 and Group 2 makes IB language option differ from language acquisition courses or exams.

On the other hand, for all IB English options, there is a strong relationship between IB predicted grades and COPE scores. "The predicted grade is the teacher's prediction of the grade the candidate is expected to achieve in the subject, based on all the evidence of the candidate's work and the teacher's knowledge of IB standards" (IBO, 2013). Therefore, the performance of students in the classroom shows how well they can perform in language proficiency exam since by careful observations a teacher can decide if the students can express themselves in verbal and written form. The strong correlation between IB predicted grades and COPE

59

scores highlights the importance of formative assessment because as Hughes (2003) points out formative assessment tests students at the end of the instructional unit by previously set criteria. Throughout two years IB English teachers can formatively assess students' performance and give predicted grades which show how well they have mastered the language. Formative assessment methods gains more importance when assessing language proficiency (Davidson & Leung, 2012). However, although they can practice the assessment requirements throughout two years, IB assessments, especially paper 1 and paper 2 can be considered as summative. Thus, a teacher can predict the students' performance to only some degree. Moreover, exam anxiety that the students feel while taking sit-in exams or recording their individual oral commentaries might have caused students to get lower grades in IB exams and they could not perform well in the target language. While explaining the disadvantages of high-stakes exams Embse and Hasson (2012) stress the factor of exam-anxiety which causes students to achieve lower grades than they usually get. As IB DP examinations can be considered as high-stakes exams which determine if the student can get high school diploma or not it is necessary to consider the exam-anxiety when evaluating the students' performance and predicting their IB DP grades.

Correlation between IB scores and COPE scores with respect to options

According to the results, for option A2, it is clear that correlation between COPE scores and IB predicted grades is more significant than the correlation between COPE scores and IB actual grades. Since predicted grades are given according to how the students perform in the classroom the teachers of A2 course were able to envision how students would perform in the COPE exam. Grouped under Group 2 Language A2 course was designed for students to improve their language skills by using literature (IBO, 2003). Therefore, the level of language proficiency that the

teachers of A2 course can measure can be aligned with the level of the achievement in a standardized language proficiency exam. On the other hand, there is a weak correlation between COPE scores and IB actual grades for option A2 which shows that the requirements and the contents of the exam might not have been made clear prior to the exam. Open ended question structure that the students had to answer in Paper 1 and Paper 2 of language A2 course create problem for the students since they lack any sort of structure.

Similarly, the correlation between COPE exam scores and IB predicted grades is more significant than the correlation between COPE score and IB actual grades for option B. Designed as language accusation course Language B aims to develop all four language skills; reading, listening, writing and speaking rather than focusing on developing literary skills (IBO,2013). Therefore, the requirements of this course are closely aligned with the requirements of COPE exam which also measures students' four skills. When the assessment criteria of Language B course and COPE are taken into consideration it is clear that they focus on the three communicative competence that Canale and Swain (1980) suggest; grammatical competence, sociolinguistic competence and strategic competence. Language B and COPE measures grammatical competence since they both focus on use of grammar as a way of communication. Language B and COPE are aligned in terms of measuring sociolinguistic competence since students are given a certain context or situation in the exam papers. Similarly, strategic competence is also assessed by both examinations since they focus more on how the test-taker express themselves rather than use of certain structures (Kathleen & Kenji, 1996). For language B course it is reasonable to claim that how well the students perform in the classroom and mock exam can be a strong determinant for the achievement in language proficiency exams such as COPE and TOEFL since the

students are already familiar with the requirements of these exams. When compared to the correlations found for A2 option the correlation between COPE scores and IB actual grades is more significant for language B option which results from the proximity between the assessment requirements of COPE exam and language B course.

Unlike for the options A2 and B, the correlation between COPE scores and IB predicted grades is weaker for option A. The correlation between COPE scores and IB actual grades is stronger than for option A when compared to the correlation between COPE scores and actual IB grades in options A2 and B. Discrepancy between the IB predicted grades and COPE scores for option A might stem from the fact that teachers made predictions according how they perform in the class and mock exams and they kept the standards higher than the standards of actual IB assessments. The year 2014 was the first year that the BELS administration made a radical decision to change the English option from second language courses such as Language A2 and Language B to a first language option Language A: Language and Literature. While keeping the standards of the assessments high the teachers actually demanded their students to work harder which resulted in a high level of achievement in English. The other reason for the discrepancy is that the requirements of Language A: Language and Literature assessments are not well defined when compared to the other courses. Finally, 2014 graduates were the first group to take a Group 1 course instead of a Group 2 course. In 9th and 10th grades the students took IGSCE English as a Second Language course which was designed for students to develop their receptive and productive skills which can be considered as simplified version of IB Language B course. Thus, it can be said that the students had difficult

times in terms of adopting themselves from a language acquisition course to a first language course.

When the mean scores of COPE exam is evaluated according to three years, it is seen that all groups of students who took different IB English options performed similarly in the exam ($\mu=105$ for option A2, $\mu=106$ for option B, $\mu=107$ for option A). When the contents and the assessment objectives of these IB English options are considered students who take Group 2 courses Language A2 and Language B were already expected to perform well in COPE exam since there is a close proximity between the assessment aims and objectives of the courses and COPE exam. However, the graduates of 2014 who took Language A: Language and Literature as their English option could perform well in the exam after they take courses about the requirements of the COPE exam. Language A students were thought to reach a certain level of English proficiency by exposing students more with the target language. That is the main four skills; writing, reading, listening and speaking, were taught students not explicitly but implicitly. The teachers of Language A course did not stress the strategies or methods of four language skills which is what Group 2 teachers did in the previous years. However, Language A students had to take courses about the parts of COPE exam in order to be aware of the skills that they had already acquired.

Correlation between COPE scores and SSE scores

In the studies that were conducted in university level, the researchers found out that there is a statistically significant correlation between academic achievement and the level of English proficiency. In the correlational study that was conducted Bayliss and Raymond (2004) discovered that the relationship between the academic achievement and English language proficiency corroborates for the university

students who study in English medium classes. Similarly, in the correlational study conducted Watt and Gressel (2009) found out that there is a strong correlation between TOEFL scores and academic success for international engineering students.

Interestingly, this study discovered that there is no statistically significant relationship between the students' COPE scores and SSE scores. SSE does not include any part related to English language proficiency. However, the results refute the conception that if a student can perform well in all subjects he/she can also perform well in English. This finding highlights the validity of different intelligence types.

Research question 2: What are the factors explaining differences between three IB English options?

Discriminant analysis across groups based on IB options

Results of the discriminant analysis on IB language options indicated that actual and predicted grades were the predictor variables most associated in explaining differences between 3 options. When all variables were considered together, predictor variables classified 80.7% (cross-validated=76.5%) of the cases correctly. Furthermore, canonical correlations were found to be high for the discriminant analysis. Actual IB scores and scaled scores differ considerably each year and for each option. Scaled score is what an IB student gets over 100 when the scores of each assessment a student get are added up. IB examinations assessments can be considered as criterion referenced assessments, the examiners evaluate students' papers and voice records according to certain criteria (IBO, 2014). However, each year students all over the world perform differently than the previous year which brings a norm- referenced aspect to IB assessment. According to how the students

taking the same course all over the world perform in the examinations IB creates grade boundaries for each course and for each assessment requirement. Scaled totals of students (out of 100) are transferred into actual grades (out of seven) according to grade boundaries prepared each year. Therefore, the students' performance and their achievement in internal and external assessments change the grade boundaries every year.

Research question 3: What are the factors explaining differences between students who are below and above COPE average score set as 75?

Discriminant analysis across groups based on COPE scores

Discriminant analysis conducted to explain differences between students whose COPE scores above and below average showed that predicted IB scores given by teachers were the predictor variable mostly differed between two student groups. Students' order of preferences was another factor differed significantly between two groups. Correct classification rate calculated for that discriminant analysis was 67.2% (Cross validated= 65.5%). Thus predicted scores did not have a high predictive power. It is important to point out again that predicted grades are given according to the performance of the students in the classroom, and how the students perform differ in every single IB English option (Bunnell, 2011). The predicted grades show us how well a student can perform in COPE exam up to some extent. If an IB DP English student demonstrates critical thinking skills in the target language he/she is likely to have a high predicted grade. On the other hand, if the students cannot adopt themselves to the requirements of the English proficiency exam they might get an unexpected grade in COPE and correct classification rate is 67.2% for this reason.

Implications for practice

After conducting this study, it can be suggested that if the students have substantial background of English it is better to offer a Group 1 course instead of a Group 2 course. As educators we are obliged to provide students with enough challenge and offering a Group 2 course is enough challenge for the students. When compared to Group two courses, Language A2 and B, Language A: Language and Literature course is more challenging for the students because in this option the students do not focus on four language skills, instead they are expected to have acquired these skills before they start studying this course. However, throughout two years of study, the students naturally develop four skills by being exposed to the target language as it is used in its original context. The texts and the assessments used in Language A: Language and Literature course aims to develop higher level of thinking skills which students benefit more when they start studying in the university level. In other words, offering group two course can seem as an easy solution for the students to pass language proficiency exams. However, in the long run the students benefit more from a Group 1 course since they become more familiar with higher level of texts written in the target language and acquire more critical thinking skills. Moreover, if a student whose first language is not English takes a Group 1 course instead Group 2 as well as his /her first language he/she becomes eligible to take a bilingual IBDP diploma. Students can use this bilingual diploma when applying universities abroad and validate their English proficiency. For the schools promoting international education a bilingual diploma is essential in terms of providing students with the opportunity to acquire proficiency in two languages.

On the other hand, a standardized language proficiency test such as TOEFL and IELTS plays an important role when applying universities, both national and

oversees, and jobs. Therefore, it is a schools duty to prepare students for their future academic and professional life. In order to prepare students for the future, the teachers of both Group 1 and Group 2 courses should offer some tutorials on the assessment types and requirements of major standardized language proficiency tests.

Implications for further research

It was suggested that the future researchers should investigate how different groups of IB DP graduates perform in the university level English medium classes. The comparison between the achievements of former Language B, A2 and A students in the university level English medium classes can be a competent topic for a new research. Also a quantitative study including a survey study to get in-depth information about the relationship between students English predicted grades and actual scores of IB DP English examinations can be investigated in new studies.

Limitations

The first limitation of this study is that the results cannot be generalized to all students who study IB DP in other private or national schools in Turkey, because BELS do not accept new students for 11th and 12th grades. Students who studied prep, 9th and 10th grades can continue to study 11th and 12th grades in BELS. All graduate students who are the samples of this study studied at BELS in prep, 9th and 10th grades.

Since BELS offered only three IB DP English options so far, the study focused only three courses in IB DP; Language A2, Language B, and Language A: Language and Literature. There are other English options in IB DP curriculum such as Literature

and Performance, Language ab intio, and Classical languages which cannot be included into this study as instruments.

REFERENCES

About the TOEFL test. (2014). Retrieved from http://www.ets.org/toefl/institutions

Allan, M. (2002). Cultural borderlands: Cultural dissonance in the international

    school.   *International School Magazine, 2*, 42.

Amrein, B., & Berliner, D (2002). High-stakes testing, uncertainty, and student

    learning. *Education Policy Analysis Archives*, *10*(18).

Armstrong, T. (1994). Tips for teaching: Differentiating instruction to include all

    students. *Preventing School Failure*, *51*(3), 49-54

Bachman, L. F. (1990) *Fundamental considerations in language testing*.

    Oxford: Oxford University Press.

Banks, C. A. M., & Banks, J. A. (1995). Equity pedagogy: An essential component

    of multicultural education. *Theory into Practice, 34*, 152-158.

Bayliss, D. D., & Raymond, P. M. (2004). The link between academic success and

    L2 proficiency in the context of two professional programs. *The Canadian*

    Modern Language Review, *61*(1), 72-76.

Bilkent History. (n.d.). In Bilkent Erzurum Laboratory School. Retrieved from

    http://bels.bilkent.edu.tr/about/prof-dr-ihsan-dogramacinin-mesaji/

Brown, J. D. (1996). *Testing in language programs*. New Jersey: Prentice Hall

    Regents.

Bunnell, T. (2011). The Growth of the International Baccalaureate Diploma

    Program: Concerns about the consistency and reliability of the assessments.

    Educational Forum, *75*(2), 174-187

Bunnell, T. (2008). The International Baccalaureate in England and Wales: the

    alternative paths for the future. *The Curriculum Journal, 19*(3), 151-160.

Burris, C., & Murphy, J. (2013). Everyone can be collage ready. *Educational*

    Leadership, *16*(3), 62-66.

Carder, M. (2007, November). The IB and languages. *International Schools Journal,*

    27(1), 73-82.

 Chalhoub-Deville, M., & Deville, C. (2006). Old, borrowed, and new thoughts in

    second language testing. In R. L. Brennan (Ed.), *Educational measurement*

    (pp. 87-114). Westport, CT: Praeger.

Cho, Y., & Bridgeman, B. (2014). Relationship of TOEFL iBT® scores to academic

    performance: some evidence from American universities. Language Testing,

    29(3), 421-442.

Coates, H., Rosicka, C., & MacMahon-Ball, M. (2007). Perceptions of the

    International

    Baccalaureate Diploma Programme among Australian and New Zealand

    universities. Camberwell, Australia: Australian Council for Educational

    Research.

COPE: How to cope. (2013). Ankara. Meteksan.

Culross, R., & Tarver, E. (2011). A summary of research on the International

    Baccalaureate Diploma Programme: Perspectives of students, teachers, and

    university admission offices in the USA. *Journal of Research in International*

    Education, 10(3), 231-243.

Davidson, C., & Leung, C. (2012). Current issues in English language teacher-based

    assessment. *TESOL Journal*, *43*(3), 393-415.

Doey, P., & Oliver, R. (2002). An investigation into the predictive validity of the

      IELTS Test as an indicator of future academic success. *Prospect, 17*(1), 36-

      54.

Dinçer, A., Takkaç M., & Akalın, S. (2010). An evaluation on English language

      education process in Turkey from the viewpoints of university preparatory

      school students. 2nd International Symposium on Sustainable Development.

      Sarajevo. Keynote speech.

Embse, N., & Hasson, R. (2012). Test anxiety and high-stakes test performance

      between school settings: implications for educators. *Preventing School*

      Failure, 56(2), 180-187.

Ekici, G. (2005). An investigation of the attitudes of students' toward the

      student selection examination through a number of variables. *Hacettepe*

      Üniversitesi Egitim Fakültesi Dergisi 28, 82-90.

Halicioglu, M. M. (2008). The IB Diploma programme in national schools: The case

      of Turkey. *Journal of Research in International Education, 7*(2), 164-183.

Hayden, M. (2006). *Introduction to international education.* London: Paul Chapman.

Hill, I. (2002). The history of international education: An International Baccalaureate

      perspective. In M.C. Hayden, J.J. Thompson, & G.Walker (Eds.),

      International education in practice (pp. 18–29). London, UK: Kogan.

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

Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: an*

      introduction (8th ed.). New York: Pearson

Güneş, S. (2011). *The background factors that influence learners' English*

      proficiency (Unpublished master's thesis). Bilkent University, Ankara

Green, F., & Vignoles, A. (2012). An empirical method for deriving grade

  equivalance for university entrance qualifications: An application to A levels

  and International Baccalaureate. *Oxford Review of Education*, *38*(4), 473-491.

IBO- International Baccalaureate Organization. (2011). *Language A: language and*

  literature guide. Retrieved from

  http://occ.ibo.org/ibis/documents/dp/gr1/language_a1/d_1_a1lan_gui_1102_2

  b_e.pdf

IBO- International Baccalaureate Organization. (2011). *Language B guide*.

  Retrieved from

  http://occ.ibo.org/ibis/documents/dp/gr2/language_a1/d_1_a1lan_gui_1102_2

  b_e.pdf

IELTS (2014). Guide for educational institutions, governments, professional bodies

  and commercial organizations. Retrieved from

  http://www.ielts.org/PDF/Guide_Edu-%20Inst_Gov_2013.pdf

İnal, S., Evin, İ., & Saracaloğlu, A. (2005). The relation between students'

  attitudes toward foreign language and foreign language achievement. *Dil*

  Dergisi, 130, 38-53.

International Baccalaureate Organisation. (2013). *The IB Diploma Programme.*

  Retrieved from http://www.ibo.org/diploma

International Baccalaureate Organisation. (2013). *Diploma programme curriculum-*

  core requirements. Retrieved from

  http://www.ibo.org/diploma/curriculum/core/essay/

International Baccalaureate Organization (2013). *Diploma Programme assessment*

  principles and practice. Retrieved from

  http://www.ibo.org/diploma/assessment

International Baccalaureate Organisation. (2014). *Find an IB world school.*
Retrieved from http://www.ibo.org/school/search.

International Baccalaureate Organization. (2003). *Diploma Programme Language A2 guide*. Retrieved from
http://occ.ibo.org/ibis/occ/Utils/getFile2.cfm?source

International Baccalaureate Organisation. (2014). *Diploma Programme Language B guide.* Retrieved from http://occ.ibo.org/ibis/occ/Utils/getFile2.cfm?source

International English Language Testing System. (n.d.). *In UCLES.* Retrieved from International English Language Testing System

IBO- International Baccalaureate Organization. (2014). About the International Baccalaureate. *International Baccalaureate*. Retrieved from
http://www.ibo.org/general/who.cfm

International Baccalaureate Organisation. (2013). *What is IB recognition?* Retrieved from http://ibo.org/recognition

Karakaya, İ., & Tavşancıl, E. (2008). The predictive validity of the university student selection examination. *Educational Sciences: Theory and Practice*. *8*(3), 1001-1019.

Kathleen, S., & Kenji, K. (1996). Testing communicative competence. *Journal of Language Education*, *65*(3), 275-290.

Kazazoğlu, S. (2013). Türkçe ve İngilizce derslerine yönelik tutumun akademik başariya   etkisi. *Eğitim ve Bilim*, *38*, 294-307.

Koç, H., Çiftçi, T., & Sönmez, F. (2013). Analysis of questions asked in OSS, LYS and YGS exams in terms of cognitive domain level according to Bloom's taxonomy. *Journal of Black Sea Studies. 36*, 257-275.

Kondakçı, N. (2014). *A comparative analysis of national and international English language curricula for high schools in Turkey* (Unpublished master's thesis). Bilkent University, Ankara.

Llamas, C., Watt, D., & Johnson, D.E. (2009). Linguistic accommodation and the salience of national identity markers in a border town. *Journal of Language and Social Psychology 28*(4): 381-407.

Llurda, E. (2000). On competence, proficiency and communicative language ability. *International Journal of Applied Linguistics*, *10*(1), 85-95

Lineham, R. (2013, June). Is the International Baccalaureate Diploma Programme effective at delivering the International Baccalaureate mission statement. *Journal of Research in International Education*, *12*(3), 259-282.

Morrow, K., 1979. Communicative language testing: revolution of evolution? In C.K. Brumfit, & K. Johnson, (Eds.), *The communicative approach to language teaching* (pp. 140-160). Oxford University Press, Oxford,

Nisbet, I. (2014, April). International education and national education- can they co-exist? *International Schools Journal*, *38*(2), 72-78.

North, B. (2012). *The development of a common framework scale of language proficiency*. New York: Peter Lang Publishing.

Oord, L. V. (2007). To westernize the nations? An analysis of the International Baccalaureate's philosophy of education. *Cambridge Journal of Education*, *37*(3), 375-390.

ÖSYM (1982). *Öğrenci seçme ve yerleştirme sınavı geçerliliğin araştırılmasında kullanılan yöntemlere ilişkin bazı sorunlar*. Ankara: (AGB-100) Öğrenci Seçme ve Yerleştirme Merkezi

Razmjoo, S. (2011). Language proficiency tests in the Iranian context: Do they represent communicative language testing model?. *Pan-Pacific Association of Applied Linguistics*, *15*(2), 85-96.

Scholarship. (n.d.). Retrieved from http://bels.bilkent.edu.tr/kayitlarimiz/burslarimiz/

Sert, N. (2006). English language proficiency and academic attainment. *Eurasian Journal of Educational Research*, *22,* 45.

Stephen, D. F., Welman, J. C., & Jordaan, W. J. English language proficiency as an indicator of academic performance at a territary instution. *Journal of Human Resource Management, 2*(3), 42-53.

Spolsky, B.(1989). *Conditions for second language learning*. Oxford: Oxford University Press.

Spolsky, B. (1995). *Measured words*. Oxford: Oxford University Press.

Stern, H. H, (1996*). Issues and options in language teaching*. Oxford: Oxford University Press.

Taras, M. (2005). Assessment- summative and formative- some theoretical reflections. *British Journal of Educational Studies*, *53*(4), 466-478.

Tarc, P., & Beatty, L. (2012). The emergence of the International Baccalaureate Diploma in Ontario. *Canadian Journal of Education*, *35*(4), 341-375.

Taylor, D. S. (1988). The meaning and the use of the term competence in linguistics and applied linguistics. *Applied Linguistics*, *9*, 148-168

Tookey, M. E. (2000). The International Baccalaureate. *Journal of Secondary Gifted Education*, *11*(2), 52.

Wait, I. W., & Gressel, J. W. (2009, October). Relationship between TOEFL score
    and academic success for international engineering students. *Journal of
    Engineering Education*, *45*(2), 18.

Vollmer, H. (1980). A study of alternatives in explaining the general language
    proficiency factor. Paper presented at language Testing Research Colloquium.
    San Francisco

Vinke, A. A., & Jochems, W. M. (1993). English proficiency and academic successin
    international postgaduate education. *Higher Education*, *26*, 275-285.

Yağız, B. (2014). *Investigating the impact of International General Certificate of
    Second Education scores and gender on the Diploma Programme scores in
    mathematics and science* (Unpublished master's thesis). Bilkent University,
    Ankara.

APPENDICES

Appendix A: IB DP language A2 course assessment criteria

PAPER 1: Comparative Commentary

Criterion A: Understanding and Comparison of the Texts

1–2 The candidate makes few comparisons of the texts.

3–4 The candidate makes superficial comparisons of the texts.

5–6 The candidate makes generally relevant comparisons of the texts.

7–8 The candidate makes relevant and mostly detailed comparisons of the texts.

9–10 The candidate makes relevant and detailed comparisons of the texts.

Criterion B: Presentation

1–2 Little organization is apparent.

3-4 Some organization is apparent.

5-6 The commentary is organized.

7-8 The commentary is well organized.

9-10 The commentary is effectively organized.

Criterion C: Language

1-2 The language lacks fluency and appropriateness.

3-4 The language sometimes lacks fluency and appropriateness.

5-6 The language is mostly fluent and appropriate.

7-8 The language is fluent and appropriate.

9-10 The language is fluent and entirely appropriate.

PAPER 2: ESSAY

Criterion A: Response to the Question

1-2 The candidate has little awareness of the implications of the question.

3-4 The candidate has a superficial awareness of the implications of the question.

5-6 The candidate has an adequate understanding of the implications of the question.

7-8 The candidate has a good understanding of the implications of the question.

9-10 The candidate has a thorough understanding of the implications of the question.

Criterion B: Presentation

1-2 Little organization is apparent.

3-4 Some organization is apparent.

5-6 The essay is organized.

7-8 The essay is well organized and persuasive.

9-10 The essay is well organized and very persuasive.

Criterion C: Language

1-2 The language lacks fluency and appropriateness.

3-4 The language sometimes lacks fluency and appropriateness.

5-6 The language is mostly fluent and appropriate.

7-8 The language is fluent and appropriate.

9-10 The language is fluent and entirely appropriate.

Written Tasks

Criterion A: Formal requirements

1 The written tasks meet few of the formal requirements.

2 The written tasks partially meet the formal requirements.

3 The written tasks generally meet the formal requirements.

4 The written tasks meet most of the formal requirements.

5 The written tasks meet fully the formal requirements.

Criterion B: Task and Content

1-2 The choice of type of text is mostly inappropriate.

3-4 The choice of type of text is partially appropriate.

5-6 The choice of type of text is appropriate.

7-8 The choice of type of text is considered and appropriate.

9-10 The choice of type of text is insightful and appropriate.

Criterion C: Language and Style

1-2 The use of language and style is rarely appropriate.

3-4 The use of language and style is sometimes appropriate.

5-6 The use of language and style is generally effective.

7-8 The use of language and style is effective.

9-10 The use of language and style is highly effective.

Oral Component

Criterion A: Quality of Ideas

1-2 The candidate has little awareness of the subject matter .

3-4 The candidate has a superficial awareness of the subject matter.

5-6 The candidate has an adequate understanding of the subject matter.

7-8 The candidate has a good understanding of the subject matter.

9-10 The candidate has an excellent understanding of the subject matter.

Criterion B: Presentation

1-2 Little organization is apparent.

3-4 Some organization is apparent.

5-6 The oral task is organized.

7-8 The oral task is well organized.

9-10 The oral task is effectively organized.

Criterion C: Language

1-2 The language lacks fluency and appropriateness.

3-4 The language sometimes lacks fluency and appropriateness.

5-6 The language is mostly fluent and appropriate.

7-8 The language is fluent and appropriate.

9-10 The language is fluent and entirely appropriate.

Appendix B: IB DP Language B course assessment criteria

Paper 2: Written productive skills
Criterion A: Language
• How effectively and accurately does the student use language?
Failure to write the minimum number of words will result in a 1-mark penalty.
1–2 Command of the language is limited and generally ineffective.
  A limited range of vocabulary is used, with many basic errors.
  Simple sentence structures are sometimes clear.
3–4 Command of the language is generally adequate, despite many inaccuracies.
  A fairly limited range of vocabulary is used, with many errors.
  Simple sentence structures are usually clear.
5–6 Command of the language is effective, despite some inaccuracies.
  A range of vocabulary is used accurately, with some errors.
  Simple sentence structures are clear.
7–8 Command of the language is effective.
  A wide range of vocabulary is used accurately, with few significant errors.
  Some complex sentence structures are clear and effective.
9–10 Command of the language is very effective.
  A wide range of vocabulary is used accurately and effectively, with very few
errors.
  Complex sentence structures are clear and effective.
Criterion B: Argument
• How skillfully does the student develop ideas?
• How clear and convincing is the argument?
• To what extent does the student react to the stimulus?
1–2 The development of ideas is very poor, and the argument is unclear and
  unconvincing.
  The structure of the argument is vague and confusing.
  The ideas are irrelevant.
3–4 The development of ideas is poor, and the argument is rarely clear and
convincing.
  The structure of the argument is sometimes apparent.
  The ideas are sometimes relevant.
5–6 The development of ideas is sometimes good, and the argument has some clarity
  and is sometimes convincing.
  The structure of the argument is evident.
  The ideas are generally relevant.
7–8 The development of ideas is good and methodical; the argument is clear and
fairly
  convincing.
  The structure of the argument is coherent and organized.
  The ideas are well expressed and relevant.

9–10 The development of ideas is very good and methodical; the argument is convincing.

   The structure of the argument is consistently coherent and organized.

   The ideas are very well expressed, relevant and engaging.

Written assignment: Receptive and written productive skills

Criterion A: Language

• How effectively and accurately does the student use language?

Failure to write the minimum number of words will result in a 1-mark penalty.

Marks Level descriptor

0   The work does not reach a standard described by the descriptors below.

1–2 Command of the language is limited and generally ineffective.

   A limited range of vocabulary is used, with many basic errors.

   Simple sentence structures are sometimes clear.

3–4 Command of the language is generally adequate, despite many inaccuracies.

   A fairly limited range of vocabulary is used, with many errors.

   Simple sentence structures are usually clear.

5–6 Command of the language is effective, despite some inaccuracies.

   A range of vocabulary is used accurately, with some errors.

   Simple sentence structures are clear.

7–8 Command of the language is effective.

   A wide range of vocabulary is used accurately, with few significant errors.

   Some complex sentence structures are clear and effective.

Criterion B: Content

• To what extent does the student show appreciation of the literary work?

• How skillfully is the task planned?

0   The work does not reach a standard described by the descriptors below.

1–2 The student does not make use of the literary work.

   Connection with the text is superficial or little developed.

   There is no evidence of organization.

3–4 The student makes little use of the literary work.

   Connection with the text is basic.

   There is little evidence of organization.

5–6 The student makes some use of the literary work.

   Connection with the text is adequate and is used fairly well.

   There is some evidence of organization.

7–8 The student makes use of the literary work.

   Connection with the text is good.

   There is evidence of organization.

9–10 The student makes good use of the literary work.

   Connection with the text is effective.

   There is clear evidence of organization.

Criterion C: Format

• How correctly does the student produce the required text type?

• To what extent are the conventions of text types appropriate?

0 The work does not reach a standard described by the descriptors below.

1 The text type is sometimes recognizable and appropriate.

   Conventions appropriate to the text type are limited.

2 The text type is generally recognizable and appropriate.

   Some conventions appropriate to the text type are evident.

3 The text type is recognizable and appropriate.

   Conventions appropriate to the text type are effective.

4 The text type is recognizable, appropriate and convincing.

   Conventions appropriate to the text type are effective and varied.

Criterion D: Rationale

• How clear and convincing is the rationale?

0 The work does not reach a standard described by the descriptors below.

1 The rationale is not clear.

2 The rationale is clear to some extent.

3 The rationale is clear, pertinent and directly linked to the literary work.

Individual oral

Criterion A: Productive skills

How successfully does the student use the language in speech?

• How fluent and clear is the student's speech?

• How accurate and varied is the language used?

• How much does the student's intonation aid communication?

0   The work does not reach a standard described by the descriptors below.

1–2 Command of spoken language is very limited.

   The production of language is very hesitant and hardly comprehensible.

   Language is often incorrect and/or very limited.

   Intonation interferes seriously with communication.

3–4 Command of spoken language is limited.

   The production of language is hesitant and not always comprehensible.

   Language is often incorrect and/or limited.

   Intonation sometimes interferes with communication.

5–6 Command of spoken language is fairly good.

   The production of language is comprehensible and fluent at times.

   Language is sometimes correct, with some idiomatic expressions.

   Intonation does not interfere with communication.

7–8 Command of spoken language is good.

   The production of language is mostly fluent.

   Language is generally correct, varied and articulate.

   Intonation contributes to communication.

9–10 Command of spoken language is very good.

   The production of language is fluent.

   Language is correct, varied and articulate; errors do not interfere with message.

   Intonation enhances communication

Criterion B: Interactive and receptive skills

To what extent does the student understand and demonstrate an ability to interact in a conversation?
• How well can the student express simple and complex ideas?
• How well can the student maintain a conversation?
0   The work does not reach a standard described by the descriptors below.
1–2 Simple ideas are understood with great difficulty and interaction is very limited.
    Simple ideas and opinions are presented incoherently.
    The conversation is disjointed.
3–4 Simple ideas are understood with difficulty and interaction is limited.
    Simple ideas and opinions are presented with difficulty, sometimes incoherently.
    The conversation does not flow coherently.
5–6 Simple ideas are understood fairly well and interaction is acceptable.
    Simple ideas and opinions are generally presented clearly.
    The conversation flows coherently at times but with some lapses.
7–8 Simple ideas are understood well and interaction is good.
    Simple ideas and opinions are presented clearly and coherently
    The conversation generally flows coherently.
9–10 Complex ideas are understood well and interaction is good.
    Both simple and complex ideas and opinions are generally presented clearly, coherently and effectively.
    The conversation flows coherently.

Interactive oral activity

Criterion A: Productive skills
How successfully does the student use the language in speech?
• How fluent and clear is the student's speech?
• How accurate and varied is the language?
• How much does the student's intonation aid communication?
0 The work does not reach a standard described by the descriptors below.
1 Command of spoken language is very limited.
   The production of language is very hesitant and hardly comprehensible.
   Language is often incorrect and/or very limited.
   Intonation interferes seriously with communication.
2 Command of spoken language is limited.
   The production of language is hesitant and not always comprehensible.
   Language is often incorrect and/or limited.
   Intonation sometimes interferes with communication.
3 Command of spoken language is fairly good.
   The production of language is comprehensible and fluent at times.
   Language is sometimes correct, with some idiomatic expressions.
   Intonation does not interfere seriously with communication.
4 Command of spoken language is good.
   The production of language is mostly fluent.
   Language is generally correct, varied and articulate.

Intonation contributes to communication.

5 Command of spoken language is very good.

The production of language is fluent.

Language is correct, varied and articulate; errors do not interfere with message.

Intonation enhances communication.

Criterion B: Interactive and receptive skills

To what extent does the student understand and demonstrate an ability to interact in a conversation?

• How well can the student express ideas and opinions?

• How well can the student maintain a conversation?

0 The work does not reach a standard described by the descriptors below.

1 Simple ideas are understood with great difficulty and interaction is very limited.

Simple ideas and opinions are presented incoherently.

The conversation is disjointed.

2 Simple ideas are understood with difficulty and interaction is limited.

Simple ideas and opinions are presented with difficulty, sometimes incoherently.

The conversation does not flow coherently.

3 Simple ideas are understood fairly well and interaction is adequate.

Simple ideas and opinions are generally presented clearly.

The conversation flows coherently at times but with some lapses.

4 Simple ideas are understood well and interaction is good.

Simple ideas and opinions are presented clearly and coherently; there is some difficulty

with complex ideas.

The conversation generally flows coherently.

5 Complex ideas are understood well and interaction is very good.

Both simple and complex ideas and opinions are generally presented clearly, coherently and effectively

Appendix C: IB DP Language A: Language and Literature course assessment criteria

Paper 1: Comparative textual analysis

Criterion A: Understanding and comparison of the texts

• To what extent does the analysis show the similarities and differences between the texts?

• To what extent does the analysis show an understanding of the texts, their type and purpose, and their possible contexts (for example, cultural, temporal, relation to audience)?

• Are the comments supported by well-chosen references to the texts?

0 The work does not reach a standard described by the descriptors below.

1 There is little understanding of the context and purpose of the texts and their similarities

  or differences; summary predominates and observations are rarely supported by

  references to the texts.

2 There is some understanding of the context and purpose of the texts, and the similarities

  or differences between them; observations are generally supported by references to the

  texts.

3 There is adequate understanding of the texts, their possible context and purpose,

  and the similarities and differences between them; comments are included, as well as

  observations that are generally supported by references to the texts.

4 There is good understanding of the texts, their context and purpose, and the similarities

  and differences between them; comments are mostly supported by well-chosen

  references to the texts.

5 There is excellent understanding of the texts, their context and purpose, and the

  similarities and differences between them; comments are fully supported by well-chosen

  references to the texts.

Criterion B: Understanding of the use and effects of stylistic features

• To what extent does the comparative analysis show awareness of how stylistic features of the texts,

such as language, structure, tone, technique and style, are used to construct meaning?

• To what extent does the comparative analysis show appreciation of the effects of stylistic features

(including the features of visual texts) on the reader?

Marks Level descriptor

0 The work does not reach a standard described by the descriptors below.

1 There is little awareness of the use of stylistic features and little or no illustration of their effects on the reader.

2 There is some awareness of the use of stylistic features, with a few references illustrating their effects on the reader.

3 There is adequate awareness of the use of stylistic features and understanding of their

  effects on the reader.

4 There is good awareness and illustration of the use of stylistic features and detailed

  understanding of their effects on the reader.

5 There is excellent awareness of the use of stylistic features, with very good understanding of their effects on the reader.

Criterion C: Organization and development

• How well organized and coherent is the comparative analysis?

• How balanced is the comparative analysis? ("Balance" here means equal treatment of the two texts.)

0 The work does not reach a standard described by the descriptors below.

1 Little organization is apparent, with no sense of balance and very little development;

  considerable emphasis is placed on one text to the detriment of the other.

2 Some organization is apparent. There is little sense of balance and some development;

  although both texts are addressed, the treatment of one is superficial.

3 The comparative analysis is organized and structured in a generally coherent way. There

  is a sense of balance and adequate development.

4 The comparative analysis is well organized and balanced. The structure is mostly coherent and there is a good sense of development.

5 The comparative analysis is well balanced and effectively organized, with a coherent and effective structure and development.

Criterion D: Language

• How clear, varied and accurate is the language?

• How appropriate is the choice of register, style and terminology? ("Register" refers, in this context,

to the student's use of elements such as vocabulary, tone, sentence structure and terminology appropriate to the task.)

0 The work does not reach a standard described by the descriptors below.

1 Language is rarely clear and appropriate; there are many errors in grammar, vocabulary

and sentence construction and little sense of register and style.

2 Language is sometimes clear and carefully chosen; grammar, vocabulary and sentence

construction are fairly accurate, although errors and inconsistencies are apparent; the register and style are to some extent appropriate to the task.

3 Language is clear and carefully chosen with an adequate degree of accuracy in grammar, vocabulary and sentence construction despite some lapses; register and style are mostly appropriate to the task.

4 Language is clear and carefully chosen, with a good degree of accuracy in grammar,

vocabulary and sentence construction; register and style are consistently appropriate to

the task.

5 Language is very clear, effective, carefully chosen and precise, with a high degree of

accuracy in grammar, vocabulary and sentence construction; register and style are effective and appropriate to the task.

Paper 2: Essay

Criterion A: Knowledge and understanding

• How much knowledge and understanding of the part 3 works and their context has the student

demonstrated in relation to the question answered?

0 The work does not reach a standard described by the descriptors below.

1 Little knowledge is shown of the part 3 works and the way context affects their meaning.

2 Knowledge of the part 3 works and the way context affects their meaning is sometimes

illustrated; understanding is superficial.

3 Knowledge of the part 3 works and the way context affects their meaning is adequately

illustrated; understanding is satisfactory.

4 Knowledge of the part 3 works and the way context affects their meaning is pertinently

illustrated and the understanding shown is good.

5 Knowledge of the part 3 works and the way context affects their meaning is thoroughly

and persuasively illustrated and the understanding shown is perceptive.

Criterion B: Response to the question

• To what extent is an understanding of the expectations of the question shown?

• How relevant is the response to these expectations, and how far does it show critical analysis?

0 The work does not reach a standard described by the descriptors below.

1 There is little awareness of the expectations of the question.

2 There is some awareness of the expectations of the question; the response is only partly

relevant and is mostly unsubstantiated generalization.

3 There is adequate awareness of the expectations of the question; the response is generally relevant and critical.

4 There is good understanding of the expectations and some of the subtleties of the

question; the response is consistently relevant and critical.

5 There is excellent understanding of the expectations and many of the subtleties of the

question; the response is relevant, focused and insightful.

Criterion C: Understanding of the use and effects of stylistic features

• To what extent does the essay show awareness of how the writer's choices of the stylistic features in the texts (for example, characterization, setting, theme, narrative point of view, structure, style and technique) are used to construct meaning?

• To what extent does the essay show understanding of the effects of stylistic features?

0 The work does not reach a standard described by the descriptors below.

1 There is limited awareness or illustration of the use of stylistic features.

2 There is some awareness and illustration of the use of stylistic features, with limited

understanding of their effects.

3 There is adequate awareness and illustration of the use of stylistic features, with adequate understanding of their effects.

4 There is good awareness and illustration of the use of stylistic features, with good understanding of their effects.

5 There is excellent awareness and illustration of the use of stylistic features, with very good understanding of their effects.

Criterion D: Organization and development

• How logical and developed is the argument of the essay?

• How coherent and effective is the formal structure of the essay?

0 The work does not reach a standard described by the descriptors below.

1 There is little focus, structure, sequencing of ideas and development.

2 There is some focus, structure, sequencing of ideas and development.

3 There is adequate focus, structure, sequencing of ideas and development.

4 There is good focus and structure, with a logical sequence and development.

5 There is precise focus and excellent structure; the work is coherently sequenced and

thoroughly developed.

Criterion E: Language

• How clear, varied and accurate is the language?

• How appropriate is the choice of register, style and terminology? ("Register" refers, in this context, to the student's use of elements such as vocabulary, tone, sentence structure and terminology appropriate to the task.)

0 The work does not reach a standard described by the descriptors below.

1 Language is rarely clear and appropriate; there are many errors in grammar, vocabulary

and sentence construction and little sense of register and style.

2 Language is sometimes clear and carefully chosen; grammar, vocabulary and sentence

construction are fairly accurate, although errors and inconsistencies are apparent; the register and style are to some extent appropriate to the task.

3 Language is clear and carefully chosen with an adequate degree of accuracy in grammar,

vocabulary and sentence construction despite some lapses; register and style are mostly

appropriate to the task.

4 Language is clear and carefully chosen, with a good degree of accuracy in grammar,

vocabulary and sentence construction; register and style are consistently appropriate to

the task.

5 Language is very clear, effective, carefully chosen and precise, with a high degree of

accuracy in grammar, vocabulary and sentence construction; register and style are effective and appropriate to the task.

Written task 1

Criterion A: Rationale

• Does the rationale for the written task explain how the task is linked to the aspects of the course being

investigated?

Note: The word length for the rationale is 200–300 words. If the word limit is exceeded, 1 mark will be deducted.

0 The work does not reach a standard described by the descriptors below.

1 The rationale shows some explanation and understanding of the aspects being investigated.

2 The rationale shows clear explanation and understanding of the aspects being investigated.

Criterion B: Task and content

• To what extent does the task show understanding of the topic(s) or text(s) to which it refers?

• How appropriate is the content to the task chosen?

• To what extent does the task show understanding of the conventions of the text type chosen?

0 The work does not reach a standard described by the descriptors below.

1–2 The task shows a superficial understanding of the topic(s) or text(s) to which it refers.

The content is generally inappropriate to the task chosen.

The task shows a superficial understanding of the conventions of the text type chosen.

3–4 The task shows a mostly adequate understanding of the topic(s) or text(s) to which it

refers.

The content is generally appropriate to the task chosen.

The task shows an adequate understanding of the conventions of the text type chosen.

5–6 The task shows a good understanding of the topic(s) or text(s) to which it refers.

The content is mostly appropriate to the task chosen.

The task shows a good understanding of the conventions of the text type chosen.

7–8 The task shows an excellent understanding of the topic(s) or text(s) to which it refers.

The content is consistently appropriate to the task chosen.

The task shows an excellent understanding of the conventions of the text type chosen.

Criterion C: Organization

• How well organized is the task?

• How coherent is the structure?

Note: The word length for the written task is 800–1,000 words. If the word limit is exceeded, 2 marks will be deducted.

0 The work does not reach a standard described by the descriptors below.

1 Little organization is apparent; the task has little structure.

2 Some organization is apparent; the task has some structure, although it is not sustained.

3 The task is organized; the structure is generally coherent.

4 The task is well organized; the structure is mostly coherent.

5 The task is effectively organized; the structure is coherent and effective.

Criterion D: Language and style

• How effective is the use of language and style?

• How appropriate to the task is the choice of register and style? ("Register" refers, in this context, to the

student's use of elements such as vocabulary, tone, sentence structure and idiom appropriate to the

task; register is assessed on the task itself.)

Note: A student who writes an appropriate rationale but fails to achieve an appropriate register in the task cannot score more than 3 marks.

0 The work does not reach a standard described by the descriptors below.

1 There is little clarity, with many basic errors; little sense of register and style.

2 There is some clarity, though grammar, spelling and sentence structure are often inaccurate; some sense of register, style and appropriate vocabulary.

3 The use of language and the style are generally clear and effective, though there are some inaccuracies in grammar, spelling and sentence construction; generally appropriate

in register, style and vocabulary.

4 The use of language and the style are clear and effective, with a good degree of accuracy;

sentence construction and vocabulary are varied, showing a growing maturity of style;

the register is appropriate.

5 The use of language and the style are very clear and effective, with a very good degree of accuracy; sentence construction and vocabulary are good; the style is confident and the

register effective

Written task 2: Critical response

Criterion A: Outline

• Does the outline of the written task clearly highlight the particular focus of the task?

Marks Level descriptor

0 The work does not reach a standard described by the descriptors below.

1 The outline partially highlights the particular focus of the task.

2 The outline clearly highlights the particular focus of the task.

Criterion B: Response to the question

• To what extent is an understanding of the expectations of the question shown?

• How relevant and focused is the response to these expectations?

• Is the response supported by well-chosen references to the text(s)?

0 The work does not reach a standard described by the descriptors below.

1–2 The student has a superficial understanding of the expectations of the question.

Ideas are frequently irrelevant and/or repetitive.

The response is not supported by references to the text(s).

3–4 There is mostly adequate understanding of the expectations of the question.

Ideas are generally relevant and focused.

The response is generally supported by references to the text(s).

5–6 There is good understanding of the expectations of the question.

Ideas are mostly relevant and focused.

The response is mostly supported by well-chosen references to the text(s).

7–8 There is thorough understanding of the expectations of the question.

Ideas are relevant and focused.

The response is fully supported by well-chosen references to the text(s).

Criterion C: Organization and argument

• How well organized is the task?

• How coherent is the structure?

• How well developed is the argument of the written task?

Note: The word length for the written task is 800–1,000 words. If the word limit is exceeded, 2 marks will be deducted.

0 The work does not reach a standard described by the descriptors below.

1 Little organization is apparent; the task has little structure and the argument is poorly

developed.

2 Some organization is apparent; the task has some structure, although it is not sustained.

The argument has some development.

90

3 The task is organized, and the structure is generally coherent. There is some development of the argument.

4 The task is well organized; the structure is mostly coherent and the argument is clearly

developed.

5 The task is effectively organized; the structure is coherent and the argument is effectively developed.

Criterion D: Language and style

• How effective is the use of language and style?

• How appropriate to the task is the choice of register and style? ("Register" refers, in this context, to the

student's use of elements such as vocabulary, tone, sentence structure and idiom appropriate to the task; register is assessed on the task itself.)

0 The work does not reach a standard described by the descriptors below.

1 There is little clarity, with many basic errors; little sense of register and style.

2 There is some clarity, though grammar, spelling and sentence structure are often inaccurate; some sense of register, style and appropriate vocabulary.

3 The use of language and the style are generally clear and effective, though there are some inaccuracies in grammar, spelling and sentence construction; generally appropriate

in register, style and vocabulary.

4 The use of language and the style are clear and effective, with a good degree of accuracy; sentence construction and vocabulary are varied, showing a growing maturity of style; the register is appropriate.

5 The use of language and the style are very clear and effective, with a very good degree of accuracy; sentence construction and vocabulary are good; the style is confident and the

register effective.

Individual oral commentary

Criterion A: Knowledge and understanding of the text or extract

• To what extent does the commentary show knowledge and understanding of the text?

• Are the comments supported by well-chosen references to the text?

Note: The extract chosen for the commentary MUST be from a work selected from the relevant prescribed list of authors (PLA). If not, the maximum mark for this criterion will be reduced to 6.

0 The work does not reach a standard described by the descriptors below.

1–2 The commentary shows limited knowledge and little or no understanding of the text;

comments are rarely supported by references to the text.

3–4 The commentary shows superficial knowledge and understanding of the text; comments are occasionally supported by references to the text.

5–6 The commentary shows adequate knowledge and understanding of the text; comments are generally supported by references to the text.

7–8 The commentary shows a very good knowledge and understanding of the text; comments are supported by well-chosen references to the text.

9–10 The commentary shows excellent knowledge and understanding of the text; comments are effectively supported by well-chosen references to the text.

Criterion B: Understanding of the use and effects of literary features

• To what extent does the commentary show an awareness of how the literary features in the text (for example, structure, technique and style) are used to construct meaning?

• To what extent does the commentary show understanding of the effects of literary features?

0 The work does not reach a standard described by the descriptors below.

1–2 There is little awareness of the use of literary features and little or no illustration of their effects on the reader.

3–4 There is some awareness of the use of literary features, with few references illustrating their effects on the reader.

5–6 There is adequate awareness and illustration of the use of literary features, with understanding of their effects on the reader.

7–8 There is good awareness and illustration of the use of literary features, with detailed

understanding of their effects on the reader.

9–10 There is excellent awareness and illustration of the use of literary features, with very good understanding of their effects on the reader.

Criterion C: Organization

• How well organized is the commentary?

• How coherent is the structure?

0 The work does not reach a standard described by the descriptors below.

1 Little organization is apparent; the commentary has little structure.

2 Some organization is apparent; the commentary has some structure.

3 The commentary is adequately organized; the structure is generally coherent.

4 The commentary is well organized; the structure is mostly coherent.

5 The commentary is very effectively organized; the structure is coherent and effective.

Criterion D: Language

• How clear, varied and accurate is the language?

• How appropriate is the choice of register and style? ("Register" refers, in this context, to the student's

use of elements such as vocabulary, tone, sentence structure and terminology appropriate to the commentary.)

0 The work does not reach a standard described by the descriptors below.

1 The language is rarely clear and appropriate, with many errors in grammar and sentence

construction and little sense of register and style.

2 The language is sometimes clear and appropriate; grammar and sentence construction

are generally accurate, although errors and inconsistencies are apparent; register and style are to some extent appropriate to the commentary.

3 The language is mostly clear and appropriate, with an adequate degree of accuracy in

grammar and sentence construction; the register and style are mostly appropriate to the

commentary.

4 The language is clear and appropriate, with a good degree of accuracy in grammar and sentence construction; register and style are effective and appropriate to the commentary.

5 The language is very clear and entirely appropriate, with a high degree of accuracy in

grammar and sentence construction; the register and style are consistently effective and

appropriate to the commentary.

Further oral activity

Criterion A: Knowledge and understanding of the text(s) and subject matter or extract

• To what extent does the activity show knowledge and understanding of the text(s) and subject chosen for the oral activity?

• Has the student shown awareness and understanding of the meaning of the text(s) in relation to the subject?

0 The work does not reach a standard described by the descriptors below.

1–2 The activity shows limited knowledge and little or no understanding of the text(s) and the subject chosen.

3–4 The activity shows some knowledge and understanding of the text(s) and some awareness of the significance of the text(s) in relation to the subject chosen.

5–6 The activity shows adequate knowledge and understanding of the text(s) and awareness of the significance of the text(s) in relation to the subject chosen.

7–8 The activity shows good knowledge and understanding of the text(s) and good awareness of the significance of the text(s) in relation to the subject chosen.

9–10 The activity shows excellent knowledge and understanding of the text(s) and excellent awareness of the significance of the text(s) in relation to the subject chosen.

Criterion B: Understanding of how language is used

• To what extent does the activity show understanding of the way language is used to create meaning?

• Has the student shown an appreciation of how language and style is used to particular effect in the text?

0 The work does not reach a standard described by the descriptors below.

1–2 The work shows a superficial understanding of the way language is used to create

meaning; there is little appreciation of the use of language and style.

3–4 The work shows some understanding of the way language is used to create meaning;

there is some appreciation of the use of language and style.

5–6 The work shows an adequate understanding of the way language is used to create

meaning and adequate appreciation of the use of language and style.

7–8 The work shows a good understanding of the way language is used to create meaning

and good appreciation of the use of language and style.

9–10 The work shows an excellent understanding of the way language is used to create

meaning. The appreciation of the use of language and style is thorough and detailed.

Criterion C: Organization

• How well organized is the oral activity?

• How coherent is the structure?

0 The work does not reach a standard described by the descriptors below.

1 Little organization is apparent; the oral activity has little structure.

2 Some organization is apparent; the oral activity has some structure.

3 The oral activity is organized; the structure is generally coherent.

4 The oral activity is well organized; the structure is mostly coherent.

5 The oral activity is effectively organized; the structure is coherent and effective.

Criterion D: Language

• How clear, varied and accurate is the language?

• How appropriate is the choice of register and style? ("Register" refers, in this context, to the student's use of elements such as vocabulary, tone, sentence structure and idiom appropriate to the further oral activity.)

0 The work does not reach a standard described by the descriptors below.

1 The language is rarely clear and appropriate, with many errors in grammar and sentence

construction and little sense of register and style.

The language is sometimes clear and appropriate; grammar and sentence construction

are generally accurate, although errors and inconsistencies are apparent; register and style are to some extent appropriate to the oral activity.

3 The language is mostly clear and appropriate, with an adequate degree of accuracy in

grammar and sentence construction; the register and style are mostly appropriate to the

oral activity.

4 The language is clear and appropriate, with a good degree of accuracy in grammar and

sentence construction; register and style are effective and appropriate to the oral activity.

5 The language is very clear and entirely appropriate, with a high degree of accuracy in

grammar and sentence construction; the register and style are consistently effective and
appropriate to the oral activity.

Appendix D: COPE exam writing and speaking criteria

WRITING CRITERIA

	1 Displays a very low level of communicative effectiveness / Is clearly not at the B2 level	2 Displays an inadequate level of communicative effectiveness for the B2 level	3 Displays an adequate level of communicative effectiveness for the B2 level	4 Displays a good level of communicative effectiveness for the B2 level	5 Displays a very high level of communicative effectiveness for the B2 level
LANGUAGE	-Grammar and/or vocabulary insufficient to handle the tasks -Meaning is usually unclear	-Grammar and vocabulary are limited and may cause problems in handling the tasks -Meaning is unclear at times	-Uses a limited range of grammar and vocabulary but generally sufficient to handle task with some unimportant errors -Meaning is clear	-Uses a range of appropriate grammar and vocabulary with few errors -Meaning is clear	-Uses a wide range of appropriate grammar and vocabulary with few errors -Meaning is clear
DISCOURSE	-Almost no expansion and/or excessive unnatural repetition -Frequently irrelevant or incoherent speech	-Limited expansion and/or unnatural repetition throughout -At times irrelevant and incoherent speech	-Frequent attempt to expand although there may be unnatural repetition -Generally relevant and coherent speech	-Responses usually expanded with almost no unnatural repetition -Almost always relevant and coherent speech	-Responses always expanded with no unnatural repetition -Relevant and coherent speech
FLUENCY	-No sign of fluency -Hesitations and/or fillers cause strain for the listener	-Speaks with some sense of fluency but with frequent inappropriate hesitations and/or fillers	-Speaks fluently but there may be some inappropriate hesitations and/or fillers	-Speaks fluently	
INTERACTION	-Interaction usually not initiated/sustained -Mostly unaware of conventions of interaction	-Interaction usually initiated/sustained -Conventions of interaction generally observed	-Interaction initiated/sustained throughout -Conventions of interaction fully observed		
PRONUNCIATION	-Pronunciation, rhythm and intonation has a negative effect on the listener	-Rhythm and intonation can sometimes have a negative effect on the listener -There may be pronunciation problems with individual words	-Pronunciation is unproblematic -Uses rhythm and intonation to convey the message clearly		

SPEAKING CRITERIA

VERY GOOD PASS 10-9	Very good coverage of topic Well-developed, coherent, clear passage with good justifications Very good use of linking devices Accurate use of a wide range of appropriate structures Good use of complex sentences Occasional unimportant language errors A wide range of appropriate vocabulary
GOOD PASS 8-7	Good coverage of topic Fairly well-developed, coherent, clear passage with good justifications Good use of linking devices Fairly accurate use of a good range of appropriate structures with some minor language errors Complex sentences are used mostly accurately Good range of vocabulary used mostly appropriately
PASS 6	Adequate coverage of topic Adequate development of ideas, generally coherent, clear passage with some justifications Uses linking devices, mostly appropriately Adequate range of structures with some error but meaning is not obscured Complex sentences are evident but not always accurate Adequate range of vocabulary used fairly appropriately
BORDERLINE FAIL 5	Limited coverage of topic Limited development of ideas and poor justification in places Linking devices used, but not always successfully Limited range of structures with some basic language errors which at times obscure meaning Limited range of vocabulary used inappropriately at times
FAIL 4-3	Poor coverage of topic Poor development of ideas and frequently incoherent Basic linkers used – not always successfully Poor range of structures used and/or with frequent language errors which makes the passage difficult to understand in places Poor range of vocabulary and often used inappropriately
BAD FAIL 2-1	Almost no relevance to the topic Frequently incoherent Uses very simple sentences with serious language errors which generally makes the passage difficult to understand No range of vocabulary