

EXCELLENCE IN  
EDUCATION JOURNAL

10<sup>th</sup> Anniversary  
2012-2022

Volume 11  
Issue 1

®EEJ

The Excellence in Education Journal  
ISSN 2474-4166

Website: [www.excellenceineducationjournal.org](http://www.excellenceineducationjournal.org)  
Email: [eejeditor@gmail.com](mailto:eejeditor@gmail.com)

Ann Gaudino, Ed.D. Editor-in-Chief  
Madison VanPelt, Assistant to the Editor  
William F. Hoge, Assistant

Copyright © 2022 Excellence in Education Journal. All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording or any information storage and retrieval system, without permission from EEJ. Readers who wish to duplicate material copyrighted by EEJ may do so by contacting the editor.

Cover art copyright © 2022 by EEJ.

EEJ publications present a variety of viewpoints. The views expressed or implied in this journal are those of the authors and should not be interpreted as official positions of the EEJ.

All Web links in this journal are correct as of the publication date but may have become inactive or otherwise modified since that time. If you notice a deactivated or changed link, please email [eejeditor@gmail.com](mailto:eejeditor@gmail.com) with the words “Link Update” in the subject line. In your message, please specify the issue.

Manuscript submission guidelines can be viewed on the website at:  
[www.excellenceineducationjournal.org](http://www.excellenceineducationjournal.org)

If you are interested in serving as a reviewer for this journal, please email your request and your curriculum vitae/resume to [eejeditor@gmail.com](mailto:eejeditor@gmail.com). A sample paper for review will be emailed to you.

## *From the Editor: Celebrating our 10<sup>th</sup> Anniversary!*

In 2012, the first issue of *The Excellence in Education Journal* was published online with the goal of promoting and disseminating international scholarly writing about excellent practices in all aspects of education. I had just returned from trips to China and Kenya and was moved by the limited access to hard copies of educational writing in both countries. Yet, even in remote areas, educators had cell phones and were using their devices to access information online.

I was also struck by the tremendous politics that were present in the book and article publishing business. It appeared that, at times, quality writing that could have a significant, positive influence in education was being denied from publication. Even more so, writing with meritorious methods, findings and recommendations that had flaws in English grammar were being systematically rejected from publication, even though such grammar errors are quite easy to correct.

*The Excellence in Education Journal* began an independent, online journal---one that could be accessed free online anywhere in the world. Online journals were a bit rare in 2012 as most journals were published in hard copy. Today, online, open access journals are much more common. Independently published journals continue to be rare even now in 2022.

Since 2012, *The Excellence in Education Journal* has had many achievements including being granted an ISSN number from the United States Library of Congress, receiving a contract from the United States Department of Education to list all articles full text in the ERIC database, and receiving a contract from EBSCOhost to list all articles full text. There are 20 reviewers from eight different countries and authors represent 50 institutions worldwide. Annual circulation exceeds 20,000 downloads.

The COVID-19 pandemic slowed the reviewing and publishing process significantly as those involved with the journal suffered loss of their and their loved ones' health. Gradually now, in the spring of 2022, production is getting back to a more normal pace.

Submissions continue to be double-blind, peer reviewed and are accepted year-round with publication occurring twice annually. There are no fees to submit or publish manuscripts so that cost will never be a barrier. Typeset and graphics are intentionally simple in order that the journal can be more easily accessed on a variety of devices worldwide to fulfill the mission of the journal.

I hope that the practices discussed in this journal will be helpful to you, our readers.

In gratitude and celebration for our 10 years of publication,

Ann C. Gaudino, Ed.D., Founder and Editor-in-Chief  
[eejeditor@gmail.com](mailto:eejeditor@gmail.com)

With gratitude to our reviewers who serve the journal:

Dr. Bundit Anuyahong, Thai-Nichi Institute of Technology, Thailand  
Dr. Evangelin Arulselvi, Princess Nora Bint Abdulrahman University, Riyadh, Saudi Arabia  
Dr. Dianbing Chin, Zhejiang Normal University, China  
Dr. Kim Creasy, University of Northern Colorado, United States  
Dr. Tiffany Flowers, Georgia State University, United States  
Rev. Dr. Walt Jagela, All Saints Catholic Church, West Virginia  
Dr. Beth Musser, Dean Emeritus, West Liberty University, United States  
Dr. Changsong Niu, Zhejiang Normal University, China  
Dr. Kakenya Ntaiya, The Kakenya Center for Excellence, Kenya  
Dr. Mustafa Ozmusul, Harran University, Turkey  
Dr. Li-Wei Peng, Indiana State University  
Dr. Tonya Perry, University of Alabama at Birmingham, United States  
Dr. Janine Wahl, Bemidji State University, United States  
Dr. Anthony Williams, Fisk University, United States  
Dr. Eleanor Wilson, The University of Virginia, United States  
Dr. Xiubin Xu, Zhejiang Normal University, China  
Dr. Yanjun Zhang, Zhejiang Normal University, China  
Prof. Joan Yakkey, The Music School of Fiesole, Italy

## TABLE OF CONTENTS

Page 1

*Motivational Correlates of Language-Specific Grit and Achievement in EFL:  
A CHAID Analysis*

Gülçin Mutlu

Page 30

*5<sup>th</sup> Grade English Language Teaching Curricula in Turkey:  
Expectations vs. Realities*

Esin Dünder and Ali Merç

Page 69

*The Assessment of Preservice Teachers' Dispositions*

David K. Griffin

Page 88

*Engagement and Big-Fish-Little-Pond Effect on Self-Concept in Science:  
Findings from TIMSS 2011*

Nurcan Kahraman, Yasemin Tas, and Sündüs Yerdelen

# Motivational Correlates of Language-Specific Grit and Achievement in EFL: A CHAID Analysis

Gülçin Mutlu

## Abstract

The first purpose of this study was to examine the associations between students' motivational characteristics and their language-specific grit for learning English. Second, this study aimed to investigate how students' language-specific grit and motivational characteristics related to their achievement in English. While examining the presence of associations sought through both of the research questions, a particular overall purpose was to identify what specific variables included in the analyses had the strongest impact on group differentiation concerning the dependent variables of grit and achievement. The participants included 182 students enrolled in a tertiary English preparatory program in Turkey. Chi-squared automatic interaction detection (CHAID) algorithm as a data mining method was used to analyze the data. The results revealed that task value, self-efficacy and intrinsic motivation had significant impacts on differentiating the students with different levels of grit. Moreover, language-specific grit, test anxiety, self-efficacy and control beliefs significantly related to students' achievement.

**Keywords:** CHAID, CHAID algorithm, Motivation, Grit, Persistence, Perseverance

*Gülçin Mutlu, Ph.D. is an Assistant Professor of Curriculum and Instruction at İzmir Democracy University in İzmir, Turkey. She can be reached at [gulcin.mutlu@idu.edu.tr](mailto:gulcin.mutlu@idu.edu.tr)*

The construct termed as *language-specific grit* in this study dates back to the construct of grit which was first operationalized by Duckworth, Peterson, Matthews and Kelly (2007) as a psychological general trait referring to one's capacity to further her or his efforts and interests for the activities taking a long time to complete. The presence of such a construct was a keen interest historically to several psychologists or researchers in the literature (e.g. Cox, 1926; Galton, 1892) a long time before Duckworth and others (2007). However, it had not been yet defined as a distinct measure of human psychology until the recent renewed interest about soft-skills and character strengths by Duckworth (2007) and others (Heckman & Rubinstein, 2001; Heckman, Stixrud, & Urzua, 2006; Lieras, 2008; Peterson & Seligman, 2004).

Duckworth et al. (2007), known with their proponent work on grit, conducted a series of studies particularly focusing on grit and the results from these studies consistently indicated that participants' grit scores were associated with various success outcomes. Following Duckworth (2007) and her associates, the links between success outcomes and grit have been firmly investigated and confirmed by several other researchers (Akos & Kretchmar, 2017; Changlek & Palanukulwong, 2015; Hagger & Hamilton, 2018; Muenks, Yang, & Wigfield, 2018; Reraki, Celik, & Sariçam, 2015); but, these pieces of evidence emphasized a more general view of grit by conceptualizing it as a stable human trait and detracting its ability to change in line with the type of the task or activity in question. In this essence, Duckworth et al. (2007) and Duckworth and Quinn (2009) warned future researchers about the possibility that the measure of grit may be a domain-specific construct. Hence, there is a need for more evidence to connect success outcomes and grit when grit is treated as a task or domain-specific construct.

Though a particular line of research on grit examined how grit was associated with several success outcomes, some other variables were also investigated regarding their links to

grit (Duckworth et al, 2007; Changlek & Palanukulwong, 2015; Hagger & Hamilton, 2018; Muenks, Yang, & Wigfield, 2018; Reraki, Celik, & Sariçam, 2015; Von Culin, Tsukayama, & Duckworth, 2014). As a proponent study, Duckworth et al. (2007), for instance, investigated how grit related to self-control, personality traits (i.e. Big Five traits of conscientiousness, extraversion, neuroticism, agreeableness, and openness to experience) and self-efficacy. Investigations into the links between several motivational variables and grit were also performed in subsequent studies (Changlek & Palanukulwong, 2015; Çelik & Sariçam, 2016; Hagger & Hamilton, 2018; Muenks, Yang, & Wigfield, 2018; Reraki, Celik, & Sariçam, 2015; Von Culin, et al., 2014). Thus, based on the attempts by several researchers above, it appears that variables other than success outcomes were also worthy of investigating to unpack the true nature of grit. It is also important to note here that in the above attempts, grit was treated as a general character strength and the participants were mostly adults who were in fact not enrolled in a particular study or a degree program.

Another investigation into the variables other than success outcomes was performed by Mutlu (2017; see also Mutlu & Yıldırım, 2019) who conducted a proponent study to date on grit in foreign language learning by investigating how language-specific grit for English related to students' learning environment perceptions, exposure to the target language and several demographics. However, this first attempt to examine the domain-specificity of the grit is not sufficient to understand the nature of this new construct. Therefore, there is more need to investigate how grit is related to success outcomes and other types of affective constructs, perhaps particularly to those known to predict achievement, such as self-efficacy or motivation. This need appears to be more obvious and crucial as to the domain-specific forms of grit since

variables like motivation and self-efficacy may show differentiated effects depending on the type of tasks or domains in question.

There has been a particular surge of interest upon the variable of motivation in the literature as a close correlate of success. However, it is also seen that it has been difficult to define or conceptualize this construct, which results in an emergence of a number of different categorizations for motivation and motivational theories. Theory of the hierarchy of needs (Maslow, 1943), expectancy-value model of motivation (Eccles, 1983; Lawler & Porter, 1967; Pintrich, 1988, 1989), achievement motivation theory (McClelland, 1961) and goal-setting theory (Locke & Latham, 1990) have been the most popular ones in the literature. Supporters or researchers of the expectancy-value model (Chen, 2002; Eccles, 1983; Eccles & Wigfield, 2002; Feather, 1992; Pintrich & De Groot, 1990; Wigfield, 1994; Wigfield & Eccles, 1992, 2000) repeatedly reported that learners' expectancies for success in a task and the subjective values they attached to the success to the task directly predicted the degree of persistence (corresponding to the recent construct of grit today) and achievement they would possess. In essence, research into the expectancy-value model provided sufficient and strong evidence about the links between persistence (recently termed as grit) shown as a type of "achievement behaviors" (Wigfield, 1994, p. 51) and motivation (Eccles, Wigfield, Harold, & Blumenfeld, 1993; Eccles & Wigfield, 1995; Wigfield, Harold, Eccles, Blumenfeld, Aberbach, Freedman-Doan, & Yoon, 1992). Thus, in line with the expectancy-value perspective and given the task of learning English, it could be expected and hypothesized that those with high expectations of being successful in English who attach positive values to achievement in English are more likely to put more grit and effort for learning the language and become more successful.

As is seen in the above account, a number of variables were investigated regarding their relationships to students' grit by previous researchers. However, in their analyses, they made no classifications or specifications regarding students' different degrees of gritty behaviors or success outcomes. Hence, research into the characteristics possessed by gritty and less gritty students or high achievers or low achievers and into the differences between these two main distinct group categories (of achievement and grittiness) are worthy of investigation. Moreover, there is a scarcity of research that links language-specific form of grit to success and affective outcomes and future investigations into these probable links are needed. In line with the above hypotheses and gaps in the literature, this study aimed to answer the following research questions:

**Research Question 1 (RQ1):** To what degree do six motivational variables affect the differences in students' language-specific grit for learning English?

**Research Question 2 (RQ2):** To what degree do students' language-specific grit and six motivational variables affect the differences in students' achievement scores in English.

## **Methods**

### **Participants**

The data were collected from the English preparatory program students enrolled at a state university in Turkey ( $n = 182$ ). The English preparatory class was a required course for the participant students. To start their main content departments, they were required to pass the English preparatory program successfully. These students were enrolled in the Faculty of Engineering ( $n = 87$ ), Faculty of Aviation and Space Sciences ( $n = 49$ ), and Faculty of Social Sciences and Humanities ( $n = 46$ ). Eighty (44 %) of the students were females and 102 (56 %) of them were males. Almost 90 % of them were aged 20 years old or less. There were only 19

students aged over 20. The average mean of age was approximately 19 years old ( $M = 19.39$ ;  $SD = 1.91$ ).

When the motivational characteristics of the students were examined (Table 1), it was seen that the participants had the highest average mean on the control beliefs sub-scale ( $M = 5.20$ ;  $SD = 1.17$ ) and the lowest mean score on test anxiety ( $M = 4.14$ ;  $SD = 1.45$ ). All of the mean scores for six motivational sub-scales corresponded to beyond-moderate degree of frequency (on a scale from 1 to 7-point Likert type). Given the students' levels of language-specific grit in English, they had a moderate level of language-specific grit on a scale from 1 to 5 ( $M = 3.28$ ;  $SD = 0.67$ ).

**Table 1**

*Motivation and Persistence-related Characteristics of the Participants (N = 182)*

<i>Motivation/Persistence Characteristics</i>	<i>M</i>	<i>SD</i>
Task value	4.84	1.16
Extrinsic motivation	5.03	1.21
Intrinsic motivation	4.98	1.21
Self-efficacy	4.86	1.18
Test anxiety	4.14	1.45
Control beliefs	5.20	1.17
Language-specific grit	3.28	0.67

## **Instruments**

### ***Language-Specific Grit Scale***

Persistence Scale for Learning English (PS) developed by Mutlu (2017) as a one-dimensional instrument with a total of 18 items was utilized to measure language-specific grit in this study. This instrument based on the goal setting theory employed a five-point Likert-type

scale ranging from *not at all true of me* (1), *slightly true of me* (2), *moderately true of me* (3), *very true of me* (4) to *completely true of me* (5). The instrument revealed an alpha reliability of .94 in Mutlu's (2017) study. In this current study, language-specific grit (persistence) scale indicated an alpha reliability of .93.

### ***Student Background Form***

A background information form was designed to elicit information concerning students' ages, genders, faculty majors and final grade scores on the preparatory program. To elicit the data about students' final grades, students were orally asked for their consents for the researcher to elicit their final grades from their class teachers and they were requested to write their names on the forms.

### ***Motivated Strategies for Learning Questionnaire (MSLQ)***

The motivation section of the MSLQ had 31 items to be responded on a 7-point Likert scale, from 1 (*not at all true for me*) to 7 (*very true for me*) was utilized in this study. This instrument was originally developed by Pintrich, Smith, Garcia and McKeachie (1991). The Turkish adaptation of the MSLQ (see Büyüköztürk, Akgün, Demirel, & Özkahveci, 2004) was used in this study. Büyüköztürk et al. (2004) reported Cronbach reliability scores ranging between .52 and .86 for the subscales of the motivation section in the Turkish version.

The theoretical framework behind the motivation section of the MSLQ is based on an adapted version of the general expectancy-value model of motivation (Eccles, 1983; Eccles & Wigfield, 2002; Wigfield, 1994). In accordance with the tenets of expectancy-value model of motivation, goals, different value beliefs and emotional reactions were theoretically represented by three main motivational components which were further categorized into six different subscales. The three main components were *expectancy*, *value* and *affect*. In line with the meanings

of these three dimensions, there were six sub-scales of the instrument named as *self-efficacy* and *control beliefs for learning* as sub-sections expectancy, *intrinsic goal orientation*, *extrinsic goal orientation* and *task value* as sub-sections of value and lastly *test anxiety* as a sub-section of affect. Table 2 depicts these sub-scales and their definitions. An alpha reliability of .89 was also found on the whole scale over 31 items in the current study. The subscales of self-efficacy, control beliefs for learning, intrinsic goal orientation, extrinsic goal orientation, task value and test anxiety respectively showed alpha reliability scores of .90, .73, .77, .67, .81 and .79 in this study.

**Table 2**

*Description of Motivation Section of the MSQI*

<i>Dimension</i>	<i>Description</i>	<i>Main Component</i>	<i>N of items</i>
Self-efficacy	The extent to which one believes her/his ability to achieve a task	E	8
Control Beliefs for Learning	The extent to which one believes in the influence of efforts to manage a task	E	4
Intrinsic Goal Orientation	The extent to which one has internal reasons (i.e. mastery or curiosity) to be engaged in a task	V	4
Extrinsic Goal Orientation	The extent to which one has external factors (i.e. rewards, competition or grades) to be engaged in a task	V	4
Task Value	The extent to which a task is perceived to be interesting, useful and important	V	6
Test Anxiety	The extent to which one is concerned about having exams.	A	5

*Note.* E= Expectancy, V= Value and A= Affect; The descriptions were derived based on Pintrich, et al., 1991.

**Data Analysis**

Decision tree as a data mining method was utilized for data analysis. CHAID (Chi-Squared Automatic Interaction Detector) algorithm was employed for decision tree development. The developer of the CHAID, Kass (1980, p. 119) briefly explained the working system of the

CHAID by saying that it “partitions the data into mutually exclusive, exhaustive, subsets that best describe the dependent variable. The subsets are constructed by using small groups of predictors.” Thus, it was believed that such type of an analysis would be suitable for this study aiming to examine the effects of motivational predictors in differentiating and classifying the learners in terms of their language-specific grit and achievement scores by means of detecting the most significant predictors describing them.

Prior to the development of classification trees via CHAID, the type and category of the variables were checked and the dependent and independent variables in continuous forms were transformed into categorical ones by using the Visual Binning option on SPSS. It is possible to have either continuous or categorical variables on CHAID. However, each form (category) of the study variables was tested with CHAID and the use of categorical variables other than continuous ones resulted in better tree development. Thus, the researcher followed with the (transformed) nominal variables for decision tree development. Table 3 shows these variables and their structure as they were used in the analyses. The alpha level was determined as .05 for the analyses in this study.

**Table 3***The Variables and Their Structure as used in CHAID Analyses*

<i>Variable Name</i>	<i>Values (Modalities)*</i>	<i>Profile of the Measurement Sample f(%)</i>	<i>Scale</i>
Task Value	<= 27	72 (39,6 %) **	CV→OV
	28 - 32	51 (28 %)	
	33 +	59 (32.4 %)	
Test Anxiety	<= 17	61 (33.5 %)	CV→OV
	18 - 24	62 (34.1 %) **	
	25 +	59 (32.4 %)	
Self-efficacy	<= 35,00	62 (34.1 %) **	CV→OV
	36 - 44	62 (34.1 %) **	
	45 +	58 (31.9 %)	
Control Beliefs for Learning	<= 19	68 (37.4 %) **	CV→OV
	20 - 23	56 (30.8 %)	
	24 +	58 (31.9 %)	
Intrinsic Goal Orientation	<= 18	71 (39 %) **	CV→OV
	19 - 22	51 (28 %)	
	23 +	60 (33 %)	
Extrinsic Goal Orientation	<= 18	63 (34.6 %)	CV→OV
	19 - 23	67 (36.8 %) **	
	24 +	52 (28.6 %)	
Language-specific Grit	<= 54	64 (35.2 %) **	CV→OV
	55 - 65	64 (35.2 %) **	
	66+	54 (29.7 %)	
Final Achievement Score in Prep Program	<= 72	64 (35.2 %) **	CV→OV
	73 - 79	61 (33.5 %)	
	80+	57 (31.3 %)	

*Note.* CV→OV= transformed from continuous variable into categorical (ordinal) variable; \*based on the calculation of the total score on each sub-scale; \*\* shows majority of the participants in relation to MSQ sub-scales, language-specific grit and achievement variables.

## Results

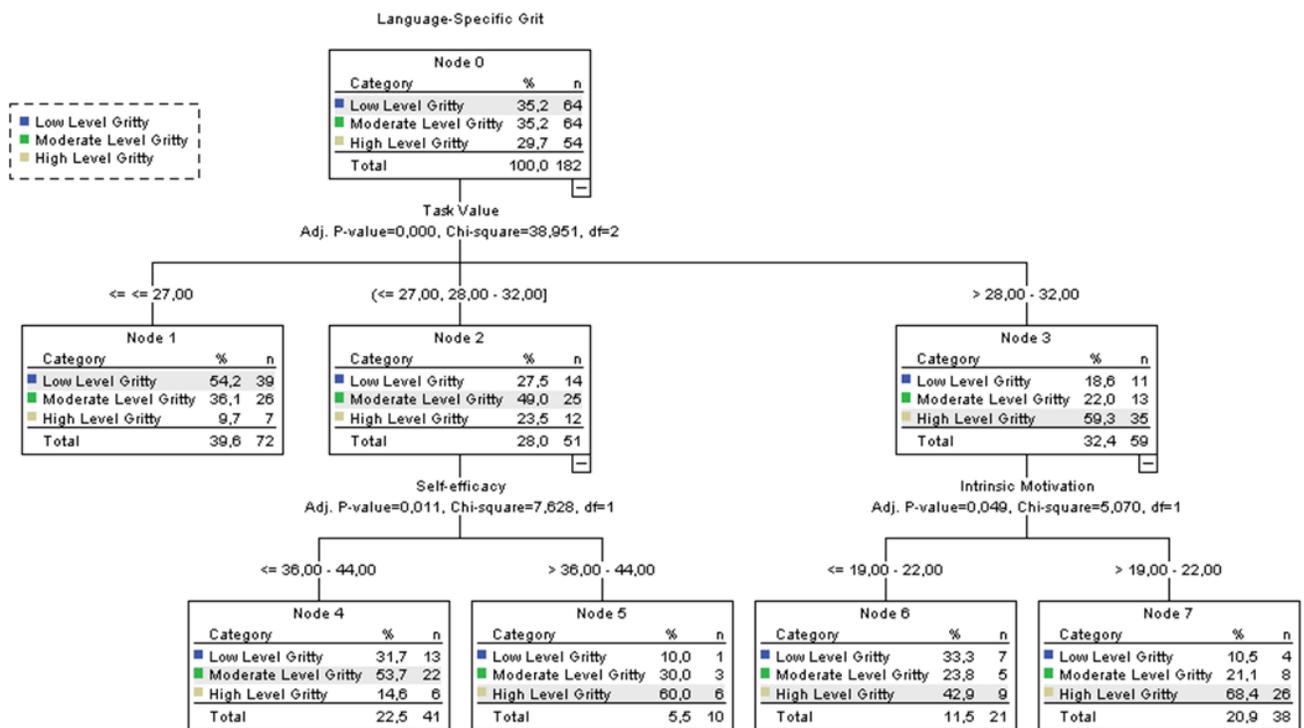
### **RQ1: Classification of Students' Level of Language-specific Grit in learning English based on Six Motivational Variables**

CHAID analysis was conducted by using the student language-specific grit as the dependent variable and six motivational variables, self-efficacy, control beliefs for learning, task

value, extrinsic motivation, intrinsic motivation, test anxiety as independent variables. As shown in Figure 1, the most significant independent motivational variable was task value,  $X^2 = 38.951$ ,  $df = 2$ ,  $p = 0.000$ , which meant that this variable had the strongest impact to differentiate and classify students into three groups concerning their language-specific grit levels.

**Figure 1**

*CHAID Decision Tree Model for Language-Specific Grit Based on Motivational Variables*



Most of the participants ( $n = 72$ ) were placed in node 1 that corresponded to a low degree of task value for learning English. The remaining 51 students belonged to node 2 and 59 students to node 3. Given the percentage distributions of the three categories of language-specific grit, high task value group (node 3) was composed of a significantly higher percentages of high-level gritty learners (59.3 %) when respectively compared to moderate (node 2; 23.5 %) and low (node

1; 9.7 %) groups. Likewise, there were more moderate level gritty learners in the moderate level task value group (node 2; 49 %) and more low-level gritty learners within the low-level task value group (node 1; 54.2 %).

When the second level of the tree was examined, self-efficacy and intrinsic motivation were found to be statistically significant. The variable of self-efficacy classified node 2 into two groups,  $X^2 = 7.628$ ,  $df = 1$ ,  $p = 0.011$ . Those with a moderate degree of self-efficacy belonged to node 4 and those with a high-level self-efficacy composed node 5. Moderate degree self-efficacy group was dominated by those students with a moderate degree of language-specific grit for English (node 4; % 53.7) followed by low level gritty students (31.7 %). Similarly, in the high degree self-efficacy group, most students (60 %) appeared to have a high level of language-specific grit while there was only one student with a low degree of language-specific grit. The variable of intrinsic motivation significantly split node 3 into two groups (node 6 and node 7),  $X^2 = 5.070$ ,  $df = 1$ ,  $p = 0.049$ . Both groups were dominated by those students with a high degree of language-specific grit (42.9 % for node 6 and 68.4 % for node 7). However, there were more low-level gritty learners in node 6, the group with moderately intrinsically motivated students compared to node 7 of highly motivated students.

The results from the tree development through CHAID algorithm also enabled the researcher to form some rules in the form of “if-then” structure. Thus, given node 7, the results revealed that if a student possessed a high level of task value and intrinsic motivation for learning, then it could be stated with 0.684 probability that this student was going to be a high-level gritty learner of English. Moreover, if this student had a high degree of task value again but a moderate degree of intrinsic motivation, it was with 0.429 probability that this student would still be a high-level gritty learner of English.

Given the overall accuracy of the model in classifying the sample of this study, it was seen that 56 % of the participants were accurately classified. The classification accuracy results here implied that 102 respondents out of 182 were classified accurately in the observed sample (Table 4). The risk that the participants would be inaccurately classified in relation to their language-specific grit levels was found to be 44 % in this sample. However, when a test sample is used for cross-validation purposes, this risk is higher with 48%.

**Table 4**

*Classification Matrix*

Observed	Predicted			% Correct
	<= 54.00	55.00 - 65.00	66.00+	
<= 54,00	39	13	12	60.9 %
55.00 - 65.00	26	22	16	34.4 %
66.00 +	7	6	41	75.9 %
Overall %	39.6 %	22.5 %	37.9 %	<b>56.0 %</b>

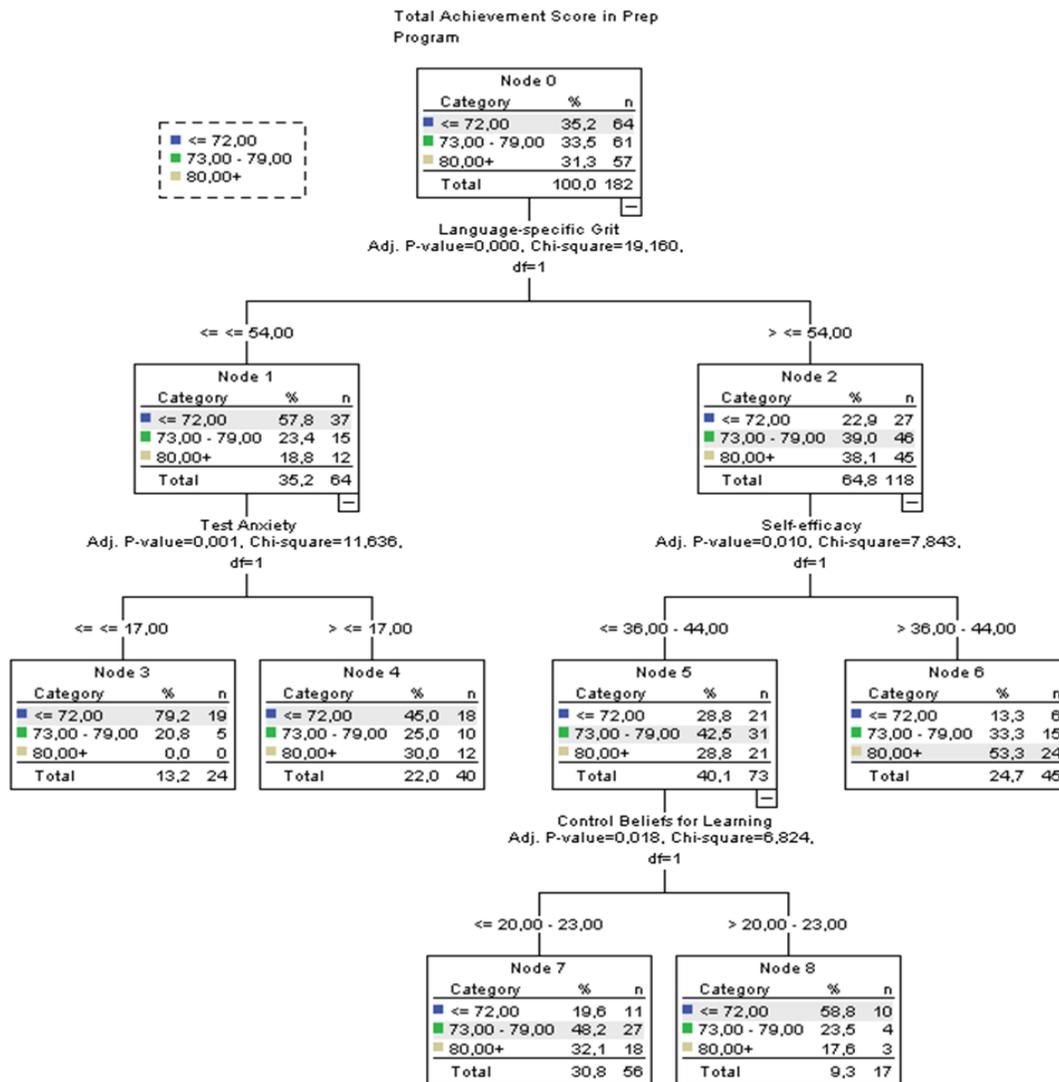
*Note.* Dependent Variable = Language-specific Grit Scores, Resubstitution = 44 %, Cross-validation = 48 %.

## **RQ2: Classification of Students' Achievement based on Language-specific Grit in Learning English and Six Motivational Variables**

CHAID analysis was conducted by using the students' achievement scores in the preparatory program as the dependent variable and six motivational variables and language-specific grit for learning English as independent variables. Figure 2 shows that the most significant and strong independent variable having the ability to classify students into three different groups of achievement was language-specific grit,  $X^2 = 19.160$ ,  $df = 1$ ,  $p = 0.000$ . The variable of language-specific grit split the sample of the study into two groups as node 1 ( $n = 64$ ) and node 2 ( $n = 118$ ). High gritty group (node 2) possessed a significantly higher percentage of top successful students (38.1 %) compared to low gritty group (node 1; 18.8 %). Furthermore, node 2 included a significantly lower percentage of bottom achievers (22.9 %) in comparison to node 1 (57.8 %).

**Figure 2**

*CHAID Decision Tree Model for Achievement Based on Language-Specific Grit and Motivational Variables*



the students with a high degree of self-efficacy in learning (node 6) appeared to be dominated by the students ( $n = 24$ ; 53.3 %) having a score 80 points and over as an achievement grade and only 13.3 % ( $n = 6$ ) of these students scored 72 and lower in the final test. In contrast, node 5 included more students ( $n = 21$ ) who scored 72 and lower but less students with 80 points and over ( $n = 21$ ; 28.8 %) compared to the same categories of node 6.

The variable of test anxiety significantly classified node 1 (low gritty learners) into two groups,  $X^2 = 11.636$ ,  $df = 1$ ,  $p = 0.001$ . One group was composed of the students with low levels of text anxiety (node 3) and the other with higher text anxiety levels (node 4). Both groups were dominated by those students getting 72 points or below; however, there were more successful students who indeed got scores between 73 and 79 points ( $n = 10$ ) or more than these ( $n=12$ ) in the group with higher text anxiety levels. In contrast, there were no students who got over 80 points in the group with lower levels of test anxiety. In this essence, node 4 possessed a significantly higher percentage of top successful students ( $n = 12$ ; 30 %) compared to node 3 ( $n = 0$ ).

Given the third level of the tree depth, control beliefs for learning was found to be significant for splitting node 5,  $X^2 = 6.824$ ,  $df = 1$ ,  $p = 0.018$  into two groups as respondents with moderate control beliefs for learning and those with high level of control beliefs. The group with higher control beliefs included significantly higher proportions of students with an average point of 72 or below (58.8 %) while the group with moderate control beliefs contained significantly higher proportions of students with an achievement score ranging between 73 and 79 points (48.2 %). Only three students in the higher control beliefs group and 18 students in the moderate control beliefs group scored as top achievers (over 80 points) as an achievement grade. Thus, those with high level control beliefs for learning appeared to be less successful in English.

Given some “if-then” type of rules using, for instance, for node 6, one can say that if a student reported a high level of language-specific grit and self-efficacy for learning, it was with 0.533 probability that this student was going to be a high achiever of English. Moreover, if a student was a low-level gritty learner of English and at the same time had low test anxiety, there was no probability that this student would become a high achiever in English (node 3). Furthermore, if a student possessed a moderate degree of control beliefs, a moderate level of self-efficacy and a high level of language-specific grit, there was only 0.196 probability that she/he would get lower scores in the final test. However, if this student had higher degrees of control beliefs for learning but not a moderate one as was given in the above case, the probability value that this student would get lower scores in the final test increased up to 0.588.

Overall accuracy and predictive potential of the model in classifying the sample of this study are shown in Table 5. It was seen that almost 54 % of the participants ( $n = 98$ ) were accurately classified in the current sample. This result also meant that when the students’ motivational characteristics (from six motivational dimensions of MSQL and one-dimensional PS) are known, the prediction risk is 46 % in this sample. However, when a test sample is used for cross-validation, this risk is 57 %.

**Table 5**

*Classification Matrix*

Observed	Predicted			% Correct
	<= 72.00	73.00 - 79.00	80.00 +	
<= 72.00	47	11	6	73.4 %
73.00 - 79.00	19	27	15	44.3 %
80.00 +	15	18	24	42.1 %
Overall %	44.5 %	30.8 %	24.7 %	<b>53.8 %</b>

*Note.* Dependent Variable = Achievement Scores, Resubstitution = 46 %, Cross-validation = 57 %.

## **Discussion and Future Directions**

The results from the CHAID for RQ1 indicated that task value, self-efficacy and intrinsic motivation were significantly associated with students' language-specific grit with task value differentiating the students with different degrees of language-specific grit as the strongest independent variable. These three variables should be particularly investigated in future language-specific grit studies. The results from the CHAID for RQ2 revealed that students with higher achievement scores had higher levels of language-specific grit for learning English, self-efficacy and text anxiety but surprisingly lower levels of control beliefs. Therefore, moderate but not high levels of control beliefs for learning is appreciated in order to promote students' achievement. However, future researchers should approach this implication with some caution in relation to the data collection instrument used. That is, the items in the control beliefs subscale might have provoked some negative opinions due to their sentence structures and negative wordings in Turkish. When the items in the control beliefs subscale were examined individually, one can easily see that these items in Turkish can also purport the idea that learners are to blame themselves or should consider their lacking related to the amount of study behaviors in the face of undesired learning outcomes. In this regard, Turkish students might have understood the items of the control beliefs dimension from a negative point of view instead of the positive connotations with the appreciation of efforts to manage an activity as was originally dedicated to them in English by the original instrument developers.

A certain number of independent variables (including demographics included in this study) or their sub-levels showed no potentials to differentiate high level gritty and successful learners of English from their less gritty or successful peers. Thus, a significant degree of reduction occurred in terms of model dimensionality as an expected result with data mining

analyses aiming to discover the real structure embedded within the data (Milanović & Stamenković, 2016). Therefore, such reductions and significant best subsets of independent variables should be considered carefully in the design of future studies.

This current study provided strong evidence for the presence of associations between grit for learning and other two other main variables, motivation and achievement. Muenks, Yang, and Wigfield (2018) previously found a similar result to the current study in that they reported strong relationships between effort sub-scale of general trait-level grit and the two variables, motivation and achievement in high school students. It was further reported in their study that self-efficacy and task value perceptions of the learners revealed strong associations with grit, which was again similar to the findings of this current study. The results from this current research also showed that grit was related to intrinsic motivation, which strongly corroborated with the results from previous studies (Changlek & Palanukulwong, 2015; Karlen, Suter, Hirta, & Maag Merki, 2019). In contrast to previous evidence about negative associations between test anxiety and grit (Changlek & Palanukulwong, 2015; Çelik & Sarıçam, 2016; Holtby, 2018), test anxiety was found to possess a facilitative role in promoting learners' grit in learning English in this study. In this essence, some researchers in the literature previously discussed the presence of two types of anxiety, facilitative and debilitating (Alpert & Haber, 1960; Carrier, Higson, Klimoski, & Peterson, 1984; Jones, 1995). Here, with this group of learners, facilitative type of anxiety might have worked in the sense that an obligatory exam to pass to the main faculty department at university posed some sort of a difficulty to do and work more for the students.

Another corroborating piece of evidence supporting the existence of links between self-efficacy perceptions and grit belonged to Rojas, Reser, Usher, and Toland (2012) who found associations between grit and self-efficacy and self-regulation scores in reading and math from

the elementary level students. The results from Reraki, Celik, and Sariçam's (2015) study also supported the current evidence as to the presence of associations between grit and academic motivation and achievement of university students. Similarly, Hagger and Hamilton (2019) found a significant association between grit (effort sub-scale) and high school students' science grades. In contrast, some research studies revealed contrasting results or weak explanations as to the associations when investigating how grit was related to achievement and motivation (Bazelaïs, Lemay & Doleck, 2016; Karlen, et al., 2019, Lumontod, 2019; Taşpınar & Külekçi, 2018).

While the above evidence supports the relationships between grit and achievement and motivational correlates, it is also important to note that effort dimension of the widely used grit scale was found to be consistently and (more) significantly related to achievement or motivation variables in comparison to the consistency of interest sub-scale by several researchers (Credé, Tynan, & Harms, 2017; Datu, Valdez, & King, 2016; Karlen et al., 2009; Muenks, Wigfield, Yang, & O'Neal, 2017; Steinmayr, Weidinger, & Wigfield, 2018; Wolters & Hussain, 2015). Such findings contributed to the use of a unidimensional instrument in this current research in that the PS instrument utilized in this study included or emphasized the effort dimension for the purposes of its content. In addition, the results from this study confirmed the use of PS as a reliable tool to investigate grit as a domain-specific construct in language learning.

One limitation could be related to sample size employed in the study for data mining analyses require the use of large sample sizes for the analyses to produce reliable results. However, when the purpose is to diagnose and identify the presence of any associations as a preliminary outlook, the use of smaller sample sizes is also appropriate (Milanović &

Stamenković, 2016). This study considered as a preliminary attempt to investigate the domain specificity of grit appear to be exempt from the limitation pertaining to sample size.

Given the pedagogical implications in the light of the results of this study, it could be recommended that educational interventions and trainings should be developed in order to promote students' grit for learning foreign languages. In such educational interventions, the constructs that are supportive of grit such as goal orientations and self-efficacy should be included and promoted. In line with the positive psychology understanding and implications supported by Seligman, Ernst, Gillham, Reivicha, and Linkins (2009), it is believed that skills that can increase grit or grit-like constructs and positive emotions can be taught and developed over time.

One comment should be made here concerning the strong influence of self-efficacy not only upon the students' grit for learning but also upon their achievement scores in English as a research finding from this study. Further research should be conducted to investigate and test the existence of multivariate causal relationships among these three variables as the literature concerning the presence of mutual associations between these variables has been already established by several researchers (Bandura, 2001; Multon, Brown, & Lent, 1991; Duckworth et al., 2007; Mohammadyari, 2012). In addition, the findings of this study implied the probability of another three-party relationship among motivation, grit and achievement for future research to investigate and the use of more advanced and varied data analysis methods is recommended to understand such multivariate relationships among the variables. It is seen that there is already a dominance of quantitative research designs in research into grit in the literature. Therefore, qualitative and mixed-design studies are especially recommended for future purposes in order to

understand the real nature of the construct in the light of learners' personal understandings and perceptions related to their grittiness.

## References

- Akos, P., and Kretchmar, J. (2017). Investigating grit at a non-cognitive predictor of college success. *The Review of Higher Education*, 40(2), 163-186. doi: 10.1353/rhe.2017.0000
- Alpert, R., and Haber, N. N. (1960). Anxiety in academic achievement situations. *Journal of Abnormal Social Psychology*, 61(2), 207-215. doi: 10.1037/h0045464
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52(1), 1-26.
- Bazelais, P., Lemay, J. D., Doleck, T. (2016). How does grit impact students' academic achievement in Science? *European Journal of Science and Mathematics Education* 4(1): 33-43.
- Büyüköztürk, Ş., Akgün, Ö. E., Demirel, F. ve Özkahveci, Ö. (2004). Gütülenme ve Öğrenme Stratejileri Ölçeği'nin Türkçe formunun geçerlik ve güvenirlik çalışması. *Kuram ve Uygulamada Eğitim Bilimleri*, 4(2), 207-239.
- Carrier, C., Higson, V., Klimoski, V. & Peterson, E. (1984). The effects of facilitative and debilitating achievement anxiety on note taking. *Journal of Education Research*, 77(3), 133-138. doi: 10.1080/00220671.1984.10885512
- Changlek, A., and Palanukulwong, T. (2015). Motivation and grit: Predictors of language learning achievement. *Veridian E-Journal*, 8(4), 23-38.
- Chen, C. S. (2002). Self-regulated learning strategies and achievement in an introduction to information systems course. *Information Technology, Learning and Performance Journal*, 20 (1), 11-25
- Cox, C. M. (1926). Genetic studies of genius: Vol. 2. *The early mental traits of three hundred geniuses*. Stanford University Press.

- Credé, M., Tynan, M., Harms, P. D. (2017). Much ado about grit: A meta-analytic synthesis of the grit literature. *Journal of Personality and Social Psychology, 113*, 492–511.
- Çelik, İ., & Sariçam, H. (2016, June). Grit and test anxiety in Turkish children and adolescents [Abstract]. Paper presented at ERPA International Congresses on Education, Sarajevo, Bosnia and Herzegovina. doi: 10.13140/RG.2.1.2750.1688
- Datu, J.A.D., Valdez, J.P.M. & King, R.B. (2016). Perseverance Counts but Consistency Does Not! Validating the Short Grit Scale in a Collectivist Setting. *Current Psychology, 35*, 121–130. doi: 10.1007/s12144-015-9374-2
- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92* (6), 1087–1101. doi: 10.1037/0022-3514.92.6.1087
- Eccles, J. (1983). Expectancies, Values, and Academic Behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives: Psychological and sociological approaches* (pp. 75–146). Freeman.
- Eccles, J., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development, 64*(3), 830–847. doi: 10.2307/1131221
- Eccles, J. S., & Wigfield, A. (1995). In the Mind of the Actor: The Structure of Adolescents' Achievement Task Values and Expectancy-Related Beliefs. *Personality and Social Psychology Bulletin, 21*(3), 215–225. <https://doi.org/10.1177/0146167295213003>
- Eccles, J.S. & Wigfield, A. (2002). Motivational Beliefs, Values and Goals. *Annual Review of Psychology, 53*. 109-132. doi: 10.1146/annurev.psych.53.100901.135153

- Feather, N. T. (1992). Values, valences, expectations, and actions. *Journal of Social Issues*, 48(2), 109–124. doi: 10.1111/j.1540-4560.1992.tb00887.x
- Galton, F. (1892). *Hereditary Genius: An inquiry into its laws and consequences*. Macmillan.
- Hagger, M.S. & Hamilton, K. (2019). Grit and self-discipline as predictors of effort and academic attainment. *British Journal of Educational Psychology*, 89(2), 324-342. doi: 10.1111/bjep.12241
- Heckman, J. J. & Rubinstein, Y. (2001). The Importance of Noncognitive Skills: Lessons from the GED Testing Program, *American Economic Review*, 91(2): 145-149. doi: 10.1257/aer.91.2.145
- Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3), 411-482. doi: 10.1086/504455
- Holtby, N. M. (2018). *Grit, coping, and math anxiety: examining the pathways through Which devotion to long-term goals may promote student well-being* (Unpublished Master's Thesis). University of Toronto, Toronto.
- Jones, G. (1995). More than just a game: Research developments and issues in competitive state anxiety in sport. *British Journal of Psychology*, 86 (4), 449-478. doi: 10.1111/j.2044-8295.1995.tb02565.x
- Karlen, Y., Suter, F., Hirt, C., & Maag Merki, K. (2019). The role of implicit theories in students' grit, achievement goals, intrinsic and extrinsic motivation, and achievement in the context of a long-term challenging task. *Learning and Individual Differences*, 74, Article 101757. doi: 10.1016/j.lindif.2019.101757

- Kass, G. V. (1980). An exploratory technique for investigating large quantities of categorical data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 29(2), 119–127.  
Retrieved from [https://www.jstor.org/stable/2986296?seq=1#metadata\\_info\\_tab\\_contents](https://www.jstor.org/stable/2986296?seq=1#metadata_info_tab_contents)
- Kurt Taspınar, H. (2018) & Külekçi, G. (2018). GRIT: An Essential Ingredient of Success in the EFL Classroom. *International Journal of Languages' Education and Teaching*. 6(3), 208-226. doi: 10.18298/ijlet.3137.
- Lawler, E. E., & Porter, L. W. (1967). Antecedent attitudes of effective managerial performance. *Organizational Behavior & Human Performance*, 2(2), 122–142. doi: 10.1016/0030-5073(67)90026-8
- Lleras, C. (2008). Do skills and behaviours in high school matter? The contribution of noncognitive factors in explaining differences in educational attainment and earnings. *Social Science Research* 37(3),888–902.
- Locke, E. A., & Latham, G. P. (1990). A theory of goal setting and task performance. Prentice Hall.
- Lumontod, R. Z., III. (2019). How grit, college adjustment, and happiness predict freshmen students' academic performance?. *International Journal of Research Studies in Psychology*. 8(2). doi: 10.5861/ijrsp.2019.4010.
- Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, 50(4), 370–396. doi: 10.1037/h0054346
- Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, 38(1), 30–38.

- Mutlu, G. (2017). *Linking EFL learning environment characteristics to persistence in EFL learning: A mixed-design study (Unpublished doctoral dissertation)*. Middle East Technical University, Ankara.
- Mutlu, G., & Yıldırım, A. (2019). Learning Environment Perceptions and Student Background Variables as Determinants of Persistence in EFL Learning. *SAGE Open*, 9(4), 1-17.  
<https://doi.org/10.1177/2158244019898805>
- McClelland, D.C. (1961). The achieving society. *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship*. Retrieved from <https://ssrn.com/abstract=1496181>
- Milanović, M. & Stamenković, M.. (2016). CHAID Decision Tree: Methodological Frame and Application. *Economic Themes*. 54(4), 563-586. doi: 10.1515/ethemes-2016-0029.
- Mohammadyari, G. (2012). Comparative study of relationship between general perceived self-efficacy and test anxiety with academic achievement of male and female students. *Procedia-Social and Behavioral Sciences*, 69, 2119–2123.
- Muenks, K., Wigfield, A., Yang, J. S., & O'Neal, C. R. (2017). How true is grit? Assessing its relations to high school and college students' personality characteristics, self-regulation, engagement, and achievement. *Journal of Educational Psychology*, 109(5), 599–620. doi: 10.1037/edu0000153.
- Muenks, K., Yang, J. S., & Wigfield, A. (2018). Associations between grit, motivation, and achievement in high school students. *Motivation Science*, 4(2), 158–176. doi: 10.1037/mot0000076.
- Peterson, C., & Seligman, M. E. P. (2004). *Character strengths and virtues: A handbook and classification*. Oxford University Press.

- Pintrich, P. R. (1988). A process-oriented view of student motivation and cognition. In J. S. Stark & L. Mets (Eds.), *Improving teaching and learning through research. New directions for institutional research*, 57 (pp. 55-70). Jossey-Bass.
- Pintrich, P. R. (1989). The dynamic interplay of student motivation and cognition in the college classroom. In C. Ames & M. Maehr (Eds.), *Advances in motivation and achievement: Vol. 6. Motivation enhancing environments* (pp. 117-160). JAI Press.
- Pintrich, P. R., & De Groot, E.V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1), 33-40.
- Pintrich, P. R., Smith, D. A. F, Garcia, T., & McKeachie, W. J. (1991). *A Manual for the use of the motivated strategies for learning questionnaire (MSLQ)*. Ann Arbor: University of Michigan, National Center for Research to Improve Secondary Teaching and Learning. Retrieved from <https://files.eric.ed.gov/fulltext/ED338122.pdf>
- Reraki, M., Çelik, İ., & Sarıçam, H. (2015). Grit as a mediator of the relationship between motivation and academic achievement. *Ozean Journal of Social Science*, 8(1), 19-32. doi: 10.13140/RG.2.2.15063.78240
- Rojas, J. P., Reser, J.A., Usher, E.L., & Toland, M.D. (2012). *Psychometric properties of the academic grit scale*. Lexington: University of Kentucky. Retrieved from <https://sites.education.uky.edu/motivation/files/2013/08/PoiasPeserTolandUsher.pdf>
- Seligman, M.E.P., Ernst, R.M., Gillham, J., Reivich, K. and Linkins, M. (2009). Positive Education: Positive Psychology and Classroom Interventions. *Oxford Review of Education*, 35(3), 293-311.

- Steinmayr, R., Weidinger, A. F., & Wigfield, A. (2018). Does students' grit predict their school achievement above and beyond their personality, motivation, and engagement? *Contemporary Educational Psychology, 53*, 106–122. doi: 10.1016/j.cedpsych.2018.02.004
- Von Culin, K. R., Tsukayama, E., & Duckworth, A. L. (2014). Unpacking grit: Motivational correlates of perseverance and passion for long-term goals. *The Journal of Positive Psychology, 9*(4), 306–312. doi: 10.1080/17439760.2014.898320
- Wigfield, A. (1994). Expectancy-Value Theory of Achievement Motivation: A Developmental Perspective. *Educational Psychology Review, 6*(1), 49-78. doi: 10.1007/BF02209024
- Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review, 12*(3), 265–310. doi: 10.1016/0273-2297(92)90011-P
- Wigfield, A., Harold, R., Eccles, J. S., Blumenfeld, P. B., Aberbach, A., Freedman-Doan, C., and Yoon, K. S. (1992). *The Structure of Children's Ability Perceptions and Achievement Values: Age, Gender, and Domain Differences*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco. Retrieved from <http://education-webfiles.s3-website-us-west-2.amazonaws.com/arp/garp/presentations/eccles92h.pdf>
- Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68–81. doi: 10.1006/ceps.1999.1015
- Wigfield, A., & Eccles, J. S. (2002). The Development of Competence Beliefs, Expectancies for Success and Achievement Values from Childhood through Adolescence. In A. Wigfield, & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 91-120). Academic Press. <http://dx.doi.org/10.1016/B978-012750053-9/50006-1>

Wolters, C.A. & Hussain, M. (2015). Investigating grit and its relations with college students' self-regulated learning and academic achievement. *Metacognition Learning*, *10*, 293–311.  
doi: 10.1007/s11409-014-9128-9

## **5<sup>th</sup> Grade English Language Teaching Curricula in Turkey: Expectations vs. Realities**

**Esin Dündar and Ali Merç**

### **Abstract**

This study aims to reveal the opinions of English teachers, advisors and a member of Curriculum Development Committee of Ministry of National Education about 5<sup>th</sup> grade English Language Curricula in Turkey. To this end, 100 English teachers were given Curriculum Opinion Scales. Fifteen English teachers, three advisors, and a member of Curriculum Development Committee were also interviewed through semi-structured interviews in order to add to the quantitative findings. The curricula were evaluated in terms of learning objectives and content, assessment and evaluation, coursebook as an instructional material, language skills, principle of appropriateness, and continuum of learning. Furthermore, participants shared their opinions on curriculum literacy, class hours, methodological perspectives, and examination system. The findings are discussed in comparison with the related literature.

**Keywords:** Curriculum evaluation, English language teaching, Turkey

*Dr. Esin Dündar is a research assistant at ELT Department of Mersin University, Turkey. She can be reached at [dundar.esin@mersin.edu.tr](mailto:dundar.esin@mersin.edu.tr)*

*Dr. Ali Merç is an Associate Professor of ELT at Anadolu University, Turkey. He can be reached at [amerc@anadolu.edu.tr](mailto:amerc@anadolu.edu.tr)*

The role of English in economy and its importance to reach information are two of the crucial factors shaping the field of language teaching (Richards & Rodgers, 2014). English curriculum has been influenced by the changing perspectives on how to teach a foreign language over time, and for today, countries expect learners to use English for communication and focus on communicative principles which causes the need for interactive and learner-centered practice (Wedell & Grassick, 2018). Although the need for learning English has led the demand for effective language curricula, which can upskill individuals in a globalized world, teaching policies are criticized for not meeting this demand (Richards & Rodgers, 2014).

In order to educate students, a carefully planned curriculum development process is crucial as curriculum is not just “a school board-approved textbook series” (Ornstein & Hunkins, 2018, p. 209). Many scholars have defined the term ‘curriculum’ by pointing out different aspects. As stated by Wiggins and McTighe (2005), curriculum is “more than a list of topics and lists of key facts and skills (the ‘inputs’)” (p. 6). While Bobbitt (2004) focuses on experience, Richards and Rodgers (2014) approach the term as a design including teaching and learning situations. Focusing specifically on language teaching curriculum, Richards (1990) believes that the outcome of a second language education has a close relation to our perspective on language curriculum development process. Curriculum development process does not just involve definitions of its elements and in-class application. We need evaluation to see whether curriculum is really applicable in a classroom environment and if we can reach expectations and learning outcomes specified (Richards, 2001a). Additionally, focusing on the actions of the teachers and learners within the classroom settings, curriculum evaluation is crucial to see to what extent curriculum serves educational development (Ornstein & Hunkins, 2018).

Regardless of how it is defined, curriculum influences the teachers, students, and society (Ornstein & Hunkins, 2018). Additionally, it is at the center of the education system as it forms a combination of theory, practice, and objective (Null, 2011). Although curriculum evaluation is not simply about the opinions of the stakeholders, their opinions and experiences have to be taken into consideration for an effective application because their perspectives and local contexts can create various interpretations and implementations of the curriculum.

ELT curriculum evaluation is not a new phenomenon for the field, the related literature provides studies conducted in various contexts from 1985 to present. For instance, in teaching English to young learners in Polish context, Stec (2011) claims that content of grammar, classroom management, insufficient time, learners' attention span, and background knowledge can be listed as some of the challenges for the teachers. Focusing on designing a curriculum guide for the Brazilian context, Gimenez and Tonelli (2013) believe that authorities should collaborate with the teachers during the curriculum development process. Moreover, in Chinese context, the study of Li (2010) investigated the implementation of English language policy and revealed that teachers were willing to share their experiences with the authorities and attend the process in order to design a curriculum which could meet the needs of the learners, however it was nearly impossible to reach the higher officials. Additionally, Salahuddin et al. (2013) evaluated the primary English curriculum in Bangladesh and pointed out lack of curriculum literacy, crowded classrooms, class hours, teaching materials, and in-service training as the factors affecting the application of the curriculum. Lastly, the study of Romero et al. (2014) evaluated ELT curricula in Mexico from the perspectives of various stakeholders and found out how uninformed the school management and teachers were about the curriculum.

Evaluation of ELT curriculum has been considerably studied in Turkish context, too. Regardless of the grade evaluated, the following points were often mentioned in ELT curriculum evaluation studies in Turkish context as the factors affecting the application of the curriculum: effect of examination system (Berkant et al., 2019; Dönmez, 2010; Karcı, 2012; Kaya, 2018), in-service training (Dinçer & Koç, 2020; Özüdoğru & Adıgüzel, 2015; Yedigöz-Kara, 2019; Zehir-Topkaya & Küçük, 2010), class hours (Cihan & Gürlen, 2013; Erdoğan, 2005; Karcı, 2012; Kaya, 2018; Zehir-Topkaya & Küçük, 2010), and teaching materials (Berkant et al., 2019; Dinçer & Koç, 2020; Erdoğan, 2005; Karcı, 2012; Zehir-Topkaya & Küçük, 2010).

### **Purpose of the Study**

Since the 2017-2018 academic year, there have been two English language curricula applied for 5<sup>th</sup> graders (young learners at the age of 10) in the Turkish education system. One of them was developed as a part of second-eighth grade English Language Curriculum (ELC). Adopting communicative principles, ELC aims for A1 proficiency level with 10 units, three weekly hours, and learning objectives specified for each language skill except writing (MNE, 2018a). After being piloted in selected schools, intensive English Language Curriculum for 5<sup>th</sup> graders (IntELC) was updated and the schools were allowed to have English courses up to 18 hours with the approval of school administration (MNE, 2018b). Just like ELC, IntELC also follows communicative principles and aims for beginning of B1 with 540 hours of classroom input, 36 units, and learning objectives specified for all language skills (MNE, 2018c).

As stated by Fullan (2007), “educational change depends on what teachers do and think- it is as simple and complex as that” (p. 129). However, authorities seem to share no or meager information with teachers in terms of curriculum changes. This lack of information causes disregard for local content realities, inconsistency between curriculum and elements such as

examinations or textbooks, and teachers do not know what they are expected to do (Wedell & Grassick, 2018). Curricular changes affect the stakeholders and it is important to see how well they adopt these changes because unadopted reforms can cause loss of resources (Bümen et al., 2014).

The present study is an attempt to evaluate ELT curricula for 5<sup>th</sup> graders from the perspectives of English teachers, advisors, and a member of Curriculum Development Committee in the country of Turkey. As they are responsible for in-class practice of the curriculum, the experiences and opinions of the teachers can provide feedback for the curriculum. However, evaluating the curricula only from the perspectives of the teachers can lead to misconceptions. Thus, getting the opinions of other parties involved can provide a better understanding of the curricula. To this end, the study searches for an answer to the following research questions:

1. What are the opinions of English teachers about IntELC and ELC?
2. What are the opinions of the advisors and the committee member about IntELC and ELC?

### **Significance of the Study**

An overview of the related literature reveals three crucial points to be considered. First, although the number of the studies investigating ELT curricula from the perspectives of the stakeholders is quite a lot, the variety of the stakeholders seems to be neglected, especially in Turkish context. Compared to the number of the studies getting the opinions of English teachers or students, studies covering the opinions of other stakeholders such as parents, supervisors, officials, inspectors, and academicians are quite rare. Second, when we investigate the data collection tools used in ELT curriculum evaluation studies, we can state that Likert type scales are considerably preferred in the Turkish context (e.g. Erdoğan, 2005; Gürsoy & Eken, 2018;

Kambur, 2018; Yedigöz-Kara, 2019). Those studies using Likert type scales only reported small-scale pilot studies or taking the opinions of the experts for the development of the scales. Only a limited number of studies reported the results of exploratory factor analysis. Claiming that there is a misconception of scale development among researchers who often develop scales by basically getting some items together, DeVellis (2017) explains the risk of this misconception as “a researcher not only may fail to exploit theory in developing a scale but also may reach erroneous conclusions about a theory by misinterpreting what a scale measures” (p. 32).

Although there are scales for the evaluation of ELT curriculum, there is still a need for valid and reliable scales, considering how those in the related literature are developed. Finally, both ELC and IntELC have been evaluated from the perspectives of teachers in many studies. However, the studies have mainly focused on ELC before its updated version in 2018. Thus, there is still a need for providing insightful information on the updated version of ELC. Although it is gaining popularity among ELT curriculum evaluation studies (e.g. Aksoy et al., 2018; Dinçer & Koç, 2020; Kambur, 2018), IntELC have been mostly evaluated from the perspectives of the teachers or students. More studies are necessary not only to compare the two curricula applied for 5<sup>th</sup> grade students but also to evaluate them from the perspectives of other stakeholders. The present study singles out by developing a valid and reliable curriculum evaluation scale as a data collection tool, focusing on the comparison of currently designed and updated curricula, and taking the opinions of not only teachers but also advisors and a committee member.

## **Methods**

### **Research Design**

Including both quantitative and qualitative approaches into the procedure, the study adopted an explanatory sequential mixed method research design to gain a general understanding

of the issue by collecting and analyzing the quantitative data first, and then, supporting it through qualitative data (Creswell, 2012).

## **Participants**

The data were collected from two participant groups. As they are in charge of the practical aspect of the curriculum, the experiences and evaluations of the teachers can provide valuable feedback for its effectiveness. Thus, one of the participant groups included 100 English teachers, working with 5<sup>th</sup> graders; 56% of the teachers had teaching experience more than 10 years and 28% of them were following IntELC. Most of the participants had a BA degree and were supporting the application of IntELC.

In the Turkish context, the Curriculum Development Committee under MNE is responsible for curriculum design and this committee is formed with academicians from the related profession and teachers from the field. In order to have a better understanding about curriculum, their experiences during the process of development and perspectives can play a crucial role. Thus, within this study, three academicians as the advisors and an English teacher as the member of the Curriculum Development Committee of MNE formed the second participant group. Advisors took charge during the updating process of ELC and development of IntELC while committee members only took part in the development process of IntELC.

## **Data Collection Tools**

### ***Curriculum Opinion Scale for English Teachers (COSET)***

The quantitative data for the study were collected through the Curriculum Opinion Scale for English Teachers (COSET), developed by the researchers. Adopting a five-point Likert type rating, COSET was formed with three parts: demographic features, 41 scale items, and an optional part for the participants to state their further opinions about the curricula. After creating

the item pool based on the review of the related literature and interviews with English teachers, a draft form was prepared and evaluated by the experts from the field of ELT, Program Development and Evaluation, and Assessment and Evaluation. A form with 67 items was prepared and piloted with 262 English teachers of 5<sup>th</sup> graders to conduct Exploratory Factor Analysis (EFA). A principal axis factor analysis was conducted on items with oblique rotation (promax). The results of Kaiser-Meyer-Olkin (KMO) and Bartlett's test of sphericity were investigated to determine the suitability of the data set for EFA and sample size was verified for the analysis (KMO= .972,  $\chi^2 = 16879.065$ ,  $df = 1711$ ,  $p < .01$ ). By taking Kaiser Criterion and Cattell Scree Plot into account, COSET was decided to have a structure of six factors. Using 0.4 as a cutoff point, items, either having a cross loading difference below .1 or with factor loadings below 0.4, were eliminated. As a result of EFA, a structure with six factors, explaining 74.580% of variance, and 41 items was determined. Reliability of the scale was measured through Cronbach's alpha coefficient, Guttman's lambda, and McDonald's omega ( $\alpha$  .985,  $\lambda-2$  .985,  $\omega$  .985). In order to validate the 6-factor structure of COSET, fit indices were investigated by conducting Confirmatory Factor Analysis (CFA) with the data collected from 153 English teachers of 5<sup>th</sup> graders.  $\chi^2/df$  (1.91), CFI (.96), IFI (.96), and NNFI (.96) model indices had perfect fit while RMSEA (.077 / 90% Confidence Interval .071; .083), SRMR (.066), and NFI (.93) showed good fit, which verified the six factors structure of COSET. The factors were labeled as follows: learning outcomes and content, assessment and evaluation, coursebook as an instructional material, principle of appropriateness, language skills, and continuum of learning.

### ***Semi-Structured Interviews***

The qualitative data were collected through semi-structured interviews with 15 English teachers, three advisors, and a member of Curriculum Development Committee of MNE. After

the analyses of the responses given to COSET, the questions of semi-structured interviews with teachers were prepared. The components of the curricula with high or low ratings were addressed through the interview questions. As the aim of interviews with advisors and committee member was to see their responses to the opinions of the teachers about the curricula, the set of questions was prepared after the analyses of the qualitative and quantitative data collected from the teachers. The interview sessions were conducted after the evaluations of the questions by the experts from the field.

### **Data Collection Procedure**

Once COSET was developed, permissions from Ethics Committee and Directorate of National Education were taken and schools were selected through the stratified random sampling method. After the middle schools in a city center in Central Anatolia Region were enlisted and classified in line with geographical and socio-economic features, they were selected randomly from the list. The administrators of the selected schools were contacted and informed about the study. The data collection took place between 2019 Spring and 2020 Fall semesters within the school environment. After the analysis of the responses given to COSET, a set of questions was prepared for the semi-structured interviews. English teachers and interview sessions were conducted with voluntary teachers. After the analyses of the data collected from English teachers, semi-structured interviews with the advisors and committee member were conducted. They were reached through email and given detailed information about the aim of the study. While one of the advisors and committee member preferred to give their answers in a written form, two online interview sessions were conducted with two of the advisors in June 2020.

## **Data Analysis**

In order to make an item-based evaluation of the responses, mean scores for each item were interpreted through the formula of “(highest point in the Likert scale – lowest point)/the number of the levels” suggested by Erkuş (2012). By adding .80 (5-1/5) the categories were defined as; 1 – 1.80 (very low), 1.81- 2.60 (low), 2.61- 3.40 (moderate), 3.41- 4.20 (high) and 4.21-5 (very high). In order to compare subtotals and total scores in line with the curriculum followed by the participants, first, the normality of the data was checked through Kolmogorov Smirnov test of normality. As the data set did not have a normal distribution ( $p < .05$ ), Mann Whitney U test was used to make a comparison between the groups. Considering the difference between statistical significance and practical significance, effect size values were also calculated through the formula “ $r = Z/\sqrt{N}$ ” and interpreted in line with the reference points:  $\pm .1$  small effect,  $\pm .3$  medium effect size, and  $\pm .5$  large effect size (Field, 2009). Qualitative data collected through semi-structured interviews were used to support quantitative data through direct quotations.

## **Results**

The opinions of the participants on ELC and IntELC are presented in the same order of the subscales in COSET and supported through responses shared during semi-structured interviews.

### **Learning Outcomes and Content**

Fourteen items of the first subscale were mainly about the reachability of the outcomes and their suitability to the learners in aspects such as the daily life of the students, their needs and interests. Descriptive statistics and Mann Whitney U test results are presented in Table 1.

**Table 1***The Results Regarding Learning Outcomes and Content*

Item	Curriculum	M	SD	Mann Whitney U			
				U	z	p	r
1	IntELC	3.85	.84	877	-1.127	.260	-0.11
	ELC	3.59	.92				
2	IntELC	3.64	.73	1005	-.027	.978	-0.002
	ELC	3.62	.84				
3	IntELC	3.89	.78	835.5	-1.515	.130	-0.15
	ELC	3.63	.79				
4	IntELC	3.57	.79	910	-.829	.407	-0.08
	ELC	3.43	.86				
5	IntELC	3.85	.93	990	-.151	.880	-0.01
	ELC	3.84	.91				
6	IntELC	3.60	.83	1002.5	-.045	.964	-0.004
	ELC	3.62	.89				
7	IntELC	3.50	.92	912.5	-.778	.437	-0.07
	ELC	3.29	1				
8	IntELC	3.60	.91	999.5	-.071	.943	-0.007
	ELC	3.59	.89				
9	IntELC	3.82	.77	842	-1.400	.162	-0.14
	ELC	3.54	.90				
10	IntELC	3.60	.83	792	-1.741	.082	-0.17
	ELC	3.16	1.1				
11	IntELC	3.85	.93	897	-.939	.348	-0.09
	ELC	3.68	.96				
12	IntELC	3.85	.80	939.5	-.605	.545	-0.06
	ELC	3.76	.83				
13	IntELC	3.71	.89	908	-.848	.396	-0.08
	ELC	3.58	.83				
14	IntELC	3.75	.88	805.5	-1.726	.084	-0.17
	ELC	3.45	.82				
Subtotal		52.14	8.8	832	-1.353	.176	-0.13
		49.8	9.2				

As presented in Table 1, except two items (7 and 10) for ELC, the teachers gave high ratings to the curricula. For IntELC the outstanding feature within the subscale was the consistency among the outcomes (Item 3, M= 3.89) while ELC got the highest ratings for reachability of the outcomes specified for vocabulary (Item 5, M= 3.84). Although IntELC got higher subtotal mean score, Mann Whitney U test results showed statistically non-significant difference ( $p > .05$ ) with small effect size ( $r < \pm .3$ ).

Although responses given to COSET could be interpreted as participants being content with the curricula, the interviewees draw a different picture about in-class practice. Except one of them, all the interviewees following IntELC stated that they could not reach the outcomes specified for language skills as grammar teaching took a great deal of their class hours. As well as criticizing the amount of new vocabulary and lack of learning outcomes specified for writing skills, the teachers following ELC believed that unbalanced distribution of the outcomes among the units and limited class hours created the difference between what was expected from the curriculum and in-class practice.

Responding to the criticisms of the teachers, the advisors and committee member insisted that neither of the curricula was grammar-based and not a single learning outcome was specified for grammar. They believed that the examination system forced teachers to focus on every single grammar structure. The point of how curriculum was presented in the coursebook was also highlighted, stating that they had to omit learning outcomes for writing skills in order to simplify ELC. Advisors also presented a theory-based rationale for the amount of new vocabulary in ELC. One of the advisors objected to the criticism in these words:

We did that on purpose because without loading the words called as language shower or input, children can't have an output. It is stated in Krashen's Natural Approach or Communicative Approach, so the coursebook writer should get that input and form a context within the coursebook and prepare activities answering what we know when we know the vocabulary. The children need the language shower otherwise how they can hear those words. Studies show that at A1 level a child needs to know 300-400 words.

(A3)

## Assessment and Evaluation

Items of this subscale evaluated the curricula in terms of considering developmental features of the students, peer and self-evaluation, providing feedback for the efficacy of the curricula, and applicability of the suggested techniques. Mean scores of the responses and Mann Whitney U test results are presented in Table 2.

**Table 2**

*The Results Regarding Assessment and Evaluation*

Item	Curriculum	M	SD	Mann Whitney U			
				U	z	p	r
15	IntELC	3.03	1	1007.5	-.004	.997	-0.0004
	ELC	2.98	.97				
16	IntELC	3.17	.86	844	-1.317	.188	-0.13
	ELC	2.87	1				
17	IntELC	3.21	.78	996.5	-.093	.926	-0.09
	ELC	3.15	.97				
18	IntELC	3.60	.68	860	-1.275	.202	-0.12
	ELC	3.33	.85				
19	IntELC	3.39	.87	768.5	-1.944	.052	-0.19
	ELC	3	.90				
20	IntELC	3.32	.90	971	-.306	.760	-0.03
	ELC	3.25	.93				
21	IntELC	3.50	.79	987.5	-.174	.862	-0.017
	ELC	3.48	.76				
22	IntELC	3.10	1.1	801	-1.638	.101	-0.16
	ELC	2.62	1.1				
23	IntELC	3.42	.83	813.5	-1.583	.114	-0.15
	ELC	3.06	1				
24	IntELC	3.46	.88	964.5	-.366	.714	-0.03
	ELC	3.38	.79				
<b>Subtotal</b>		33.2	7	832	-1.353	.176	-0.13
		31.1	7				

As it can be understood from Table 2, teachers were content with the applicability of the suggested evaluation techniques. Consistency between the content and evaluation techniques, guiding teachers in terms of assessment and evaluation, and relevancy of the assessment and

evaluation to the developmental features of the students were found to be other outstanding features of IntELC while ELC got moderate ratings. However, Mann Whitney U test results showed statistically non-significant differences with small effect sizes ( $p > .05$ ,  $r < \pm .3$ ).

The leading topics discussed during the interviews with the teachers were applicability of the suggested evaluation techniques and evaluation of language skills. Although applicability of the suggested techniques got high ratings in COSET, the interviewees questioned the capability of those techniques in evaluating language skills. They also mentioned disregarding the level of the students, limited class hours, and the amount of the outcomes as problems.

When it comes to the evaluation of language skills, interviews revealed that schools had different procedures regardless of the curriculum they followed. While some of the teachers, following IntELC, were evaluating language skills during in-class activities, one of the schools was conducting skills-based exams. For the teachers following ELC, limited class hours and lack of technical equipment forced teachers to focus more on grammar and vocabulary rather than language skills and this resulted in students being inexperienced in skills-based activities. Sharing an information given during the district group teacher meeting, one of the teachers stated:

It was mentioned during district group teachers meeting that if we decided as group for teachers to conduct listening exam for 5<sup>th</sup> graders, we could, but none of the teachers in our district was willing to take that risk because in order to test the listening skills of the students. First, I should really concentrate on those skills during the lessons. As I can't do that, there is no point in making a listening exam. (ELC, T2)

Responding to the opinions of the teachers, advisors highlighted the importance of evaluating language skills and they explained the issue through two main factors: lack of

knowledge on how to conduct skills-based exams and conformism. They stated that as well as being stated in Basic Law of Education, conducting skills-based exam was also crucial to reach the main objective of the curricula.

### Coursebook as an Instructional Material

The third subscale evaluates MNE-approved coursebooks. Descriptive statistics and Mann Whitney U test results are presented in Table 3.

**Table 3**

*The Results Regarding Coursebook as an Instructional Material*

Item	Curriculum	M	SD	Mann Whitney U			
				U	z	p	r
25	IntELC	3.03	1	910	-.786	.432	-0.07
	ELC	3.13	1				
26	IntELC	3.10	1	848.5	-1.274	.203	-0.12
	ELC	2.76	1				
27	IntELC	2.78	1	998	-.080	.936	-0.008
	ELC	2.73	.99				
28	IntELC	3.14	1	998.5	-.077	.939	-0.007
	ELC	3.11	1				
29	IntELC	3.28	1	1006.5	-.012	.990	-0.001
	ELC	3.26	.99				
Subtotal		15.3	4.4	995	-.100	.920	-0.01
		15	4.2				

Based on the mean scores presented in Table 3, we can state that participants were not content with MNE-approved coursebook regardless of their curriculum. The capability of the coursebook in terms of improving speaking skills got the lowest scores from the participants. Opinions shared during the interviews were consistent with the ratings. The teachers following ELC found the material consistent with the objectives of the curriculum, but they criticized it in terms of organization of the units, number of the activities, and improving language skills. As the teachers were unaware of the coursebook sets prepared specifically for IntELC, they were using

the coursebook designed for ELC as well as one published by international publishers. The main motive behind using two coursebooks was to prepare students for high-stakes exams by following an MNE-approved one and to use a supportive coursebook suitable to the content of IntELC. The problems with MNE-approved coursebook were functionality, unbalanced organization of the units, the number of the activities, and not appealing the interest of the students. Thus, they felt the need to follow an additional coursebook to support their students. One of the teachers expressed her opinions in these words:

I don't think that the coursebook is efficient, it was partially prepared in line with the outcomes. In the first unit of MNE-approved coursebook there are too many grammar structures to cover, I dealt with the first unit for weeks, there were activities following one another and students got bored with so much grammar points. There is nothing fun, there are so many structures in the first unit, it gets simple with the second unit. In the extra coursebook we use there is a song at the end of each unit, the students are having fun but in MNE-approved coursebook there is no place to relax. (IntELC, T2)

Pointing out the importance of selecting the best option possible, advisors disagreed with the opinions of the teachers regarding MNE-approved coursebooks. One of the advisors stated that the source of the problem was not the coursebook but the methodological perspective of the teachers and believed that MNE should charge publishing houses for the design of the coursebooks instead of preparing them itself. It was also stated that there were indeed two online coursebook sets prepared for IntELC and they were labeled as supportive materials to eliminate the time consuming approval process of MNE. However, the teachers were unaware of them because of a miscommunication between the schools and MNE. Explaining the development process of the alternative coursebooks, committee member commented that teachers did not need

to follow the coursebook designed for a different curriculum because of the examination as students following IntELC efficiently could be successful at the high-stakes exams.

### Principle of Appropriateness

Within this subscale, four items focus on the consistency of content and outcomes with the class hours and readiness level of the students. The results of the analysis are presented in Table 4.

**Table 4**

*The Results Regarding Principle of Appropriateness*

Item	Curriculum	M	SD	Mann Whitney U			
				U	z	p	r
30	IntELC	3.07	1.1	905	-.823	.411	-0.08
	ELC	3.23	1				
31	IntELC	3.25	1	1006.5	-.012	.990	-0.001
	ELC	3.23	1.1				
32	IntELC	3.53	1.1	719	-2.308	.021	-0.23
	ELC	2.97	1				
33	IntELC	3.32	.98	883	-1.009	.313	-0.10
	ELC	3.08	1				
Subtotal		13.1	3.3	939.5	-.529	.597	-0.05
		12.5	3.3				

Based on the mean scores presented in Table 4, it can be stated that both of the curricula got moderate ratings, except item 32. Considering weekly hours for the specification of the content (item 32) was the only item with statistically significant mean difference ( $p < .05$ ). However, the effect size for this item was found to be small ( $r = -0.23$ ). Although responses given to COSET revealed that teachers following IntELC were content with the weekly class hours, except one of them, all of the interviewees believed that class hours were not enough to cover the whole content. Complaining about the limited weekly class hours and seeing them as the source of their struggles, interviewees following ELC shared that they were having some difficulties to

explain some of the structures to the students. They stated that no matter how many times they explained some structures, students did not understand. Suggesting covering one grammar structure throughout each grade, some of the teachers believed that students could not understand the structures when they had no knowledge about them in their native language.

On the other hand, the advisors and committee member believed that teachers were teaching grammar in a wrong way. Furthermore, they stated that IntELC was prepared based on 20 class hours and what teachers should do was to redesign the content in line with their class hours, which could also be an opportunity for teacher autonomy. When it comes to the limited class hours of ELC, they agreed with the teachers and shared their struggle to explain the importance of more class hours to the authorities who took the class hours in the countries of Europe as a base. Comparing English exposure rate between European countries and Turkish context, they also admitted the difficulty of increasing class hours as it was directly related to the number of the teachers assigned.

### **Language Skills**

The items of this subscale evaluate the importance given to language skills within the curricula. Mann Whitney U test results are presented in Table 5.

**Table 5***The Results Regarding Language Skills*

Item	Curriculum	M	SD	<i>Mann Whitney U</i>			
				<i>U</i>	<i>z</i>	<i>p</i>	<i>r</i>
34	IntELC	3.35	1	907.5	-.810	.418	-0.08
	ELC	3.12	.97				
35	IntELC	3.64	.95	763	-2.010	.044	-0.20
	ELC	3.19	1				
36	IntELC	3.57	.95	958	-.419	.675	-0.04
	ELC	3.47	.99				
37	IntELC	3.28	1	708	-2.388	.017	-0.23
	ELC	2.72	1				
38	IntELC	3.78	.99	791	-1.877	.060	-0.18
	ELC	3.44	.91				
Subtotal		17.6	4.4	734.5	-2.115	.034	-0.21
		15.9	4				

Considering the mean scores, presented in Table 5, we can claim that according to teachers, both of the curricula gave less importance to speaking skills and pronunciation (Item 34 and 37). Although Mann Whitey U analysis showed statistically significant differences between the two curricula in terms of writing skills (item 35), pronunciation (item 37), and subtotals ( $p < .05$ ), the effect sizes were found to be small ( $r < \pm .3$ ). Making a general evaluation of the curricula in terms of language skills, interviewees following IntELC supported two different opinions. Two of the teachers found IntELC more efficient in terms of improving language skills, while three of them believed that overloaded grammar content constituted an impediment to focus on language skills. For the interviewees following ELC, the issue was not about focusing on language skills. There were some practical obstacles such as the number of the activities, overcrowded classes, and class hours for the teachers to reach those outcomes.

Insisting on remembering the fact that both ELC and IntELC were only focusing on language skills rather than grammar, the advisors and committee member stated that teachers

were dealing with grammar more than needed, and both of the curricula could be effective if applied correctly. Regarding the outcomes of language skills, one of the advisors responded:

There is no separate outcome for pronunciation because we integrated the language elements, grammar, vocabulary, and pronunciation into speaking. Actually, we are emphasizing it in some of the speaking outcomes. We are not stating it explicitly as ‘children know how to intone’ but it is included in the skills. (A1)

### Continuum of Learning

Three items of the continuum of learning subscale were about the consistency between activities and outcomes, student-centered activities, and preparing students for the next grades. Mean scores and Mann Whitney U results are presented in Table 6.

**Table 6**

*The Results Regarding Continuum of Learning*

Item	Curriculum	M	SD	<i>Mann Whitney U</i>			
				<i>U</i>	<i>z</i>	<i>p</i>	<i>r</i>
39	IntELC	3.92	.76	768.5	-2.085	.037	-0.20
	ELC	3.51	.93				
40	IntELC	4.07	.76	663	-2.949	.003	-0.29
	ELC	3.47	.93				
41	IntELC	4	.66	749	-2.145	.032	-0.21
	ELC	3.52	.96				
<b>Subtotal</b>		12	1.8	692	-2.491	.013	-0.24
		10.5	2.4				

As it can be understood from Table 6, all item-based and subtotal mean score differences were found to be statistically significant ( $p < .05$ ). However, all effect size values were small ( $r < \pm .3$ ). Mean scores can be interpreted as teachers were content with their curricula in terms of continuum of learning, but interviewees shared quite different opinions. They blamed the education system for making students expect everything to be handed on a silver platter. IntELC

was believed to provide a solid base for the students, but information overload could cause problems. While some of the teachers praised ELC for being consistent with the following grades, some of them criticized it as it could not provide a solid base for the students.

Criticizing the perspectives of the teachers, committee member made the following comment:

... when they see “talking about past time events”, the teachers start to make students memorize the second form of the verbs. But the curriculum doesn’t want this. Also the outcomes stated in the curriculum are all collected under four language skills. Students need to elicit grammar structures with language skills. This is stated at the introduction section of the curriculum. But as I said before, the reason behind this is the lack of curriculum literacy of the teachers. (CM)

### **Overall Evaluation**

Lastly, total score mean values were compared. Based on the results, it can be stated that total score mean value of IntELC (M= 143.57) was higher than ELC (M= 135.02). However, the difference was non-significant and effect size was small ( $U= 830.5$ ,  $z= -1.363$ ,  $p>.05$ ,  $r= -0.13$ ). At the end of the interview sessions, teachers following IntELC mentioned that they had to find their own way of applying the curriculum and were not provided any guidance, which in a way caused different applications among schools. Supporting the application, teachers also highlighted the importance of providing a coherent coursebook with IntELC and improving physical conditions in order to get better results. As they wanted more class hours, interviewees following ELC were also supporting IntELC. Moreover, they felt the need of guidance, shared their expectations from the curriculum as a document, and as the appliers of the curriculum, they

wanted to have an active role in the process of curriculum development. They believed that teachers were left out of the process.

According to the advisors and committee members, teachers had active roles in the development process as they were part of the committee as a member and MNE also got feedback from other teachers from the field. Responding to the criticism of the teachers about the curriculum document, they stated that within centralized education system, it was not possible to prepare a detailed curriculum and it should not provide solutions to the in-class problems because it was impossible to foresee every possible problem. Misconception of the coursebook as the curriculum was also pointed out. Reminding the relation between the curriculum and the coursebook, one of the advisors made the following comment:

It is really important how the curriculum is reflected in the material. At this point what stands out is what material designers understand from the curriculum and how they reflect it. Proficiency of those people should be questioned. (A2)

Criticizing the perspective of curriculum as thinking about what to include in a unit of coursebook, one of the advisors suggested designing the whole education system in line with the curriculum and providing local constant in-service training in order to improve the application of the curricula. He also highlighted the importance of making skills-based exams by associating the current examination system with conducting a swimming exam through pen and paper.

### **Discussion and Conclusions**

The results of the present study showed that teachers had some difficulties in reaching the learning outcomes for different reasons. Teachers following IntELC showed the overloaded grammar content and intensity of the learning outcomes as the reasons behind their struggle, which was also mentioned in the studies evaluating IntELC (Berkant et al., 2019; Dilekli, 2018;

Dinçer & Koç, 2020; Yedigöz-Kara, 2019). Being the source of many troubles in practice, limited class hours were the factors hindering teachers to reach the outcomes in ELC, which was consistent with the study of Cihan and Gürten (2013). Regardless of their curriculum, teachers were focusing on grammar more than needed. Not only did the advisors and committee member persistently pointed that two curricula did not focus on grammar but also it was clearly stated in the documents of the curricula that they were based on communicative perspectives. As explained in IntELC:

Parallel to the English Language Teaching Program (for the 2nd – 8th grades) published by Board of Education in 2017, the present program for the lower secondary education (5<sup>th</sup>, 6<sup>th</sup>, 7<sup>th</sup> and 8<sup>th</sup>) is based on communicative principles to English language teaching. It organizes the curricular objectives with regard to language skills and functions (uses), not in terms of any dichotomy or classification on grammatical forms (usages). Thus, the present program is based on a series of thematic and communicative syllabi, success of which is inevitably dependent upon the analogous approach in teachers' instructional choices, in measurement repertoire and in learners' practice (MNE, 2018c, pp. 3-4).

There can be two explanations for the mismatch between what IntELC and ELC intended to do and the in-class practice of the teachers. The first possible reason stated was curriculum literacy. According to Sural and Dedeşali (2018), curriculum literacy has a crucial role in reaching the expected learning outcomes. It is deemed to be important for saving teachers from the captivity of the coursebook and preserving teacher autonomy (Ben-Peretz, 1990). In the Turkish context, teachers saw the coursebook as the curriculum and did not follow the updates in the curriculum (TED, 2009). Moreover, Saral (2019) found out that English teachers at the state schools in Turkey had a moderate level of curriculum literacy. Teachers need curriculum literacy

to decode the official document and understand the objectives and philosophy behind to reflect the curriculum into practice (Kahramanoğlu, 2019). Although 71% of the teachers in the present study claimed reading the updated curriculum document, their responses might indicate contrary.

The effect of examination system can be another explanation. Students are expected to have detailed knowledge on grammar and vocabulary instead of language skills to be able to pass high-stakes exams. Exam-related concerns of the teachers were consistent with the findings of the studies evaluating IntELC from teachers' perspectives (Berkant et al., 2019; Yedigöz-Kara, 2019). Studies from various contexts have also indicated the effect of the national assessment system on the application of the curriculum (e.g. Al-Darwish, 2006; Alwan, 2006; Glasgow, 2014; Li, 2010; Nothaisong, 2015; Tsai, 2007; Wu, 2013). Additionally, teachers following ELC mentioned their struggles between the theory and practice and the burdensome examination system. High-stakes exams changed the role of tests within the teaching process and this created high expectations for school-based stakeholders (Cheng & Curtis, 2008). This situation is called as 'washback effect', a term which connotes uncalled influence on learning-teaching situations (Alderson & Wall, 1993). Within Turkish education system, learners from all levels have to deal with high-stakes exams for better education (Özmen, 2012). For instance, at the end of middle school, students take the secondary education placement exam and in the English section they are responsible for grammar, vocabulary, and reading skills, which is quite contrary to the objectives of the curriculum (Kılıçkaya, 2016). The mismatch between the national examination system and curriculum is deemed to cause the following: teaching to test, wasting sources, disregarding the objectives of the curricula, and increase of inequality of opportunity in education (Wedell, 2014). Not including other language skills except reading can cause disregarding those skills during the in-class practice. Additionally, teachers' concern for the

secondary education placement examination can be quite early considering the grade. Besides, shaping the lessons as instructed in the curricula can help students to be successful at English sections of any exams as claimed by the committee member.

One of the points about which advisors and teachers shared different opinions was the amount of new vocabulary in ELC. Although it was not an issue considering the responses given to COSET, the interviewees criticized ELC harshly because of the amount of new vocabulary. On the other hand, advisors explained their intention through the term ‘language shower’ and pointed out the importance of how vocabulary was taught. Wilkins (as cited in Thornbury, 2002, p. 13) claims that “without grammar very little can be conveyed, without vocabulary nothing can be conveyed”. As well as its role in language learning, vocabulary is also necessary to learn and use grammar (Cameron, 2001). The studies in the related literature have mentioned 2000 words as a threshold (Schmitt, 2000; Thornbury, 2002). Moreover, explaining the learner’s vocabulary size, Nation (2013) notes that “the larger the vocabulary size, the greater the quantity of language that needs to be processed in order to meet the words to be learned again” (p.108). According to Thornbury (2002), what learners need is to be exposed to the words as much as possible as well as repeated memory retrieval. Instead of the amount of the vocabulary, the attention should be paid on how to teach learners new vocabulary in a meaningful context and how effective the coursebooks are in this regard.

Another important point was the teachers’ obligation or preference to conduct pen and paper exams mainly because of their prejudgment about their students’ level and technical obstacles. The advisors listed conformity, class hours, and teachers’ lack of knowledge of how to conduct skills-based exams as the reasons. Conformity can be defined as adopting the behavior performed by the majority of group members (van Leeuwen & Haun, 2014). Discussing the

impact of conformity on teacher autonomy, Saldana (2013) points out that prospective teachers have the tendency to choose the preferred method within the system. Moreover, excluding language skills from high-stakes exams can mislead the learners about the importance of these skills (Yaman, 2018). As explained by McEwen (1995), “what is assessed becomes what is valued, which becomes what is taught” (p. 42). The way English is evaluated in high-stakes tests should be changed. However, it is not easy. Wedell (2014) believes that exam results are important to all stakeholders for different reasons, and if the national assessment became consistent with the communicative principles, most of the students could not pass those exams as they were not getting the necessary support. Therefore, it is easier to pretend that the curriculum is taught and assessed as it is intended.

Regardless of their curriculum, the teachers criticized MNE-approved coursebooks from different aspects. Coursebook as an instructional material is a crucial element for most of the language classes, even sometimes the only source of input and practice (Richards, 2001b). According to McGrath (2013), these materials are called coursebook because they are “the foundation for a course” (p. 5). However, the result of heavy reliance on the coursebooks can be their control on the instruction (Kitao & Kitao, 1997). Although none of these materials can be a perfect fit for the language classes, an efficient coursebook can reflect the curriculum by providing a variety of context and activities that teachers cannot design on their own (Richards, 2001b). One of the reasons behind the reliance on coursebooks was the examination system. Teachers felt the obligation to cover the materials in detail in order to make their students get higher scores from high-stakes exams. Many teachers work under the pressure of the theory encouraging them to adjust the curriculum into their practices and the demand of school-based stakeholders to cover the coursebook in line with the assessment system (Wedell, 2014).

Limited class hours were one of the long-lasting problems. The insufficient class hours have also been documented by the studies evaluating the 5<sup>th</sup> grade English curriculum (Cihan & Gürten, 2013; Işık, 2019). Moreover, it has been cited as one of the factors affecting the application of the curriculum in other contexts (e.g. Al-Darwish, 2006; Altaieb, 2013; Sun, 2007; Tsai, 2007). Class hours seem to be one of the points on which teachers and advisors agreed. While teachers were sharing the difficulties they had because of limited class hours, the advisors shared their struggle to explain the importance of having more class hours to the authorities. Considering the opportunity of experiencing real use of language in a European context, the situation is quite different for the learners in a Turkish context. As stated by Lightbown and Spada (2013), learners in a classroom are more likely to be exposed to the new language and discourse types less. Considering the points discussed so far, the core problem seems to be the inconsistency between what is expected and what is practiced in local contexts. Pointing out the discrepancy between macro-level objectives and micro-level in practice, Kırkgöz (2009) suggests that Turkey needs to realize the discrepancy between macro-level objectives and micro-level in-practice and needs to have a coherent language policy for ELT to become more effective.

Curriculum development is a dynamic process and improving the factors hindering the effective application of the curriculum is also an important part of this process. The success of a curriculum is related to sharing a common perspective and the strong interaction among MNE, coursebook writers, school management, and teachers. Based on the findings, some implications and suggestions can be provided. The need for in-service training has been indicated by various studies not only in the Turkish context (e.g. Çankaya, 2015; Dinçer & Koç, 2020; Dönmez, 2010; İnam, 2009; Örmeci, 2009; Yedigöz-Kara, 2019; Zehir-Topkaya & Küçük, 2010) but also

in other contexts (e.g. Almalki, 2014; Altaieb, 2013; Burgos, 2012; Harris, 2010; Hillberry, 2008; Nothaisong, 2015; Powell, 2008; Tsai, 2007). Thus, providing in-service training about the content, methodological perspective, and outcomes of the curricula in order to overcome the problem of curriculum literacy among the teachers can be effective for the application of the curricula. There should also be a consistency between the evaluation system adopted in high-stakes exams and the curricula. As claimed by McGrath (2013), “syllabus development, textbook production and examinations need to be a part of an integrated operation. It helps if they are housed in the same building, but regular coordination meetings should be a sine qua non” (p. 193). Lastly, as being one of the main problems, class hours should be increased mainly because few hours limit the language exposure rate, which may result in not reaching the expected outcomes (Moon, 2005) and affecting the way teachers design their in-class practices.

## References

- Aksoy, E., Bozdoğan, D., Akbaş, U., & Seferoğlu, G. (2018). Old wine in a new bottle: Implementation of intensive language program in the 5<sup>th</sup> grade in Turkey. *Eurasian Journal of Applied Linguistics*, 4(2), 301-324. <http://dx.doi.org/10.32601/ejal464187>
- Al-Darwish, S. (2006). *An investigation of teachers' perceptions of the English language curriculum in Kuwaiti elementary schools* (Publication No.3218972) [Doctoral dissertation, University of Denver]. ProQuest Dissertations and Theses Global.
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115-129. <https://doi.org/10.1093/applin/14.2.115>
- Almalki, M. M. (2014). *Teacher perception of a new English as a foreign language (EFL) curriculum in Saudi Arabia* [Master's thesis, State University of New York]. SUNY Digital Repository. <http://hdl.handle.net/1951/65435>
- Altaieb, S. R. (2013). *Teachers' perception of the English language curriculum in Libyan public schools: An investigation and assessment of implementation process of English curriculum in Libyan public high schools* (Publication No. 3607636) [Doctoral dissertation, University of Denver]. ProQuest Dissertations and Theses Global.
- Alwan, F. H. (2006). *An analysis of English language teachers' perceptions of curriculum change in the United Arab Emirates* [Doctoral dissertation, University of Exeter]. ERIC. <https://files.eric.ed.gov/fulltext/ED503767.pdf>
- Ben-Peretz, M. (1990). *The teacher curriculum encounter: Freeing teachers from the tyranny of texts*. State University of New York Press.

- Berkant, H. G., Özaslan, D., & Doğan, E. (2019). Yabancı dil ağırlıklı eğitim uygulamasına ilişkin İngilizce öğretmenlerinin görüşleri [The views of English language teachers about intensive foreign language education program]. *Milli Eğitim Dergisi*, 48(1), 553-570.  
<https://dergipark.org.tr/tr/pub/milliegitim/issue/51765/674255>
- Bobbitt, F. (2004) Scientific method in curriculum-making. In D. J. Flinders and S. J. Thornton (Eds.), *The Curriculum studies reader* (pp. 9-16). RoutledgeFalmer.
- Burgos, S. (2012). *Behind classroom doors: Consistency between policy and practice in the English as a second language classroom* (Publication No. 3511515) [Doctoral dissertation, Universidad Del Turabo]. ProQuest Dissertations and Theses Global.
- Bümen, N. T., Çakar, E., & Yıldız, D. G. (2014). Türkiye’de öğretim programına bağlılık ve bağlılığı etkileyen etkenler [Curriculum fidelity and factors affecting fidelity in the Turkish context]. *Educational Sciences: Theory and Practice*, 14(1), 203-228.  
<https://doi.org/10.12738/estp.2014.1.2020>
- Cameron, L. (2001). *Teaching languages to young learners*. Cambridge University Press.
- Cihan, T., & Gürten, E. (2013). İlköğretim 5. sınıf İngilizce dersi öğretim programına ilişkin öğretmen görüşleri [Teachers’ opinions on the English language curriculum of the 5<sup>th</sup> grade of primary education]. *Anadolu University Journal of Social Sciences*, 13(1), 131-146. <https://hdl.handle.net/11421/127>
- Cheng, L. & Curtis, A. (2008). Washback or backwash: A review of the impact of testing on teaching and learning. In L. Cheng, Y. Watanabe & A. Curtis (Eds.), *Washback in language testing* (pp. 3-18). Lawrence Erlbaum.
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4<sup>th</sup> ed.). Pearson.

- Çankaya, P. (2015). *An evaluation of the primary 3rd grade English language teaching program: Tekirdağ case* [Master's thesis, Çanakkale 18 Mart University]. Council of Higher Education Thesis Center.  
[https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=Wbc656i315e2eV6-EZV1og\\_EtJCOiUo6XgfCSwy7mufrXDv4ENh3PFHSTm9JvASf](https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=Wbc656i315e2eV6-EZV1og_EtJCOiUo6XgfCSwy7mufrXDv4ENh3PFHSTm9JvASf)
- DeVellis, R. F. (2017). *Scale development: Theory and applications* (4<sup>th</sup> ed.). SAGE
- Dincer, A. & Koç, H. K. (2020). The implementation of an intensive English language program in the fifth grade in Turkey: A qualitative evaluation. *Journal of Theoretical Educational Science*, 13(1), 25-43. <https://doi.org/10.30831/akukeg.532955>
- Dilekli, Y. (2018). Ortaokul İngilizce hazırlık sınıfı programı pilot uygulamasının öğretmen görüşlerine göre değerlendirilmesi [Evaluation of pilot scheme of secondary schools' English preparatory class curriculum according to teachers' views]. *OPUS-International Journal of Society Researchers*, 8(15), 1399-1425. <https://doi.org/10.26466/opus.445713>
- Dönmez, Ö. (2010). *Implementation of the new eighth grade English language curriculum from the perspectives of teachers and students* [Master's thesis, Middle East Technical University]. METU Library. <https://etd.lib.metu.edu.tr/upload/12612264/index.pdf>
- Erdoğan, V. (2005). *An evaluation of the English curriculum implemented at the 4<sup>th</sup> and 5<sup>th</sup> grade primary state schools: The views of the teachers and the students* [Master's thesis, Mersin University]. Council of Higher Education Thesis Center.  
[https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=-L8ilcwn9ZRRc\\_YMKxXW1lzVK4vfkNJo6x-BTGHAUqJ8C6Nf72gdZ7p4M1nl-knA](https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=-L8ilcwn9ZRRc_YMKxXW1lzVK4vfkNJo6x-BTGHAUqJ8C6Nf72gdZ7p4M1nl-knA)

- Erkuş, A. (2016). *Psikolojide ölçme ve ölçek geliştirme-I: Temel kavramlar ve işlemler* [Measurement and scale development in Psychology-I: Basic concepts and operations]. Pegem Publishing.
- Field, A. (2009). *Discovering statistics using SPSS (and sex, drugs and rock'n'roll)* (3<sup>rd</sup> ed.). SAGE.
- Fullan, M. (2007). *The new meaning of educational change* (4<sup>th</sup> ed.). Teacher College Press.
- Gimenez, T. & Tonelli, J. R. A. (2013). Building and EFL curriculum for young learners: A Brazilian experience. *International Journal of English Language & Translation Studies*, 1(3), 88-97. <http://www.eltsjournal.org/archive.html>
- Glasgow, G. P. (2014). Teaching English in English, 'in principle': The national foreign language curriculum for Japanese senior high schools. *International Journal of Pedagogies and Learning*, 9(2), 152-161. <https://doi.org/10.1080/18334105.2014.11082028>
- Gürsoy, E., & Eken, E. (2018). English teachers' understanding of the new English language teaching program and their classroom implementations. *Journal of Language and Linguistic Studies*, 14(2), 18-33. <https://dergipark.org.tr/en/pub/jlls/issue/43364/527917>
- Harris, L. S. (2010). *A case study of the English second language programs of a North Carolina school district* (Publication No. 3439262) [Doctoral dissertation, The University of North Carolina]. ProQuest Dissertations and Theses Global.
- Hillberry, M. M. (2008). *An examination of the effectiveness of English language learner programs at the elementary level* (Publication No. 3340059) [Doctoral dissertation, Oakland University]. ProQuest Dissertations and Theses Global.

Işık, E. (2019). *Güncellenen 5. sınıf İngilizce öğretim programının (2017) eğitsel eleştiri modeliyle değerlendirilmesi: Bir durum çalışması [The evaluation of the updated 5<sup>th</sup> grade English language curriculum (2017) with educational criticism model: A case study]* [Doctoral dissertation, Fırat University]. Council of Higher Education Thesis Center.

[https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=npGs9H39x7G6401x51yqpFBqxd5mvs47Cb49S5Ef\\_duwsP-jCpzflbtd\\_R\\_H0kVE](https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=npGs9H39x7G6401x51yqpFBqxd5mvs47Cb49S5Ef_duwsP-jCpzflbtd_R_H0kVE)

İnam, G. (2009). *İlköğretim okulları 4. sınıf İngilizce dersi öğretim programının değerlendirilmesine ilişkin öğretmen görüşleri [Teacher opinions about 4th classes English program in primary schools]* [Master's thesis, Adnan Menderes University]. Council of Higher Education Thesis Center.

[https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=CwVIqqBuz1VkysVpueogAQHET2\\_GOtC11tpq04ZrFzD2HHGzJyYgFHvM13nyzjQc](https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=CwVIqqBuz1VkysVpueogAQHET2_GOtC11tpq04ZrFzD2HHGzJyYgFHvM13nyzjQc)

Kahramanoğlu, R. (2019). Öğretmenlerin öğretim programı okuryazarlığına yönelik yeterlik düzeyleri üzerine bir inceleme [A study on teachers' levels of curriculum literacy]. *Journal of International Social Research*, 12(65), 827-840.

<http://dx.doi.org/10.17719/jisr.2019.3495>

Kambur, S. (2018). *An Evaluation of 5<sup>th</sup> grade intensive English language curriculum in terms of teacher opinions* [Master's thesis, Yeditepe University]. Council of Higher Education Thesis Center.

<https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=hcgrYffRbz0Z44UJEuLtwYPBeetWINbtTHRX4Mwk2vVZNQNeAbOq43Vu1xY54Shx>

- Karcı, C. (2012). *Orta öğretim dokuzuncu sınıf İngilizce öğretim programının öğretmen görüşlerine göre değerlendirilmesi [Evaluation of English curriculum in 9<sup>th</sup> grade of secondary schools through teachers' views]* [Master's thesis, Adnan Menderes University]. Council of Higher Education Thesis Center.  
[https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=RYan9\\_S-Z7Eir3xdWGXBIMYTZqacD6SSqRhxR1oPhhxcI5Xu4fiLIYpk1ACw7If6](https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=RYan9_S-Z7Eir3xdWGXBIMYTZqacD6SSqRhxR1oPhhxcI5Xu4fiLIYpk1ACw7If6)
- Kaya, S. (2018). *Evaluation of the middle school English language curriculum developed in 2012 utilizing Stake's countenance evaluation model* [Doctoral dissertation, Middle East Technical University]. Council of Higher Education Thesis Center.  
<https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=MzP7PYssFqdb3WIjroAktU9R5pKvBpuL2LDJdZ5P5pFUO6vwlUP2mDBGPsOHTBN>
- Kılıçkaya, F. (2016). Washback effects of a high-stakes exam on lower secondary school English teachers' practices in the classroom. *Lublin Studies in Modern Languages and Literature*, 40(1), 116-134. <https://files.eric.ed.gov/fulltext/ED569921.pdf>
- Kırkgöz, Y. (2009). Globalization and English language policy in Turkey. *Educational Policy*, 23(5), 663-684. <https://doi.org/10.1177/0895904808316319>
- Kitao, K., & Kitao, S. K. (1997). Selecting and developing teaching/learning materials. *The Internet TESL Journal*, 4(4). <http://iteslj.org/Articles/Kitao-Materials.html>
- Li, M. (2010). EFL teachers and English language education in the PRC: Are they the policy makers? *The Asia-Pacific Education Researcher*, 19(3), 439-451.  
<https://doi.org/10.3860/taper.v19i3.1852>
- Lightbown, P. M. & Spada, N. (2013). *How languages are learned* (4<sup>th</sup> ed.). Oxford University Press.

- McEwen, N. (1995). Educational accountability in Alberta. *Canadian Journal of Education*, 20(1), 27-44. <https://doi.org/10.2307/1495050>
- McGrath, I. (2013). *Teaching materials and the role of EFL/ESL teachers: Practice and theory*. Bloomsbury.
- Ministry of National Education (MNE). (2018a). *İngilizce dersi öğretim programı (ilkokul ve ortaokul) [English language teaching curriculum (primary and secondary school)]*. <http://mufredat.meb.gov.tr>
- Ministry of National Education (MNE). (2018b). *43769797-10.07.01-E.14674863 sayılı resmi yazı [Official letter]*
- Ministry of National Education (MNE). (2018c). *Yabancı dil ağırlıklı 5. sınıf İngilizce dersi öğretim programı [Intensive English language teaching curriculum for 5<sup>th</sup> grade]*. <http://mufredat.meb.gov.tr>
- Moon, J. (2005). *Teaching English to young learners: The challenges and the benefits*. British Council. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.483.7932&rep=rep1&type=pdf>
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2<sup>nd</sup> ed.). Cambridge University Press.
- Nonthaisong, K. (2015). *The impact of the English language policy in the basic education core curriculum BE 2551 (AD 2008) on English instruction of Thai secondary EFL teachers: A multiple case study in Northeast Thailand* [Doctoral dissertation, The University of Alabama]. The University of Alabama DSpace. <https://ir.ua.edu/handle/123456789/2417>
- Null, W. (2011). *Curriculum: From theory to practice*. Rowman & Littlefield Publishers.

Ornstein, A., & Hunkins, F. (2018). *Curriculum: Foundations, principles and issues* (7<sup>th</sup> ed.). Pearson.

Örmeci, D. N. (2009). *An evaluation of English language curricula implemented at the 4<sup>th</sup>, 5<sup>th</sup>, and 6<sup>th</sup> grades in respect of teachers' opinions* [Master's thesis, Trakya University]. Council of Higher Education Thesis Center.

[https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=CwVIqqBuz1VkysVpueogATE\\_Y8HEeIHyoNERHa4Z0fn1wlpqyFRAdwx\\_pGoMW1h](https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=CwVIqqBuz1VkysVpueogATE_Y8HEeIHyoNERHa4Z0fn1wlpqyFRAdwx_pGoMW1h)

Özmen, K. S. (2011). Analyzing washback effect of SEPPPO on prospective English teachers. *Journal of Language and Linguistic Studies*, 7(2), 24-52.

<https://dergipark.org.tr/en/pub/jlls/issue/9934/122899>

Özüdoğru, F., & Adıgüzel, O. C. (2015). İlkokul 2. sınıf İngilizce öğretim programının değerlendirilmesi [Evaluation of primary school 2<sup>nd</sup> grade English language teaching curriculum]. *Turkish Studies*, 10(11), 1251-1276.

<http://dx.doi.org/10.7827/TurkishStudies.8692>

Powell, J. M. (2008). *English language learner programs and services: A case study of Nebraska middle schools* (Publication No. 3315207) [Doctoral dissertation, University of Nebraska]. ProQuest Dissertations and Theses Global.

Richards, J. C. (1990). *The language teaching matrix*. Cambridge University Press

Richards, J. C. (2001a). *Curriculum development in language teaching*. Cambridge University Press.

Richards, J. C. (2001b). The role of textbooks in a language program. *RELC Guidelines*, 23(2), 12-16. <https://www.professorjackrichards.com/wpcontent/uploads/roleof-textbooks.pdf>

- Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching* (3<sup>rd</sup> ed.). Cambridge University Press.
- Romero, J.L.R., Sayer, P. & Irigoyen, E.N.P. (2014). English language teaching in public primary schools in Mexico: The practices and challenges of implementing a national language education program. *International Journal of Qualitative Studies in Education*, 27(8), 1020-1043. <https://doi.org/10.1080/09518398.2014.924638>
- Salahuddin, A.N.M., Khan, M.M.R., & Rahman, M.A. (2013). Challenges of implementing English curriculum at rural primary schools of Bangladesh. *The International Journal of Social Sciences*, 7(1), 34-51.
- Saldana, J. (2013). Power and conformity in today's schools. *International Journal of Humanities and Social Science*, 3(1), 228-232.  
[http://ijhssnet.com/journals/Vol\\_3\\_No\\_1\\_January\\_2013/27.pdf](http://ijhssnet.com/journals/Vol_3_No_1_January_2013/27.pdf)
- Saral, N. Ç. (2019). *Exploring curriculum literacy level of English language teachers in Turkey* [Master's thesis, Gazi University]. Council of Higher Education Thesis Center.  
<https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=vjszP7PzV0HebcjFEvDfwMbP7MdAy4S6Fj-jhEsD89tXnkri4klhQyDyISw218De>
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge University Press
- Stec, M. (2011). Early language teaching and syllabuses. *Procedia-Social and Behavioral Sciences*, 29, 1123-1132. <https://doi.org/10.1016/j.sbspro.2011.11.346>
- Sun, L. (2007). *Multiple case studies of elementary English teachers' implementation of the grade 1–9 English curriculum in Taiwan* (Publication No. 3280272) [Doctoral dissertation, University of Northern Colorado]. ProQuest Dissertations and Theses Global.

- Sural, S., & Dedeali, N. C. (2018). A study of curriculum literacy and information literacy levels of teacher candidates in department of social sciences education. *European Journal of Educational Research*, 7(2), 301-315. <https://doi.org/10.12973/eu-jer.7.2.303>
- TED (2009). *Öğretmen Yeterlikleri [Teacher qualifications]*. Türk Eğitim Derneği.
- Thornbury, S. (2002). *How to teach vocabulary*. Pearson.
- Tsai, T. H. (2007). *Taiwanese educators' perspectives on the implementation of the new English education policy* (Publication No. 3273298) [Doctoral dissertation, Alliant International University]. ProQuest Dissertations and Theses Global.
- Van Leeuwen, E. J., & Haun, D. B. (2014). Conformity without majority? The case for demarcating social from majority influences. *Animal Behaviour*, 96, 187-194. <http://dx.doi.org/10.1016/j.anbehav.2014.08.004>
- Wedell, M. (2014). The national curriculum and the national English assessment system: Close relations or total strangers? Presented at *10th International Bandung Conference, Bandung, Indonesia: Bandung Institute of Technology*. [https://www.researchgate.net/publication/272822663\\_the\\_national\\_curriculum\\_and\\_the\\_national\\_English\\_assessment\\_system\\_Close\\_relations\\_or\\_total\\_strangers](https://www.researchgate.net/publication/272822663_the_national_curriculum_and_the_national_English_assessment_system_Close_relations_or_total_strangers)
- Wedell, M., & Grassick, L. (Eds.). (2018). *International perspectives on teachers living with curriculum change*. Palgrave Macmillan.
- Wiggins, G., & McTighe, J. (2005). *Understanding by design*. Association for Supervision and Curriculum Development.
- Wu, W. (2013). *Towards a new intercultural methodology in response to teachers' challenges from the new curriculum innovation in China* (Publication No.NR92602) [Doctoral dissertation, University of Alberta]. ProQuest Dissertations and Theses Global.

Yaman, İ. (2018). Türkiye’de İngilizce öğrenmek: Zorluklar ve fırsatlar [Learning English in Turkey: Challenges and opportunities]. *RumeliDE Journal of Language and Literature Studies*, 11, 161-175. <https://doi.org/10.29000/rumelide.417491>

Yedigöz-Kara, Z. (2019). *Öğretmenlerin yabancı dil ağırlıklı beşinci sınıf İngilizce dersi öğretim programıyla ilgili endişelerinin incelenmesi [Teachers’ concerns about the intensive English language curriculum for the 5<sup>th</sup> grade]* [Master’s thesis, Ege University]. Council of Higher Education Thesis Center.

<https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=npGs9H39x7G6401x51yqpJZR8YOfEHZxa8DSYnYXEQbmtooPSUmBtrVK2yO9XRq9>

Zehir-Topkaya, E., & Küçük, Ö. (2010). An evaluation of 4<sup>th</sup> and 5<sup>th</sup> grade English language teaching program. *Elementary Education Online*, 9(1), 52-65.

<https://dergipark.org.tr/tr/pub/ilkonline/issue/8596/106953>

# **The Assessment of Preservice Teachers' Dispositions**

**David K. Griffin**

## **Abstract**

This paper reviews some of the approaches to defining and assessing dispositions in preservice teachers. A review of the literature is presented, followed by a discussion of issues relating to defining dispositions, and the importance of assessing dispositions in teacher candidates.

Finally, the results of a study that included a sample of 15 preservice teachers who were assessed on the Educator Disposition Assessment by two independent reviewers are presented. Study limitations and possible future research methodologies are also discussed.

**Keywords:** dispositions, preservice teacher

*David K. Griffin, Ed.D., BCBA-D is an Associate Professor at Nova Southeastern University, a Licensed School Psychologist and a Board Certified Behavior Analyst. He can be reached at [griffida@nova.edu](mailto:griffida@nova.edu)*

With the seemingly ever-changing make-up of today's classrooms, there is an increasing focus on university programs' curricula that prepare future teachers to meet the unique needs of diverse students. Traditionally, these programs have been required to document that those who graduate demonstrate competence in basic standards, with the focus on the graduating candidates' abilities to understand and provide instruction relating to key academic areas.

In more recent times, however, the focus has broadened to include not only the assessment of pedagogy and academic preparation but that of teacher dispositions as well. Many accreditation organizations are now including the assessment of dispositions as part of the accreditation process. The Council for the Accreditation of Educator Professionals (CAEP), for example, requires that teacher education programs assess candidates' dispositions regularly (CAEP 2018), and demonstrate that candidates demonstrate various dispositions prior to graduation. Likewise, the Interstate New Teacher Assessment and Support Consortium (InTASC) Core Standards (2018) include several key elements that specifically address dispositions. For example, the focus of Standard 2 is on Learning Differences; "The teacher uses understanding of individual differences and diverse cultures and communities to ensure inclusive learning environments that enable each learner to meet high standards" (InTASC Model Core Teaching Standards: A Resource for State Dialogue, 2011). Thus, with the ever-increasing focus on dispositions, teacher preparation programs recognize the importance of not only modeling appropriate dispositions but including dispositions at various stages in the overall curriculum (Da Ros-Voseles & Moss, 2007). To add to this, many teacher and preservice teacher evaluation instruments contain elements of dispositions (Marzano & Brown, 2009; Danielson, Axtell &

McKay, 2009). Finally, Nixon Dam, and Packard (2010) reported that over 50 percent of non-contract teacher renewal was specifically related to the teacher not demonstrating appropriate, positive dispositional behaviors.

The assessment of dispositions is not as clear cut as assessing a candidate's ability to teach academic skills, since there are various definitions of dispositions, and lack of agreement regarding which specific dispositions are important and need to be assessed. As such, the focus of dispositions to be assessed is usually the responsibility of the institution overseeing the teacher preparation program, or the accreditation agency. Furthermore, as Notar, Riley, and Taylor, (2009) point out, since dispositions are not part of content knowledge, they are not always detectable.

The actual assessment of dispositions has been widely debated in the teacher education profession over the past few years (Welch, Pitts, Tenini, Kuenlen, & Wood, 2010). Thornton (2006) points out that despite the fact that dispositions are essential elements of teacher preparation, they are frequently neglected and not incorporated into teacher education programs. Yet, as Cummins and Asempapa (2013) point out, if teacher education programs first clearly define crucial teacher and learning dispositions, courses could be tailored to include the fostering and assessment of such dispositions. At the same time, Osguthorpe (2013) stated "It is perplexing that many teacher education programs assess teacher candidates on something (dispositions) that such programs might not seek to actually teach or develop" (p 21). Furthermore, Stewart and Davis (2005) indicate even though colleges have identified important dispositions, there is little evidence to show that programs actually teach candidates how to further develop or enhance these dispositions. Since the goal of teacher preparation

programs is to produce highly qualified teachers, Dottin, (2009) reports that dispositions, in addition to knowledge and skills, should be considered an important component of the preparation program. However, as Osguthorpe (2013) states, despite the fact that teacher preparation programs are required to follow various standards that include dispositions, these programs have yet to develop a consistent approach to disposition development and assessment.

Another challenge faced by many teacher preparation programs is that many, if not all courses are taken in an online format. A teacher candidate may successfully complete the academic requirements of the program online, yet the opportunity to observe the candidate teach in the classroom does not occur until the internship phase. This is often the final requirement for graduation, and it is possible for a candidate to be placed in a classroom who does not demonstrate appropriate teaching dispositions. As such it is the responsibility of the institution to develop creative ways to assess (and hopefully make positive contributions to) the teacher candidate's teaching performance which includes a wide variety of desired classroom dispositions.

### **Defining Dispositions**

One of the first tasks teacher preparation programs must address prior to assessing dispositions is to identify and define the specific dispositions of interest and importance. A review of the literature provides one with an abundance of terms, definitions, and descriptions of what are considered to be desirable teacher dispositions. However, as Welch et al., (2010) indicate, there is a lack of operational definitions of behaviors thought to be related to teacher dispositions which in turn limits teacher preparation programs from implementing effective assessment and evaluation.

In 2008, The National Council for Accreditation of Teacher Education (NCATE) published a set of professional standards as a guide for teacher education programs which included references to professional dispositions. The manual stipulates that graduating teacher candidates need to possess dispositions to work successfully with all children and demonstrate professional dispositions that facilitates student learning.

Even though specific measurable, operational definitions were not included, NCATE provided a general description of important dispositions that teacher candidates should demonstrate caring, honesty, fairness, empathy respectfulness, responsibility, and thoughtfulness. Included also is the belief that all children can learn, and the creation of caring and supportive learning environments (The National Council for Accreditation of Teacher Education, 2008).

Jensen, Whiting, and Chapman (2018) summarized their research regarding dispositions teacher educators should display when working with children from diverse backgrounds. These include empathy, meekness, social awareness, inclusion, and advocacy. They summarized, in their study, research that defines these 5 disposition areas and how they can be identified in teachers. Based on their review of the research, these authors developed the “Multicultural Teacher Dispositions Scale (MTDS)”. This tool is discussed later in this paper.

Several researchers attempt to provide definitions of teacher-related dispositions. Katz and Raths (1986), for example, define dispositions as "attributions which summarize a trend of a teacher's actions across similar contexts" (p. 3). Taylor and Wasicsko, (2000) define dispositions as the personal qualities or characteristics that are possessed by individuals, including attitudes, beliefs, interests, appreciations, values, and modes of adjustments. Notar, Riley, and Taylor (2009) include enthusiasm, sensitivity, responsibility, commitment, professionalism, skillful

preparation, a sense of respect for others, communication, and appropriate dress, deportment and demeanor. They add patience, determination, and courage to the list of important dispositions teacher candidates should demonstrate. Bauer & Thornton (2013) included several dispositions that were present in successful teachers, including creativity, and the ability to be critical. Adding to this, Choi, Benson, and Shudak (2016) state "...dispositions are those internal conditions (attitudes, values, beliefs, thoughts, etc.) that influence our external behaviors (actions and interactions with students and others) (p 72). Finally, Usher (2002) included the following when discussing dispositions; having empathy, or the ability to understand and accept another person's point of view, and a positive view of others, or believing in the worth, ability, and potential of others.

One issue is, again, not all agree on how to define the dispositions thought to be important in teacher education programs or provide a summary of observable behaviors that can be linked to specific dispositions. The assessment of "professionalism" for example, might include certain behaviors for one observer, but not for another. Thus, in order to obtain valid and reliable measures of teacher dispositions, specific behaviors need to be identified and operationally defined, allowing the evaluator to document if the specific behavior is demonstrated or not.

Even though not all agree on specific definitions of important teacher dispositions, as Cummins and Asempapa (2013) remind us there is agreement that certain dispositions are essential for teachers to be successful in the teaching profession. According to Stewart and Davis (2005) a common thread seems to be that of respect, enthusiasm, efficacy, communication skills, and grammar correctness (p 37). Finally, since definitions and descriptions of dispositions seem

to be left up to the teaching institution, and a document listing specific, appropriate teacher dispositions is not provided by accrediting agencies, many teacher preparation programs have chosen to identify a set of dispositions thought to be crucial in the classroom and have attempted to develop an approach to assessing them. However, once again, as Murray (2007) points out, the assessment of dispositions will have little value without valid and reliable measures. Issues relating to the reliability/validity of the instrument used in the assessment of dispositions will often lead to questionable results.

### **The Importance of Assessing Dispositions**

Apart from complying with various accreditation standards, one key reason for addressing and assessing dispositions for teacher education candidates is that there is a relationship between teachers' dispositions and the quality of student learning (Notar, Riley, and Taylor, 2009). These authors feel it is important for teacher preparation programs to objectively assess candidates' dispositions both in internships and coursework. Likewise, Chickering and Gameson (1987) indicate that teacher dispositions have a direct relationship to student learning. One specific disposition mentioned was having respect for diverse ways of learning. Related to this, Dee and Henkins (2002) concluded that understanding preservice teachers' perceptions of diversity is the first step towards driving programmatic change.

Continuing with the importance of assessing classroom dispositions, Wilkerson (2006) went as far to indicate "dispositions are, in the long run, more important than knowledge and skills" (p. 2), while Sherman (2006) suggested that a teacher's overall classroom approach may be more important than his/her pedagogical skills and knowledge when it comes to learning. And, as Johnston, Almerico, Henriotte, & Shapiro (2011) indicate, teacher candidates' professional dispositions are predictors of overall teaching effectiveness.

## **Approaches to Assessing Dispositions**

The first question one might ask is "Is it possible to assess dispositions?".

An accurate assessment of anything begins with a clear, precise definition of the phenomena being measured. Before specific definitions are agreed upon, there must be an agreement on which dispositions are to be measured. Add to this the ability to actually observe what one wants to measure, and an assessment/data collection tool that is valid and reliable. Various assessment approaches have been successfully utilized to assess a variety of constructs and observable behavior, including surveys, checklists, interviews, self-reports, portfolios, standardized assessments/test scores, and so on. Several of these assessment approaches have been used in a variety of teacher preparation programs as well, specifically relating to assessing dispositions (Conderman & Walker, 2015, Notar, Riley, & Taylor, 2009, Choi, Benson, & Shudak, 2016, DiGiacinto, Bugler & Wayda, 2017). Phelps (2006) reports that the best approach to assessing dispositions is to physically observe preservice teachers in the classroom. She feels that indicators of appropriate teaching dispositions include caring about student learning and planning creative lessons. Finally, Wasicsko (2007) discusses the importance of using a self-evaluation approach in combination with an external evaluator. Here, the preservice teacher can compare his or her perceptions of his or her own dispositions to what an evaluator observes.

Within the past several years, many teacher preparation programs have developed surveys and/or checklists/rating scales to focus on the assessment of pre-service/teacher dispositions. Choi, Benson, & Shudak, (2016) used a 19-item rating scale that included a 3-point rubric (1 = below expectations to 3 = exceeds expectations) in their approach to assessing dispositions. Here, candidates were rated twice by both university and field supervisors. They conclude in their findings that unless the term "dispositions" is clearly defined, it cannot be

reliably and validly assessed. Frederiksen, Cooner, and Stevenson (2012) used a quantitative non-experimental approach to assess perceived dispositions of preservice teachers towards urban and nonurban settings in addition to measuring the effectiveness of the internship in these settings.

Welch, Napoleon, Hill, and Roumell (2014) developed The Virtual Teaching Dispositions Scale (VTDS) to assess teacher dispositions in a virtual setting. The VTDS assesses teacher-related dispositions in three basic domains: Social Presence, Pedagogical Presence, and Expert/Cognitive Presence. For each domain, teachers rate themselves on several descriptive items according to a 4-point Likert scale (1 = Very untrue of me (includes having no experience in that situation), 2 = Not really true of me, 3 = Somewhat true of me, 4 = Very true of me). For example, one specific item in the Social Presence domain is “I am empathetic to the needs of my students”. The authors discussed their pilot study using the VTDS and reliability measures for each domain.

As mentioned prior, Jensen, Whiting, and Chapman (2018) developed the Multicultural Teacher Dispositions Scale (MTDS), Phase 2. This edition of the assessment contained 26 “agreement” items relating to the dispositions of empathy, meekness, social awareness, inclusion, and advocacy. After piloting this version, and compiling the data collected, the authors reported that continued refinement is ongoing.

After considerable research, Almerico, Johnston, Henriott, and Shapiro (2010) developed the Educator Disposition Assessment (EDA). This evaluation tool uses specific identifiers to assess candidates on several disposition areas including: Effective Oral and Written Communication Skills, Displaying a Positive and Enthusiastic Attitude, Preparedness in Teaching and Learning, Appreciation for Cultural/Academic Diversity, Initiative, and

Social/Emotional Intelligence. Each of these categories is further broken down using a 3-point Likert-scale with clearly defined descriptive indicators as part of a scoring rubric that are rated as; “0 (Needs Improvement)”, “1 (Developing)” or “2 (Meets Expectations)”. The rater uses the scoring rubric as a descriptive guide to evaluate the various disposition elements in each category.

The psychometric properties relating to this instrument (and other disposition assessment tools) are discussed in detail in other studies (Johnston, Almerico, Henriott, & Shapiro, 2011), (Johnston, Wilson, & Almerico, 2018).

### **Purpose of the Present Study**

This purpose of this study is to present the results of a study that included 15 preservice teachers who were assessed on the Educator Disposition Assessment (EDA) (Almerico, Johnston, Henriott, & Shapiro, 2010) by two independent raters between February 2020 and April 2020. Additionally, Inter-Rater Reliability measures evaluate the degree to which the raters agree/disagree on the student assessments, and the individual EDA categories.

### **Method**

#### **Sample**

All preservice teachers enrolled in the Undergraduate Teacher Education Program at a private, not-for-profit University in South Florida are required complete, as a final part of their preparation program, a 14-week student teaching internship. This capstone event occurs just prior to graduation. They work collaboratively with the classroom cooperating teacher and receive ongoing supervision from a clinically trained University Supervisor. Initially, they observe and take notes in the classroom, and gradually take on more classroom responsibilities until they finally take full responsibility for the classroom. They are evaluated at several points during the internship by both the cooperating teacher and University supervisor. Additionally,

the cooperating teacher and University supervisor meet to discuss their individual evaluations, and again meet with each teacher candidate to review the evaluation results. This study presents the evaluation results of 15 teacher candidates at the final stage of their internship as related to dispositions.

### **Instrument**

The Educator Disposition Assessment (EDA) was used as a comprehensive measure of dispositions (Almerico, Johnston, Henriott, & Shapiro, 2010). This tool is summarized in the above section. The cooperating teacher rated each intern on the EDA during the final week of internship, and the University Supervisor rated the intern during his/her final observation. Both the cooperating teacher and University Supervisor received training on the EDA prior to administration. All assessment data were entered into LiveText, which is password protected. No identifying information was entered or recorded into LiveText, only EDA evaluation results. It should be noted that the EDA Likert categories were recoded to “1 (Needs Improvement)”, “2 (Developing)”, and “3 (Meets Expectations)” for the purpose of analysis.

### **Results**

Review of the data collected by the University Supervisor indicated that most teacher candidates (14) scored “Meets Expectations” on the many elements of this assessment. The exceptions were that one student scored “Developing” on three areas of this assessment (Positive and Enthusiastic Attitude, Self-regulated Learner Behavior/Initiative, and Social and Emotional Intelligence).

Review of the data collected by the Cooperating teacher indicated that, again, most teacher candidates (13) scored “Meets Expectations” on the many elements of this assessment. The exceptions were that one teacher candidate scored “Developing” on Appreciation for Culture

and Academic Diversity, and a second teacher candidate scored “Developing” on Self-regulated Learner Behavior/Initiative.

Next, evaluation results were compared between the cooperating teacher and University Supervisor to compile inter-rated reliability. The table below summarizes the results.

**Table 1**

*Data Summary*

EDA Category	University Supervisor Rating (Mean/Median) (n= 15)	Cooperating Teacher Rating (Mean/Median) (n= 15)
Demonstrates Effective Oral Communication Skills	3.00/3.00	3.00/3.00
Demonstrates Effective Communication Skills	3.00/3.00	3.00/3.00
Demonstrates Professionalism	3.00/3.00	3.00/3.00
Demonstrates a Positive and Enthusiastic Attitude	2.93/3.00	3.00/3.00
Demonstrates Preparedness in Teaching and Learning	3.00/3.00	3.00/3.00
Exhibits an Appreciation of and Value for Culture and Academic Diversity	3.00/3.00	2.933/3.00
Collaborates Effectively with Stakeholders	3.00/3.00	3.00/3.00
Demonstrates Self-regulated Learner Behavior/Takes Initiative	2.933/3.00	2.933/3.00
Exhibits the Social and Emotional Intelligence to Promote Personal and Educational Goals/Stability	2.933/3.00	3.00/3.00

When looking at the Inter-rater reliability on each EDA category for individual preservice teachers, there was full agreement by both raters on most assessment items. The exceptions were as follows:

For teacher candidate 6, the cooperating teacher rated all item a score of 3 (“Meets expectations”), while the University supervisor rated the same student a score of 2 (“Developing”) on three of the EDA categories (Positive and Enthusiastic Attitude, Demonstrates Self-regulated Learner Behavior/Takes Initiative, and Exhibits the Social and Emotional Intelligence).

For teacher candidate 9, the University supervisor rated all EDA items a score of 3 (“Meets expectations”), while the cooperating teacher agreed, with the exception of EDA category 6. Here s/he rated this item (Exhibits an Appreciation of and Value for Culture and Academic Diversity) a score of 2 (“Developing”). Additional statistics computed on the data did not reveal any statistically significant differences when looking at overall Inter-rater reliability.

### **Discussion**

As can be seen in the data summary and table, all teacher candidates received high ratings from both the cooperating teacher and University Supervisor. In only five occasions were ratings less than 3 given; one preservice teacher received ratings of 2 (Developing) on Positive and Enthusiastic Attitude, Self-Regulated Learner Behavior/Initiative, and Social/Emotional Intelligence from the University Supervisor; a second received a rating of 2 (Developing) on Demonstrates Self-regulated Learner Behavior/Takes Initiative from the University Supervisor; a third preservice teacher received a rating of 2 (Developing) on Appreciation of and Value for Culture and Academic Diversity from the cooperating teacher. These three preservice teachers received ratings of “3” in all areas by the second rater. No preservice teacher received a rating of “Needs Improvement” on any of the EDA categories.

## **Study Limitations**

One limitation of this study is that the preservice teachers who were assessed on the EDA only had one assessment point (at the end of internship). It is possible that they demonstrated the various dispositions (as measured by the EDA) prior to beginning their education program or developed them as they progressed through the program. It would be beneficial to administer the EDA at program entrance point and then again at exit point to evaluate any changes in dispositions. Thus, it is not possible to state that the education program and/or internship was the result of the high ratings they received overall. Additionally, it would be beneficial to have the preservice teachers rate themselves on the EDA and compare their responses to those from the cooperating teacher and University Supervisor. Furthermore, it would be important to recruit a control group (perhaps students in programs other than teacher education) and compare pre/post results between groups. Finally, this study evaluated 15 preservice teachers in one localized area. This is a very small sample size, and external validity is an issue. A larger sample size, and wider study region may produce different results.

## References

- Almerico, G., Johnston, P., Henriott, D., & Shapiro, M. (2011). Dispositions assessment in teacher education: developing an assessment instrument for the college classroom and the field. *Research in Higher Education Journal*, Vol 11. Retrieved from <http://aabri.com/manuscripts/10665.pdf>
- Bauer, D., & Thornton, H. (2013). A case analysis of middle level teacher preparation and long-term teacher dispositions middle level teacher preparation and responsive dispositions. *Research in Middle Level Education Online*, 37(3), 1–19. Retrieved from <http://files.eric.ed.gov/fulltext/EJ1032360.pdf>
- Borko, H., Liston, D., & Whitcomb, J. A. (2007). Apples and fishes: The debate over dispositions in teacher education. *Journal of Teacher Education*, 58, 359-364.
- Chickering, A. W., Gamson, Z. F., & American Association for Higher Education, W. C. (1987). Seven principles for good practice in undergraduate Eeducation. *AAHE Bulletin*, 3-7.
- Choi, H., Benson, N. F., & Shudak, N. J. (2016). Assessment of teacher candidate dispositions: Evidence of reliability and validity. *Teacher Education Quarterly*, Summer, 71-89.
- Conderman, G., & Walker, D. A., (2015). Assessing dispositions in teacher preparation programs: Are candidates seeing the same thing? *The Teacher Educator*, 50, 215-231
- Council for the Accreditation of Educator Professionals. (2018). Retrieved from <http://caepnet.org/standards/standard-1>
- Cummins, L. & Asempapa, B. (2013). Fostering teacher candidate dispositions in teacher education programs. *Journal of the Scholarship of Teaching and Learning*, 13, 3, 99-119.

- Dee, Jay R & Henkin, Alan B. (2002). Assessing dispositions toward cultural diversity among Preservice teachers. *Urban Education; Thousand Oaks*, 37, 22-40.
- Da Ros-Voseles, D., & Moss, L. (2007). The role of dispositions in the education of future teachers. *Young Children*, 62, 5, 90-98.
- DiGiacinto, K. L., Bugler, S. M., & Wayda, V. (2017). Rethinking PETE Program admissions to include teacher candidate dispositions. *The Physical Educator*, 74, 63-84.
- Dottin ES (2009). Professional judgment and dispositions in teacher education. *Teacher Education*, 25(1):83-88.
- Frederiksen, H., Cooner, D., & Stevenson, C. (2012). Assessing Teacher Dispositions in pre-service teachers. *Journal of College Teaching & Learning (Online)*; Littleton. Vol. 9, Is. 1, 39.
- InTASC Model Core Teaching Standards: A Resource for State Dialogue, (2011). Retrieved from [https://www.ccsso.org/sites/default/files/2017-11/InTASC\\_Model\\_Core\\_Teaching\\_Standards\\_2011.pdf](https://www.ccsso.org/sites/default/files/2017-11/InTASC_Model_Core_Teaching_Standards_2011.pdf)
- Interstate New Teacher Assessment and Support Consortium (2018). Retrieved from <http://www.ecu.edu/cs-educ/teached/upload/INTASCStandardsIndicators.pdf>
- Jensen, B; Whiting, E. F., & Chapman, S. *Journal of Psychoeducational Assessment* Vol. 36, Iss. 2, (Mar 2018): 120-135. DOI:10.1177/0734282916662426
- Johnston, P, Wilson, A., & Almerico, G., (2018). Meeting psychometric requirements for disposition assessment: Valid and reliable indicators of teacher dispositions. *Journal of Instructional Pedagogies*, 21. Retrieved from <http://www.aabri.com/manuscripts/182847.pdf>

- Johnston, P, Almerico, G., Henriott, D., & Shapiro, M. (2011). Descriptions of dispositions for assessment in pre-service teacher education field experiences. *Education, 132, 2*. Retrieved from <https://www.questia.com/library/journal/1G1-278759337/descriptions-of-dispositions-for-assessment-in-pre-service>
- Katz, L. G., & Raths, J. D. (1986). Dispositional goals for teacher education: Problems of identification and assessment. (ERIC Document Reproduction Service No. ED2722470)
- Martins, N., & Ungerer, L. (2017) *Africa Education Review (14)* 156-73  
DOI:10.1080/18146627.2017.1294971
- Marzano, R. & Brown, J. (2009). *A handbook for the art and science of teaching*. Association for Supervision and Curriculum Development.
- Murray, F. B. (2007). Disposition: A superfluous construct in teacher education. *Journal of Teacher Education, 58*, 381-387
- National Council for Accreditation of Teacher Education (NCATE) (2001). *Professional standards for the accreditation of schools, colleges, and departments of education*. Washington, DC: Author
- National Council for Accreditation of Teacher Education (NCATE). *Professional standards 2008*. (2008) Retrieved from <http://www.ncate.org/~media/Files/caep/accreditation-resources/ncate-standards-2008.pdf?la=en>
- Nixon, A., Dam, M. & Packard, A. (2010). Teacher disposition and contract non-renewal. *Planning and Changing, 41*, 210-219.
- Notar, C. E., Riley, G. W., & Taylor, P. W. (2009). Dispositions: Ability and assessment. *International Journal of Education, 1, 1*, 1-14.

- Osguthorpe, R. D. (2013). Attending to ethical and moral dispositions in teacher education. *Issues in Teacher Education*, 22, 1, 17-28.
- Phelps, P. H. (2006). The Dilemma of Dispositions. *The Clearing House: A Journal of Educational Strategies, issues, and Ideas*. 79:4, 174-178, DOI 10.3200/TCHS.79.4.174-178.
- Sherman, S. (2006). Moral dispositions in teacher education: Making them matter. *Teacher Education Quarterly*, Fall, 41–57. Retrieved from <http://files.eric.ed.gov/fulltext/EJ795225.pdf>
- Stewart, P. & Davis, S. (2005) Developing Dispositions of Preservice teachers through Membership in Professional Organizations. *Journal of Authentic Learning*, 2, 1, 37-46.
- Taylor, R. & Wasicsko, M. (2000). *The dispositions to teach*. Paper presented at SRATE, Kentucky.
- Thornton, H. (2006). Dispositions in Action: Do Dispositions make a difference in Practice? *Teacher Education Quarterly*, 33, 2, 53-68.
- Usher, D. (2002, November). *Arthur Combs' five dimensions of helper belief reformulated as five dispositions of teacher effectiveness*. Paper presented at the meeting of the First Annual Symposium on Educator Dispositions: Effective Teacher--Effective Person. Eastern Kentucky University.
- Wasicsko, M. M. (2007). Dispositions—A perceptual approach. In M.E. Diez & J. Rath (Eds.), *Dispositions in Teacher Education* (pp. 3-29). Information Age Publishing, Inc.
- Welch, A. G., Napoleon, L., Hill, B., & Roumell, E. (2014). Virtual teaching dispositions scale (VTDS): A multi-dimensional instrument to assess teaching dispositions in virtual classrooms. *MERLOT Journal of Online Learning and Teaching*, 10 (3) 446-467.

Welch, F. C., Pitts, R. E., Tenini, K. J., Kuenlen, M. G., & Wood, S. G. (2010). Significant issues in defining and assessing teacher dispositions. *The Teacher Educator*, 45(3), 179–201. <http://doi.org/10.1080/08878730.2010.489992>

Wilkerson, J. (2006). Measuring teacher dispositions: Standards-based or morality-based? *The Teachers College Record*. Retrieved from <http://www.tcrecord.org/content.asp?contentid=12493>

## **Engagement and Big-Fish-Little-Pond Effect on Self-Concept in Science: Findings from TIMSS 2011**

Nurcan Kahraman, Yasemin Tas, and Sündüs Yerdelen

### **Abstract**

Big-Fish-Little-Pond-Effect (BFLPE) suggests that students form their own self-concept by comparing their abilities to those of their peers. This study aims to examine how students' gender, socioeconomic status, science achievement, and engagement dimensions (motivational and behavioral engagement) predict students' self-concept in science by considering BFLPE. This study used TIMSS 2011 data set including 6928 eighth grade Turkish students from 239 schools. Since the data were in a nested structure (students nested in classrooms) HLM analysis was preferred to minimize the dependences of the data. HLM analyses showed that about 10% of the variance in students' self-concept in science learning was between classes while the rest of the variance was within classes. Findings suggest that students' motivational and behavioral engagements are important indicators of students' self-concept as well as science achievement. Additionally, at the class level, aggregated science achievement was found to be significant and a negative predictor of self-concept, which supported BFLPE.

**Keywords:** TIMMS, science education, Academic self-concept, Big-Fish-Little-Pond Effect

*Dr. Nurcan Kahraman is an Associate Professor in the Department of Mathematics and Science Education at Bursa Uludag University. He can be reached at [nurcankahraman@uludag.edu.tr](mailto:nurcankahraman@uludag.edu.tr)*

*Dr. Yasemin Tas is an Associate Professor in the Department of Mathematics and Science Education at Ataturk Education. He can be reached at [tasyase@gmail.com](mailto:tasyase@gmail.com)*

*Dr. Sündüs Yerdelen is an Associate Professor in the Department of Mathematics and Science Education at Kafkas University. He can be reached at [suyerdelen@gmail.com](mailto:suyerdelen@gmail.com)*

Academic self-concept is one of the constructs of self-belief that attracted researchers' attention (Bong & Skaalvik, 2003). Individuals' self-perceived ability in a domain is the central element of self-concept (Bong & Skaalvik, 2003; Ferla, Valcke, & Cai, 2009). Self-concept beliefs are formed based on individuals' past experiences such as accomplishments and failures. To illustrate, individuals compare their self-perceived competence with others, thus individuals' evaluation of their competence includes some normative criteria (Bong & Skaalvik, 2003; Ferla et al., 2009; Wigfield & Eccles, 2000). Self-concept is important because "[o]ne's perceptions of himself are thought to influence the ways in which he acts, and his acts in turn influence the ways in which he perceives himself" (Shavelson, Hubner, & Stanton, 1976, p. 411). Several studies have demonstrated that self-concept is positively correlated to desirable student outcomes such as self-efficacy and interest (e.g. Ferla et al., 2009), student motivation to aspire a career in science (e.g. Jansen, Scherer, & Schroeders, 2015), and academic achievement (e.g. Suárez-Álvarez, Fernández-Alonso, & Muñiz, 2014). Examining the factors influencing students' self-concept is important to enhance quality of education and, in turn, students' learning outcomes (Bong & Skaalvik, 2003). In the present study, students' self-concept in science is explored considering the Big-Fish-Little-Pond Effect (BFLPE).

According to BFLPE, the frame of reference for students to compare their abilities is their classmates' abilities. Based on this comparison, students form their own academic self-concept (Marsh et al., 2015). The BFLPE mainly focused on the relationship between academic self-concept and achievement which revealed differential effects at different levels. Studies pointed out that student achievement and self-concept was positively related at the student level while the effect of average achievement on student self-concept was negative at the class or school level (e.g., Marsh et al., 2008a; Seaton, Marsh, & Craven, 2010). This effect is known as

BFLPE. According to BFLPE, students attending high achieving classes or schools have lower academic self-concept than students with equal achievement who are attending low or average achieving classes or schools (e.g., Marsh, 2004; Seaton et al., 2010). For instance, Marsh (2004), utilizing Australian students' responses to The Program for International Student Assessment (PISA) 2000 survey, found that individual achievement positively predicted self-concept whereas school-average achievement was a negative predictor of self-concept.

Most of the research on BFLPE was conducted with secondary school students in Western countries and there is a need to investigate this effect with younger students in different countries (Marsh et al., 2015). Some distinct findings regarding self-concept and achievement were reported by studies conducted in East Asian countries (e.g., Kung, 2009; Wilkins, 2004; Yoshino, 2012). For instance, utilizing TIMSS 2003 data Kung (2009) found that although Taiwanese students have higher mathematics achievement, they reported lower mathematics self-concepts than their Western counterparts. Similarly, Yoshino (2012) exploring TIMSS 2007 data found that Japanese students have higher achievement but lower mathematics self-concept than American students. Turkey, located in the Middle East, is a bridge between Europe and Asia. In this sense, the Turkish culture is distinct from other nations' cultures which reflect both Western and Eastern values (Yalvac, Tekkaya, Cakiroglu, & Kahyaoglu, 2007). Besides that, the education system in Turkey is competitive; students take elimination exams not only to enter high school but also to enter a college. Therefore, due to such exams, students in Turkey frequently compare themselves with their peers and make social comparisons. Furthermore, supporting these comparisons, the evaluative feedback provided to students by teachers (Odabaşı Çimer, Bütüner, & Yiğit, 2010) may also influence students' academic self-concept beliefs. Therefore, the present study is an attempt to better understand the influence of Turkish students'

perceptions of their classmates' achievement on their own self-concept beliefs and this study aims to contribute to the generalizability of BFLPE by studying it within the Turkish context. Furthermore, the present study uses The Trends in International Mathematics and Science Study (TIMSS) 2011 data set. TIMSS data were most appropriate to study BFLPE, since the unit of analysis is classroom (Marsh et al., 2015). Parallel to BFLPE assertions, we hypothesize that students' self-concept beliefs in science will be positively predicted by individual achievement and negatively predicted by class-average science achievement. In other words, students' beliefs about their competence to do well in science will be high if they have high achievement. On the other hand, students will feel less competent if their classmates have high achievement than students with the same ability level but having classmates with low achievement. Besides investigating BFLPE, this study also aims to examine the relationship between student engagement and academic self-concept in science. The role of engagement in self-concept also deserves investigation as will be discussed in the next section.

### **Self-concept and Engagement**

Engagement attracts educational researchers' attention (Sinatra, Heddy, & Lombardi, 2015) and is shown to be closely related to students' achievement (e.g., Bae & DeBusk-Lane, 2019; Schnitzler, Holzberger, & Seidel, 2021). In the relevant literature, there are various definitions about school engagement (Jimerson, Campos, & Greif, 2003). Linnenbrink and Pintrich (2003) addressed engagement as a multifaceted construct in terms of behavioral, motivational, and cognitive engagement. Behavioral engagement is interested in students' observable behaviors like their effort for, or persistence at a task. This type of engagement can be easily observed by watching students. The second component of engagement is motivational engagement which is concerned with students' interest in and value attached to learning the

material. The last component of engagement, cognitive engagement, refers to students' investment in learning. In other words, students who are engaged cognitively think about the content and the learning process and use self-regulated strategies to understand the content. Since TIMSS 2011 only provided data related to behavioral and motivational engagement, this study is limited with only these two dimensions of engagement.

The behavioral engagement dimension is focused on students' observable behaviors, in other words, students' participation, effort and persistence in school activities (Finn, Pannozzo & Voelkl, 1995). Students who are behaviorally "engaged in" tend to do the tasks, show effort, be persistent on the task, and seek help when they need it (Fredricks et al., 2004). Connell and Wellborn (1991) proposed "context-self-action-outcome" model and suggested that students' self-beliefs, like self-concept, directly affect their engagement. In particular, students who show effort and persistence at a task are those who have high self-concept, namely who think that they can do the task. There are many researchers that confirm predictor effect of self-concept on students' engagement in the literature (e.g. Bush, 2005; Schnitzler et al., 2021). However, the relevant literature also suggests that the relation between self-beliefs and engagement may be reciprocal; engagement can also help students believe their abilities (Salanova, Llorens & Schaufeli, 2011). Supporting this idea, Badiozaman (2012) investigated the relationship between students' self-concept and engagement and indicated that not only students' self-concept affects their engagement, but students' engagement may also affect their self-concept. Moreover, Green et al. (2012) investigated the relationship between high school students' self-concept and their school engagement and suggested that students' self-concept is a significant predictor of their engagement. The researchers offered that future studies may consider engagement as a predictor of students' self-concept. However, according to the authors' knowledge, there is not much

research investigating behavioral engagement as a predictor of self-concept. Hence, the current study aims to examine how behavioral engagement predicts students' self-concept in science.

The other component of engagement, motivational engagement, refers to students' interest in the task. There are at least three aspects of motivational engagement: interest, affect, and value beliefs (Linnenbrink & Pintrich, 2003). Students' interest refers to their liking or disliking the subject; while, utility value refers to students' thoughts about the usefulness or importance of the task. Lastly, affect is concerned with students' affective or emotional experiences (Pintrich & Schunk, 1996). These aspects of motivational engagement are positively related to students' self-concept and this relation is reciprocal. In other words, students who find tasks useful or interesting tend to have high self-concept and vice versa (e.g., Eccles & Wigfield, 1995; Eccles & Wigfield 2002). Empirical evidence exists for these relationships (e.g., Bakadorova & Raufelder, 2017; Cheung, 2018; Raufelder et al., 2015). For instance, in a longitudinal study Bakadorova and Raufelder (2017) examined the relationships among German students' emotional and behavioral school engagement, self-concept, and with peers and teachers as motivators from beginning of 8th grade to the end of 9th grade. Over time, associations demonstrated that self-concept was positively related to both behavioral and emotional engagement. On the other hand, self-concept negatively predicted students' emotional engagement over time. The authors explained that this might be due to suppression effect since the correlations between the variables were positive. In another study conducted in the United States, Wang and Eccles (2013) found that middle school students' self-concept and task value beliefs mediated the relationship between perceived school environment and student engagement. More specifically, self-concepts and task value significantly and positively predicted behavioral, emotional, and cognitive engagement. In the present study, liking science

and valuing science will be examined as predictors of self-concept beliefs in science. Empirical evidence also suggests that the strength of the relationship between science self-concept and value of science may change from culture to culture (Schütte, 2015). This study also attempts to understand the associations between self-concept and motivational engagement in the Turkish culture.

### **The Role of Gender and SES in Self-Concept**

Previous research showed that gender was associated with academic self-concept (e.g., Cooper, Krieg, & Brownell, 2018; Jansen, Schroeders, & Lüdtke, 2014). For example, Wilkins (2004) examined the gender difference in students' science self-concept for different countries based on the TIMSS data. The author found that a gender gap existed between males and females in the favor of the former. Additionally, Wilkins stated that this gap was observed at different magnitudes across the countries and, in fact, females had higher self-concept in science in few countries. In another study, Jansen et al. (2014) found that female students have lower self-concept in chemistry and physics in a sample of 10<sup>th</sup> grade German students. Nevertheless, gender difference in students' science self-concept has been relatively less studied in Turkey and studies have revealed inconsistent findings. For example, Senler and Sungur (2009) examined elementary and middle school Turkish students' self-concept in science regarding gender. Their results revealed no difference between girls and boys in their science self-concept. On the other hand, in their study with high school students taking biology course, Pehlivan and Köseoğlu (2010) found that girls had higher self-concepts than boys.

Socioeconomic status is another variable that may influence academic self-concept (e.g., Easterbrook, Kuppens, & Manstead, 2020). For example, Maqsd and Rouhani (1991) found a positive correlation between socioeconomic status and self-concept of secondary school students.

Additionally, some researchers who focused on whether the BFLPE applied to all students in the same way examined if socioeconomic status influenced the magnitude of the negative association between the group self-concept and individual self-concept (e.g., Seaton, et al., 2010; Trautwein, Lüdtke, Marsh, & Nagy, 2009). Therefore, in this study, it seems that controlling for the SES variable is necessary when examining the association between engagement and self-concept.

### **Significance and Purpose of the Study**

In line of the aforementioned studies, investigating the factors influencing students' self-concept is important since it is related to several adaptive student outcomes (e.g., Jansen et al., 2015; Suárez-Álvarez et al., 2014). In the present study, Turkish students' self-concept in science is explored from the perspective of BFLPE. Parker et al. (2021) calls for educational policy makers to consider BFLPE while making decisions. Although generalizability of negative effect of the class or school achievement on students' self-concept have been demonstrated by several studies (e.g., Guo et al., 2018), some cultural differences are also reported regarding self-concept and achievement (e.g., Yoshino, 2012). Reflecting both Western and Eastern values, Turkish culture is unique (Yalvac et al., 2007). The examination oriented and competitive Turkish education system may influence students' self-concept, as well (Senler & Sungur, 2009). Utilizing TIMSS 2011 data, this study attempts to provide empirical evidence about these relations in Turkey. Furthermore, predictive power of emotional and behavioral engagement on students' self-concept will be examined. According to our knowledge, there is limited research investigating behavioral engagement as a predictor of self-concept and this study will contribute to this issue. Additionally, due to their potential influences, gender and socioeconomic status were entered in the model to control their effects. In the light of the mentioned literature, the

present study has two major aims: first it aims to investigate the relation of middle school students' self-concept to student engagement and achievement in science. Second, considering BFLPE on students' self-concept. Utilizing Turkish 8th grade students' responses to TIMSS 2011 survey the present study aims to answer the following research questions:

1. To what extent students' science achievement and engagement dimensions (i.e., motivational and behavioral engagement) predict students' self-concept in science while controlling for gender and socioeconomic status?
2. Is there a BFLPE (class average achievement) on students' self-concept in science?

## **Method**

### **Design of the Study**

This study is a correlational quantitative study (Frankel & Wallen, 2006) which aims to examine the relation between engagement, achievement and self-concept of middle school students in science.

### **Sample**

In the present study, TIMSS 2011 data set that was obtained from Turkey was used. In the data collection procedure of TIMSS, the stratified two-stage sampling design was used. Namely, in the first stage of this sampling procedure, schools were sorted by considering their sizes and important demographic variables and then systematic random sampling method was applied. In the second stage, one or two intact classes with the students at the target grade levels were selected by using systematic random sampling method (Joncas & Foy, 2012).

Sample of the present study includes 6928 (49.3% girls and 50.7 % boys) eighth grade students from 239 classrooms. Number of participants in each classroom ranged between 10 and 56. Mean age of students is 14.06 (SD= .67). Students mostly have computer at home (58.2%),

own room (52.9%) and less than 100 books at home (83.7%); but do not have internet connection (54.4%).

## **Instruments**

Instruments used in the present study were obtained from TIMSS 2011 questionnaire. TIMSS examined the validity of the instruments and suggested both criterion- related, and comparative validity (Martin & Mullis, 2012). Characteristics of the instruments are summarized in Table 1.

### ***Self-concept in Science***

Students' self-concept scores were computed by using nine items that were included in Student Confident in Science (SCS) Scale. (e.g., "I am good at working out difficult science problems" and "Science is more difficult for me than for many of my classmates"). This scale was based on 4-point Likert response scale ranging from 1 (Agree a lot) to 4 (Disagree a lot). Cronbach's alpha internal consistency for the self-concept scale was high (.85). Principal component analysis of the items suggested that item loadings in Turkish data set varied between .58 to .76 for SCS scale (Martin & Mullis, 2012).

### ***Students' Value Science (SVS) Scale***

In TIMSS 2011, students' value science was measured by students' responses to six items including a sample item of "I need science to learn other school subjects". Students' responses ranged from 1 (Agree a lot) to 4 (Disagree a lot). Reliability (Cronbach's alpha) of this scale was found to be .82 indicating high internal consistency. Factor loadings of items varied between .59 and .84 for Turkish data (Martin & Mullis, 2012).

### ***Students Like Learning Science (SLS) Scale***

The extent to which students liked learning science was assessed by students' responses to five items. A sample item was: "I enjoy learning science". Responses were collected using a 4-point Likert scale with anchors ranging from 1 (Agree a lot) to 4 (disagree a lot). Cronbach's alpha internal consistency was calculated as .79 indicating good reliability. Factor loadings of items varied between .61 and .87 for Turkish data (Martin & Mullis, 2012).

### ***Students' Engaged in Learning Science (SES) Scale***

In TIMSS 2011, the level of students' engagement in learning science was measured by five items. A sample item was "I am interested in what my teacher says". Response options in the scale included four possibilities on a Likert scale ranging from 1 (Agree a lot) to 4 (Disagree a lot). Cronbach's alpha internal consistency of this scale was calculated as .66 indicating acceptable internal consistency (Özdamar, 1997). Factor loadings of items varied between .42 and .76 for Turkish data (Martin & Mullis, 2012).

### ***Science Achievement Scores (SAS)***

In TIMSS 2011, science achievement questions included four science domains: biology, chemistry, physics, and earth science. The general science achievement scores were transformed into five plausible values by using Item Response Theory. Besides using achievement at level-1 (student level), these 5 plausible values were also aggregated for use at second level (class level).

### ***Socioeconomic Status***

Students' socioeconomic status points were computed by using the items of Home Educational Resources (HER) scale. HER is a composite variable created based on students' responses on the 3 questions concerned with the availability of home resources. These items include the number of books at home, number of home study supports, and highest level of

education of either parent. Cronbach's Alpha was .63 indicating acceptable reliability (Özdamar, 1997). Factor loadings of items varied between .75 and .80 for Turkish data (Martin & Mullis, 2012).

**Table 1**

*TIMSS 2011's scales that were used in this study*

Measured construct	Scale of TIMSS 2011	Number of items	Sample item	Cronbach alpha	Response scale
Self-concept	Students confidence in science (SCS)	9	I usually do well in science	.85	1=agree a lot 2=agree a little
Motivational engagement	Students Like Learning Science (SLS)	5	I enjoy learning science	.79	3=disagree a little
	Students Value Science (SVS) Scale	6	I need science to learn other school subjects	.82	4=disagree a lot
Behavioral engagement	Students Engaged in Science Learning (ESL)	5	I am interested in what my teacher says	.66	
Socioeconomic status	Home Educational Resources (HER)	3	Number of books at home	.63	Different for each question

## Procedures

The data were obtained from TIMSS 2011's official website. Source: TIMSS 2007 Assessment. Copyright © 2009 International Association for the Evaluation of Educational Achievement (IEA). Publisher: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. This website includes an international database of data which was obtained from 45 participating countries, and anybody who is interested in examining the TIMSS data can access and download the data set. In line with the research questions, only the data collected from 8th grade students in Turkey was used in the study. Since the student sample of the present study was clustered within classes, the responses students provided cannot be regarded as independent from each other. Therefore, dependencies among student responses

should be taken into consideration to obtain more robust results from analyses. Thus, HLM analysis method was utilized with HLM 6 software (Raudenbush, Bryk, Cheong, & Congdon, 2004). Another advantage of using HLM analysis is that variables at student level and teacher/school level can be analyzed simultaneously. In this study, both student and class level variables were tested. In order to find out which variables significantly predicted students' self-concept in science, students' science achievement scores, motivational engagement (i.e., liking science and value science), behavioral engagement, student gender, and socioeconomic status were incorporated into the HLM analysis at level-1, and aggregated classroom achievement scores were incorporated as level-2 variables. We considered weighing variables while conducting the HLM analysis. All predictors (except gender) were grand mean centered. Moreover, Raudenbush, Bryk, and Congdon (2005) suggested that multilevel modeling regression analysis should be conducted for each of the plausible value and averaging coefficients obtained throughout the analyses. Considering their recommendation, we ran the HLM model five times for each plausible value and then calculated averages.

## **Results**

### **Preliminary Analyses**

Before performing HLM preliminary analyses were conducted. There was no missing value more than 3% which is inconsequential for the analysis (Schafer, 1999). The variables showed normal distribution. The skewness and kurtosis values of the variables were the range of between -1 to +1. Additionally, there were no violation of the assumptions of HLM. Besides, means, standard deviations and bivariate correlation coefficients are presented in Table 2. Results showed that all bivariate correlations were significant. The highest correlation was between liking science and self-concept ( $r = .65, p < .01$ ). The descriptive statistics suggest that

students are at moderate level on liking science, value science, self-concept in science, and behavioral engagement. Mean science achievement score of eight grade Turkish students was 478.98 which was below the TIMSS scale average score of 500 (Oral & McGivney, 2013).

**Table 2**

*Descriptive statistics and bivariate correlations among variables*

	Min.	Max.	Mean	SD	Cut points		Bivariate correlations				
					Low- High	Middle - 2	3	4	5	6	
<b>1. Socioeconomic status</b>	4.32	14.02	8.34	2.07	8.2-12.5	.05**	.04**	.17**	.12**	.42**	
<b>2. Liking Science</b>	3.51	12.94	10.60	1.89	8.4-10.8		.48**	.65**	.60**	.24**	
<b>3. Value science</b>	4.14	13.1	10.01	1.84	8.6-10.5			.48**	.49**	.15**	
<b>4. Self-concept in science</b>	3.09	14.97	10.27	2.01	9-11.5				.60**	.40**	
<b>5. Behavioral engagement</b>	3.56	13.83	10.38	1.85	8.4-11.2					.26**	
<b>6. Science achievement</b>			478.98	97.17							

\*\* p<.01

### **Predicting Self-Concept in Science**

First, intraclass correlation coefficient (ICC) was computed by using the variance components obtained from one-way random effect ANOVA model. ICC is an indicator of expected (population) correlation between two randomly selected students within the same class (Hox, 2010). Besides, ICC can also be considered as the magnitude of random effect in multilevel models (Lorah, 2018). For student self-concept, ICC indicated that approximately 10% of the variance in students' self-concept in science learning can be explained by class level variables.

As for the prediction of self-concept, five multilevel modeling regression analyses were conducted with each of the achievement plausible values. Achievement was handled as the

achievement score of students, without group-mean centering, at level-1, and as the achievement of the classroom at level-2. Afterwards, based on these results, parameter estimates, standard errors, t values and between and within variances were calculated by averaging results obtained from results derived for each plausible value (Raudenbush et al., 2005). Results revealed that, at level-1, gender, science achievement, socioeconomic status, liking science, value science and behavioral engagement significantly predicted students' self-concept in science, while at level-2, aggregated classroom science achievement was a significant and negative predictor of self-concept in science. The regression coefficients were presented in Table 3. Furthermore, to investigate the explained variance  $R^2$  was calculated for both level 1 and level 2 by using the equation (1) and (2) respectively:

$$R^2 \text{ at level 1} = 1 - (\sigma^2_{\text{cond}} + \tau^2_{\text{cond}}) / (\sigma^2_{\text{uncond}} + \tau^2_{\text{uncond}}) \quad (1)$$

$$R^2 \text{ at level 2} = 1 - [(\sigma^2_{\text{cond}}/nh) + \tau^2_{\text{cond}}] / [(\sigma^2_{\text{uncond}}/nh) + \tau^2_{\text{uncond}}] \quad (2)$$

All level-1 predictors accounted for approximately 57% of the student level variance in students' self-concept in science. Moreover, aggregated classroom science achievement explained approximately 7% of the between class variance of self-concept. Based on the t values, it can be argued that the best predictor of students' self-concept in learning science was liking science ( $t=25.82, p<.001$ ) which was followed by behavioral engagement ( $t=17.60, p<.001$ ). The equations (3) and (4) represent the final full model as presented below:

$$\gamma_{ij} = \beta_{0j} + \beta_{1j}(\text{Gender}) + \beta_{2j}(\text{Science Achievement}) + \beta_{3j}(\text{Socioeconomic Status}) + \beta_{4j}(\text{Liking Science}) + \beta_{5j}(\text{Value Science}) + \beta_{6j}(\text{Behavioral Engagement}) + r_{ij} \quad (3)$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{Aggregated Classroom Science Achievement}) + u_{0j} \quad (4)$$

**Table 3***Predicting Students' Self-concept in Science*

<b>Students' Confidence in Learning Science</b>			
<b>Predictor</b>	<b>Coefficient</b>	<b>SE</b>	<b>t</b>
<b>Student level</b>			
<b>Intercept, <math>\gamma_{00}</math></b>	9.9225	.0745	133.26*
<b>Gender, <math>\gamma_{10}</math></b>	.2286	.0443	5.17*
<b>Science achievement<sub>within-class</sub>, <math>\gamma_{20}</math></b>	.0046	.0003	15.37*
<b>Socioeconomic status, <math>\gamma_{30}</math></b>	.0674	.0110	6.12*
<b>Liking science, <math>\gamma_{40}</math></b>	.4077	.0158	25.82*
<b>Value science, <math>\gamma_{50}</math></b>	.1413	.0169	8.37*
<b>Behavioral engagement, <math>\gamma_{60}</math></b>	.2869	.0163	17,60*
<b>Class level</b>			
<b>Aggregated science achievement<sub>between-class</sub>, <math>\gamma_{01}</math></b>	-.0042	.0008	-5.61*
<b>R<sup>2</sup> between-class</b>		.07	
<b>R<sup>2</sup> within-class</b>		.57	

\*: p&lt;.001

The equation (5) is used to calculate the effects size of BFLPE (Marsh, Lüdtke, et al. (2009). The effect size was found to be .44 for level 1, and -.24 for level 2.

$$\text{BFLPE ES} = 2 * \beta * \frac{\sigma_p}{\sigma_y} \quad (5)$$

In this equation, while  $\beta$  represents unstandardized regression coefficient,  $\sigma_p$  and  $\sigma_y$  represent standard deviation of predictor and outcome variables, respectively.

### **Discussion**

The present study aimed to investigate: (1) the relation of middle school students' self-concept to student engagement and achievement in science and (2) the Big Fish Little Pond Effect (BFLPE) on students' science self-concept. Additionally, gender and socioeconomic status were included in the study in order to control for their potential effects on students' self-concept. TIMSS 2011 data obtained from Turkish eight grade students were utilized for this purpose and

HLM analysis was conducted. Analysis results showed that at the student level, girls and students with higher socioeconomic status reported higher levels of science self-concept than boys and students with lower socioeconomic status respectively. These findings are consistent with previous research findings (e.g., Maqsd & Rouhani, 1991; Pehlivan & Köseoğlu, 2010). Furthermore, science achievement was a significant and positive predictor of self-concept. This finding indicates that students with high performance in science tend to have more positive beliefs about their abilities to perform well in science. The positive relationship between academic self-concept and student achievement was anticipated as previous research also linked higher self-concept with higher achievement at the student level (e.g., Marsh, 2004; Marsh et al., 2014).

Two dimensions of engagement, behavioral engagement and motivational engagement, were included in the present study. Students' attendance to the task, making an effort to complete the task or persistence on the task were treated as students' behavioral engagement. The results suggested that behavioral engagement was a significant and positive predictor of students' self-concept in science. This is an important finding because previous research generally investigated the relationship in the opposite direction; students with positive judgments about their competence (i.e., have high self-concept) tend to be more engaged in behaviorally than individuals with low levels of self-concept beliefs (e.g., Pintrich & Schunk, 2002). Although a number of researchers suggest that behavioral engagement can also affect students' self-concept (Badiozaman, 2012; Bakadorova & Raufelder, 2017), according to the authors' knowledge, there are not many studies investigating students' behavioral engagement as a predictor of self-concept beliefs. The current study suggests that students, who attend to their classes and make an effort to complete the task or persist on the task, tend to have higher levels of self-concept beliefs than

others. Actually, the predictive effect of behavioral engagement on students' self-concept is rational, since students' self-concept beliefs can occur based on their past experiences (e.g. Wigfield & Eccles, 2000). Hence, students' attendance, persistence or effort in science may support their perceptions about their ability in science. However, further studies can examine the predictive effect of students' behavioral engagement on their self-concept beliefs in detail. Besides that, student's value in science and liking science were treated as trajectories of motivational engagement in this study. The results suggested that students who had high interest in science and who had high levels of science value tend to have higher levels of self-concept beliefs. In other words, both of the dimensions of the motivational engagement had significant predictive effects on students' self-concept beliefs. This was an expected result since Salanova et al. (2011) underline the reciprocal relation between motivational engagement and self-concept. In this sense, students who have high self-concept beliefs are those who also have high interest and value, or vice versa. The findings of the present study contribute to the relevant literature; although research has generally linked behavioral engagement to self-concept, in the present study behavioral engagement predicted self-concept.

The second purpose of this study was to examine the BFLPE on Turkish students' self-concept beliefs over and above gender, socioeconomic status, science achievement, and engagement dimensions of interest. Class average science achievement was entered in the model as a class level predictor. Analysis results showed that class average science achievement was a negative and significant predictor of science self-concept. Thus, students reported lower self-concept when classmates' achievement was high, and higher self-concept were reported when classmates' achievement was low. This finding implies social comparison effects on students' perceptions of their ability to perform well in science. Students may compare their achievement

with their classmates and if other students' achievement is high, they may question their competence which may have an unfavorable effect on their self-perceived ability. This negative effect is in line with previous research which also has revealed a negative predictive effect of average class achievement on self-concept (e.g., Marsh et al., 2008; Seaton et al., 2010). Therefore, findings of the current study are consistent with previous research which was mostly conducted in Western countries. To our knowledge, no study has investigated BFLPE with Turkish sample and findings support that BFLPE also exists in the Turkish context. Our interpretation is that examination oriented Turkish education system and prevalence of evaluative feedback provided by teachers may have contributed to such findings. Students are frequently given tests and their performance on those tests is highlighted in the classrooms. Both teachers and families emphasize the importance of getting high scores on tests since eighth grade students take a national examination which determines their placement in high schools. Furthermore, teachers' provision of evaluative feedback may help students focus on their strengths and weaknesses. In Turkey, teachers mostly give evaluative oral and written feedback where, for example, they make judgements about their students' performance (Odabaşı Çimer et al., 2010). Supporting this idea, Marsh et al. (2015) investigated BFLPE for different countries and suggested that it was smaller for Middle Eastern Islamic countries than Asian and Western countries. The researchers also discussed that the students of these countries don't receive evaluative feedback as much as the students of Asians or Westerns do. Hence, in a further study, investigating the relation between BFLPE and feedback can be useful for the relevant literature.

Based on the findings of the present study, we suggest that science teachers may try to support students' motivational and behavioral engagement which in turn may increase self-concept beliefs in science. For instance, teachers can assign different roles which foster them to

engage in the task. In other words, teachers can encourage their students to engage behaviorally in science courses. Furthermore, gaining further understanding about how contextual factors (i.e., average class achievement) affect students' self-concept beliefs may be useful to create more supportive learning environments for students. Rather than making comparisons among students' achievement, which may focus students' attention to classmates' performance, self-improvement may be emphasized by educators. Rather than concentrating on performance, appreciating student effort during learning process may help the creation of more supportive learning environments. Teacher's provisioning of less evaluative feedback but more informative feedback may also encourage students to focus on their own progress.

Besides its contribution to the field, this study has several limitations too. First, the data are limited to TIMSS 2011 Turkey study. This data set is important for Turkey because, it is potential to reflect the results of the first attempts of the constructivist approach which has firstly form the basis of the science education curriculum in 2005 (Ministry of Education, 2005). Although there are several attempts of researchers to figure out the results of the new science curriculum by considering various variables since 2005, these studies are generally lack in leading or providing basis for longitudinal studies. Additionally, longitudinal investigation of engagement and BFLPE will provide new perspective about the development of students' self-concept in the classrooms. By taking attention to this starting point, this research has the potential to initiate the longitudinal investigation of the role of engagement and BFLPE in self-concept in science education by investigating the same subjects from TIMSS Turkey studies being conducted in the following years. The second limitation of this study is that although student engagement includes cognitive engagement dimension, because of the TIMSS data set content, liking science and value science were examined as predictors of self-concept beliefs in

science. Lastly, the present study limited to 8<sup>th</sup> grade students. Although TIMSS study mostly surveyed 4<sup>th</sup> and 8<sup>th</sup> graders for same variables consistently, survey of 4<sup>th</sup> graders did not include value science variable. Therefore, 4<sup>th</sup> graders data were not sufficient to represent the motivational engagement dimension. Therefore, when evaluating the findings of this study, these issues should be considered.

## References

- Areepattamannil, S., Freeman, J. G., & Klinger, D. A. (2011). Influence of motivation, self-beliefs, and instructional practices on science achievement of adolescents in Canada. *Social Psychology of Education: An International Journal*, 14(2), 233–259. <https://doi.org/10.1007/s11218-010-9144-9>
- Badiozaman, A. (2012). *A study on the relationship between Malaysian learners' self-concept in academic writing and their engagement in one higher learning institution* [Unpublished Doctoral dissertation], Massey University, Palmerston North, New Zealand.
- Bandalos, D. L., Yates, K., & Thorndike-Christ, T. (1995). Effects of math self-concept, perceived self-efficacy, and attributions for failure and success on test anxiety. *Journal of Educational Psychology*, 87(4), 611–623. <https://doi.org/10.1037/0022-0663.87.4.611>
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Bae, C. L., & DeBusk-Lane, M. (2019). Middle school engagement profiles: Implications for motivation and achievement in science. *Learning and Individual Differences*, 74, 101753. <https://doi.org/10.1016/j.lindif.2019.101753>
- Britner, S. L. (2008). Motivation in high school science students: A comparison of gender differences in life, physical, and earth science class. *Journal of Research in Science Teaching*, 45, 955-970. <https://doi.org/10.1002/tea.20249>
- Britner, S. L., & Pajares, F. (2001). Self-efficacy beliefs, motivation, race, and gender in middle school science. *Journal of Women and Minorities in Science and Engineering*, 7(4), 271-285. [Doi:10.1615/JWomenMinorScienEng.v7.i4.10](https://doi.org/10.1615/JWomenMinorScienEng.v7.i4.10)

- Britner, S. L., & Pajares, F. (2006). Sources of science self-efficacy beliefs of middle school students. *Journal of Research in Science Teaching*, 43, 485-499.  
<https://doi.org/10.1002/tea.20131>
- Bong, M., Cho, C., Ahn, H. S., & Kim, H. J. (2012). Comparison of self-beliefs for predicting student motivation and achievement. *The Journal of Educational Research*, 105, 336-352. <https://doi.org/10.1080/00220671.2011.627401>
- Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, 15(1), 1-40.  
<https://doi.org/10.1023/A:1021302408382>
- Buhs, E. S. (2005). Peer rejection, negative peer treatment, and school adjustment: Self-concept and classroom engagement as mediating processes. *Journal of School Psychology*, 43(5), 407-424. <https://doi.org/10.1016/j.jsp.2005.09.001>
- Chen, J. A., & Usher, E. L. (2013). Profiles of the sources of science self-efficacy. *Learning and Individual Differences*, 24, 11-21. <https://doi.org/10.1016/j.lindif.2012.11.002>
- Cheung, D. (2018). The key factors affecting students' individual interest in school science lessons. *International Journal of Science Education*, 40(1), 1-23.  
<https://doi.org/10.1080/09500693.2017.1362711>
- Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy, and relatedness: A motivational analysis of self-system processes. In M. R. Gunnar & L. A. Sroufe (Eds.), *Self processes and development* (pp. 43–77). Lawrence Erlbaum Associates, Inc.

- Cooper, K. M., Krieg, A., & Brownell, S. E. (2018). Who perceives they are smarter? Exploring the influence of student characteristics on student academic self-concept in physiology. *Advances in Physiology Education, 42*(2), 200-208.  
<https://doi.org/10.1152/advan.00085.2017>
- Easterbrook, M. J., Kuppens, T., & Manstead, A. S. (2020). Socioeconomic status and the structure of the self-concept. *British Journal of Social Psychology, 59*(1), 66-86.  
<https://doi.org/10.1111/bjso.12334>
- Eccles, J. S. (2005). Influences of parents' education on their children's educational attainments: The role of parent and child perceptions. *London Review of Education, 3*, 191-204.
- Eccles, J. S., & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin, 21*(3), 215-225. <https://doi.org/10.1177/0146167295213003>
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*(1), 109-132
- Ferla, J., Valcke, M., & Cai, Y. (2009). Academic self-efficacy and academic self-concept: Reconsidering structural relationships. *Learning and Individual Differences, 19*, 499-405.  
<https://doi.org/10.1016/j.lindif.2009.05.004>
- Finn, J. D., Pannozzo, G. M., & Voelkl, K. E. (1995). Disruptive and inattentive- withdrawn behavior and achievement among fourth graders. *Elementary School Journal, 95*, 421–454. <http://www.jstor.org/stable/1001656>
- Fraenkel, J. R., & Wallen, N. E. (2006). *How to design and evaluate research in education* (6th ed.). New York, NY: McGraw-Hill.

- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59-109.  
<https://doi.org/10.3102/00346543074001059>
- Green, J., Liem, G. A. D., Martin, A. J., Colmar, S., Marsh, H. W., & McInerney, D. (2012). Academic motivation, self-concept, engagement, and performance in high school: Key processes from a longitudinal perspective. *Journal of Adolescence*, 35(5), 1111-1122.  
<https://doi.org/10.1016/j.adolescence.2012.02.016>
- Guo, J., Marsh, H. W., Parker, P. D., & Dicke, T. (2018). Cross-cultural generalizability of social and dimensional comparison effects on reading, math, and science self-concepts for primary school students using the combined PIRLS and TIMSS data. *Learning and Instruction*, 58, 210-219. <https://doi.org/10.1016/j.learninstruc.2018.07.007>
- Guvercin, O., Tekkaya, C., & Sungur, S. (2010). A cross age study of elementary students' motivation towards science learning. *Hacettepe University Journal of Education*, 39, 233-243.
- Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). Routledge.
- Jansen, M., Scherer, R., & Schroeders, U. (2015). Students' self-concept and self-efficacy in the sciences: Differential relations to antecedents and educational outcomes. *Contemporary Educational Psychology*, 41, 13-24. <https://doi.org/10.1016/j.cedpsych.2014.11.002>
- Jansen, M., Schroeders, U., & Lüdtke, O. (2014). Academic self-concept in science: Multidimensionality, relations to achievement measures, and gender differences. *Learning and Individual Differences*, 30, 11-21.  
<https://doi.org/10.1016/j.lindif.2013.12.003>

- Jimerson, S. R., Campos, E., & Greif, J. L. (2003). Toward an understanding of definitions and measures of school engagement and related terms. *The California School Psychologist*, 8(1), 7-27. <https://doi.org/10.1007/BF03340893>
- Joncas, M., & Foy, P. (2012). Sample design in TIMSS and PIRLS. Methods and procedures. TIMSS and PIRLS International Study Center: Lynch School of Education, Boston College. Retrieved on August 19th, 2015 from [http://timssandpirls.bc.edu/methods/pdf/TP\\_Sampling\\_Design.pdf](http://timssandpirls.bc.edu/methods/pdf/TP_Sampling_Design.pdf).
- Karaarslan, G., & Sungur, S. (2011). Elementary students' self-efficacy beliefs in science: Role of grade level, gender, and socio-economic status. *Science Education International*, 22(1), 72-79.
- Kung, H. Y. (2009). Perception or confidence? Self-concept, self-efficacy and achievement in mathematics: A longitudinal study. *Policy Futures in Education*, 7(4), 387-398. <https://doi.org/10.2304/pfie.2009.7.4.387>
- Linnenbrink, E. A., & Pintrich, P. R. (2003). The role of self-efficacy beliefs in student engagement and learning in the classroom. *Reading & Writing Quarterly*, 19(2), 119-137. <https://doi.org/10.1080/10573560308223>
- Lorah, J. (2018). Effect size measures for multilevel models: definition, interpretation, and TIMSS example. *Large-scale Assessments in Education*, 6, 8. <https://doi.org/10.1186/s40536-018-0061-2>
- Maqsd, M., & Rouhani, S. (1991). Relationships between socioeconomic status, locus of control, self-concept, and academic achievement of Batswana adolescents. *Journal of Youth and Adolescence*, 20(1), <https://doi.org/10.1007/BF01537354>

- Marsh, H. W. (2004). Negative effects of school-average achievement on academic self-concept: A comparison of the big-fish-little-pond effect across Australian states and territories. *Australian Journal of Education, 48*(1), 5-26.  
<https://doi.org/10.1177/000494410404800102>
- Marsh, H. W., Abduljabbar, A. S., Morin, A. J., Parker, P., Abdelfattah, F., Nagengast, B., & Abu-Hilal, M. M. (2015). The big-fish-little-pond effect: Generalizability of social comparison processes over two age cohorts from Western, Asian, and Middle Eastern Islamic countries. *Journal of Educational Psychology, 107*(1), 258-271.  
<https://doi.org/10.1037/a0037485>
- Marsh H. W., Abduljabbar, A. S., Parker, P. D., Morin, A. J. S., Abdelfattah, F., & Nagengast, F. (2014). The big-fish-little-pond effect in mathematics: A cross-cultural comparison of U.S. and Saudi Arabian TIMSS responses. *Journal of Cross-Cultural Psychology, 45*(5), 777-804. <https://doi.org/10.1177/0022022113519858>
- Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthen, B., & Nagengast, B. (2009). Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioral Research, 44*, 764-802.  
<https://doi.org/10.1080/00273170903333665>
- Marsh, H. W., Seaton, M., Trautwein, U., Lüdtke, O., Hau, K. T., O'Mara A. J., & Craven, R. G. (2008a). The big-fish–little-pond-effect stands up to critical scrutiny: Implications for theory, methodology, and future research. *Educational Psychological Review, 20*, 319-350. <https://doi.org/10.1007/s10648-008-9075-6>

- Marsh, H. W., Trautwein, U., Lüdtke, O., & Koller, O. (2008b). Social comparison and big-fish–little-pond-effect on self-concept and other self-belief constructs: Role of generalized and specific others. *Journal of Educational Psychology*, 100, 510-524.  
<https://doi.org/10.1037/0022-0663.100.3.510>
- Martin, M.O. & Mullis, I.V.S. (Eds.). (2012). Methods and procedures in TIMSS and PIRLS 2011. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mau, W. C. (2003). Factors that influence persistence in science and engineering career aspirations. *The Career Development Quarterly*, 51, 234-243. [doi:10.1002/j.2161-0045.2003.tb00604.x](https://doi.org/10.1002/j.2161-0045.2003.tb00604.x)
- Ministry of Education (2005). *İlköğretim fen ve teknoloji dersi (6, 7 ve 8. sınıflar) öğretim programı [Elementary education science and technology course (6th, 7th and 8th grades) curriculum]*, Ankara.
- Odabaşı Çimer, S., Bütüner, S. Ö., & Yiğit, N. (2010). Öğretmenlerin öğrencilerine verdikleri dönütlerin tiplerinin ve niteliklerinin incelenmesi [An investigation of the feedback types and qualities provided by teachers to their students]. *Uludağ Üniversitesi Eğitim Fakültesi Dergisi*, 23(2), 517-538. Retrieved from  
<https://dergipark.org.tr/tr/pub/uefad/issue/16692/173496>
- Oral, I., & McGivney, E. (2013). Türkiye’de matematik ve fen bilimleri alanlarında öğrenci performansı ve başarının belirleyicileri [The predictors of student performance in mathematics and science subjects in Turkey]. Analysis report retrieved on August 23th, 2015 from <http://erg.sabanciuniv.edu/sites/erg.sabanciuniv.edu/files/ERG%20-TIMSS%202011%20Analiz%20Raporu-03.09.2013.pdf>

- Özdamar, K. (1997). Paket Programlar ile istatistiksel Veri Analizi I. Anadolu Üniversitesi Fen Fakültesi Yayınları, Eskişehir.
- Pajares, F., & Schunk, D. H. (2002). Self and self-belief in psychology and education: A historical perspective. In J. Aronson (Ed.), *Improving Academic Achievement* (pp. 1-31). Academic Press.
- Pajares, M. F., & Johnson, M. J. (1993). Confidence and competence in writing: the role of self-efficacy, outcome expectancy, and apprehension. Paper presented at the Annual Meeting of the American Educational Research Association, April 1993, Atlanta, GA.
- Parker, P., Dicke, T., Guo, J., Basarkod, G., & Marsh, H. (2021). Ability stratification predicts the size of the big-fish-little-pond effect. *Educational Researcher*, 50(6), 334-344.  
<https://doi.org/10.3102/0013189X20986176>
- Pehlivan, H., & Köseoğlu, P. (2010). Ankara fen lisesi öğrencilerinin biyoloji dersine yönelik tutumları ile akademik benlik tasarımları [Ankara science high school students' attitudes towards biology and their depiction of academic self-concept]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 38, 225-235.
- Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, 33-40. <https://doi.org/10.1037/0022-0663.82.1.33>
- Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education: Theory, research, and practice*. Englewood Cliffs, NJ: Merrill.

- Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). A Manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ). National Center for Research to Improve Postsecondary Teaching and Learning., The University of Michigan.
- Raudenbush, S. W., Bryk, A., & Congdon, R. (2005). HLM 6. Lincolnwood, IL: Scientific Software International.
- Raufelder, D., Sahabandu, D., Martínez, G. S., & Escobar, V. (2015). The mediating role of social relationships in the association of adolescents' individual school self-concept and their school engagement, belonging and helplessness in school. *Educational Psychology*, 35(2), 137-157. <https://doi.org/10.1080/01443410.2013.849327>
- Salanova, M., Llorens, S., & Schaufeli, W. B. (2011). “Yes, I can, I feel good, and I just do it!” On gain cycles and spirals of efficacy beliefs, affect, and engagement. *Applied Psychology*, 60(2), 255-285. <https://doi.org/10.1111/j.1464-0597.2010.00435.x>
- Schnitzler, K., Holzberger, D., & Seidel, T. (2021). All better than being disengaged: Student engagement patterns and their relations to academic self-concept and achievement. *European Journal of Psychology of Education*, 36(3), 627-652. <https://doi.org/10.1007/s10212-020-00500-6>
- Schütte, K. (2015). Science self-concept and valuing science: a cross-cultural analysis of their relation among students from Western and East Asian countries. *Social Psychology of Education*, 18(4), 635-652. <https://doi.org/10.1007/s11218-015-9311-0>
- Seaton, M., Marsh, H. W., & Craven, R. G. (2010). Big-fish-little-pond effect: Generalizability and moderation-two sides of the same coin. *American Educational Research Journal*, 47(2), 390-433. <https://doi.org/10.3102/0002831209350493>

Senler, B., & Sungur, S. (2009). Parental influences on students' self-concept, task value beliefs, and achievement in science. *Spanish Journal of Psychology, 12*, 106-117.

[10.1017/s1138741600001529](https://doi.org/10.1017/s1138741600001529)

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46*(3), 407-441.

<https://doi.org/10.2307/1170010>

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychology, 50*, 1–13.

<https://doi.org/10.1080/00461520.2014.1002924>

Suárez-Álvarez, J., Fernández-Alonso, R., & Muñoz, J. (2014). Self-concept, motivation, expectations, and socioeconomic level as predictors of academic performance in mathematics. *Learning and Individual Differences, 30*, 118-123.

<https://doi.org/10.1016/j.lindif.2013.10.019>

Trautwein, U., Lüdtke, O., Marsh, H. W., & Nagy, G. (2009). Within-school social comparison: How students perceive the standing of their class predicts academic self-concept. *Journal of Educational Psychology, 101*(4), 853–866. <https://doi.org/10.1037/a0016306>

Wang, J., & Staver, J. R. (1997). An empirical study of gender differences in Chinese students' science achievement. *The Journal of Educational Research, 90*(4), 252-255.

<https://doi.org/10.1080/00220671.1997.10544579>

Wang, M. T., & Eccles, J. S. (2013). School context, achievement motivation, and academic engagement: A longitudinal study of school engagement using a multidimensional perspective. *Learning and Instruction, 28*, 12-23.

<https://doi.org/10.1016/j.learninstruc.2013.04.002>

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation.

*Contemporary Educational Psychology*, 25, 68–81.

<https://doi.org/10.1006/ceps.1999.1015>

Wilkins, J. L. (2004). Mathematics and science self-concept: An international investigation. *The*

*Journal of Experimental Education*, 72(4), 331-346.

<https://doi.org/10.3200/JEXE.72.4.331-346>

Yalvac, B., Tekkaya, C., Cakiroglu, J., & Kahyaoglu, E. (2007). Turkish pre-service science

teachers' views on science-technology-society issues. *International Journal of Science*

*Education*, 29, 331-348. <https://doi.org/10.1080/09500690600708667>

Yoshino, A. (2012). The relationship between self-concept and achievement in TIMSS 2007: A

comparison between American and Japanese students. *International Review of Education*,

58, 199–219. <https://doi.org/10.1007/s11159-012-9283-7>

**EEJ<sup>®</sup>**