

## Research Article

# The impact of high-stakes testing on the teaching and learning processes of mathematics

Cennet Göloğlu Demir<sup>1</sup> and Özlem Kaplan Keleş<sup>2</sup>

<sup>1</sup>Bandırma Onyedi Eylül University, Turkey (ORCID: 0000-0002-8770-6107)

<sup>2</sup>Ministry of National Education, Turkey (ORCID: 0000-0002-5028-7850)

The impacts of high-stakes testing administered within the framework of accountability policies concern all the stakeholders in the field of education, particularly the policymakers. As such, in the present study, which is specific to the mathematics course, the aim was to examine the impacts of high-stakes testing on teaching and learning processes, measurement and evaluation, communication, and motivation based on teacher perceptions. The study employed one of the qualitative research designs, namely phenomenology. The participants of the study were comprised of 13 middle school mathematics teachers identified via the maximum sampling technique. The data in the study were collected through the semi-structured interview technique and analyzed by using the content analysis method. The results of the study revealed that the teachers were inclined to use the traditional teaching method and solve multiple-choice tests, that the questions they used for the measurement of performance resembled those in high-stakes testing, and that the motivation of both teachers and the students were negatively affected. These findings were all similar to those frequently reported in the literature. Moreover, it was revealed that high-stakes testing prevented the consideration of individual differences and the teaching of values, that high-stakes testing was the focal point in communication, and that the impacts of high-risk testing showed a widespread impact upon different grades ranging from grade 8, during which the testing was done, down onto lower grades. Within the scope of these results, policymakers were recommended to take action toward improving the teaching process by acting together with the teachers to minimize the negative impacts of high-stakes testing on the process.

Keywords: Education policy; High-stakes testing; Teaching experience; Educational assessment; Motivation; Communication

Article History: Submitted 24 January 2021; Revised 29 April 2021; Published online 4 June 2021

## 1. Introduction

Assessment and accountability policies regarding high-stakes testing continue to affect many people and institutions, such as the education process, students, teachers, parents, and politicians, briefly education stakeholders. If the results of a test are used to make important decisions that affect students, teachers, administrators, and schools, then these tests are characterized as high-stakes tests (Madaus, 1988). Johnson et al. (2008) explained that a test or a test program would be high-stakes if the results of the test were used for failing or passing a student onto a higher grade, for enabling or preventing the graduation of a student from school, for teachers' accountability, for

---

### Address of Corresponding Author

Cennet Göloğlu Demir, PhD, Bandırma Onyedi Eylül University, Faculty of Health Sciences, 10200, Bandırma/Balıkesir, Turkey.

✉ [gologlu.cennet@gmail.com](mailto:gologlu.cennet@gmail.com)

**How to cite:** Göloğlu-Demir, C., & Kaplan-Keleş, Ö. (2021). The impact of high-stakes testing on the teaching and learning processes of mathematics. *Journal of Pedagogical Research*, 5(2), 119-137. <http://dx.doi.org/10.33902/JPR.2021269677>

establishing school images and for managing funds. Today, these kinds of tests are administered with the purpose of assessment under different names, such as “high-stakes testing” or “standardized testing” (Zucker, 2003). Thus, for tests that are administered nationwide in a standardized way, the concept of “high-stakes testing” is used in the related literature (Hamilton et al., 2002; Jones et al., 2003; Kumandaş & Kutlu, 2010). High-stakes tests are considerably effective on educational policies. Breakspear (2012) examined the impact the programme for international student assessment (PISA) had on the educational reform policies and practices across various countries. In research, respondents were asked how influential specific PISA results and analyses have been in informing the policy-making process in their country. The respondents from England-UK, Denmark, and Japan rated PISA as ‘extremely’ influential, whilst respondents from Finland, France, Indonesia, Luxembourg, and Turkey rated PISA as ‘not very’ influential. In the context of national high-stakes testing, policymakers use national high-stakes tests to change the behavior of teachers and students in desirable ways (Amrein & Berliner, 2002). In addition to this, policymakers can use test scores to spend educational resources to low achieving schools or geographic areas and to support cross-national comparisons which can lead governments to commit a larger share of the national budget to education (Chapman & Snyder, 2000).

High-stakes testing is advocated based on arguments that it motivates teachers to teach more effectively, and directs students to study harder and to learn more effectively. Moreover, the advocates of high-stakes testing believe that test results contribute to the improvement of teaching and the professional development of teachers (Amrein & Berliner, 2002). In contrast to these arguments, researchers present numerous views and proof regarding the negative impacts of high-stakes testing.

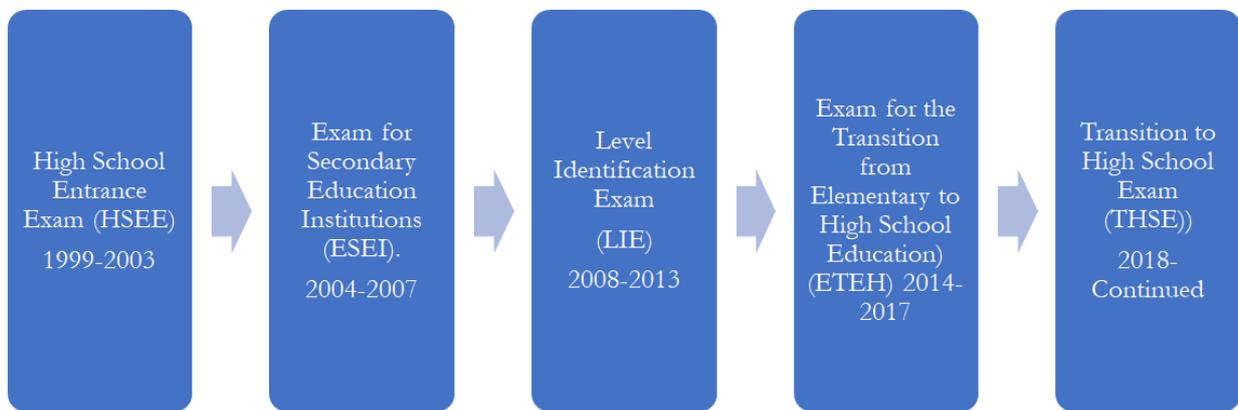
High-stakes testing impacts the teaching practices of teachers (Madaus, 1988; Smith, 1991). Jonsson and Leden (2019) stated that high-stakes testing affected teachers’ pedagogical practices and gave some messages about what to teach and how to assess. The impact of these tests on teaching and learning are labeled as “washback”. At the macro level, washback is related to the degree at which a test affects government policies, school directors, and parents’ expectations from their parents. On the other hand, at the micro-level, washback refers to the degree at which a test affects students’ learning performances and the teaching method and techniques of teachers in class (Chapman & Snyder, 2000). Alderson and Wall (1993) define washback as things that teachers and students would not normally do but do so owing to the testing. Hence, it is important to reveal the impact of high-stakes testing with respect to washback. The results of a meta-synthesis study, which was conducted by Au (2007) and based on 49 qualitative studies in relation to this topic, revealed that the primary impact of high-stakes testing was basing the teaching points on tests and teachers being more inclined to adopt teacher-centered teaching. Only in very few studies was an adverse condition, specifically student-centered and collaborative practices, identified. The results of a study conducted by Klein et al. (2006) with teachers at 20 schools revealed that high-stakes testing affected the teaching process, i.e. teachers engaged in test-oriented teaching. Barnes (2005) stated that instead of increasing the quality of education, high-stakes testing caused students to focus on the parts, which would make students successful in tests and ignore the other parts in the course syllabus. In a study by Zhao et al. (2016), it was revealed that high-stakes testing had a ferocious control over learning and teaching and students in the class were passive learners and memorizers of facts. In studies where the impacts of high-stakes testing were investigated, it was found that high-stakes testing increased students’ levels of stress and anxiety, while it decreased their levels of intrinsic motivation and self-oriented learning (Ashadi & Rice, 2016). Similarly, Thompson (2013) stated that high-stakes testing increased students’ anxiety, stress, and the pressure they felt upon them. Researchers reported that high-stakes testing was also a source of pressure for teachers (Barnes, 2005; Buyruk, 2014; Gündoğdu et al., 2010; Looney, 2009; Popham, 2001), increased teachers’ levels of anxiety and decreased their mood and motivation (Abrams, 2004). In contrast to these negative impacts, Cizek (2001) highlighted that high-stakes testing had a positive impact on teachers as it supported their

professional development. These tests do not just affect learners and teachers but affect parents. In research studies conducted in the literature so far, it was revealed that due to high-stakes testing, parents imposed pressure on their children, on teachers, and the school (Can, 2017; Şahin et al. 2012, TEDMEM, 2014; Thompson, 2013). Similarly, in a survey study carried out by Jones and Egley (2004) with 708 teachers, it was found that high-stakes testing had negative impacts on the teaching program, teaching, learning, and the motivation of both learners and teachers.

Despite the adverse aspects of and many arguments related to high-stakes testing, it is widely utilized in many developed and developing countries, such as England, the United States of America, Australia, Sweden, Indonesia, and Chile. High-stakes testing has been implemented in Turkey as well, particularly in the last 20 years. High-stakes testing plays an important role in Turkey as it is a highly populated country (Ekiz, 2019). Students are administered centralized tests nationwide by the Ministry of Education for admission into different types of high schools in Turkey. The scope, the difficulty level, the quality of the items, the assessment procedures, and the method of administration of these tests have changed five times in the last 20 years. Figure 1 displays the chronological timeline of the history of the high school entrance tests in Turkey.

Figure 1

*High-Stakes Testing in the Transition from Middle School to High School in Turkey in the last 20 years*



As of the end of the year 1999, a centralized High School Entrance Exam (HSEE) started to be administered once a year at the end of grade 8 in Turkey. In 2004, it continued to be administered under the name Exam for Secondary Education Institutions (ESEI). ESEI was criticized based on such justifications as students' levels of interest in school learning decreased, 8th-grade students were more inclined to take private lessons and courses, and students found the lessons that were not within the scope of ESEI unimportant (Gür et al., 2013). Owing to these criticisms, in the year 2007, in addition to the ESEI score, the elementary grade point average was also taken into consideration in the calculation of the high school entrance score. Because ESEI was a test solely administered in grade 8 and that students were not interested in grade 6 and grade 7 lessons, in the year 2008, ESEI was changed into a test referred to as LIE (Level Identification Exam), which was administered in grades 6, 7 and 8, rather than solely in grade 8. The high school placement score in the LIE administration was calculated based on the LIE score, elementary GPA, and the overall attitude score with the following weightings: 70%, 25%, and 5%, respectively. However, the practice of considering an overall attitude score in the calculation of the high school placement score was stopped by the State Council in the year 2010 (Atılgan, 2018; Gür et al., 2013). One of the greatest criticisms made toward LIE was that while students were attending private lessons and courses only during grade 8 in the years when the former exams were administered, with the onset of LIE, students started to attend private lessons and courses throughout grades 6, 7, and 8.

Since LIE was claimed to have increased inequality in opportunity, changes were initially made in the LIE during the year 2009, and then its administration was discontinued altogether in the year 2013. Instead, the ETEH (Exam for the Transition from Elementary to High School Education)

was introduced. ETEH was a practice whereby one of the school exams was administered as a centralized and multiple-choice test each term. The use of ETEH for student selection was also criticized based on claims that students developed psychological problems stemming from having to sit a test each term in grade 8 and that the test yielded many first place scores (Atılgan, 2018; Bađcı, 2016; Education Reform Initiative, 2013; Directorate General for Secondary Education, 2016). Finally, with a decision taken in the year 2017, the administration of ETEH was discontinued. Instead, the Transition to High School Exam (THSE) was introduced (Atılgan, 2018). The most prominent feature that has made this test different from the others was the quality of the questions. The questions are in line with the learning outcomes of the education program, but at the same time require higher-order thinking skills as only 10% of the students can be placed in high schools with a THSE score. The students who cannot enter a high school with a THSE score are placed in schools via a system based on the residential address. Accordingly, even though politicians have announced that not all students need to sit the THSE, almost all the students take the test. During the administration of THSE, between the years 2003 and 2020, the minister of national education has changed six times, all being appointed by the same government. Hence, it is believed that populist approaches have been effective in the expression of positive and negative aspects of changes rather than scientific data (Güler et al., 2019). The high number of students and the limited number of quotas in high-quality schools make THSE, which is administered with the purpose of selecting and placing students in a higher educational institution, important for administrators, teachers, students, and parents. In all the above-mentioned tests, mathematics has been the most fundamental and important subject. Parents have their children take private lessons, particularly in mathematics.

A literature review of related research studies has shown that there is a limited number of studies conducted on the impact of high-stakes testing on the mathematics teaching process. The present study possesses the potential of presenting significant results regarding the impacts of high-stakes testing on the learning-teaching process, measurement and assessment, communication, and motivation based on the perceptions of mathematics teachers. Moreover, the results of the present study will contribute to assessment policies and to the decision-makers in Turkey and other countries where there are similar practices with respect to strengthening the positive impacts of high-stakes testing on the teaching process and eliminating the negative impacts.

### **1.1. The Aim**

The aim of the present research study was to investigate the impacts of high-stakes testing on middle school mathematics education based on teacher perceptions. Thus, the answers to the following research questions were sought:

- What is the impact of high-stakes testing on the learning-teaching process?
- What is the impact of high-stakes testing on measurement and assessment?
- What is the impact of high-stakes testing on learner and teacher motivation?
- What is the impact of high-stakes testing on communication?

## **2. Method**

In the present study, the method of phenomenology, one of the qualitative research designs, was employed. In phenomenology, the phenomena that we are familiar with but do not have profound and in-depth information on are interpreted and described based on individuals' experiences (Merriam, 2009). As stated by Patton (2002), phenomenology studies focus on how people perceive the phenomenon in question, how they describe it, how they feel about it, how they judge it, how they remember it, what sense they make of it, and what they talk about it with other people. In the present study, the phenomenological research design was employed with the aim of identifying the impacts of high-stakes testing on the mathematics teaching process from the perspective of teachers' perceptions.

## 2.1. Participants

The participants of the study were selected by means of the maximum variation sampling method. In this method, which is used within the scope of purposive sampling, initially, the characteristic features or criteria are specified (Patton, 2002). In the selection of the teachers, the first criterion that was used as the teachers' having teaching experience in grade eight (when the high school entrance test is administered). Subsequently, the teachers were selected from different regions, the aim of which was to provide variety within the scope of the research. The variety was also sought in terms of location of work, namely in either the city center or in regional areas, and in terms of teaching experiences. In qualitative studies, there are no firm rules regarding the number of participants needed in a research study; it can vary depending on the scope of the research (Patton, 2002). In the present study, "the participants were continuously approached until information redundancy was reached" (Lincoln & Guba, 1985). Thus, the interviews were held with 13 teachers, the characteristics of whom are presented in Table 1.

Table 1

*The Characteristics of the Participants*

<i>Participant</i>	<i>Sex</i>	<i>Professional Experience (years)</i>	<i>Location of the School</i>	<i>Region</i>
Teacher T1	Female	12	City Center	Marmara
Teacher T2	Female	5	Regional	Central Anatolia
Teacher T3	Male	8	Regional	Central Anatolia
Teacher T4	Female	17	Regional	Eastern Anatolia
Teacher T5	Male	14	City Center	Central Anatolia
Teacher T6	Female	5	Regional	Aegean
Teacher T7	Female	7	Regional	Mediterranean
Teacher T8	Female	6	Regional	Eastern Anatolia
Teacher T9	Male	14	City Center	Marmara
Teacher T10	Female	10	City Center	Mediterranean
Teacher T11	Male	11	City Center	Central Anatolia
Teacher T12	Female	22	City Center	Marmara
Teacher T13	Male	18	City Center	Black Sea

Eight of the participants were female, whereas five were male, as shown in Table 1. In terms of professional experience, five of the participants had between 5 and 9 years of experience, five of them had between 10 and 14 years of experience, and three of them had 15 or more years of experience. The teachers worked in 6 different regions in Turkey. Seven of the teachers worked in the city center, whereas six of them worked in regional areas (town or village).

## 2.2. Data Collection Tool

The research data were collected via a semi-structured interview protocol. The semi-structured interview is employed when the questions are flexible in terms of not being based on predetermined words and order (Merriam, 2009). During the development of the semi-structured interview protocols, the views of two experts on education programs and instruction, one expert on measurement and assessment, and two mathematics teachers holding master's degrees were consulted. After the essential corrections were made in line with the experts' views, finally a Turkish teacher was asked to check the language used in the interview protocol. The draft interview protocol that was prepared was used to conduct an interview with two mathematics teachers. Both of the teachers were found to have understood the questions in the same way. The teachers stated that they did not experience difficulty in answering the questions. Subsequent to the pilot implementation, the semi-structured interview protocol consisting of five questions based on the aim and the sub-aims of the study was finalized. In addition, alternative and follow-up

questions were placed in the protocol to help the participants understand the questions and provide more in-depth answers (Maykut & Morehouse, 2005).

### 2.2.1. Interview Procedures

Semi-structured interviews were held in an online environment. Prior to the interviews, the participants were e-mailed an information document, on which the aim of the research and how the interviews would be held were written. The participants were asked to sign and send back a permission form. The interviews lasted approximately 20-25 minutes. The questions were as follows:

1. What are the impacts of the high school entrance test on the learning-teaching process in your class?
2. What are the impacts of the high school entrance test on your approach toward measurement and assessment?
3. What do you think about the impacts of the high school entrance test on student motivation?
4. What do you think about the impacts of the high school entrance test on your motivation?
5. What are the impacts of the high school entrance test on communication?

### 2.3. Data Analysis

The data of the study were analyzed using the content analysis technique. Accordingly, an in-depth analysis of the data was conducted by following the stages of coding the data, finding the categories, organizing the data in accordance with the codes and categories, and interpreting the findings (Corbin & Strauss, 2015; Moustakas, 1994). An inductive approach was adopted during the coding of the data. In other words, an initial code list was not established. All the codes were formed after the data were obtained. During the analysis of the data, a detailed written account was initially made of all the interviews and notes. Subsequently, the coding and theme forming stage was realized by using Microsoft Excel and Word office programs. The data were coded via the eclectic coding method, by which more than one method is used in combination. Accordingly, in the first cycle, *emotion coding*, *in vivo coding*, and *descriptive coding* were done simultaneously. In emotion coding, the emotions that the teachers remembered or were experiencing or the emotions that the researcher inferred were coded. In vivo coding, the participants' own actual statements were coded. Finally, in descriptive coding, words or short phrases were determined by the researcher to summarize the main topic in an excerpt or a paragraph of the qualitative data. Subsequently, via pattern coding, one of the second cycle coding methods, the codes were transformed into meaningful and restricted higher codes (Saldaña, 2016). Pattern coding is a method of grouping the codes obtained in the first cycle into a small number of clusters, themes, or structures (Miles & Huberman, 1994). Finally, the themes and sub-themes were portrayed visually on a tree diagram.

### 2.4. The Trustworthiness of the Study

Various strategies are recommended in the literature to ensure the credibility, dependability, transferability, and confirmability of the findings obtained in qualitative studies (Krefting, 1991; Lincoln & Guba, 1985). Within the scope of the present study, the methods of member checking, peer debriefing, deep focus data collection, detailed description, and data consistency analysis were utilized. At the end of the interviews, the data were summarized by the researcher, and feedback was received from the interviewee. Thus, confirmation by the participant was obtained for the plausibility of the study. In the present study, data were collected via in-depth interviews. By means of a detailed description, which allows for the evaluation of the data analysis results for accuracy and transferability (Creswell & Miller, 2000), the study sought to provide as many details as possible about the themes and codes. Finally, for the dependability of the qualitative study, confirmation was received for one part of the codes by two experienced experts on education programs and instruction. To this end, the raw data, codes, and themes of two of the interviews were shared with the expert researchers, who examined the consistency between the data and the

codes, and the consistency between the codes and the themes. The Miles and Huberman (1994) formula was used to calculate the consistency among the codes. The inter-rater reliability was found to be 83%. Inter-rater reliability values above 80% are accepted as reliable (Miles & Huberman, 1994).

## 2.5. The Role of the Researcher

The researchers had more than 10 years of experience in teaching mathematics in middle school. This provided the researchers with the opportunity to observe and experience the impacts of national tests on teaching, students, teachers, parents, and administrators. The researchers being experienced teachers had a positive impact on the interview process. Learning that the researchers were also engaged in teaching like themselves made the participants feel more comfortable in providing genuine answers to the interview questions. Moreover, the experiences of the researchers enabled them to understand the data more effectively and to ask better questions. These experiences prevented the researchers from leading a biased approach in analyzing and interpreting the data. Moreover, the researchers did not deviate from the phenomenological approach during the interpretation of the data.

## 2.6. Ethical Procedures

Ethical principles and rules were followed during the planning of the research, data collection, analysis, and reporting. This research was found ethically appropriate with the decision at the 2021-1 meeting held by Bandırma Onyedi Eylül University Social and Humanities Ethics Committee.

## 3. Findings

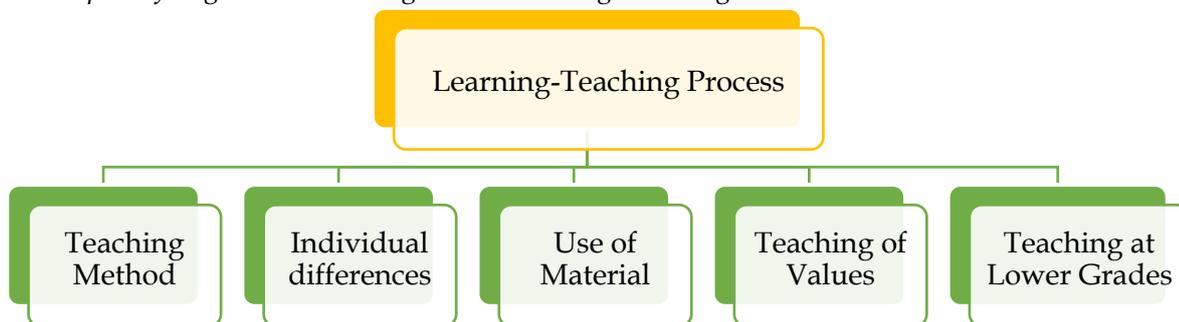
The findings obtained from the interviews held with the teachers revealed that high-stakes testing affected the learning-teaching process, measurement and assessment, communication, and motivation. Furthermore, they revealed how these impacts would be different if there were no high-stakes testing. The themes and sub-themes that were identified are presented below through a tree diagram and descriptive statements.

### 3.1. The Impact of High-Stakes Testing on the Learning-Teaching Process

The theme referring to the impact of high-stakes testing on the “learning-teaching process” consists of five sub-themes. The theme and sub-themes are presented in Figure 2.

Figure 2

*The Impact of High-Stakes Testing on the Learning-Teaching Process*



As can be observed in Figure 2, “learning-teaching process” theme consists of sub-themes which are “teaching method”, “individual differences”, “use of material”, “teaching of values”, and “teaching at lower grades”.

#### 3.1.1. The Teaching Method

The teachers stated that high-stakes testing negatively affected the methods and techniques they used in teaching mathematics. They reported that due to these tests, they placed more emphasis on

the traditional teaching method and were unable to do revisions owing to their anxiety of covering the content in time. Teacher T3 expresses this condition as follows: *"As we need to cover the topic in the fastest possible way, most of the time I have to use the traditional teaching method."* The teachers stated that they mostly spent class time on solving test-related questions and providing students with tactics about how to solve these questions. Teacher T1 expresses his/her views regarding this issue as follows: *"As the THSE consists of multiple-choice questions, I have to do lessons based more on solving tests, solving questions to this end."* One other negative impact of high-stakes testing is not providing the opportunity for students to engage in active learning due to the lack of or very few activities they can engage in. Teacher T1 said, *"We did some small drama activities during some topics but I kept the number of these down because we have many topics but not much time."* Unlike the other teachers, in support of the condition, only one of the teachers (teacher T6) stated that these tests enabled him/her to do more activities by stating the following: *"I increased [the number of] my activities to enable the students to participate in the lesson actively."*

Along with these negative impacts, the teachers reported that they increased their use of technology. The teachers stated that they used technology to reinforce content, to associate content with daily life, or to solve test-related questions on the smartboard. The teachers were using the smart board and the online education platform provided by the National Ministry of Education called EBA (education information network) as technological aids. How technology was used and for what purposes were expressed by teachers T7 *"It is I who explains the topic again, but at the point of reinforcing and associating it with daily life, I get support from technology and EBA"* and teacher T11 stated *"I project the quality questions on the smartboard to save time or by photocopying them."*

The teachers claimed that although the test questions were in line with the learning outcomes of their own teaching program, owing to the nature of the test questions, they had a hard time trying to teach students the reasoning skill. For instance, explaining how they tried to equip students with the systematic thinking skill, T4 stated:

*"... I think that the questions are in line with the learning outcomes but they require thinking, they are time-consuming and they do not consist of questions that all students can solve ...we help [students] by simplifying the topic to the level that all students can understand, how logical reasoning needs be done... after students gain an understanding of the learning outcome, we help them to reason logically."*

### 3.1.2. Individual Differences

The teachers reported that they either did not take into consideration individual differences or did so only partially during the teaching process. They stated that if the students lacked the preliminary knowledge, they disregarded this, and nor did they do any additional planning for those students who lacked the reasoning skill. In relation to this point, T5 asserted that:

*"To be able to take into consideration individual differences, there needs to be more class time, and hands-on lessons consisting of few students. In a class of 40 students with individual differences, you can only do an exam-oriented lesson using a common method."*

Within this statement, it can be observed that the teacher refers to the impact of not just testing but also the physical learning environment on consideration of individual differences. On the other hand, teacher T1 mentions THSE and the associated time problem that prevent teachers from taking into account individual differences by saying, *"Unfortunately when preparing for THSE, doing work in relation to individual differences is not much possible because there is not enough time to do a special activity or other work"*. Similarly, T10 expressed his/her views as follows: *"I do my lessons considering that they have an average level of readiness. Unfortunately, I cannot take into consideration individual differences."* Only two of the teachers reported that they took into account individual differences during the teaching process.

### 3.1.3. Use of Material

The teachers reported that they either did not use any materials or rarely used them in the teaching process. To exemplify, teacher T1 stated, *“As I have to do work toward THSE, I have to postpone the use of materials and the implementation of in-class activities because it feels like it is a waste of time.”* The statements of T1 indicate that s/he has concerns for solving tests for the THSE and that the problem of time prevents the use of materials.

On the other hand, teachers expressed the need for materials in addition to the textbooks given to them by the government. The materials referred to are supplementary textbooks or test books used for test preparation. One of the teachers stated that the question bankbook prepared by the Ministry of National Education partially met this need. In terms of materials, teacher T4 said, *“Textbooks are not sufficient for THSE. There are no questions resembling those on the THSE; a supplementary source is definitely needed.”* Teacher T9 emphasized the need for test preparation books, saying, *“The presence of the test necessitates a supplementary book.”*

### 3.1.4. Teaching of Values

Despite the existence of value teaching and competence development in the mathematics syllabus, the teachers reported that they did not give place to the teaching of values in their classes due to high-stakes testing. For instance, teacher T5 said, *“It is very difficult to give [teach] values when one is rushing to get things done in mathematics for THSE.”* Similarly, teacher T7 reported, *“My effort to teach values decreased due to rushing to cover the test topics.”* On the other hand, there were also teachers who claimed that the test did not have any impact on the teaching of values. For example, teacher T11 said, *“The influence of any of the values in the entire mathematics program on children is nowhere near their interest in a pen used by any YouTuber, which is very sad. I don’t think anything would change if there were no [high-stakes] testing.”* Teacher T8 expressed similar views by saying, *“As it is a numerical subject, values would not be focused on even if there were no [high-stakes] testing.”*

### 3.1.5. Teaching at Lower Grades

The teachers stated that the impact of the test on the learning and teaching process was not only felt on grade 8, which is the only grade level when the test is taken, but also on the learning-teaching process as of grades 4 and 5. That high-stakes testing affects the learning-teaching process not only of the grade when the testing is administered but also of lower grades is expressed by teacher T10, who said, *“Even a child who is still at grade 3 or 4 is educated completely according to this system”* and teacher T6, who stated, *“It is essential to follow the education program and to prepare for the test as of grade 5.”*

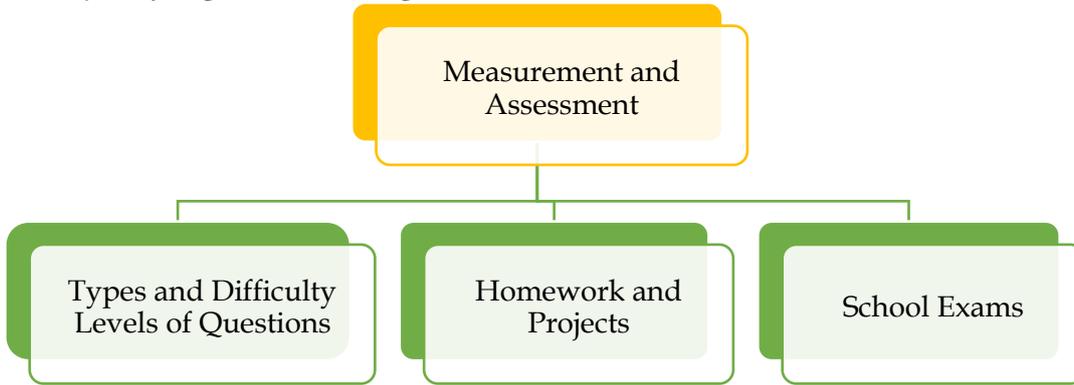
The teachers reported that they solved test-oriented questions as of lower grades, and made changes in especially seventh-grade topics, which they completed earlier to start the 8th-grade topics. Expressing how the THSE impacted the learning-teaching process, teacher T12 said, *“While teaching in different grade levels, I feel the need to make changes in the content in accordance with THSE and change the course of the lessons accordingly,”* and teacher T11 stated, *“I explain a little of how they [students] would see the topic in coming years ...I make an introduction to the THSE topics in grade 7 by completing the 2nd term topics earlier.”*

## 3.2. The Impact of High-Stakes Testing on Measurement and Assessment

The theme of the impact of high-stakes testing on “measurement and assessment” consists of three sub-themes. The theme and sub-themes are presented in Figure 2.

Figure 3

*The Impact of High-Stakes Testing on Measurement and Assessment*



As can be seen in Figure 3, “measurement and assessment” theme consists of sub-themes which are “types and difficulty levels of questions”, “homework and projects” and “school exams” sub-themes.

### 3.2.1. *Types and Difficulty Levels of the Questions*

The teachers claimed that the factor that mostly affected the measurement and assessment was the question type and the difficulty of the test questions. What teachers referred to by “question type” was the multiple-choice questions in the test measuring the reasoning skill. Thus, these questions were defined by the teachers as “quality questions or new generation questions”.

The teachers reported that they paid attention to resemble the questions they used to measure student performance to those in THSE and even adjusted the difficulty levels of the questions to the difficulty level of those in the THSE. The teachers asserted that these questions were quite difficult and that very few students could answer these questions. For example, T13 mentioned the difficulty of the test question: “*I believe that a majority of the 14-year old millions of children would not be able to and cannot solve the questions which even we cannot solve very easily.*” Finally, teacher T7, who asserted that the process of measurement and assessment was affected by the test questions, said, “*...but I experienced a change in the area of measurement and assessment. With questions and examples resembling the test questions, I prepare the students for the test.*”

### 3.2.2. *Homework and Projects*

The teachers' approach to homework and projects was again found to be under the influence of high-stakes testing. Teachers reported that they assigned homework that consisted of multiple-choice tests comprising difficult questions that measured the reasoning skill, similar to the ones in the [LGS] test. In relation to homework assignments, teacher T10 said, “*The types of homework I assigned changed with respect to question type and difficulty level*” and T4 stated, “*The type of the homework assignments is in line with the questions to be asked in the THSE [test].*” As for teacher T13, s/he pointed out:

“*At the end of the lesson, I assign the children homework from test books because the questions in these books consist of quality questions... And as project homework, I assign project work that is associated with daily life as in the [THSE] test questions which are adjusted to fit the students' level.*”

### 3.2.3. *School Exams*

The teachers reported that the achievement exams they administered at school to evaluate student performance were partially affected by high-stakes testing. They claimed that they did not base the entire achievement exam on THSE type of questions as they believed that doing so would make a majority of the students unsuccessful. For example, to indicate that s/he displayed behavior contrary to high-stakes testing, teacher T3 asserted that:

“If I had asked THSE type of questions in my [achievement] tests, my students’ levels of success and naturally their motivation would have decreased. To avoid distancing them [from the lesson], I ask them questions that I think are easier than THSE and which I think they can do.” (T3)

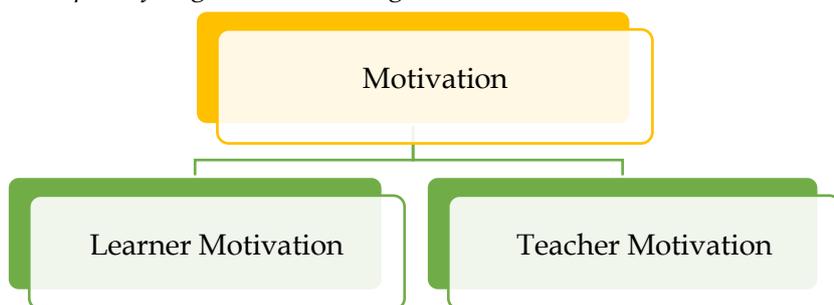
Similarly, teacher T11 argued that achievement exams in schools were partially affected: “Having [achievement] exams equivalent to THSE will cause [students’] scores to decrease. By necessity, I prepare for average-level exams. I try to assess [the performance of] students at higher levels by means of one or two quality questions.”

### 3.3. The Impact of High-Stakes Testing on Motivation

The theme of the impact of high-stakes testing on “motivation” consists of two sub-themes. The theme and sub-themes are presented in Figure 4.

Figure 4

*The Impact of High-Stakes Testing on Motivation*



The issue that the teachers focused on the most was motivation, as shown in Figure 4. The teachers addressed motivation in different ways: “Learner motivation” and their own motivation; that is, “teacher motivation”.

#### 3.3.1. Learner Motivation

While mentioning the impacts of high-stakes testing on learner motivation, the teachers categorized students as successful and unsuccessful or the 10% group and the 90% group. The former group refers to the group consisting of successful students who have the potential to pass the test because high-stakes testing has a different impact on successful and unsuccessful students. The teachers reported that students with a low level of achievement developed a dislike for the subject, and the remaining ones lost their interest and withdrew themselves from the subject. In contrast, they asserted that the motivation level of successful students increased. For instance, teacher T7 stated, “While good students did try to be better, average or lower-level ones let it go all together and started to distance themselves from mathematics.”

One of the issues that the teachers mostly dwelled on was the anxiety and stress that high-stakes testing caused in students. For example, the account of teacher T5 as follows: “As mathematics is selective and difficult in the THSE, which is being administered for two years, children are more anxious and stressed.”

The teachers reported that high-stakes testing influenced students’ emotions toward the subject. It particularly developed negative emotions toward the subject. Furthermore, the teachers expressed their view that high-stakes testing caused hopelessness and fatigue in students. Teacher T5 explained this view as follows: “...but the students who find the format of the test difficult experience feelings of hopelessness, anxiety, and distancing from the mathematics subject.”

#### 3.3.2. Teacher Motivation

The teachers also expressed that high-stakes testing had a negative impact on themselves as well and that this reflected directly or indirectly on their teaching. They highlighted that they felt stress and anxiety. They categorized the anxiety they felt into two: time-related anxiety and the anxiety

of not being able to complete the coverage in due time. These two types of anxiety are intertwined. Teacher T1 expressed the anxiety they felt by saying,

“I have the anxiety of not being able to complete the topics in due time. I mean I have the anxiety that if I get sick and have to get a sick report, I will fall behind in my lessons by not being able to do them.”

Teacher T5 described the stress s/he felt as follows: *“Mathematics teachers are in difficult conditions in terms of emotional stress...The fact that our students spend many minutes trying to solve a single question and then get low net scores to make us very sad.”* What teacher T2 claimed shows that the stress that the teachers felt was actually associated with the students: *“As students get anxious, we get more stressed.”*

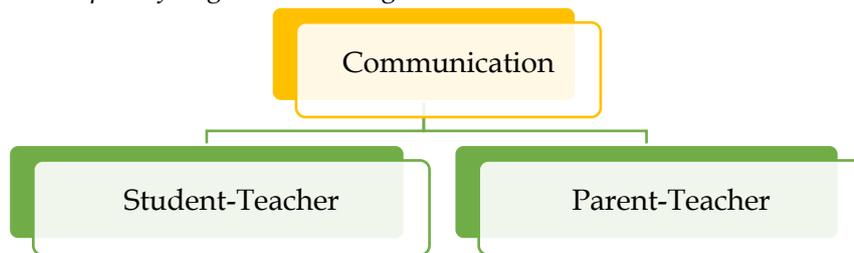
The teachers stated that they became demotivated or their level of motivation toward teaching decreased. For instance, T11 mentioned this by saying, *“I feel like I am solving some questions and listening to myself. This decreases my motivation.”* In addition, the pressure that the teachers felt upon them is revealed in the statements expressed by teacher T10: *“I am highly burdened. My subject is a foundation lesson and the new generation questions being at a high level exert an extreme level of burden upon me.”*

### 3.4. The Impact of High-Stakes Testing on Communication

The theme of the impact of high-stakes testing on “communication” consists of two sub-themes. The theme and sub-themes are presented in Figure 5.

Figure 5

*The Impact of High-Risk Testing on Communication*



As it is seen in Figure 5, the “student-teacher” communication and the “parent-teacher” communication revealed as sub-themes of “communication” theme.

#### 3.4.1. Student-Teacher Communication

The teachers reported that their discussions with the students were generally based on the test [THSE]. It is deduced that the content of these talks consisted of statements that frightened and pressurized the students or increased their motivation. For example, teacher T2 and teacher T1 reveal how the pressure on teachers impacts their communication with students. Teacher T2 said, *“whether we like it or not, we can frighten students to listen by saying this topic will appear in THSE,”* and teacher T1 stated, *“Unfortunately because students’ success is indirectly associated with teachers’ success too, I feel forced to get angry at them when they do not attach the necessary importance to their studies.”* The teachers claimed that they made motivation boosting talks to make students like the subject, to prevent them from becoming embittered against the subject, and to decrease the negative perception about the subject. One exceptional comment was made only by teacher T6, who claimed that s/he did not mention the test in his/her communication during the learning-teaching process: *“... I do not talk about THSE so as to teach the subject without making students feel embittered against the subject.”*

#### 3.4.2. Parent-Teacher Communication

The teachers stated that high-stakes testing increased communication between parents and teachers. However, they highlighted that it was the parents of successful students or interested parents that they communicated more with. The statements of teacher T7 depict the situation

clearly: *“The frequency of meetings with parents increased. The parents of good students started to visit more often, but again the parents of lower-performing students did not visit very frequently.”*

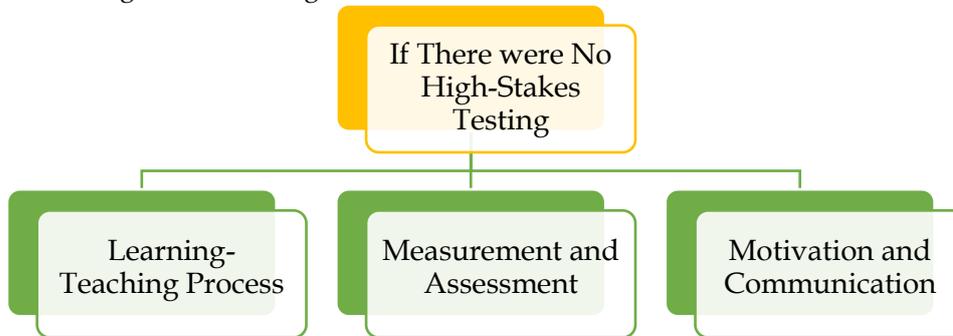
The teachers claimed that the content of the meetings with the parents was mostly the test [THSE], while issues such as student behaviors, students’ interests, and talents were disregarded. An issue that the teachers complained about was parents’ misvaluations of and accusations against the teachers. They stated that a teacher who provided students with tests was considered a good teacher, while they accused the teacher when a student performed poorly. The statements of teacher T9 explained the communication conflicts they had with parents: *“The parents of low performing students accuse the teacher.... they experience difficulty in explaining to the parents that the test would select a limited number of students”*

### 3.5. If There Were No High-Stakes Testing

The theme of “If there were no high-states testing” is actually a condition that the teachers expressed under all the themes. The reason why this theme was presented separately and not under the other themes is that it reveals how teacher behaviors would be affected if there were no high-risk testing. The theme and sub-themes are presented in Figure 6.

Figure 6

*If there were no high-stakes testing*



As it is seen in Figure 6, this theme consists of the three sub-themes: “learning- teaching process”, “measurement and assessment”, “motivation and communication”.

#### 3.5.1. The Learning-Teaching Process

The teachers expressed that if there were no high-stakes testing, they would engage in more active teaching. In other words, they mentioned that they would increase the use of activities, materials, and technology and enable their students to engage in hands-on experiential learning. They claimed that in this way their lessons would be more fun. For example, teacher T13 asserted *“If there were no THSE, I would perform a lesson focused completely on hands-on experiential learning. I would prepare materials and activities to suit each lesson.”*

Similarly, teacher T3 stated, *“If there were no THSE, I would base the lessons not on solving questions in a fast way, but on internalizing and understanding the topics.”* Both teachers claimed that interest and participation in the lessons would increase in this way. With regard to this issue, teacher T9 explained his/her views that there would be a change in the conditions where unsuccessful students lost interest in the subject, which was mentioned earlier: *“If there were no THSE, the level of interest and participation of mediocre students would increase.”* Finally, the teachers emphasized that teaching of values would also increase. For instance, teacher T10 asserted, *“If there were no THSE, perhaps values could be taught because currently, students do not want to hear any knowledge other than THSE.”* Only one of the teachers held an opinion different from the others. Teacher T6 claimed that the learning-teaching process would be negatively impacted: *“If there were no THSE, I would use the traditional teaching method and decrease the number of activities”*

### 3.5.2. Measurement and Assessment

The teachers asserted that if there were no high-stakes testing, measurement and assessment processes would change. They stated that they would particularly use question types different from those in the (THSE) test. In relation to this, T10 stated,

“If there were no THSE, I would make more use of non-test related tasks in measurement and assessment”, and teacher T1 stated, “... I would probably prepare my written achievement test questions as if I were preparing crossword puzzle questions – to make them both thought-provoking and fun.”

In addition, *teachers* T11 and T7 also referred to the fact that the question types would change in the measurement and assessment process. Teacher T11 said, “*If there were no THSE, the question type would certainly be different. These children are unable to engage in questioning and evaluation processes in this system of education. These skills cannot be acquired in the final grade by covering 1 or 2 test books*” and finally, teacher T7 expressed him/herself as follows: “*If there were no THSE, the measurement and assessment process would be a little more different. That is, I would not force the students in the field of skill-based questions this much.*”

### 3.5.3. Motivation and Communication

The teachers dwelled on the fact that if there were no THSE, they would not have time-related anxiety and thus not feel the need for holding extra lessons as the topics could easily be covered in the time available. In addition to this, the teachers highlighted the fact that if there were no THSE, student stress and anxiety would decrease and that this would have a positive affect communication and the way the lessons are implemented. Another issue that teachers highlighted in the area of communication was freedom from parent pressure. In the speech belonging to teacher T13, it can be observed that T13 was of the opinion that this was possible: “... *Perhaps parents wouldn't attribute their children's failure to the teachers.*”

## 4. Discussion

The present study provides evidence-based data to policymakers concerning the impacts of high-stakes testing on the teaching process specific to the mathematics subject in Turkey and other similar countries where high-stakes testing has been administered for approximately 30 years under different names. This study aimed to guide decision-makers in Turkey and other countries where there are similar practices in making improvements to negate the adverse impacts of high-stakes testing on the teaching process and to make the necessary readjustments.

The investigation of the impact of high-stakes testing on the “*learning-teaching process*” revealed that the teachers resorted more to traditional methods in teaching, and they either rarely used or did not use activities nor materials to make the lessons active. Likewise, it is reported in the related literature that high-stakes testing led teachers to be inclined toward using traditional ways of teaching methods (Au, 2007; Minarechova, 2012; Ramezaney, 2014). A study by Klein et al. (2006) revealed that standardized tests affected teachers' teaching methods and caused teachers to implement test-oriented lessons. In another study, conducted by Pedulla et al. (2003), it was reported that owing to high-stakes testing, teachers spent less time on teaching-enriching activities, such as excursions. Musoleno and White (2010), who conducted a study on the middle school level, found that after high-stakes testing started to be implemented, teachers started to decrease their use of teaching practices entailing approaches based on collaboration and questioning. Similarly, Cranley (2018) put forward that high-stakes testing impacted teachers' pedagogical approaches in their mathematics lessons. As a result of this study, it was revealed, based on teachers' statements, that teachers did not take into account individual differences and did not pay sufficient attention to the teaching of values due to their anxiety to cover the content in time and to solve more tests. A study by Scot et al. (2008) revealed that high-stakes testing prevented individual differences (e.g. giftedness) from being taken into consideration. In one other study, conducted by Jones and Egley (2004), it was reported that high-stakes testing prevented teachers

from meeting students' learning needs. The fact that teachers disregarded the teaching of values that existed in the mathematics teaching program is an indication that teachers focused on the topics addressed in high-stakes testing, but ignored the other parts (Barnes, 2005). This could also be regarded as a "Narrowed Curriculum", which is frequently mentioned in the related literature on the exclusion of value teaching (Jones & Egley, 2004). Moreover, the teachers claimed that high-stakes testing did not only impact the learning-teaching process in grade 8, the syllabus for which the students were responsible, and the only grade at which they sat the test, but also the learning-teaching process in the lower grades. This result that the study yielded shows that high-stakes testing has widespread impacts. In other words, this means that teachers are inclined to adopt traditional teaching and pursue test-oriented teaching not only in grade 8 but also in grades 5, 6, and 7. However, in a study by Li and Xiong (2018), it was revealed that the time spent on test preparation by students in class negatively impacted their performance in high-stakes testing. In conclusion, it can be asserted that the impacts of high-stakes testing on the learning-teaching process will prevent the development of students in the cognitive, affective, and psychomotor domains and the realization of education policies regarding students' acquisition of 21st-century skills.

The examination of the impacts of high-stakes testing on "measurement and assessment" revealed teachers' views that the difficulty level and question types in high-stakes testing reflected on their own assessment tools as well. The type and difficulty level of the questions used by the teachers consisted of difficult questions measuring reasoning skills similar to those in the (THSE) test. This shows consistency with the findings reported in a study by Au (2007), which showed that curricular control caused by high-stakes testing resembles tests being dependent on their own structures. While the reflection of high-stakes testing on measurement and assessment is similar in the sub-theme of "homework and projects", it is varied in "school exams". The teachers stated that they assigned homework that consisted of multiple-choice tests comprising THSE-like difficult questions measuring the reasoning skill, while the projects they assigned were based on daily life themes. As for the school exams that the teachers administered, they were only partially impacted by high-stakes testing. The teachers were of the opinion that if they prepared achievement test questions in accordance with high-stakes testing questions, their students would be unsuccessful. This situation shows the teachers' disbelief that most of the students could be successful in high-stakes testing.

It has been revealed under the theme of "motivation" that high-stakes testing impacted the motivation of both teachers and students. The teachers stated that it affected the motivation of successful and unsuccessful students in different ways. While the level of motivation of successful students increased, that of unsuccessful students increasingly decreased. Lingard (2009) argued that high-stakes testing would have negative impacts leading to a low level of self-esteem in students with a low level of success and a decrease in commitment to learning. The teachers participating in the present study reported that students experienced stress and anxiety, developed negative emotions against the subject, displayed fatigue, and sank into despair. This finding is consistent with those reported in previous studies revealing that high-stakes testing caused stress and anxiety in students (Büyükoztürk, 2016; Casbarro, 2004; Dinç et al., 2014). In addition to stress and anxiety, another negative impact of high-stakes testing is causing students to perceive the learning process as a boring activity and causing them to become distanced from the learning process (Şahin, 2008). High-stakes testing has a negative impact on the motivation of not only students but also teachers. The teachers in the present study reported that they themselves also experienced stress and anxiety, that their eagerness to teach decreased and they felt pressure on them. In a study by Lynn diFate (2008), in which the stress factors regarding high-stakes tests were revealed, it was identified that time management and motivation were high-stress sources. Similarly, one of the findings of the present study that teachers experienced the anxiety of not being able to cover the course content in due time shows consistency with this finding.

Under the theme of “*communication*”, the teachers put forward that high-stakes testing affected the “*learner-student*” and “*parent-teacher*” communication. The teachers stated that they generally talked to students about high-stakes testing and that the content of these talks consisted of chats that aimed to place fear and pressure on the students or to increase their motivation. In a study with teachers conducted by Pedulla et al. (2003), it was reported that teachers felt intense pressure upon them for their students to perform well. The teachers stated that they made motivational talks in order to enable students to develop positive feelings toward the subject. High-stakes testing impacts not only the teacher-student communication but also the teacher-parent communication in schools. The teachers expressed that high-stakes testing increased their communication with the parents of successful and interested students. It was revealed that the content of the meetings between teachers and parents was based more on high-stakes testing and that other issues, such as student behaviors, interests, and talents were disregarded. The fact that parents regarded teachers who had their students solve multiple-choice tests as good teachers, while they accused teachers when students were unsuccessful caused communication conflicts between teachers and parents. In a study carried out by Pedulla et al. (2003), it was reported that parents pressurized teachers to increase students’ test scores. Similarly, in a study by Jones and Egley (2004), it was revealed that parents pressurized parents. Accordingly, it can be concluded that high-stakes testing generally affects the communication process negatively.

Under the theme of “*If there were no high-stakes testing*,” the teachers in the present study claimed that they would behave differently in the areas of “*the learning-teaching process*”, “*measurement and assessment*”, “*motivation*” and “*communication*” if there were no high-stakes testing. They stated that in such a condition, they would engage more in active teaching and students’ interests and participation in the lessons would increase. They emphasized that if there were no high-stakes testing, they would use question types different from those in the (THSE) test, that they would not have time concerns, and thus would be able to cover the course content with ease. They believed that student’ levels of stress and anxiety would decrease and that this would have positive impacts on the teaching methodology and communication. Finally, they claimed that they would be free of parent pressure.

These findings highlight once more that high-stakes testing has negative affects and indicate that if there were no high-stakes testing, the teaching process would be impacted positively. What teachers and students would actually not do if there were no test, but do so due to the test is referred to as backwash (impact) (Alderson & Wall, 1993). Hence, the results that the present study yielded serve as evidence to the backwash (impact) that high-stakes testing causes.

## 5. Conclusion and Recommendations

Teachers addressed the impacts of high-stakes testing on the process of mathematics teaching in terms of “*the learning-teaching process*”, “*motivation*”, “*communication*”, and “*if there were no THSE*”. The results of the study revealed that the inclination of teachers to adapt traditional teaching and to solve class tests, and their use of assessment questions that resembled those in high-stakes testing had a negative impact on teacher and learner motivation. Similar results were reported in numerous other studies in the related literature. On the other hand, the related literature had no or very few similar results regarding the fact that high-stakes testing prevented the consideration of individual differences in class, the lack of value teaching, the central place that high-stakes testing has in messages within the communication process, and the presence of the widespread influence of high-stakes testing on the teaching process starting at a certain grade and leading down to lower grades, which were results that the present study also revealed. Thus, these results open up a new area of discussion on the impacts of high-stakes testing on the teaching process. Researchers can be recommended to investigate these results more specifically and comprehensively. Policymakers need to take steps toward improving the process, together with the teachers, by minimizing the negative impacts of high-stakes testing on teaching. The elimination of the factors causing problems of stress, pressure, and anxiety, which negatively affect

the psychology of particular students and teachers, would lead to a more positive learning-teaching process. However, conducting studies on eliminating the negativities within the process without the support and approval of teachers will enable the discussion of high-stakes testing to continue because the negative impacts of high-stakes testing will negatively affect the outcomes of education as well.

## References

- Abrams, L. M. (2004). *Teachers' views on high-stakes testing: Implications for the eclassroom*. Arizona State University, Education Policy Research Unit.
- Alderson, C., & Wall, D. (1993) Does washback exist? *Applied Linguistics*, 14(2), 115-12
- Amrein, A. L., & Berliner, D.C. (2002). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives*, 10(18). <http://epaa.asu.edu/epaa/v10n18/>
- Ashadi, A., & Rice, S. (2016). High stakes testing and teacher access to professional opportunities: lessons from Indonesia. *Journal of Education Policy*, 31(6), 727-741. <https://doi.org/10.1080/02680939.2016.1193901>
- Atılgan, H. (2018). Transition among education levels in Turkey: past-present and a recommended model. *Ege Journal of Education*, 19(1), 1-18. <https://doi.org/10.12984/egeefd.363268>
- Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, 36(5), 258-267. <https://doi.org/10.3102/0013189x07306523>
- Bağcı, E. (2016). *Alignment of teog mathematics questions to the mathematics curriculum and the level of attainment the goals of teog system* [Unpublished Master's Thesis]. Ankara University, Institute of Educational Sciences, Ankara.
- Barnes, M. (2005). *The Discriminatory Effects of High-Stakes Testing in Georgia: Exploring Causes and Solutions*. Education Law and Policy Forum, Education Law Consortium, The University of Georgia, Athens.
- Breakspear, S. (2012). *The policy impact of pisa: an exploration of the normative effects of international benchmarking in school system performance*. OECD Publishing. <https://doi.org/10.1787/5k9fdffr28-en>
- Buyruk, H. (2014). Standardized examinations as a teacher performance indicator and performance evaluation in education. *Trakya University Journal of Education*, 4(2), 28-42.
- Büyüköztürk, Ş. (2016). Sınavlar üzerine düşünceler [Thoughts on exams]. *Kalem International Journal of Education and Human Sciences*, 6(2), 345-356.
- Can, E. (2017). Determination of the effects of central exams according to the view of students, *The Journal of Academic Social Science*. 5(58), 108-122. <http://doi.org/10.16992/ASOS.12842>
- Casbarro, J. (2004). Reducing anxiety in the era of high stakes testing. *Principals*, 83(5), 36-38.
- Chapman, D., & Snyder, C. W. (2000) Can high stakes national testing improve instruction: Reexamining conventional wisdom, *International Journal of Educational Development*, 20(6) 457-474. [https://doi.org/10.1016/S0738-0593\(00\)00020-1](https://doi.org/10.1016/S0738-0593(00)00020-1)
- Cizek, G. J. (2001). More unintended consequences of high-stakes testing. *Educational Measurement: Issues and Practice*, 20(4), 19-27. <https://doi.org/10.1111/j.1745-3992.2001.tb00072.x>
- Corbin, J., & Strauss, A. (2015). *Basics of qualitative research: techniques and procedures for developing grounded theory*. Sage.
- Cranley, L. (2018). *An investigation into the impact of high-stakes testing, through the naplan assessment, on the teaching and learning of mathematics in one primary school* [Unpublished Master's Theses]. University of Notre Dame, Australia.
- Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory into Practice*, 39(3), 124-130. [https://doi.org/10.1207/s15430421tip3903\\_2](https://doi.org/10.1207/s15430421tip3903_2)
- Dinç, E., Dere, İ., & Koluman, S. (2014). An investigation of lay people's experiences and view points on the practices of the transition between schooling levels. *Adiyaman University Journal of Social Sciences* 7(17), 397-423. <https://doi.org/10.14520/adyusbd.761>
- Directorate General for Secondary Education (2016). *Ortaöğretim izleme ve değerlendirme raporu* [Secondary education monitoring and evaluation report]. Ministry of National Education. [http://ogm.meb.gov.tr/meb\\_iys\\_dosyalar/2017\\_06/13152815\\_izlemeraporubirlestirilmis\\_3\\_ekim1.pdf](http://ogm.meb.gov.tr/meb_iys_dosyalar/2017_06/13152815_izlemeraporubirlestirilmis_3_ekim1.pdf)
- Education Reform Initiative [ERI] (2013). *Yeni ortaöğretim geçiş sistemi üzerine değerlendirmeler* [Evaluations on the new secondary education transition system] <http://www.egitimreformugirisimi.org/yayin/yeni-ortaogretime-gecis->

- Ekiz, H. (2019). *Standardized testing for transition to high schools: teachers perceptions of how national tests influence teaching and learning in middle schools english classes* [Unblished Master's Thesis]. Middle East Technical University, Ankara.
- Güler, M., Arslan, Z., & Çelik, D. (2019). Mathematics teachers' views on the 2018 entrance exam for high schools. *YYU Journal of Education Faculty*, 16(1), 337-363. <http://doi.org/10.23891/efdyyu.2019.128>
- Gündoğdu, K., Kızıldaş, E. & Çimen, N. (2010). Opinions of students and teachers in relation to high school entrance exam-sbs (case of Erzurum). *Elementary Education Online*, 9(1), 316-330
- Gür, B. S., Çelik, Z., & Coşkun, İ. (2013). *Türkiye'de ortaöğretim geleceği: hiyerarşi mi, eşitlik mi?* [Future of secondary education in Turkey: Is hierarchy or equality?] SETA Analiz.
- Hamilton, L. S., Stecher, B. M., & Klein, S. P. (2002). *Making sense of test based accountability in education*. Rand Corporation.
- Johnson, D. D., Johnson, B., Farenga, S., & Ness D. (2008). *Stop High Stakes Testing: An Appeal to America's Conscience*. Rowman & Littlefield.
- Jones, B. D., & Egley, R. J. (2004). Voices from the frontlines: teachers' perceptions of high-stakes testing, *Education Policy Analysis Archives*, 12(39). <https://doi.org/10.14507/epaa.v12n39.2004>
- Jones, M. D., Jones, B. D., & Hargrove, T. (2003). *The unintended consequences of high-stakes testing*. Rowmanand Little field Publishers.
- Jonsson A., & Leden L. (2019) The ambiguous influence of highstakes testing on science teaching in Sweden, *International Journal of Science Education*, 41(14), 1926-1943. <https://doi.org/10.1080/09500693.2019.1647474>
- Klein, A. M., Zevenbergen, A. A., & Brown, N. (2006). Managing standardized testing in today's schools. *Journal of Educational Thought*, 40(2), 145.
- Krefting, L. (1991). Rigor in qualitative research: The assessment of trustworthiness. *American Journal of Occupational Therapy*, 45(3), 214-222. <https://doi.org/10.5014/ajot.45.3.214>
- Kumandaş, H., & Kutlu, Ö. (2010). High stakes testing: do secondary education examination involve any risks? *Procedia- Social and Behavioral Sciences*, 9, 758-764. <https://doi.org/10.1016/j.sbspro.2010.12.230>
- Li, H., & Xiong, Y. (2018). The relationship between test preparation and state test performance: Evidence from the Measure of Effective Teaching (MET) project. *Education Policy Analysis Archives*, 26(64). <http://dx.doi.org/10.14507/epaa.26.3530>
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage Publications.
- Lingard, B. (2009). Testing times: The need for new intelligent accountabilities for schooling. *QTU Professional Magazine*, 24, 13-19.
- Looney, J. W. (2009). *Assessment and innovation in education*. OECD Publishing. <http://dx.doi.org/10.1787/222814543073>
- Lynn diFate, T. (2008). *Stress factors of elementary and middle school teachers associated with high stakes testing as required by No Child Left Behind* [Unpublished Doctoral Dissertation]. University of Bridgeport, Connecticut.
- Madaus, G. F. (1988). The influence of testing on the curriculum. In L. N. Tanner (Ed.), *Critical issues in curriculum: Eighty-seventh yearbook of the national society for the study of education* (pp. 83-121). University of Chicago Press.
- Maykut P.& Morehouse R. (2005). *Beginning qualitative research. A philosophic and practical guide*. The Falmer Press
- Merriam, S. (2009). *Qualitative research: A guide to design and implementation*. Jossey-Bass.
- Miles, M. B.,& Huberman, A.M. (1994). *Qualitative data analysis: an expanded sourcebook*. Sage.
- Minarechová, M. (2012). Negative impacts of high-stakes testing. *Journal of Pedagogy/Pedagogický Casopis*, 3(1), 82-100. <https://doi.org/10.2478/v10159-012-0004-x>
- Moustakas, C. (1994). *Phenomenological research methods*. Sage.
- Musoleno, R. R., & White, G. P. (2010). Influences of high stakes testing on middle school mission and practice, *RMLE Online*, 34(3), 1-10. <https://doi.org/10.1080/19404476.2010.11462076>
- Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Sage Publications.
- Pedulla, J. J., Abrams, L. M., Madaus, G. F., Russell, M. K., Ramos, M. A., & Miao, J. (2003). *Perceived Effects of State-Mandated Testing Programs on Teaching and Learning: Findings from a National Survey of Teachers*. National Board on Educational Testing and Public Policy, Lynch School of Education, Boston College. <https://www.bc.edu/research/nbetpp/statements/nbr2summary.pdf>
- Popham, W. J. (2001). Teaching to the test. *Educational Leadership*, 58(6), 16-20.

- Ramezaney, M. (2014). The washback effects of university entrance exam on Iranian EFL Teachers' curricular planning and instruction techniques. *Procedia-Social and Behavioral Sciences*, 98, 1508-1517. <https://doi.org/10.1016/j.sbspro.2014.03.572>
- Saldaña, J. (2016). *The coding manual for qualitative researchers*. Sage.
- Scot, T. P., Callahan C. M., & Urquhart J. (2008). Paint-by number teachers and cookie-cutter students: the unintended effects of high-stakes testing on the education of gifted students. *Roeper Review*, 31(1), 40-52. <https://doi.org/10.1080/02783190802527364>
- Smith, M. L. (1991). Put to the test: The effects of external testing on teachers. *Educational Researcher*, 20(5), 8-11. <https://doi.org/10.3102/0013189X020005008>
- Şahin, S. (2008). The Effects of high school entrance examinations on the primary schools and students. *Çagdaş Journal of Education*, 34(362), 15-21.
- Şahin, S., Uz Baş, A., Şahin Fırat, N., & Sucuođlu, H. (2012). Middle school teachers' and students' view about the secondary school entrance examinations. *International Journal of Human Sciences*, (9)2, 847-878
- TEDMEM (2014). *Eđitim deđerlendirme raporu* [Education evaluation report]. Turkish Education Association. <https://tedmem.org/yayin/2014-egitim-degerlendirme-raporu>
- Thompson, G. (2013). NAPLAN, My School and accountability: Teacher perceptions of the effects of testing. *International Education Journal: Comparative Perspectives*, 12(2), 62-84.
- Zhao, M. R., Mu, B. L., & Lu, C. P. (2016). Teaching to the test: approaches to teaching in senior secondary schools in the context of curriculum reform in China. *Creative Education*, 7, 32-43. <http://dx.doi.org/10.4236/ce.2016.71004>
- Zucker, S. (2003). *Fundamentals of standardized testing*. Pearson Education.