USING ARTIFICIAL INTELLIGENCE ASSISTED DIGITAL TOOL SOCRATIVE WITHIN THE SCOPE OF MEASUREMENT AND EVALUATION IN TEACHING TURKISH AS A FOREIGN LANGUAGE

Research Article

Rabia DEMIRKOL[*] / Merve SUROGLU SOFU[**]

Abstract: This research investigates the effectiveness of the artificial intelligence supported Socrative digital tool in terms of assessment and evaluation during the teaching process of Turkish as a foreign language. The integration of such tools is inevitable, considering the rapid development of artificial intelligence technologies. In most institutions, traditional methods are widely preferred in exam techniques, which prolongs the assessment process in general. Socrative is a platform which gives feedback to the instructor and learner instantly. It is assumed that Socrative might be more effective than classical assessment methods. The research group is composed of 100 B1-level students attending a foundation university in Istanbul who learn Turkish as a foreign language. As a result, in classical and Socrative tests, it was observed that there was no significant difference among gender, age, and language groups. Students scored approximately three points higher in Socrative tests compared to the classical test, and a medium positive relationship at the level of $r = .48$ was determined between the two types of tests. Both tests were administered on the same day. Quantitative data were analyzed using SPSS, and qualitative data were evaluated through content analysis. The results revealed a significant difference in favor of the Socrative-based test. The study recommends wider use of artificial intelligence-supported tools like Socrative in the assessment and evaluation processes of language education.

Keywords: Artificial intelligence, artificial intelligence tools, measurement and evaluation, Socrative, teaching Turkish as a foreign language.

YABANCI DİL OLARAK TÜRKÇE ÖĞRETİMİNDE YAPAY ZEKÂ DESTEKLİ DİJİTAL ARAÇ SOCRATIVE'İN ÖLÇME VE DEĞERLENDİRME KAPSAMINDA KULLANILMASI

Araştırma Makalesi

Özet: Bu araştırmanın amacı, Socrative isimli yapay zekâ destekli dijital aracın yabancı dil olarak Türkçe öğretiminde ölçme ve değerlendirme bağlamında ne kadar etkili olduğunu saptamaktır. Yapay zekâ alanındaki gelişmeler, bu araçların eğitimde kullanılmasını gerekli kılmaktadır. Yabancı dil olarak Türkçe öğretiminde çağdaş yöntem ve tekniklerin kullanılması büyük önem taşımaktadır. Birçok alanda olduğu gibi bu alanda da hâlen birçok kurumda geleneksel sınav yöntemleri tercih edilmektedir. Bu durum ölçme ve değerlendirmeye ayrılan zamanı artırmaktadır. Socrative, öğreticilere ve eğitmenlere anında geri bildirim sunan bir platformdur. Socrative'in klasik ölçme ve değerlendirme yöntemlerine oranla daha etkili olduğu hipotezinden varsayılmaktadır. Bu nedenle bu araştırmada Socrative'in yabancı dil olarak Türkçe öğretiminde ölçme ve değerlendirme bağlamında etkili olup olmadığının araştırılmasına gerek duyulmuştur. Araştırmanın çalışma grubu İstanbul'da bir vakıf üniversitesinde eğitim gören B1 düzeyinde yabancı dil olarak Türkçe öğrenen 100 öğrencidir. Bulgular; cinsiyet, yaş veya dil grupları arasında klasik ve Socrative test puanlarında önemli bir fark olmadığını ortaya koymuştur. Ancak, öğrenciler Socrative testinde klasik teste kıyasla yaklaşık üç puan daha yüksek puan almış ve iki test türü arasında orta derecede pozitif bir korelasyon ($r = .48$) bulunmuştur. Her iki test

* Dr.; Istanbul Nisantasi University; demirkolrabiaa@gmail.com 0000-0003-0998-7393

** Asst. Prof.; Istanbul Nisantasi University, Rectorate, Turkish Language Department; merve.suroglusofu@nisantasi.edu.tr 0000-0001-9101-8211

de çalışma grubuna aynı gün uygulanmıştır. Nicel verilerin analizi için SPSS kullanılırken nitel verilerin analizi için içerik analizi kullanılmıştır. Araştırma sonucunda klasik testlerle Socrative'de hazırlanmış testler arasında Socrative lehine anlamlı fark tespit edilmiştir. Çalışma, yabancı dil olarak Türkçe öğretiminde Socrative ve buna benzer yapay zekâ araçlarının ölçme ve değerlendirme kapsamında kullanımının artırılmasını önermektedir.

Anahtar Kelimeler: Ölçme ve değerlendirme, Socrative, yabancı dil olarak Türkçe öğretimi, yapay zekâ, yapay zekâ araçları.

Introduction

Artificial intelligence has gained increasing popularity in recent years and has begun to spread to many areas of life. OpenAI's launch of ChatGPT has greatly increased the prevalence of artificial intelligence and made this technology an integral part of people's daily lives (Wiredu, 2023). The potential of artificial intelligence to simplify daily tasks, solve complex problems and improve the quality of life without requiring technical and theoretical knowledge has made it popular among a wide range of users (Ng et al., 2021). Artificial intelligence now continues to develop effectively in every area of life, from education to the healthcare sector. The importance of artificial intelligence stems from its potential to offer innovative solutions to many problems. The term AI was first defined by John McCarthy at the Dortmund conference (Arslan, 2020). It is defined as an information technology that has the ability to perceive reason, comprehend, make sense, generalize, infer, learn and successfully perform multiple tasks at the same time (Gondal, 2018). In other words, artificial intelligence means the use of reasoning ability and predictive power, which are specific to human intelligence, by machines in solving complex problems and making decisions considering changing conditions (Obschonka & Audretsch, 2020).
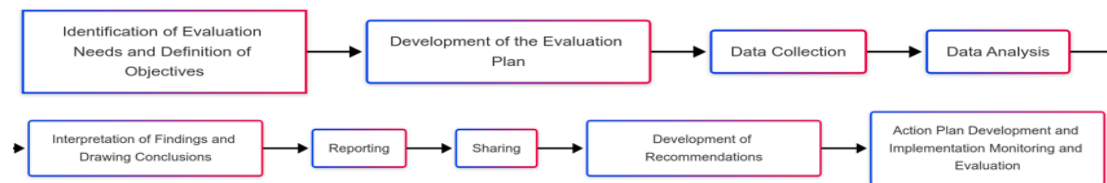
According to the research of Holmes et al. (2019), artificial intelligence can help in areas such as curriculum planning, personnel arrangements, examination management, cybersecurity, facility management, and security in school management, and can make indirect contributions to educational processes. Researchers who apply artificial intelligence early in education state that multiple-choice tests can be used as an effective tool in evaluating students' success levels and strengthening the teaching process.

1. Process-Based Measurement and Evaluation

Process-based measurement and evaluation provide the opportunity to detect and correct problems that may arise at every stage of the educational process. In addition, it enables timely and appropriate interventions to be made, when necessary, by determining the students' achievement levels (Akçadağ, 2010; Çeçen, 2011). Process-oriented assessment methods also allow students to actively construct knowledge (Shepard, 2000).

Figure 1

Process-Based Evaluation Stages

Process-based assessment cannot be considered as a measurement-evaluation method independent of the teaching-learning process; on the contrary, this approach is an approach that actively contributes to the teaching-learning process (Gipps & Stobart, 2003). According to Pierce and O'Malley (1992), process-based assessment is an authentic approach and is based on real-life experiences to achieve educational goals. A process-based assessment framework requires the determination of the issues that should be taken into consideration by the teacher when students are assessed in the classroom (Van Niekerk et al., 2010). The elements of a process-based assessment framework should include outcomes, content, assessment methodology, and the context in which the assessment takes place (Van Niekerk, 2002). When viewed from a national and international perspective, it is seen that there is a definite need to create a process-based assessment framework for technology education. Moreland and Jones (2000) expressed this as follows: "Teachers' formative interaction with students distanced learning from the procedural and conceptual aspects of the subject and learning and formative assessment interactions focused on general skills rather than students' technological understanding."

2. The Contribution of AI-Assisted Digital Tools to Process-Based Assessment

Literature reviews have revealed that Turkish teachers generally prefer traditional assessment and evaluation approaches, are reluctant to apply supplementary assessment methods and stay away from process-oriented assessment techniques. These findings are frequently encountered in the literature (Akata, 2009; Aktürk, 2012; Gelbal & Kelecioğlu, 2007; Metin & Demiryürek, 2009; Yiğit & Kırımlı, 2014). There are various artificial intelligence tools that can be used in education, especially in the field of process-based measurement and evaluation, with the rapid feedback system and advanced algorithms it offers. This study bridges a knowledge gap in the current literature on the use of digital assessment tools in teaching Turkish as a foreign language. Although tools like Socrative have been highly utilized in other foreign language teaching contexts, they are not thoroughly investigated in Turkish foreign language teaching. Moreover, very few studies have compared AI-assisted and traditional assessment tools in this field. One of these is Socrative. Socrative is an AI-powered digital tool that helps educators create questions in just a few seconds with automatic AI support by entering prompts in different question types. Socrative allows instructors to see students' answers instantly and thus supports the learning process by providing rapid feedback. It allows different question types such as multiple choice, true-false, and short answer. Thus, various assessment methods can be included in the same exam with artificial intelligence support. Socrative collects and analyzes exam results. It presents learners' performance with detailed reports and helps to observe trends.

3. The Purpose of the Study

The aim of this study is to determine how effective the artificial intelligence tool called Socrative is in the context of measurement and evaluation in teaching Turkish as a foreign language. It was hypothesized that Socrative-based assessments would result in higher scores compared to traditional (classical) assessments. The aim is to measure the effectiveness of Socrative in terms of student achievement. The research questions are formed as follows:

1. Do the classical and Socrative test scores of students learning Turkish as a foreign language show a significant difference in terms of gender variable?
2. Do the classical and Socrative test scores of students learning Turkish as a foreign language show a significant difference in terms of age range variable?
3. Do the classical and Socrative test scores differ significantly among students whose native languages are different? (e.g., Arabic, Kazakh, Persian)
4. Do the classical and Socrative test scores of students learning Turkish as a foreign language show a significant difference in terms of average both scores?
5. What are the views of students learning Turkish as a foreign language on classical and Socrative tests?

4. Method

4.1. Research Model

This study employs an explanatory mixed-methods design. Mixed-methods research refers to the process by which a researcher integrates qualitative and quantitative methods, approaches, and concepts within a specific research process or across multiple studies. This method aims to develop a more comprehensive and in-depth understanding by using both qualitative and quantitative data collection and analysis techniques together (Tashakkori & Teddlie, 1998; Creswell, 2003; Johnson & Onwuegbuzie, 2004). In the application of the explanatory mixed design, quantitative data are primarily collected and analyzed first, followed by qualitative data. The priority in this approach is generally on quantitative data, and qualitative data are collected primarily to enhance the scope and depth of quantitative data (Baki & Gökçek, 2012). Ethical approval for this study was obtained from the Ethics Committee of Istanbul Nişantaşı University Rectorate, with decision number 20241121-061 dated 21.11.2024.

4.2. Study Group

The study group of the research consisted of 100 students at B1 language level studying at XXX University Turkish Language Teaching Application and Research Center (TÖMER). The study employed purposive sampling, focusing specifically on B1-level students. Participants outside this proficiency level were excluded to ensure comparability in language competence.

Table 1

Gender, Age and Countries of Language Learners

Group		*n*	*%*
Gender	Male	50	50.00
	Female	50	50.00
Age	17-25	65	65.00
	26-34	25	25.00
	35+	10	10.00
Country	Iran	50	50.00
	Egypt	15	15.00
	Kazakhstan	10	10.00
	Afghanistan	5	5.00
	Syria	5	5.00

Lebanon	5	5.00
Morocco	5	5.00
Algeria	5	5.00

100 students at B1 language level participated in the research. 50 of the students were male and 50 were female. The proportional equivalent of this was determined as 50% female and 50% male.

Of the 100 students who participated in the study, 65 stated that they were aged 17-25, 25 stated that they were aged 26-35, and 10 stated that they were 35 years old or older. The proportional equivalent of this was determined as 65% for the 17-25 age range, 25% for the 26-35 age range, and 10% for those over 35.

Of the 100 students who participated in the research, 50 were Iranian, 15 were Egyptian, 10 were Kazakh, 5 were Afghan, 5 were Lebanese, 5 were Syrian, 5 were Moroccan, and 5 were Algerian.

4.3. Data Collection Tool

In this study, the data collection process was carried out using two different tools. One of the tools is a test prepared with classical methods, and the other is a test developed using the Socrative artificial intelligence based digital tool. Each measurement tool is a test consisting of the same number of questions on the subject of the imperative mood.

4.3.1. Test Prepared with Classical Methods

This test was prepared by the researchers using traditional test development methods. The content of the test was specifically designed to measure knowledge and skills on the subject of the imperative mood. 15 questions of various types (multiple choice, fill in the blank, true/false) were prepared in line with the specified topics. The questions aimed to measure how much the students understood and could apply the imperative mood. The test was piloted on a small sample group and validity and reliability analysis was performed. Necessary revisions were made in line with the feedback. The test results were evaluated and analysed in line with the specified criteria. These criteria are that each question was assigned one point and the highest score that can be obtained from the test is 15.

4.3.2. Test Prepared with Socrative Artificial Intelligence Tool

Socrative is an artificial intelligence tool used to collect and evaluate data in the learning process. In this research, a test on the imperative mood was prepared using the Socrative tool. The features of the Socrative platform were examined and appropriate features were selected to meet the requirements of the test. Socrative was preferred for its ability to provide real-time feedback and support various test types. A test was created on Socrative using 3 different question types (multiple choice, short answer, matching) on the imperative mood. The content of the test was designed to complement the scope of the test prepared with the classical method. The test was presented to students through Socrative, and the results were collected instantly. Thanks to Socrative's automatic data collection and reporting features, the data was obtained quickly. The collected data was examined using the analysis tools provided by the

Socrative platform and the results were evaluated in detail. Each question was assigned one point and the highest score that can be obtained from the test is 15.

The data on the qualitative aspect of the research was obtained through focus semi-structured focus group interviews. As part of the focus semi-structured focus group interviews, students were asked the following questions: "What do you think about the use of technology in learning a foreign language? What tools did you use and how did they help you?" and "What are the advantages and disadvantages of learning a language with technological tools compared to traditional methods?" These questions were created by the researchers and approved by three field experts. Two of the field experts were from Marmara University and one was from Atatürk University.

4.4. Data Analysis

Quantitative data were analysed using SPSS software. Qualitative data were collected through a semi-structured focus group interviews and examined in detail using the content analysis method. The qualitative data were analysed through thematic analysis, and codes and overarching themes were identified based on students' responses. Content analysis is "a method used to provide an objective, measurable and verifiable explanation of the manifest content of messages" (Fiske, 1996). Krippendorff defines content analysis as "a research technique that makes repeatable and meaningful inferences from the data in a message" (Aziz, 2015). A normality test was conducted to evaluate the classical and Socrative test scores of students learning Turkish as a foreign language. In this analysis, it was determined that the skewness and kurtosis coefficients were between -1.5 and +1.5, and it was decided that parametric tests could be used in data analysis (George & Mallery, 2010). To ensure the quality of the qualitative findings, credibility was enhanced with systematic coding and theme check, transferability was attended to by accurately describing the setting of participants, dependability was ensured by having an explicit audit trail of the analysis process, and confirmability was ensured by grounded interpretations directly from the students' responses. Inter-rater reliability was established through Cohen's Kappa, and a result of 0.82 indicated high agreement between the two coders in applying the codes and identifying dominant themes in the students' responses.

Table 2

Standard Deviation, Skewness and Kurtosis Values of Groups

Groups			SD	Skewness	Kurtosis
Gender	Classical	Male	1.664	.239	-.791
		Female	1.527	-.290	-.550
	Socrative	Male	1.839	.021	-.914
		Female	2.206	-.652	-.978
Age	Classical	17-25	1.444	-.490	-.136
		26-34	1.697	1.230	.659
		35+	.527	-.050	-1.177
	Socrative	17-25	2.082	.309	.608
		26-34	1.850	-.485	-.832
		35+	.876	-.223	-1.134
Language	Classical	Arabic	1.222	-.016	-.085
		Kazakh	.816	.687	1.334

	Persian	1.556	-.193	-.804
	Arabic	2.318	.199	-1.388
Socrative	Kazakh	.738	-.166	-.734
	Persian	1.858	-.370	-1.035

5. Findings

5.1. Findings Related to the First Research Question

In order to understand whether the classical and Socrative test scores of students learning Turkish as a foreign language show a significant difference in terms of gender variable, the average scores of the groups were compared with the t-test in independent samples and the findings obtained are shown in Table 3. The classic test and the Socrative test are structurally different measurement tools. Due to different application conditions, the two measurements are considered independent. Therefore, the independent samples t-test was preferred.

Table 3

Comparison of the Classical and Socrative Test Scores of the Participants According to the Gender Variable

	Gender	*N*	*Mean*	*SD*	*t*	*Sig.*
Classical	Male	50	8.50	1.716	.262	.79
	Female	50	8.30	1.703		
Socrative	Male	50	11.50	2.068	.096	.92
	Female	50	11.40	2.547		

When Table 3 is examined, the results indicate that the average scores of the students learning Turkish as a foreign language in the classical and Socrative tests do not show a statistically significant difference in terms of gender, *t(98) = 0.26, p = .79 for the classical test, and t(98) = 0.10, p = .92* for the Socrative test. In other words, the average scores of the male and female participants in both the Classical and Socrative tests did not differ statistically.

5.2. Findings Related to the Second Research Question

In order to understand whether the classical and Socrative test scores of students learning Turkish as a foreign language show a significant difference in terms of the age range variable, the average scores of the groups were compared with the one-way ANOVA test and the findings are shown in Table 4.

Table 4

Comparison of the Classical and Socrative Test Scores of the Participants According to the Age Range Variable

		N	*Mean*	*SD*	*F*	*Sig.*
Classical	17-25	60	8.33	1.557	2.196	.143
	26-34	30	7.83	1.722		
	35+	10	10.50	.707		
Socrative	17-25	60	11.00	2.296	1.611	.229
	26-34	30	11.50	2.168		
	35+	10	14.00	.000		

When Table 4 is examined, the results indicate that the average scores of the students learning Turkish as a foreign language in the classical Test and the Socrative test do not show a

statistically significant difference in terms of the age range variable *F(2,97) = 2.20, p = .14* for the classical test, and *F(2,97) = 1.61, p = .23* for the Socrative test. In other words, the average scores of the participants in both the classical and Socrative tests are very similar in terms of age ranges.

5.3. Findings Related to the Third Research Question

In order to understand whether the classical and Socrative test scores of students learning Turkish as a foreign language show a significant difference in terms of the language variable, the average scores of the groups were compared with the One-Way ANOVA test and the findings are shown in Table 5.

Table 5

Comparison of the Classical and Socrative Test Scores of the Participants According to the Language Variable

		N	*Mean*	*SD*	*F*	*Sig.*
Classical	Arabic	35	7.57	1.813	1.396	.275
	Kazakh	10	9.00	.489		
	Persian	55	8.82	1.601		
Socrative	Arabic	35	10.00	2.646	2.782	.090
	Kazakh	10	11.50	.707		
	Persian	55	12.36	1.748		

When Table 5 is examined, the results indicate that the average scores of the Classical Test and Socrative Test of the students who learn Turkish as a foreign language do not show a statistically significant difference in terms of the language variable *(F(2,97) = 1.40, p = .27* for the classical test, and *F(2,97) = 2.78, p = .09* for the Socrative test. In other words, the average scores of the participants for both the classical and Socrative tests are very similar in terms of the language variable.

5.4. Findings Related to the Fourth Research Question

The average scores of students learning Turkish as a foreign language from the classical test and the Socrative test were compared with the t-test in independent samples and the findings are shown in Table 6. Table 6 presents data from 100 students who took both the classical and Socrative tests.

Table 6

Comparison of the Average Scores of Participants from the Classical and Socrative Tests

	Test	*N*	*Mean*	*SD*	*t*	*Sig.*
Points	Classical Test	100	8.40	1.667	-12,96	.00
	Socrative Test	100	11.45	2.259		

Upon examining Table 6, it is observed that there is a statistically significant difference in favour of the Socrative test between the average scores obtained by students learning Turkish as a foreign language on the traditional test and the Socrative test, according to the t-test results for related samples, t(99) = -12.96, p < .001. The average score students obtained on the Socrative test (x̄ = 11.45) is approximately 3.05 points higher than the average score they obtained on the traditional test (x̄ = 8.40). In order to determine the direction and strength of

the relationship between students' traditional test and Socrative test scores, the correlation between the scores obtained from the tests was examined, and the findings are presented in Table 7.

Table 7

Correlation of Participants' Scores from Classical and Socrative Tests

		Classical	*Socrative*
Classical	Pearson Correlation	1	.481[**]
	Sig. (2-tailed)		.000
	N	100	100
Socrative	Pearson Correlation	.481[**]	1
	Sig. (2-tailed)	.000	
	N	100	100

The level of the relationship between the variables can be interpreted as weak if the correlation coefficient is between 0-0.29; moderate if it is between 0.30-0.64; strong if it is between 0.65-0.84; and very strong if it is between 0.85-1 (Ural & Kılıç, p. 244). When Table 5 is examined, the results indicate that the scores that the students received from the classical and Socrative tests have a moderately positive relationship (r: ,481; p=,00<.05). This situation can also be seen in Table 7. Based on Table 7, it can be assumed that there is a linear relationship between the classical and Socrative test scores.

5.5. Findings Related to the Fifth Research Question

Within the scope of this research question, a semi-structured focus group interviews were applied to 10 randomly selected students. The answers given to the first interview question are interpreted by dividing them into themes below. The themes and subthemes were organized using a clear hierarchical structure, with main themes presented as primary headings and related subthemes listed beneath each to enhance readability and clarity.

Table 8

Data Obtained from Responses to the First Interview Question

IQ1: *"How do you view the use of technology in foreign language learning? Which tools have you used and how have they helped you?"*	
	Codes
P1, P2, P7: "I find the use of technology very useful. Language learning apps, especially Duolingo, helped me improve my language through daily repetition. Also, my vocabulary has expanded thanks to online dictionaries."	Language Learning Apps and Tools
P4, P5, P6: "The use of technology accelerated my learning process. Translation tools like Google Translate and Reverso supported me when I had difficulties. Also, I can evaluate my level thanks to online language exams."	Online Courses and Resources
P3, P8: "Online language exchange platforms were very useful. With the Tandem app, I was able to practice speaking with native Turkish speakers, which played a big role in improving my speaking skills."	Listening and Speaking Practice
P9, P10: "I used the checking feature in Word for written exercises, which helped me a lot in correcting my writing. I also join language learning communities and interact with other students on social media."	Developing Writing and Reading Skills

P1, P2, P7 accentuated the opportunities provided by the technological tools within language learning processes. It was underlined that language learning applications, especially Duolingo,

Memrise, and Babbel, provide important facilities in expanding the students' vocabularies and practicing the language besides facilitating interactive learning. The students indicate that these technological tools transform language learning into a fun, user-friendly process. This shows that technological tools create a motivating effect through individualization of the learning process.

P4, P5, P6 highlighted the flexibility and easy access to various resources with the response *"The use of technology accelerated my learning process. Translation tools like Google Translate and Reverso"*. Online platforms like Zoom, Quizlet, and Khan Academy are seen to be supportive for learning processes out of the classroom. However, students' need for teacher support reveals that the traditional approach is still very significant. This case shows that online courses are used only as a supportive tool in learning, and teacher guidance plays a critical role in it.

P3, P8 highlighted the advantages of technology in listening and speaking practices. Such tools include Spotify and LingQ, which actively contribute to the development of language skills. In such tools, students are able to increase their language use in real-life situations. This situation underlines the importance of the flexibility and accessibility that technological tools provide, especially in listening and speaking practices.

P9, P10 reported that technological tools contribute to the development of writing and reading skills. The spelling checker features in software tools like Microsoft Word and online language forums would significantly help students develop the ability of using written language with the response *"I also join language learning communities and interact with other students on social media"*. However, the lack of detailed information and personal feedback that the traditional methods provide them with in this regard can be counted as a field where technology is still lagging behind.

According to the answers given by the students, the results reflect that technological tools are generally rated positively in the language learning process, but they also have certain shortcomings compared to traditional methods. Technology provides time and space independent learning, interactive and motivating learning opportunities. However, the fact that technology does not possess elements such as face-to-face feedback, social interaction, and discipline that traditional methods provided reveals that it cannot totally replace traditional methods either. It follows from here that technological tools play only a supportive role in the learning process, and teacher guidance and structured learning processes retain their importance.

The answers to the second interview question are interpreted below by dividing them into themes.

Table 9

Data Obtained from Responses to the Second Interview Question

IQ2: *"What are the advantages and disadvantages of learning languages with technological tools compared to traditional methods?"*	
	Codes
P1, P3, P6: "Technological tools allow us to work independently of time and space in	Flexibility and Ease

language learning, which is a great advantage. However, the face-to-face feedback that traditional methods offer can be lacking, which can have a negative impact on the learning process."	of Access
P2, P4, P7: "Language learning apps make the learning process more interactive and fun. However, one-to-one teacher support and personal feedback are limited in traditional methods."	Interactive and Personalized Learning
P5, P8: "Technological tools are very useful, especially for listening and speaking practice. However, the lack of direct teacher supervision, as in traditional methods, can make it difficult to recognize mistakes."	Developing Language Skills
P9, P10: "Online learning spaces allow language learners to connect with each other. However, the opportunities for direct interaction and collaboration provided in a traditional classroom setting are limited on these platforms."	Social Interaction and Group Work

Another advantage which students consider is the flexibility and ease of access provided by technological tools. P1 and P3 find technology to be effective in language learning because it permits working independently of time and space and also ensures instant access to a wide range of materials. According to P6, the ability to record and re-watch lessons contributes to learning. However, in addition to the advantages, disadvantages such as a lack of face-to-face feedback and the reduction of discipline and interaction opportunities in the traditional classroom environment were also mentioned. In this case, while technology increases flexibility, it is also evident that, conversely, it cannot sufficiently meet the instant and direct feedback mechanisms provided by traditional methods. A very strong consensus prevails on the interactive and personalized nature of language learning through technological tools. While P2 explained, *"Language learning apps make the learning process more interactive"*, for P4 the personalized learning opportunities are what really count. According to P7, technology that enables one to work at an individual learning speed cannot replace the competitive and motivational elements existing within the traditional classroom environment. These answers highlight the importance of the way in which technology can shape learning according to individual needs but at the same time reveal limitations that may arise from lack of social interaction and motivation sources in this process.

The students consider the role of technological tools in developing the language skills to be an important advantage. While P5 draws attention to the opportunities offered by technology, especially with regard to listening and speaking practices, for P8, the ease of frequent repetition and reinforcement of what has been learned is underlined. Still, it was claimed in this context that it would be difficult to notice the mistakes due to the lack of direct teacher control and that it may not be sufficient to reach the deep knowledge provided by the traditional method. This shows that technology contributes significantly to the development of language skills, but the lack of guidance by a teacher and structured learning processes is also a disadvantage.

Criticism about technology use by students is also linked to social interaction and group work. While P9 values the opportunity to establish a global connection by means of online platforms, he criticizes the limited opportunities for direct interaction and collaboration provided in a traditional classroom environment. While P10 admits the advantage of access to a number of resources and materials, he reported that without guidance from the teacher and structured learning processes, the information tends to be disorganized with the *"Online learning spaces allow language learners to connect with each other. The opportunities for*

direct interaction and collaboration provided in a traditional classroom setting are, however, limited on these platforms".

This situation illustrates that technology brings some limitations in terms of social interaction and group work; therefore, the structured and guidance-based processes of traditional methods maintain their importance in language learning. The responses of the students have highlighted the advantages brought by the technological tools into the language learning processes like flexibility, ease of access, interactivity, and personalized learning opportunities. However, besides these advantages, the lack of face-to-face feedback, discipline, social interaction, and structured guidance processes provided by traditional methods also come to the fore as important disadvantages. This result means that even though it is possible to consider technological tools as important supporting tools in language learning, they are never capable of replacing traditional methods entirely. Therefore, using technology together with traditional methods in language learning processes can make it possible to benefit from the strengths of both approaches.

Conclusion and Suggestions

The study aimed to determine how effective the artificial intelligence tool called Socrative is in terms of measurement and evaluation in teaching Turkish as a foreign language. The research results showed that there was a significant difference between classical tests and tests created using Socrative, to the advantage of Socrative. The study has several limitations that should be acknowledged. The sample was limited to B1-level students, which may restrict the generalizability of the findings. Additionally, the study evaluated only one digital assessment tool (Socrative), which may affect the robustness of the results. There is no statistically significant difference between male and female participants in terms of classical and Socrative test scores.

Similar results are reached when the literature is reviewed. Awedh et al. (2015) investigated the effect of using Socrative with smartphones on students' learning performance; they observed the benefits of interaction between teachers and students and among classmates, which positively affect collaborative learning and students' participation in the lesson. The results of the study reveal that collaborative learning and student participation in the lesson increase students' learning performance.

El Shaban (2017) investigated the perceptions of English as a second language learners of Socrative integrated with active learning activities regarding the use of this tool; it was found that it contributed to increasing the level of students' participation, supported their critical thinking and encouraged collaboration.

Farrow (2016) states that the Internet enables access to information and digital content for reuse in other contexts, which poses significant ethical challenges. Open educational practices are based on four basic principles: free access to information, collaboration, rewarding sharing, and community support for educational innovation (Valverde, 2010). Two of the main issues in Digital Ethics resulting from the widespread diffusion of ICT are privacy and protection of intellectual property, which can be included as part of ethics and digital competences (Maggiolini, 2014). In contrast, students at various educational levels, including

prospective teachers, need to acquire such competences. New teachers, who need to be equipped with 21st century competences, must have a certain level of use in order to integrate the digital world into the classroom environment and, at the same time, must act in a way that is compatible with theoretical knowledge (Buguet & Buxarrais, 2013; García-Gutiérrez, 2013). One of the techniques that can be used in the digital classroom environment is gamification. The main advantages of gamification in the educational context are affects students' behavior, commitment and motivation, which can lead to the development of knowledge and skills (Hsin Yuan Huang & Soman, 2013). Socrative also allows for spontaneous polling by choosing between multiple choice, true/false, and short answer question formats. It allows for the preparation of practice questions for mandatory certification exams, student control over critical thinking questions that allow students to respond anonymously, formal assessment checkpoints, and opportunities for review of content material (Wash, 2014).

In his study, Dakka (2015) stated that using student-paced assessment using Socrative increased student performance. The results showed that 53% of students improved their performance, while 23% neither improved nor underperformed. Qualitative data showed that students felt an improvement in their learning experience. Pryke (2020) also stated that Socrative tests increased students' academic performance.

Guarascio et al. (2017) evaluated Socrative in terms of 'traditional student response systems', that is, the teacher asking the class questions and students raising their hands to answer. Dervan (2014) examined the use of Socrative as an alternative to handheld clickers. In both cases, the comparisons are in favor of Socrative: Socrative is better; it constitutes an improvement. In conclusion, despite its limitations, this study provides valuable insights into the comparative effectiveness of classical and Socrative-based assessments for B1-level Turkish learners, highlighting both the benefits of digital testing and areas for future research. Overall, the findings suggest that Socrative can be a more effective and engaging assessment tool than traditional tests for B1-level Turkish learners, while providing insights for future research on digital assessment in language education.

The recommendations prepared based on the research results are as follows:

- Educational institutions can modernize their assessment and evaluation processes by integrating AI tools.
- In order to fully benefit from the effectiveness of tools such as Socrative, training programs should be organized for instructors and trainers.
- Socrative and other AI-based tools should be integrated by instructors in a way that is compatible with existing traditional exam methods. This integration can create more comprehensive and effective assessment processes by combining the advantages of both methods.
- The effects of AI-based tools should be examined on a wider group of students and different language levels by curriculum planners and researchers. In addition, studies should be conducted to compare different AI-based tools and evaluate their long-term effects.

References

Akata, A. (2009). *Türkçe programıyla ilgili ölçme ve değerlendirme sürecinin işlevselliği üzerine bir araştırma* [A study on the functionality of the assessment and evaluation process regarding the Turkish program] [Unpublished master's thesis]. Abant Izzet Baysal University.

Akçadağ, T. (2010). Öğretmenlerin ilköğretim programındaki yöntem teknik ölçme ve değerlendirme konularına ilişkin eğitim ihtiyaçları [Educational needs of teachers regarding method, technique, measurement and evaluation issues in the primary education curriculum]. *Bilig, 53*(1), 29-50.

Aktürk, A. O. (2012). Öğretmen adaylarının öğrenmeye ilişkin tutumlarının öğretmenlik mesleğini tercih etme nedenlerine ve akademik başarılarına göre incelenmesi [Investigation of pre-service teachers' attitudes towards learning according to their reasons for preferring the teaching profession and their academic achievements]. *Ahi Evran University Journal of Kırşehir Education Faculty, 13*(3), 283-297.

Arslan, K. (2020). Eğitimde yapay zekâ ve uygulamaları [Artificial intelligence in education and its applications]. *Batı Anadolu Eğitim Bilimleri Dergisi, 11*(1), 71-88.

Awedh, M., Mueen, A., Zafar, B., & Manzoor, U. (2015). Using Socrative and Smartphones for the support of collaborative learning. *International Journal on Integrating Technology in Education (IJITE) 3*,(4), 17-24.

Aziz, A. (2015). *Sosyal bilimlerde araştırma yöntemleri ve teknikleri* [Research methods and techniques in social sciences]. Nobel.

Baki, A., & Gökçek, T. (2012). Karma yöntem araştırmalarına genel bir bakış [A general overview of mixed methods research]. *Electronic Journal of Social Sciences, 11*(42), 1-21.

Buguet, M., & Buxarrais, M. R. (2013). La eticidad de las TIC. Las competencias transversales y sus paradojas. *Revista Teoría de la Educación: Educación y Cultura en la Sociedad de la Información, 14*(3), 87-100.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Sage.

Çeçen, M. A. (2011). Türkçe öğretmenlerinin seviye belirleme sınavı ve Türkçe sorularına ilişkin görüşleri [Turkish language teachers' views on the placement test and Turkish questions]. *Mustafa Kemal University Journal of Social Sciences Institute, 8*(15), 201-211.

Dervan, P. (2014), Increasing in-class student engagement using Socrative (An online student response system). *The All Ireland Journal of Teaching and Learning in Higher Education 6*(3), 1801-1813.

El Shaban, A. (2017). The use of Socrative in ESL classrooms: Towards active learning. *Teaching English with Technology*, *17*(4), 64-77.

Farrow, R. (2016). A framework for the ethics of open education. *Open Praxis, 8(*2), 93-109.

Fiske, J. (1996). *İletişim çalışmalarına giriş* [Introduction to communication studies] (S. İrvan, Trans.). Bilim ve Sanat.

García-Gutiérrez, J. (2013). Aproximación ética a la competencia digital. Los niveles de uso y sentido en ámbitos educativos virtuales. *Revista Teoría de la Educación: Educación y Cultura en la Sociedad de la Información, 14*(3), 121-145.

Gelbal, S., & Kelecioğlu, H. (2007). Öğretmenlerin ölçme ve değerlendirme yöntemleri hakkındaki yeterlik algıları ve karşılaştıkları sorunlar [Teachers' proficiency perceptions of measurement and evaluation methods and the problems they encounter]. *Hacettepe University Journal of Education, 33*, 135-145.

George, D., & Mallery, M. (2010). *SPSS for Windows step by step: A simple guide and reference* (17.0 update). Pearson.

Gipps, C., & Stobart, G. (2003). Alternative assessment. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 549–575). Kluwer Academic.

Gondal, K. M. (2018). Artificial intelligence and educational leadership. *Annals of King Edward Medical University, 24*(4), 1-2.

Guarascio, A. J., Nemecek, B. D., & Zimmerman, D. E. (2017). Evaluation of students' perceptions of the Socrative application versus a traditional student response system and its impact on classroom engagement. *Currents in Pharmacy Teaching and Learning, 9*(5), 808–812.

Hancock, D. R., & Algozzine, B. (2006). *Doing case study research: A practical guide for beginning researchers*. Teachers College Press.

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

Hsin Yuan Huang, W., & Soman, D. (2013). *A practitioner's guide to gamification of education*. University of Toronto. http://inside.rotman.utoronto.ca/behaviouraleconomicsinaction/files/2013/09/GuideGamificationEducationDec2013.pdf

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14-26.

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Sage.

Maggiolini, P. (2014). A deep study on the concept of digital ethics. *Revista de Administração de Empresas, 54*(5), 585-591.

Metin, M., & Demiryürek, G. (2009). Türkçe öğretmenlerinin yenilenen Türkçe öğretim programlarının ölçme-değerlendirme anlayışı hakkındaki düşünceleri [Turkish

teachers' views on measurement and evaluation approach of the new Turkish curriculum]. *Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi, 28*, 37-51.

Moreland, J., & Jones, A. (2000). Emerging assessment practices in an emergent curriculum: Implications for technology. *International Journal of Technology and Design Education, 10*, 283-305.

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence, 2*, 100041.

Obschonka, M., & Audretsch, D. B. (2020). Artificial intelligence and big data entrepreneurship: A new era has begun. *Small Business Economics, 55*, 529-539.

Pierce, L. V., & O'malley, J. M. (1992). *Performance and portfolio assessment for language minority students*. National Clearing House for Bilingual Education.

Pryke, S. (2020). The use of Socrative in university social science teaching. *Learning and Teaching, 13*(1), 67-86.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4-14.

Tashakkori, A., & Teddlie, C. (1998). Mixed methodology: Combining qualitative and quantitative approaches. *Applied Social Research Methods Series (Vol.46)*. Sage.

Ural, A., & Kılıç, İ. (2013). *Bilimsel araştırma süreci ve SPSS ile veri analizi* [Scientific research process and data analysis with SPSS]. Detay.

Valverde, J. (2010). El movimiento de "educación abierta" y la "universidad expandida" [The "open education" movement and the "expanded university"]. *Tendencias Pedagógicas, 16*, 157-180.

Van Niekerk, E. F. (2002). *'n Prosesgebaseerde assesseringsraamwerk vir Tegnologie-onderwys: 'n Gevalstudie* [A process-based assessment framework for technology education: A case study] [Unpublished master's thesis]. Randse Afrikaanse Universiteit.

Van Niekerk, E., Ankiewicz, P., & de Swardt, E. (2010). A process-based assessment framework for technology education: A case study. *International Journal of Technology and Design Education, 20*, 191–215. https://doi.org/10.1007/s10798-008-9070-8

Wash, P. D. (2014). Taking advantage of mobile devices: Using Socrative in the classroom. *Journal of Teaching and Learning With Technology, 3*(1), 99–101. https://doi.org/10.14434/jotlt.v3n1.5016

Wiredu, J. (2023). An investigation on the characteristics, abilities, constraints, and functions of artificial intelligence (AI): The age of ChatGPT as an essential. *Information and Management, 108*(3), 62614-62620.

Yiğit, F., & Kırımlı, B. (2014). Türkçe öğretmenlerinin alternatif ölçme-değerlendirme yöntemlerini uygulama biçimleri ve uygulamada karşılaştıkları sorunlar [Turkish teachers' styles of applying alternative measurement-evaluation methods and problems encountered in practice]. *Turkish Studies - International Periodical For The Languages, Literature and History of Turkish or Turkic, 9*(3), 1621-1639.

Genişletilmiş Özet

Yapay zekâ, son yıllarda önem kazanmış ve hayatın birçok alanında kullanılmaya başlamıştır. Yapay zekâ; insan bilişini algılama, akıl yürütme, kavrama, anlamlandırma, genelleştirme, çıkarımda bulunma, öğrenme ve aynı anda birden fazla görevi başarıyla yerine getirme yeteneklerine sahip bir bilişim teknolojisi olarak tanımlanmaktadır (Gondal, 2018). Bir diğer ifadeyle yapay zekâ, karmaşık problemlerin çözümünde ve değişen koşullar göz önünde bulundurularak karar verilmesinde, insan zekâsına özgü olan muhakeme yeteneği ve tahmin gücünün makineler tarafından kullanılması anlamına gelmektedir (Obschonka & Audretsch, 2020).

Holmes ve diğerlerine (2019) göre yapay zekâ; okul yönetiminde müfredat planlaması, personel düzenlemeleri, sınavların yönetimi, siber güvenlik, tesis yönetimi ve güvenlik gibi alanlarda yardımcı olabilmekte ve eğitim süreçlerine dolaylı katkılar sağlayabilmektedir. Eğitimde yapay zekâyı kullanan araştırmacılar, çoktan seçmeli testlerin öğrencilerin başarı düzeylerini değerlendirmede ve öğretim sürecini güçlendirmede etkili bir araç olarak kullanılabileceğini belirtmektedir.

Süreç temelli ölçme ve değerlendirme; eğitim sürecindeki her aşamada ortaya çıkabilecek sorunları tespit etme ve bunları düzeltme imkânı sağlamaktadır. Buna ek olarak öğrencilerin kazanım düzeylerini belirleyerek gerektiğinde zamanında ve yerinde müdahalelerde bulunmayı mümkün kılmaktadır (Akçadağ, 2010; Çeçen, 2011). Süreç odaklı değerlendirme yöntemleri ayrıca öğrencilerin bilgiyi aktif bir şekilde yapılandırmalarına imkân tanımaktadır (Shepard, 2000).

Bu araştırma bir durum çalışmasıdır. Creswell (2003), durum çalışmalarını, araştırmacının belirli bir zaman diliminde veya dar bir bağlamda, tek bir ya da birden fazla durumu çeşitli veri toplama teknikleri (görüşme, gözlem, belge, görsel materyal vb.) kullanarak kapsamlı bir şekilde analiz ettiği ve bu süreçte ilgili temaların ve özelliklerin ortaya konulduğu bir nitel araştırma yaklaşımı olarak tanımlar. Hancock ve Algozzine (2006) ise durum çalışmasını, olayların doğal ortamlarında, zaman ve mekân sınırlamaları göz önünde bulundurularak farklı veri toplama araçlarıyla detaylı bir şekilde incelenmesi ve açıklanması olarak ifade etmektedir. Nicel veriler, SPSS yazılımı kullanılarak analiz edilmiştir. Nitel veriler ise yarı yapılandırılmış görüşme formu aracılığıyla toplanmış ve içerik analizi yöntemiyle ayrıntılı olarak incelenmiştir. İçerik analizi, "iletilerin açık içeriğinin nesnel, ölçülebilir ve doğrulanabilir bir açıklamasını sağlamak amacıyla kullanılan bir yöntemdir" (Fiske, 1996). Krippendorff (2004'ten aktaran Aziz 2015) ise içerik analizini, "bir mesajdaki verilerden tekrarlanabilir ve anlamlı çıkarımlar yapan bir araştırma tekniği" olarak tanımlamaktadır.

Türkçeyi yabancı dil olarak öğrenen öğrencilerin klasik ve Socrative test puanlarının değerlendirilmesi amacıyla normallik testi yapılmıştır. Yapılan bu analizde çarpıklık ve basıklık (skewness ve kurtisos) katsayılarının -1,5 ile +1,5 arasında oldukları belirlenmiş ve veri analizinde parametrik testlerin kullanılabileceğine karar verilmiştir (George & Mallery, 2010).

Araştırmada Socrative adlı yapay zekâ aracının yabancı dil olarak Türkçe öğretiminde ölçme ve değerlendirme bağlamında ne kadar etkili olduğu saptanmaya çalışılmıştır. Araştırma sonuçları; klasik testler ile Socrative kullanılarak oluşturulan testler arasında, Socrative'in avantajına anlamlı bir fark bulunduğunu göstermiştir.

Klasik ve Socrative test puanları açısından erkek ve kadın katılımcılar arasında istatistiksel olarak anlamlı bir fark bulunmamaktadır. Her iki test için erkek ve kadın öğrencilerin ortalama puanları birbirine yakındır. Yaş grupları arasında klasik ve Socrative test puanları açısından anlamlı bir fark yoktur. 17-25, 26-34 ve 35+ yaş gruplarındaki katılımcıların puan ortalamaları birbirine benzemektedir. Farklı dillerden (Arapça, Kazakça, Farsça) gelen öğrencilerin klasik ve Socrative test puanları açısından da anlamlı bir fark tespit edilmemiştir. Her dil grubunun ortalama puanları birbirine yakın değerler göstermektedir. Klasik test ve Socrative test puanları arasında istatistiksel olarak anlamlı bir fark bulunmaktadır; öğrenciler Socrative testte daha yüksek ortalama puanlar almışlardır. Bu fark yaklaşık 3 puan olarak belirlenmiştir. Klasik ve Socrative test puanları arasında orta derecede pozitif bir ilişki vardır (korelasyon katsayısı 0,481). Bu, iki testten alınan puanların birbirini etkilediği anlamına gelmektedir.

Alan yazını tarandığında benzer sonuçlara ulaşılmaktadır. Awedh ve diğerleri (2015) Socrative'in akıllı telefonlarla birlikte kullanımının öğrencilerin öğrenme performansı üzerindeki etkisini araştırmışlar; öğretmen, öğrenci ve sınıf arkadaşları arasında etkileşimin, işbirlikçi öğrenmeyi ve öğrencilerin derse katılımını olumlu yönde etkileyen faydalarını gözlemlemişlerdir. Çalışmanın sonuçları; işbirlikçi öğrenmenin ve öğrencinin derse aktif katılımının öğrenme performansını artırdığını ortaya koymaktadır.

El Shaban (2017) aktif öğrenme aktiviteleri ile entegre edilmiş olan Socrative'in İngilizceyi ikinci dil olarak öğrenenlerin bu aracın kullanımına ilişkin algılarını araştırmış; öğrencilerin katılım düzeyini artırmaya katkıda bulunduğunu, eleştirel düşünmelerini desteklediğini ve iş birliğini teşvik ettiğini tespit etmiştir.

Socrative; soru formatını çoktan seçmeli, doğru/yanlış ve kısa yanıt arasından seçerek kendiliğinden yoklama yapılmasına da olanak tanımaktadır. Zorunlu sertifika sınavları için alıştırma sorularının hazırlanmasına, öğrencilerin anonim olarak yanıt vermesine olanak tanıyan eleştirel düşünme soruları üzerinde kontrollerine, resmî değerlendirme kontrol noktalarına ve içerik materyalinin gözden geçirilmesi fırsatlarına imkân sağlamaktadır (Wash, 2014). Araştırma kapsamında geliştirilen öneriler şu şekildedir:

- Socrative gibi araçların etkinliğinden tam olarak yararlanabilmek için, öğreticilere ve eğitmenlere yönelik eğitim programları düzenlenmelidir.
- Socrative ve diğer yapay zekâ araçlarının, mevcut geleneksel sınav yöntemleriyle uyumlu bir şekilde sürece entegre edilmesi sağlanmalıdır. Bu entegrasyon, her iki yöntemin avantajlarını bir araya getirerek daha kapsamlı ve etkili değerlendirme süreçleri oluşturabilir.
- Yapay zekâ araçlarının etkisinin daha geniş öğrenci grupları ve farklı dil seviyeleri üzerinde incelenmesi gerekmektedir. Ayrıca, farklı yapay zekâ araçlarının karşılaştırıldığı ve uzun vadeli etkilerinin değerlendirildiği çalışmalar yapılmalıdır.