International Journal of Excellent Leadership (IJEL)

www.ijel.org

Investigation of Teacher Attitudes Towards Artificial Intelligence -Based Measurement and Evaluation in Some Variables[1]

İbrahim Karamanlıoğlu[2], Afet Sözen Özden[3], Bertan Akyol[4]

Abstract

In this article, it was aimed to determine whether teachers' "artificial intelligence (AI)-based measurement and evaluation" attitudes differ by socio-demographic characteristics. The sample of this study consisted of 339 teachers who were actively working in the 2023-24 academic year and who had experienced AI-based measurement and evaluation tools and methods at least once. As a result of factor analysis, teachers' attitudes towards AI-based measurement and evaluation were categorized under 4 factors: "Giving importance", "Being active", "Developing negative emotions" and "Experience". It was determined that the attitudes of the teachers attending in the study towards giving importance to AI-based measurement and evaluation differed according to their "age", "professional seniority" and "in-service training on AI-based measurement and evaluation". It was determined that their attitudes towards being active differed according to their "age", "the level they worked at" and "the type of school". It was determined that the attitudes of developing negative emotions differed according to their "gender", "age", "professional seniority", and "in-service training on AI-based measurement and evaluation". It was determined that experience attitudes differed according to their "age" and "in-service training on AI-based measurement and evaluation". This study will raise awareness about the importance of AI-based measurement and evaluation in the digital age and teachers' attitudes towards AI-based measurement and evaluation.

Keywords: *Measurement and Evaluation, Artificial Intelligence, Teacher, Attitude*

Article History: Received: May 18, 2024 Accepted: June 28, 2024 Published: June 30, 2024

Article Type Research Article

Recommended Citation: Karamanlıoğlu, İ. & Sözen Özden, A. & Akyol, B. (2024). Investigation of Teacher Attitudes Towards Artificial Intelligence-Based Measurement and Evaluation in Some Variables. *International Journal of Excellent Leadership (IJEL), 4*(1), 40-53

[1] This study was presented as an oral presentation at the IV. International Congress on Excellence in Education.

[2] Corresponding author: School Manager, Fevzipaşa Akbudak Primary School, Aydın/Türkiye, i.ka09@hotmail.com , ORCID: 0000-0002-4779-8174

[3] Independent Scientist, afetsozen@gmail.com , ORCID: 0000-0001-9821-7029

[4] Prof. Dr., Aydın Adnan Menderes University, Aydın/Türkiye, bertan.akyol@adu.edu.tr ORCID: 0000-0002-1513-1885

Introduction

Measurement and evaluation involves the process of determining the quantities of certain characteristics, expressing them in numbers and making sense of these expressions. The importance of measurement-evaluation in education in the digital age stems from its potential to make learning processes more effective, efficient and fair. In this process, measurement and evaluation must be strong, relevant, valid, reliable and focused on strengthening learning. Today, measurement and evaluation in the field of education has undergone a great transformation with the integration of technological innovations and digital tools. This transformation has made teaching and learning processes more effective, accessible and student-focused. This situation has brought new responsibilities to teachers and revealed the need for teachers with sufficient digital measurement and evaluation knowledge in our age. One of the digital measurement and evaluation tools used in the field of education is "Artificial Intelligence (AI)" based tools and methods. AI-based tools and methods are now used in many sectors, including education (Suh and Ahn, 2022), to increase the efficiency and quality of services provided to students and teachers. Vasconcelos and dos Santos (2023) state that AI tools enable students to develop their "reflective" and "critical" thinking skills, reveal their creativity, acquire problem-solving skills and understand concepts effectively, while Huang (2018) states that the integration of AI into teaching enables student-centered learning and is an important step in its effective implementation.

AI-based tools and methods it improving measurement in education, including measurement and testing. These tools and methods can provide teachers with valuable information about students' learning outcomes, performance, and teaching effectiveness. Nazaretsky et al. (2022) stated this as follows: "AI-based assessment tools can analyze students' responses to assignments and provide personalized feedback to help students identify their strengths and weaknesses." In this process, teachers' perspectives on AI-based measurement and evaluation tools and methods are important for the process.

"Measurement and evaluation" is a critical stage of the teaching and learning process. According to Stiggins (2014), teachers spend almost one third of their professional time on measurement and evaluation practices. This points to the importance of measurement-evaluation in particular, and hence the critical role of technological developments and the use of AI in measurement and evaluation in teaching and learning. Research shows that, although extremely important, teachers in general need more support and training than is currently available for measurement and evaluation, and new generation measurement tools/methods that focus on AI-based measurement (Tsagari, 2011; Mede and Atay, 2017).

Even though teachers have sufficient knowledge about the concepts and practices related to measurement and evaluation and know how to use measurement and evaluation effectively to improve teaching and learning, this does not mean that teachers are effective evaluators in the digital environment. Therefore, it has become inevitable for teachers to have skills related to measurement and evaluation practices in digital environment. For this reason, determining teachers' attitudes towards AI-based tools and methods in measurement and evaluation practices carried out in the educational environment is very important in terms of keeping up with the age, saving time in terms of evaluation, increasing interaction, and being able to make individual and mass evaluations.

In this perspective, the research's problem statement was determined as "Do teachers' attitudes towards AI-based measurement and evaluation differ according to their sociodemographic characteristics?". The sub-problems created to answer the main problem statement are as follows:

1. "Do teachers' attitudes towards AI-based measurement and evaluation differ according to their gender?"
2. "Do teachers' attitudes towards AI-based measurement and evaluation differ according to their ages?"
3. "Do teachers' attitudes towards AI-based measurement and evaluation differ according to their professional seniority?"

4. "Do teachers' attitudes towards AI-based measurement and evaluation differ according to the level they work at?"
5. "Do teachers' attitudes towards AI-based measurement and evaluation differ according to the type of school they work in?"
6. "Do teachers' attitudes towards AI-based measurement and evaluation differ according to taking measurement and evaluation courses at university?"
7. "Do teachers' attitudes towards AI-based measurement and evaluation differ according to their in-service training on measurement and evaluation?"
8. "Do teachers' attitudes towards AI-based measurement and evaluation differ according to their in-service training on AI-based measurement and evaluation?"

It is clear that the use of measurement-assessment AI applications in the area of education will have benefits in the future in terms of time, labor force and objectivity of the measurement and evaluation results obtained. In this context, the study aimed to find out whether teachers' attitudes about "AI-based measurement and evaluation" differ by teachers' socio-demographic characteristics.

Considering the purpose of this study under all this information, a study to determine teachers' attitudes towards "AI-based measurement and evaluation" is not available in the literature to the best of our knowledge. In this respect, the research is important in terms of contributing to the literature, educators and policy makers. In addition, a scale was used to find teachers' attitudes about AI-based measurement and evaluation. This scale is the "Measurement and Evaluation Attitude Scale for Teachers" developed by Tezci (2019) and adapted for AI-based measurement and evaluation and reliability analysis was performed. In this respect, it is thought that the inclusion of the scale as "AI-based Measurement and Evaluation Attitude Scale for Teachers" will contribute to both the literature and researchers. As a result, this study will raise awareness about the importance of AI-based measurement and evaluation in the digital age and teachers' attitudes towards AI-based measurement and evaluation. It is also expected that this study will respond to teachers' perceived needs in AI-based measurement and evaluation and thus help schools within the "Ministry of National Education" in Türkiye to take action to train teachers on this subject if necessary.

Method

Research Model

This research is a quantitative research and research model is a survey model. According to Christensen et al. (2015: 368), "the survey model aims to reveal the changes that occur over time or the inner face of a particular situation". Karasar (2012) stated that the survey model aims to determine the occurrence of variables separately, as type, variety or quantity.

Research Population and Sample

The population of this study included of teachers who were actively working in the 2023-24 academic year and who had experienced AI-based measurement and evaluation tools and methods at least once. In this context, it is not possible to reach the entire research universe. For this reason, random sampling method was used to reach the research sample. The sample of this study included of 339 teachers who were actively working in the 2023-2024 academic year and who had experienced AI-based measurement and evaluation tools and methods at least once. The distribution of the sample according to socio-demographic characteristics is given in Table 1.

Table 1.
Frequency-Percentage Value According to Socio-demographic Characteristics of Teachers

Variable	Category	Frequency (n)	Percent (%)
Gender	"1: Female"	250	73,7
	"2: Male"	89	26,3
Age	"1: 21-25 years old"	6	1,8
	"2: 26-30 years old"	28	8,3
	"3: 31-35 years old"	134	39,5
	"4: 36-40 years old"	96	28,3
	"5: 41-45 years old"	44	13,0
	"6: 46-50 years old"	24	7,1
	"7: 51 years and above"	7	2,1
Duration of professional seniority	"1: Less than 1 year"	19	5,6
	"2: 1-5 years"	15	4,4
	"3: 6-10 years"	110	32,4
	"4: 11-15 years"	108	31,9
	"5: 16-20 years"	38	11,2
	"6: 21 years and above"	49	14,5
Level of school attended	"1: Pre-school"	38	11,2
	"2: Primary school"	83	24,5
	"3: Middle school"	165	48,7
	"4: High school"	53	15,6
Type of school served	"1: Public school"	312	92,0
	"2: Private school"	27	8,0
Taking a measurement and evaluation course at the university	"1: Yes"	324	95,6
	"2: No"	15	4,4
Receiving in-service training on measurement and evaluation	"1: Yes"	163	48,1
	"2: No"	176	51,9
Receiving in-service training on AI-based measurement and evaluation	"1: Yes"	31	9,1
	"2: No"	308	90,9

Table 1 shows that 73.7% of the teachers in the study were female and 26.3% were male. When the age ranges were analyzed, it was determined that the most (n=134, 39.5%) teachers between the ages of 31-35 participated in the study and the least (n=6, 1.9%) teachers between the "ages of 21-25" participated in the study. In terms of professional seniority, 32.4% of the teachers in the study had 6-10 years of seniority, 31.9% "11-15 years of seniority", 14.5% "21 years and above", 11.2% "16-20 years of seniority", 5.6% had "less than 1 year" and 4.4% "1-5 years of seniority". It was determined that 48.7% of the teachers participating in the study worked in secondary school, 24.5% in "primary school", 15.6% in "high school" and 11.2% in "pre-school". While 92% of the teachers in the research stated that they work in public schools, 8% stated that they worked in private schools. Among the teachers attending the study, 95.6% reported that they "took a measurement and evaluation course at the university", while 4.4% reported that they "took no measurement and evaluation course". While 48.1% of the teachers "took in-service training on measurement and evaluation", 51.9% reported that they "took no in-service training on measurement and evaluation". Finally, 9.1% of the teachers reported that they "took in-service training on AI-based measurement and evaluation", while 90.9% reported that they "took no in-service training on AI-based measurement and evaluation".

Data Collection Tool and Data Analysis

The "Measurement and Evaluation Attitude Scale for Teachers" developed by Tezci (2019) was used in the study. This scale developed by Tezci (2019) was organized by the researchers in accordance with AI-based measurement and evaluation within the scope of the research topic. Data were collected online via Google Forms. Factor analysis was applied to the scale used in the study and similar results were obtained with the factor analysis conducted by Tezci (2019) and it was determined that the 22-item scale consisted of 4 factors. The results of the factor analysis are shown in Table 2.

Table 2.
Results Related to Factor Analysis

	Factor 1	Factor 2	Factor 3	Factor 4
M5	,865			
M3	,823			
M8	,798			
M6	,798			
M4	,697			
M12	,682			
M11	,641			
M7	,633			
M13		,861		
M21		,850		
M20		,849		
M22		,829		
M10		,810		
M18		,792		
M15			,805	
M16			,790	
M17			,775	
M14			,724	
M1				,757
M2				,640
M19				,603
M9				,517

KMO: ,831 Bartlett's Test Statistic: 1791,511 p=0,000

The KMO and Bartlett test values in Table 2 show that the data obtained in the study are suitable for factor analysis. According to factor analysis, it was found that the 22 scale items used in the research could be grouped under 4 factors, and it was observed that the scale items in each factor were found with the results of the study by Tezci (2019). Therefore, each factor was named as in Tezci (2019) study. In this context, it was determined that the scale statements were grouped under 4 factors: "Giving importance", "Being active", "Developing negative emotions" and "Experience". "Giving importance" factor expresses the importance teachers give to AI-based measurement and evaluation. Under this factor, teachers are presented with statements regarding their intention to use tools and methods related to AI-based measurement and evaluation, and the benefits and importance of these new methods. "Being active" factor refers to teachers' active participation in the process of AI-based measurement and evaluation. Under this factor, statements regarding teachers' active use of these new measurement and evaluation tools and methods as part of the digital transformation process in education were collected. "Developing negative emotions" factor refers to teachers' negative emotional states regarding the process related to AI-based measurement and evaluation. Under this factor, teachers were presented with statements regarding their fears and concerns about loss of control related to AI-based measurement and evaluation. "Experience" factor refers to teachers' past experiences with AI-based measurement and evaluation. Under this factor, teachers were presented with statements regarding their knowledge levels and experiences with AI-based measurement and evaluation. Hypotheses were used to answer the research questions under these 4 factors. In order to choose which hypothesis tests to use, the normal distribution of the data was tested with the "Kolmogorov-Smirnov (KS) test" and the result is given in Table 3.

Table 3
Results of K-S Normality Test

	Test Statistic	Degrees of Freedom	P-value (p)
Giving importance (Factor 1)	,093	339	,000
Being active (Factor 2)	,123	339	,000
Developing negative emotions (Factor 3)	,140	339	,000
Experience (Factor 4)	,091	339	,000

* p<0,05

It was determined that the attitude scores related to the 4 factors examined in the study were not show normal distribution and non-parametric hypothesis tests were used to answer the research questions. Statistical significance level $p<0.05$ was accepted. IBM SPSS 28.0 program was used to analyze the data obtained. Frequency-percentage values were used to determine the distribution of the teachers attending the research according to their sociodemographic characteristics.

Validity and Reliability

A pilot study was conducted for the validity-reliability of the prepared scale statements. Reliability for the factors and the whole scale was evaluated with Cronbach's Alpha coefficient. Statistics related to Cronbach's Alpha coefficient are given in Table 4.

Table 4
Results of Reliability Analysis

Scale	Cronbach's Alfa Coefficient	Number of items
All items	0,851	22
Giving importance	0,721	8
Being active	0,827	6
Developing negative emotions	0,824	4
Experience	0,795	4

Table 4 shows that the reliability coefficients ranged between 0.721-0.851 for the factors. It was determined that this value was 0.851 for the whole scale. Since these values are above 0.70, it can be said that the scale and sub-factors used in the study are reliable in terms of internal consistency.

Findings

In the study, "Mann-Whitney U test", was used to test whether the attitudes of the teachers participating in the study towards AI-based measurement and evaluation differ by gender. Related findings are given in Table 5.

Table 5
Differences Teacher Attitudes Towards AI-Based Measurement and Evaluation by Gender

Factor	Gender	N	Mean Rank.	Test Statistic	p
Mean attitude score for Giving importance	"1: Female"	250	163,79	U=9571,500	0,050
	"2: Male"	89	187,46	Z=-1,963	
Mean attitude score for being active	"1: Female"	250	172,92	U=10395,500	0,357
	"2: Male"	89	161,80	Z=-,921	
Mean attitude score for developing negative emotions	"1: Female"	250	163,49	U=9497,000	0,039*
	"2: Male"	89	188,29	Z=-2,067	
Mean attitude towards the experience	"1: Female"	250	166,64	U=10285,500	0,287
	"2: Male"	89	179,43	Z=-1,064	

As seen in Table 5, it was found that the probability values of the "Attitude towards developing negative emotions" points of the teachers attending in the study towards AI-based measurement and evaluation are less than 0.05 significant value ($p<0.05$). It was seen that the probability values of the teachers' "Attitude towards Giving importance", "Attitude towards being active" and "Attitude towards experience" scores were greater than 0.05 significant value ($p>0.05$). For this reason, it was determined that the "Developing negative emotions" attitudes of the teachers participating in the research regarding AI-based measurement and evaluation differed according to their gender, while the "Giving importance", "Being active" and "Experience" attitudes did not differ according to the gender of the teachers. When the results of the attitude of developing negative emotions were examined, it was determined that male teachers developed more negative emotions than female teachers about AI-based measurement and evaluation.

"The Kruskal-Wallis test" was used to test whether the attitudes of the teachers participating in the study towards AI-based measurement and evaluation differed according to their ages. Related findings are given in Table 6.

Table 6

Differences in Teacher Attitudes Towards AI-Based Measurement and Evaluation by Age

Factor	Age	N	Mean Rank	χ^2	p	Different Group
Mean attitude score for giving importance	"1: 21-25 years old"	6	55,25	16,705	0,010*	1<4
	"2: 26-30 years old"	28	162,41			
	"3: 31-35 years old"	134	169,53			
	"4: 36-40 years old"	96	187,79			
	"5: 41-45 years old"	44	164,30			
	"6: 46-50 years old"	24	134,92			
	"7: 51 years and above"	7	219,86			
Mean attitude score for being active	"1: 21-25 years old"	6	257,75	14,847	0,021*	6<3
	"2: 26-30 years old"	28	173,79			
	"3: 31-35 years old"	134	183,09			
	"4: 36-40 years old"	96	161,32			
	"5: 41-45 years old"	44	136,70			
	"6: 46-50 years old"	24	179,40			
	"7: 51 years and above"	7	125,07			
Mean attitude score for developing negative emotions	"1: 21-25 years old"	6	128,25	29,792	0,000*	6>3, 4
	"2: 26-30 years old"	28	157,70			
	"3: 31-35 years old"	134	167,42			
	"4: 36-40 years old"	96	84,63			
	"5: 41-45 years old"	44	192,75			
	"6: 46-50 years old"	24	179,41			
	"7: 51 years and above"	7	260,00			
Mean attitude towards the experience	"1: 21-25 years old"	6	222,83	14,133	0,028*	5<3 3<4
	"2: 26-30 years old"	28	151,95			
	"3: 31-35 years old"	134	179,13			
	"4: 36-40 years old"	96	173,05			
	"5: 41-45 years old"	44	129,81			
	"6: 46-50 years old"	24	197,46			
	"7: 51 years and above"	7	138,71			

As seen in Table 6, it was found, that the probability values of the "Attitude towards giving importance", "Attitude towards being active", "Attitude towards developing negative emotions" and "Attitude towards experience" points of the teachers attending in the research are less than 0.05 significant value ($p<0.05$). Therefore, it was determined that teachers' attitudes towards AI-based measurement and evaluation differed according to their ages for all factors. When the results of giving importance attitude were analyzed, it was determined that teachers aged 21-25 gave less importance to AI-based measurement and evaluation than teachers aged 36-40. When the results of the attitude of being active were examined, it was determined that teachers aged 46-50 were less active in AI-based measurement and evaluation than teachers aged 31-35. When the results of the attitude of developing negative emotions were analyzed, it was determined that teachers aged 46-50 developed more negative emotions about AI-based measurement and evaluation than teachers aged 31-35 and teachers aged 36-40. When the results of experience attitude were analyzed, it was determined that teachers aged 31-35 had more experience in AI-based measurement and evaluation than teachers aged 41-45, but they had less experience than teachers aged 36-40.

"The Kruskal-Wallis test" was used to test whether the attitudes of the teachers participating in the study towards AI-based measurement and evaluation differed according to their professional seniority. Related findings are given in Table 7.

Table 7

Differences in Teacher Attitudes Towards AI-Based Measurement and Evaluation by Duration of Professional Seniority

Factor	Duration of professional seniority	N	Mean Rank.	χ^2	p	Different Group
Mean attitude score for giving importance	"1: Less than 1 year"	19	137,63	12,736	0,026*	3> 1, 2 5> 1, 2
	"2: 1-5 years"	15	138,63			
	"3: 6-10 years"	110	182,47			
	"4: 11-15 years"	108	169,98			
	"5: 16-20 years"	38	198,68			
	"6: 21 years and above"	49	141,97			
Mean attitude score for being active	"1: Less than 1 year"	19	199,03	3,577	0,612	-
	"2: 1-5 years"	15	174,40			
	"3: 6-10 years"	110	175,94			
	"4: 11-15 years"	108	160,71			
	"5: 16-20 years"	38	174,49			
	"6: 21 years and above"	49	161,06			
Mean attitude score for developing negative emotions	"1: Less than 1 year"	19	128,86	24,640	0,000*	6>1, 3
	"2: 1-5 years"	15	212,04			
	"3: 6-10 years"	110	152,38			
	"4: 11-15 years"	108	190,73			
	"5: 16-20 years"	38	162,70			
	"6: 21 years and above"	49	177,92			
Mean attitude towards the experience	"1: Less than 1 year"	19	190,45	1,975	0,853	-
	"2: 1-5 years"	15	148,77			
	"3: 6-10 years"	110	170,97			
	"4: 11-15 years"	108	169,10			
	"5: 16-20 years"	38	176,54			
	"6: 21 years and above"	49	163,31			

As seen in Table 7, it was found that the probability values of the "Attitude towards giving importance" and "Attitude towards developing negative emotions" points of the teachers attending in the study towards AI-based measurement and evaluation are less than 0.05 significant value ($p<0.05$). The probability values of teachers' "Attitude towards being active" and "Attitude towards experience" scores are greater than 0.05 significant value ($p>0.05$). For this reason, it was determined that the attitudes of the teachers attending the research about AI-based measurement and evaluation "Giving importance" and "Developing negative emotions" differed according to their professional seniority, while the attitudes of "Being active" and "Experience" did not differ according to the professional seniority of the teachers. When the results related to the attitude of giving importance were examined, it was determined that teachers with 16-20 years and 6-10 years of seniority gave more importance to AI-based measurement and evaluation than teachers with less than 1 year and 1-5 years of professional seniority. When the results of the attitude of developing negative emotions were examined, it was found that teachers who had 20 years and above of professional seniority developed more negative emotions about AI-based measurement and evaluation than teachers who had 1-5 years and 6-10 years of professional seniority.

"The Kruskal-Wallis test" was used to test whether the attitudes of the teachers participating in the study towards AI-based measurement and evaluation differed according to the level they worked at. Related findings are given in Table 8.

Table 8

Differences in Teacher Attitudes Towards AI-Based Measurement and Evaluation by Level of Employment

Factor	Level of school attended	N	Mean Rank	χ^2	p	Different Group
Mean attitude score for giving importance	"1: Pre-school"	38	167,53	3,980	0,264	-
	"2: Primary school"	83	152,45			
	"3: Middle school"	165	178,42			
	"4: High school"	53	173,04			
Mean attitude score for being active	"1: Pre-school"	38	183,92	8,561	0,036*	4<1,2,3
	"2: Primary school"	83	174,04			
	"3: Middle school"	165	176,17			
	"4: High school"	53	134,50			
Mean attitude score for developing negative emotions	"1: Pre-school"	38	170,36	0,140	0,987	-
	"2: Primary school"	83	167,13			
	"3: Middle school"	165	170,26			
	"4: High school"	53	173,42			
Mean attitude towards the experience	"1: Pre-school"	38	166,28	1,302	0,729	-
	"2: Primary school"	83	170,75			
	"3: Middle school"	165	174,52			
	"4: High school"	53	157,42			

As seen in Table 8, it was found that the probability values of the "Attitude towards giving importance", "Attitude towards developing negative emotions" and "Attitude towards experience" points of the teachers attending in the study towards AI-based measurement and evaluation are greater than 0.05 significant value (p>0.05). For this reason, it was determined that teachers' attitudes towards AI-based measurement and evaluation "Giving importance", "Developing negative emotions" and "Experience" did not differ according to the level they worked at, while their attitudes towards "Being active" differed according to the level they worked at. When the results of the attitude of being active were examined, it was found that high school teachers were less active in AI-based measurement and evaluation than pre-school, primary and secondary school teachers.

"The Mann Whitney U test" was used to test whether the attitudes of the teachers attending in the study towards AI-based measurement and evaluation differ by the type of school served. Related findings are given in Table 9.

Table 9

Differences in Teacher Attitudes Towards AI-Based Measurement and Evaluation by Seniority Type of School

Factor	Type of school served	N	Mean Rank	Test Statistic	p	Different Group
Mean attitude score for giving importance	"1: Public school"	312	168,01	U=3591,500 Z=-1,274	0,203	2<1
	"2: Private school"	27	192,98			
Mean attitude score for being active	"1: Public school"	312	173,45	U=3134,500 Z=-2,212	0,027*	
	"2: Private school"	27	130,09			
Mean attitude score for developing negative emotions	"1: Public school"	312	168,09	U=3617,500 Z=-1,227	0,220	
	"2: Private school"	27	192,02			
Mean attitude towards the experience	"1: Public school"	312	170,44	U=4074,500 Z=-0,283	0,777	
	"2: Private school"	27	164,91			

As seen in Table 9, it was found that the probability values of the "Attitude towards giving importance", "Attitude towards developing negative emotions" and "Attitude towards experience" points of the teachers attending in the study towards AI-based measurement and evaluation are greater than 0.05 significant value (p>0.05). It was seen that the probability values of the "Attitude towards being active" scores of the teachers attending in the study towards AI-based measurement and evaluation were less than 0.05 significant value (p<0.05). For this reason, it was determined that teachers' attitudes towards AI-based measurement and evaluation "Giving importance", "Developing negative emotions" and "Experience" did not differ according to the type of school they worked in, while their attitudes towards "Being active" differed according to the type of school they worked in.

When the results of the attitude of being active were examined, it was determined that teachers working in public schools were more active in AI-based measurement and evaluation than teachers working in private schools.

"The Mann Whitney U test" was used to test whether the attitudes of the teachers attending in the study towards AI-based measurement and evaluation differed by "situation taking measurement and evaluation courses at university". Related findings are given in Table 10.

Table 10

Differences in Teacher Attitudes Towards AI-Based Measurement and Evaluation by Situation of Taking Measurement-Evaluation Courses at University

Factor	Situation of taking measurement-evaluation courses at university	N	Mean Rank	Test Statistic	p	Different Group
Mean attitude score for giving importance	"1: Yes"	103	54,28	U=234,5	0,736	-
	"2: No"	5	59,10	Z=-0,337		
Mean attitude score for being active	"1: Yes"	103	54,59	U=248,5	0,895	-
	"2: No"	5	52,70	Z=-0,132		
Mean attitude score for developing negative emotions	"1: Yes"	103	53,75	U=180,0	0,257	-
	"2: No"	5	69,90	Z=-1,135		
Mean attitude towards the experience	"1: Yes"	103	54,48	U=255,5	0,977	-
	"2: No"	5	54,90	Z=-0,029		

As seen in Table 10, it was found that the probability values of the "Attitude towards giving importance", "Attitude towards being active", "Attitude towards developing negative emotions" and "Attitude towards experience" points of the teachers attending in the study are greater than 0.05 significant value (p>0.05). Therefore, it was determined that teachers' attitudes towards AI-based measurement and evaluation did not differ for all factors according to the status of taking measurement and evaluation courses at university.

"The Mann Whitney U test" was used to test whether the attitudes of the teachers attending in the study towards AI-based measurement and evaluation differed by "status of receiving in-service training on measurement and evaluation". Related findings are given in Table 11.

Table 11

Differences in Teacher Attitudes Towards AI-Based Measurement and Evaluation by Status of Receiving In-Service Training Related to Measurement and Evaluation

Factor	Status of receiving in-service training on measurement and evaluation	N	Mean Rank	Test Statistic	p	Different Group
Mean attitude score for giving importance	"1: Yes"	163	168,19	U=14049,500	0,743	-
	"2: No"	176	171,67	Z=-0,328		
Mean attitude score for being active	"1: Yes"	163	165,16	U=13555,500	0,381	-
	"2: No"	176	174,48	Z=-0,877		
Mean attitude score for developing negative emotions	"1: Yes"	163	158,58	U=12483,000	0,037*	1<2
	"2: No"	176	180,57	Z=-2,081		
Mean attitude towards the experience	"1: Yes"	163	171,62	U=14080,000	0,768	-
	"2: No"	176	168,50	Z=-0,295		

As seen in Table 11, it was found that the probability values of the "Attitude towards giving importance", "Attitude towards being active", and "Attitude towards experience" points of the teachers attending in the study towards AI-based measurement and evaluation are greater than 0.05 significant value (p>0.05). It was seen that the probability values of the "Attitude towards developing negative emotions" scores of the teachers attending in the study towards AI-based measurement and evaluation were less than 0.05 significant value (p<0.05). Therefore, it was determined that teachers' attitudes towards developing negative emotions towards AI-based measurement and evaluation differed

according to their "in-service training on measurement and evaluation". When the results of developing negative emotions were examined, it was determined that teachers who "received in-service training on measurement and evaluation" developed less negative emotions about AI-based measurement and evaluation than teachers who "did not receive in-service training on measurement and evaluation".

"The Mann Whitney U test" was used to test whether the attitudes of the teachers attending in the study towards AI-based measurement and evaluation differed by "status of receiving in-service training on AI-based measurement and evaluation". Related findings are given in Table 12.

Table 12

Differences in Teacher Attitudes Towards AI-Based Measurement and Evaluation by Status Of Receiving In-Service Training on AI-Based Measurement and Evaluation

Factor	Status of receiving in-service training on AI-based measurement and evaluation	N	Mean Rank	Test Statistic	p	Different Group
Mean attitude score for giving importance	"1: Yes"	31	173,43	U=3718,500	0,042	2<1
	"2: No"	308	135,95	Z=-2,036	*	
Mean attitude score for being active	"1: Yes"	31	170,31	U=4764,500	0,985	-
	"2: No"	308	169,97	Z=-0,018		
Mean attitude score for developing negative emotions	"1: Yes"	31	112,92	U=3004,500	0,001	1<2
	"2: No"	308	175,75	Z=-3,429	*	
Mean attitude towards the experience	"1: Yes"	31	217,16	U=3312,000	0,005	2<1
	"2: No"	308	165,25	Z=-1,724	*	

As seen in Table 12, it was found that the probability values of the "Attitude towards giving importance", "Attitude towards developing negative emotions" and "Attitude towards experience" scores of the teachers attending in the study towards AI-based measurement and evaluation are less than 0.05 significant value ($p<0.05$). It was seen that the probability values of the "Attitude towards being active" score of the teachers attending in the study towards AI-based measurement and evaluation were greater than 0.05 significant value ($p>0.05$). Therefore, it was concluded that teachers' attitudes towards AI-based measurement and evaluation differed according to the status of receiving in-service training on AI-based measurement and evaluation for all factors except the factor of "being active". When the results of giving importance were analyzed, it was determined that teachers who "received in-service training on AI-based measurement and evaluation" gave more importance to AI-based measurement and evaluation than teachers who "did not receive in-service training on AI-based measurement and evaluation". When the results of developing negative emotions were examined, it was determined that the teachers who "received in-service training on AI-based measurement and evaluation" developed less negative emotions about AI-based measurement and evaluation than the teachers who "did not receive in-service training on AI-based measurement and evaluation". When the experience results were analyzed, it was found that teachers who "received in-service training on AI-based measurement and evaluation" had more experience in AI-based measurement and evaluation than teachers who "did not receive in-service training on AI-based measurement and evaluation".

Conclusion and Recommendations

The aim of the study was to determine whether teachers' attitudes towards AI-based measurement and evaluation differ by their socio-demographic characteristics. The result of the study revealed that teachers' attitudes towards the sub-factors of the scale of attitudes towards AI-based measurement and evaluation, namely "Giving importance", "Being active", "Developing negative emotions" and "Experience", varied by some socio-demographic characteristic. It was determined that teachers aged 21-25 attach less importance to AI-based measurement and evaluation than teachers aged 36-40. It was determined that teachers with "16-20 years" and "6-10 years of seniority" gave more importance to AI-based measurement and evaluation than teachers with less than 1 year and 1-5

years of professional seniority. It was concluded that teachers "receiving in-service training on AI-based measurement and evaluation" gave more importance to AI-based measurement and evaluation than teachers who "did not receive in-service training on AI-based measurement and evaluation". It is an expected result of the study that teachers who "received in-service training on AI-based measurement and evaluation" attach importance to this issue. However, it is an unexpected result that teachers in the younger age group attach less importance to AI-based measurement and evaluation than older teachers. Considering that young teachers can adapt to current and contemporary methods more easily, it is expected that they would attach more importance to this issue. The result obtained in the research reflects the opposite situation. This result can be interpreted as an indicator of young teachers' lack of experience. This interpretation also supports the conclusion that teachers with less than 1-5 years of professional seniority attach less importance to AI-based measurement and evaluation than teachers with "6-10 years" and "16-20 years" of professional seniority. This result differs from the findings of the study conducted by Seyrek et al. (2024) in the literature. Seyrek et al. (2024) revealed in their study that young teachers use artificial intelligence tools more frequently in their classes.

It was determined that teachers between the ages of 46-50 were less effective in AI-based measurement and evaluation than teachers between the ages of 31-35. It is an expected result that teachers aged 46-50 are less effective in AI-based measurement and evaluation than teachers aged 31-35. This can be interpreted as a result of the fact that teachers between the ages of 46-50 have more difficulty in adapting to technology than teachers between the ages of 31-35, prefer to stick to traditional methods, and are reluctant to participate in relevant in-service trainings. High school teachers were found to be less effective in AI-based measurement and evaluation than preschool, primary and secondary school teachers. High school teachers' exam-oriented studies with a more intense curriculum that prepares students for university may cause them to be more cautious about using new technologies effectively. Therefore, this result is an expected result in the study. It was determined that teachers "working in public schools" were more active in AI-based measurement and evaluation than teachers "working in private schools". Considering that private schools generally have better technological infrastructure and more financial resources, this result is unexpected. However, the larger and more diverse student population in public schools compared to private schools may have encouraged teachers to use AI-based measurement and evaluation methods effectively. Teachers who effectively use AI-based measurement and evaluation tools and methods in public schools where class size and student diversity are high will both save time and respond to the individual needs of students.

It was determined that male teachers developed more negative feelings about AI-based measurement and evaluation than female teachers. This result may be due to the lower number of male teachers participating in the study. One of the conditions determined while collecting the research data was the participation of teachers who had experienced AI-based measurement and evaluation practices at least once. Depending on this condition, it can be thought that male teachers participated in the research less than female teachers. This situation can be explained by the fact that male teachers develop more negative feelings about the subject due to the fact that they have less experience in AI-based measurement and evaluation than female teachers. It was determined that teachers aged between 46-50 years developed more negative feelings about AI-based measurement and evaluation than teachers aged between "31-35 years" and "36-40 years". It was found that teachers with 20 years and above of professional seniority developed more negative feelings towards AI-based measurement and evaluation than teachers who had a professional seniority of "1-5 years" and "6-10 years". These two results related to age and professional seniority are expected. The fact that teachers with a higher age range and professional seniority prefer to stick to more traditional methods and their lack of experience in AI may cause them to develop a negative feeling on this issue. It was determined that teachers who "received in-service training on measurement and evaluation" developed less negative feelings towards AI-based measurement and evaluation than teachers who "did not receive in-service training" on this subject. It was concluded that teachers who "received in-service training on AI-based measurement and evaluation" developed less negative feelings about AI-based measurement and evaluation than teachers who "did not receive in-service training on AI-based measurement and evaluation". Both of these results are expected in this study. Teachers receiving both "measurement

and evaluation" and "AI-based measurement and evaluation" training can exhibit a positive approach to the developments and innovations in the areas of measurement and evaluation thanks to the training they receive.

It was concluded that teachers aged 31-35 had more experience in AI-based measurement and evaluation than teachers aged 41-45, but they had less experience than teachers aged 36-40. It is an expected result that teachers between the ages of 31-35 have gained more experience in this subject compared to teachers between the ages of 41-45 in line with the reasons such as the fact that they have been more acquainted with technology during their university education and that educational technologies and AI-based educational tools have become widespread when they started their profession. However, the fact that teachers aged 36-40 have more experience in AI-based measurement and evaluation than teachers aged 31-35 may be an indication that teachers aged 36-40 have gained experience by participating in in-service trainings, development programs and seminars on the subject at the beginning and various periods of their careers. It was determined that teachers who "received in-service training on AI-based measurement and evaluation" had more experience in AI-based measurement and evaluation than teachers who "did not receive in-service training" on this subject. This result is an expected result in the research. Because it is clear that teachers who receive in-service training based on AI would gain experience by putting the training they receive into practice.

When the research conclusions are taken into consideration, it is proposed that teachers should be encouraged to use AI-based measurement and evaluation methods and training and seminars should be organized on the process. In his study, Çavuş (2024) emphasized that teachers may face various difficulties such as lack of transparency, prejudice, ethical concerns, cost, integration with existing systems, technical difficulties, student motivation, and resistance to change regarding the use of artificial intelligence in educational evaluation, and that this situation can be eliminated through pre-service and in-service training on the subject. Studies on the problems and solutions that teachers face when using AI-based measurement and evaluation tools and methods will also contribute to the process. In addition, teachers should be given the necessary support regarding the implementation phase of AI-based measurement and evaluation tools and methods developed in accordance with the age in order to make an effective and fair measurement and evaluation in teaching and learning. Considering the results for teachers with the least teaching experience in particular, the following suggestions can be considered to ensure positive attitudes towards AI-based measurement and evaluation in pre-service teacher education:

- Organizing trainings that explain how AI-based measurement and evaluation tools work, their advantages, and how they can be used in classroom practices can increase teacher candidates' awareness of AI.
- Practical lessons that demonstrate how AI-based measurement and evaluation tools can be used in the classroom can help prospective teachers develop a positive attitude by increasing their familiarity with these technologies.
- Case studies featuring successful AI-based measurement and evaluation applications can help prospective teachers concretely see the potential benefits of these technologies.
- In pre-service teacher education, it is important to understand the concerns and negative attitudes experienced by teachers regarding these new technologies and to establish feedback mechanisms to produce solutions. It may also be useful to provide guides and resources that would provide support in the use of technology.
- Ethics, privacy and security issues of AI-based measurement and evaluation tools and methods should be included in pre-service teacher training. This can reduce teachers' potential concerns and enable them to use these tools and methods consciously.
- It should be emphasized that AI-based measurement and evaluation tools can provide a wide range of fair evaluations, and the features of these tools that can evaluate students by taking into account their individual differences should be introduced.

Limitations

This study is limited to the responses of 339 teachers to the data collection tool. The findings of the research are limited to the measurement power of the scale used in the research. The research is limited to the time period in which the scale was applied.

References

Christensen, L. B., Johnson, R. B., & Turner, L. A. (2015). Tarama araştırması. In A. Aypay (Ed.), *Araştırma yöntemleri desen ve analiz* E. Gümüş, Çev. (Trans. Edts.), (pp. 367-399). Anı.

Çavuş, M. N. (2024). Eğitimde yapay zeka tabanlı ölçme ve değerlendirme üzerine bir derleme. *International Journal of English for Specific Purposes, 2*(1), 39-54.

Huang, S.-P. (2018). Effects of using artificial intelligence teaching system for environmental education on environmental knowledge and attitude. *EURASIA Journal of Mathematics, Science and Technology Education, 14*(7), 3277-3284. https://doi.org/10.29333/ejmste/91248

Karasar, N. (2012). *Bilimsel araştırma Yöntemi.* Bilim Kitap Kırtasiye Yayınevi.

Mede, E., & Atay, D. (2017). English language teachers' measurement literacy: The Turkish context. *Dil Dergisi, 168*(1), 43-60.

Nazaretsky, T., Ariely, M., Cukurova, M., & Alexandron, G. (2022). Teachers' trust in AI-powered educational technology and a professional development program to improve it. *British Journal of Educational Technology, 53*(4), 914-931. https://doi.org/10.1111/bjet.13232

Seyrek, M., Yıldız, S., Emeksiz, H., Şahin, A., & Türkmen, M. T. (2024). Öğretmenlerin eğitimde yapay zeka kullanımına yönelik algıları. *International Journal of Social and Humanities Sciences Research (JSHSR), 11*(106), 845-856.

Stiggins, R. (2014). Improve measurement literacy outside of schools too. *Phi Delta Kappan, 96*(2), 67-72.

Suh, W., & Ahn, S. (2022). Development and validation of a scale measuring student attitudes toward artificial intelligence. *Sage Open, 12*(2). https://doi.org/10.1177/21582440221100463

Tsagari, D. (2011). Investigating the 'measurement literacy' of EFL state school teachers in Greece. In D. Tsagari, & I. Csépes (eds.), *Classroom-based language measurement*, Peter Lang, (pp. 169-190).

Tezci, İ. H. (2019). Öğretmenlere yönelik ölçme ve değerlendirme tutum ölçeği geliştirme çalışması. *I. Uluslararası Bilim, Eğitim, Sanat ve Teknoloji (UBEST) Sempozyumu Tam Metin Kitabı*, İzmir, (pp. 408-417).

Vasconcelos, M. A., & dos Santos, R. P. (2023). Enhancing STEM learning with ChatGPT and Bing Chat as objects to think with: A case study. *EURASIA Journal of Mathematics, Science and Technology Education, 19*(7). https://doi.org/10.29333/ejmste/13313

****

Ethical Declaration and Committee Approval

In this research, the principles of scientific research and publication ethics were followed.

****

Proportion of the Author Contribution

All researchers contributed equally to this study.