Üniversite Araştırmaları Dergisi, 2026, Cilt 9, Sayı 2, Sayfa: 225-241
Journal of University Research, 2026, Volume 9, Issue 2, Page: 225-241

Araştırma Makalesi
Research Article

Artificial Intelligence Visual Literacy Scale: Scale Development Study

Yapay Zekâ Görsel Okuryazarlığı Ölçeği: Ölçek Geliştirme Çalışması

Abdulkerim Turkaya[1*], Ali Ertuğrul Küpeli[2], Adnan Kan[3]

[1]Ankara Music and Fine Arts University, Vocational School of MFA, Handicrafts Department, Ankara/Türkiye
[2]Gazi University, Gazi Faculty of Education, Department of Fine Arts Education, Ankara/Türkiye
[3]Gazi Üniversitesi, Gazi Faculty of Education, Department of Educational Sciences, Ankara/Türkiye

Abstract: In this study, the Artificial Intelligence Visual Literacy Scale (AIVLS) was designed and developed by the researchers to analyze individuals' competencies in understanding, evaluating, and producing AI-supported visual content in today's digital age. The scale development process was carried out based on classical measurement theory, and validity and reliability analyses were meticulously performed during this process. The study group consisted of a total of 622 students who continued their education in fine arts-related fields at associate, undergraduate, and graduate levels in universities as of the spring semester of the 2024-2025 academic year. The "Demographic Characteristics Questionnaire" and the "Artificial Intelligence Visual Literacy Scale," developed by the researchers as data collection tools, were administered via an online survey using Google Forms. To examine students' artificial intelligence visual literacy levels, 18 five-point Likert-type items were developed. The developed scale was first applied as a pilot study to 63 students; following the data obtained, Exploratory Factor Analysis (EFA) was tested on 295 students, and Confirmatory Factor Analysis (CFA) on 264 students. The study employed a quantitative research design. In the validity and reliability tests of the scale items, descriptive statistics, validity, and reliability analyses were conducted using SPSS and AMOS, and the obtained data were included in the research. As a result of the research, it was confirmed that the 18-item scale has a 4-dimensional structure explaining 67.738% of the total variance. Within the scope of the reliability study, the overall Cronbach's Alpha (α) reliability coefficient of the scale was found to be .881, and the fit index values for confirmatory factor analysis were found to be within "acceptable" limits. With these results, it can be said that the developed scale is a valid and reliable scale that can be applied in the field of artificial intelligence visual literacy. It is believed that by using the developed scale, the gap in the literature regarding the evaluation of artificial intelligence visual literacy in Turkey will be filled, individuals' visual literacy levels regarding artificial intelligence will be better understood, and it will contribute to the realization of many future studies.

Keywords: Artificial Intelligence Visual Literacy Scale, Scale Development, Art Education, Graphic Design

Özet: Bu çalışmada, günümüz dijital çağında bireylerin yapay zekâ destekli görsel içerikleri anlama, değerlendirme ve üretme yeterliklerini çözümlemek amacıyla araştırmacılar tarafından Yapay Zekâ Görsel Okuryazarlığı Ölçeği (YZGÖ) tasarlanıp geliştirilmiştir. Ölçek geliştirme süreci, klasik ölçme kuramına dayalı olarak yürütülmüş ve bu süreçte geçerlik ve güvenirlik analizleri titizlikle gerçekleştirilmiştir. Araştırmanın çalışma grubu, 2024-2025 eğitim-öğretim yılı bahar yarıyılı itibariyle üniversitelerin ön lisans, lisans ve lisansüstü düzeylerinde güzel sanatlar eğitimi alanıyla ilgili öğrenimlerine devam eden toplam 622 öğrenciden oluşmaktadır. Araştırmacılar tarafından veri toplama aracı olarak geliştirilen "Demografik Özellik Anketi" ile "Yapay Zekâ Görsel Okuryazarlığı Ölçeği" Google form üzerinden çevrimiçi anket yoluyla yapılmıştır. Öğrencilerin yapay zekâ görsel okuryazarlık düzeylerini incelemek amacıyla 18 maddelik 5'li Likert tipi ölçek maddeleri oluşturulmuştur. Geliştirilen ölçek önce pilot uygulama olarak 63 öğrenci üzerinde uygulanmış, elde edilen verilerin ardından 295 öğrenci üzerinde AFA, 264 öğrenci üzerinde de DFA test edilmiştir. Araştırma yöntemi nicel olarak tasarlanmıştır. Ölçek maddelerinin geçerlilik ve güvenirlik testlerinde SPSS ve SPSS AMOS programları kullanılarak frekans ve yüzde hesaplamaları yapılmış ve elde edilen veriler araştıa dâhil edilmiştir. Araştırma sonucunda, 18 maddelik ölçeğin toplam varyansın %67,738'ini açıklayan 4 boyutlu bir yapıya sahip olduğu teyit edilmiştir. Güvenirlik çalışması kapsamında ölçeğin genelinde Cronbach Alfa (α) güvenirlik katsayısı ,881 olduğu ve doğrulayıcı faktör analizi için uyum indeksi değerlerinin "kabul edilebilir" sınırlar içinde olduğu bulunmuştur. Bu sonuçlarla, geliştirilen ölçeğin yapay zekâ görsel okuryazarlığı alanında uygulanabilecek geçerli ve güvenilir bir ölçek olduğu söylenebilir. Geliştirilen ölçek kullanılarak, Türkiye'de yapay zekâ görsel okuryazarlığının değerlendirilmesinde literatürdeki boşluğun doldurulacağı, bireylerin yapay zekâ konusundaki görsel okuryazarlık düzeylerinin daha iyi anlaşılacağı ve gelecekte birçok çalışmanın gerçekleştirilmesine katkı sunulacağı düşünülmektedir.

Anahtar Kelimeler: Yapay Zekâ Görsel Okuryazarlığı Ölçeği, Ölçek Geliştirme, Sanat Eğitimi, Grafik Tasarım

* İletişim Yazarı / *Corresponding author.*
akturkaya@gmail.com

Geliş Tarihi / *Received Date*: 28.02.2026
Revizyon Talebi Tarihi / *Revision Request Date*: 08.05.2026
Son Revizyonun Geliş Tarihi / *Last Revised Version Received Date*: 31.05.2026
Revizyon Sonrası Kabul Tarihi / *Accepted After Revision Date*: 04.06.2026

https://doi.org/10.32329/uad.1899455

1. Introduction

In the digitalized world of the 21st century, the ways in which information is produced, transmitted, and consumed have undergone a profound transformation. One of the most significant dimensions of this transformation is the widespread integration of artificial intelligence (AI) technologies into nearly every aspect of social life. Creative fields such as education, media, art, design, and communication are increasingly being reshaped by these technologies, resulting in accelerated and diversified production processes. In particular, the rapid proliferation of AI-powered visual production tools has fundamentally transformed not only how individuals consume visual content but also how they produce, analyze, and evaluate such content. In this context, the ability to accurately interpret, analyze, and reproduce AI-generated visual content should be considered not merely a technical skill, but also a cognitive and critical competence.

Visual literacy refers to individuals' ability to perceive, interpret, analyze, and derive meaning from visual information. Although the concept has traditionally been discussed within the scope of media literacy and art education, it has gained broader significance with the advancement of digital technologies (Avgerinou & Pettersson, 2011). The literature clearly indicates the need for valid and reliable measurement tools focusing specifically on AI-supported visual literacy. Although the AI Media Literacy Scale developed by Lin (2021) represents an important contribution to the field, it does not sufficiently address dimensions such as visual storytelling, ethical awareness, and critical visual interpretation. Today, visual content is no longer produced solely by humans; AI algorithms are increasingly capable of generating, transforming, and manipulating visual materials. This transformation has rendered traditional understandings of visual literacy insufficient and has created a need for new assessment tools capable of measuring individuals' abilities to interpret and critically evaluate AI-generated visual content.

Artificial intelligence technologies have developed rapidly in recent years, particularly in the field of educational technologies (Luckin et al., 2016). AI-based systems capable of generating text, images, videos, music, and software code have attracted considerable attention due to their ability to simulate human-like production processes. These developments have substantial implications for education. Students studying in disciplines such as design, media, communication, and art education are increasingly encountering AI-supported tools throughout their learning processes. Applications such as Midjourney, DALL·E, and Adobe Firefly can generate sophisticated visuals based solely on text prompts, thereby influencing students' creative production processes, design skills, and critical evaluation capacities. However, despite the opportunities provided by these technologies, they also introduce important concerns related to ethics, originality, manipulation, misinformation, and algorithmic bias (Tiernan et al., 2023).

Visual literacy is closely associated with disciplines such as media literacy, digital literacy, and critical thinking, as it encompasses the ability to analyze, interpret, and produce visual information (Avgerinou & Pettersson, 2011). Although various scale development studies related to visual literacy and digital media literacy exist in the literature (Hobbs, 2010; Serafini, 2013), the number of comprehensive and valid instruments specifically addressing the relationship between artificial intelligence and visual literacy remains limited. Studies conducted in Türkiye indicate that AI literacy levels are generally at moderate to high levels; however, these levels may vary depending on variables such as gender, grade level, internet usage habits, and the frequency of following AI-related developments (Ayçiçek, 2025; Çam et al., 2021; Çelebi et al., 2023). Existing studies primarily provide psychometric frameworks guiding methodological approaches to AI literacy measurement. Nevertheless, current instruments generally focus on traditional media content and conventional visual elements, failing to adequately address the new layers of meaning, production, and interpretation introduced by artificial intelligence technologies. Consequently, significant gaps remain in the evaluation of AI-supported visual meaning-making processes.

1.1. Purpose and Significance of the Research

The primary purpose of this study is to develop an original measurement instrument called the Artificial Intelligence Visual Literacy Scale (AIVLS) and to examine its psychometric validity and reliability. The scale aims to measure individuals' abilities to understand, analyze, critically evaluate, and ethically interpret AI-supported visual content. Accordingly, the scale was designed as a multidimensional structure consisting of subdimensions such as interpretation and analysis, design and creation, ethical sensitivity, and familiarity with artificial intelligence technologies.

The scale may be useful for educational institutions, as it provides an effective assessment tool for determining students' competencies regarding AI-generated visual content within media, communication, art, and design-oriented educational programs. Furthermore, the data obtained through this instrument may contribute

https://doi.org/10.32329/uad.1899455

to the restructuring of educational curricula and pedagogical approaches, the development of teacher training programs, and the enhancement of digital ethics awareness.

This study proposes a holistic approach to measuring the new generation of literacy skills required in an increasingly AI-dominated visual culture. Through item analyses as well as exploratory and confirmatory factor analyses, the developed scale not only addresses a significant gap in the literature but also contributes to fostering critical visual thinking in education during the age of artificial intelligence.

2. Method

This section provides information and explanations regarding the research model, the determination of the study group (population-sample) according to the research model, the data collection tools used in the research, how the data collection process was carried out, and the analysis of the data obtained from the research.

2.1. Research Design

Scales are measurement instruments developed to classify, rank, and determine the quantity or degree of specific characteristics to be measured (Karakoç & Dönmez, 2014). A review of the scale development literature reveals numerous methodological guidelines proposed by researchers and experts regarding the development of reliable and valid measurement instruments. Within the scope of the present study, a scale was developed to measure individuals' abilities to understand, analyze, critically evaluate, and ethically interpret AI-supported visual content. The study was designed using a quantitative research methodology. To examine the construct validity of the developed scale, Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) were conducted. In addition, Cronbach's Alpha coefficient was calculated to determine the internal consistency and reliability of the scale.

2.2. Study Group

Sampling in educational research is generally conducted to allow detailed investigation of a subset of the population rather than examining the entire population. Data obtained from the sample are then used to make generalizations about the target population (Cemaloğlu, 2023). The study group consists of a total of 622 students studying at 45 universities related to the field of fine arts education in the spring semester of the 2024-2025 academic year. Official permissions were obtained from the relevant institutional authorities of all par-

Table 1. Working Group

No	University Name	City
1	Ağrı İbrahim Çeçen University	Ağrı
2	Alanya Alaaddin Keykubat University	Antalya
3	Ankara Music and Fine Arts University	Ankara
4	Ankara Social Sciences University	Ankara
5	Ankara Yıldırım Beyazıt University	Ankara
6	Ardahan University	Ardahan
7	Atatürk University	Erzurum
8	Aydın Adnan Menderes University	Aydın
9	Bahçeşehir University	İstanbul
10	Bartın University	Bartın
11	Başkent University	Ankara
12	Bayburt University	Bayburt
13	Bilkent University	Ankara
14	Bingöl University	Bingöl
15	Bitlis Eren University	Bitlis
16	Burdur Mehmet Akif Ersoy University	Burdur
17	Bursa Uludağ University	Bursa
18	Çanakkale Onsekiz Mart University	Çanakkale
19	Çukurova University	Adana
20	Doğuş University	İstanbul
21	Dokuz Eylül University	İzmir
22	Düzce University	Düzce
23	Fırat University	Elazığ
24	Gazi University	Ankara
25	Hasan Kalyoncu University	Gaziantep
26	Iğdır University	Iğdır
27	İstanbul Okan University	İstanbul
28	İstanbul Sabahattin Zaim University	İstanbul
29	İzmir Demokrasi University	İzmir
30	Kafkas University	Kars
31	Kastamonu University	Kastamonu
32	Konya Technical University	Konya
33	Mersin University	Mersin
34	Niğde Ömer Halisdemir University	Niğde
35	Ondokuz Mayıs University	Samsun
36	Middle East Technical University	Ankara
37	Pamukkale University	Denizli
38	Sakarya University of Applied Sciences	Sakarya
39	Selçuk University	Konya
40	Sinop University	Sinop
41	Sivas Cumhuriyet University	Sivas
42	Tokat Gaziosmanpaşa University	Tokat
43	Trabzon University	Trabzon
44	Uşak University	Uşak
45	Yaşar University	İzmir

Official permits have been obtained from the relevant authorities of all the universities listed in Table 1.

https://doi.org/10.32329/uad.1899455

ticipating universities prior to data collection. ▶**Table 1** lists the names of universities that conducted online surveys after obtaining official permission from their respective authorities.

A simple random sampling method, one of the probability-based sampling techniques, was employed in the selection of participants. Random sampling is regarded as one of the most widely recognized sampling strategies in educational research (Teddlie & Tashakkori, 2009, as cited in Özmantar, 2018). The scale was initially administered to 63 students as part of a pilot study. Following the pilot implementation, Exploratory Factor Analysis (EFA) was conducted with data obtained from 295 students to determine the construct validity and factor structure of the scale. Subsequently, Confirmatory Factor Analysis (CFA) was performed with data collected from 264 students to test the relationships between the observed variables and the latent constructs represented by these variables.

Demographic characteristics of the participants included variables such as gender, age, educational level, academic field, internet usage duration, technology usage habits, and artificial intelligence application usage. ▶**Table 2** contains demographic information about the participants in the working group.

Table 2. Demographic Information About the Working Group

Variables		F (n)	%
Gender	Female	364	58,5%
	Male	258	41,5%
Age	18-20 years	131	21,1%
	21-24 years	386	62,1%
	25-34 years	53	8,6%
	35-44 years	37	5,9%
	45-54 years	9	1,4%
	55 years and above	6	0,9%
Grade level	Associate Degree	216	34,7%
	Bachelor's Degree	328	52,7%
	Master's Degree	46	7,3%
	Doctorate	19	3,1%
	Proficiency in Arts	13	2,2%
Marital status	Married	74	11,9%
	Single	548	88,1%
What is your field of study?	Industrial Design	57	9,1%
	Traditional Turkish Arts	90	14,5%
	Graphic Design	312	50,1%
	Sculpture	16	2,6%
	Painting	147	23,7%
How much time do you spend online? (24 hours)	1 hours and under	21	3,4%
	1-3 hours	143	23%
	4-6 hours	325	52,2%
	7-9 hours	97	15,6%
	10 hours and above	36	5,8%
Do you have a personal computer that you use?	Yes	537	86,3%
	No	85	13,7%
Do you use a smartphone?	Yes	622	100%
	No	0	0%
Do you keep up with technological developments?	Yes	546	87,7%
	No	76	12,3%
Do you use artificial intelligence applications?	Yes	519	83,4%
	No	103	16,6%
Which artificial intelligence application do you use most frequently?	ChatGPT	436	70,9%
	DeepSeek	23	3,6%
	Google Gemini	49	7,6%
	Midjourney	11	1,5%
	I don't use	103	16,4%

2.3. Data Collection Instruments

An online questionnaire consisting of two sections was used as the primary data collection tool in the study. The first section included a "Demographic Information Questionnaire" developed by the researchers to collect participants' demographic information. The second section consisted of the "Artificial Intelligence Visual Literacy Scale (AIVLS)," which was developed by the researchers using a five-point Likert-type format. During the scale development process, expert opinions were consulted in accordance with the objectives of the study.

2.3.1. Demographic Information Questionnaire

The Demographic Information Questionnaire consisted of 11 multiple-choice questions designed to obtain preliminary information regarding participants' demographic characteristics, including age, gender, educational level, academic discipline, technology-following habits, duration of internet usage, and use of artificial intelligence applications.

2.3.2. Artificial Intelligence Visual Literacy Scale (AIVLS)

The Artificial Intelligence Visual Literacy Scale developed by the researchers is a four-dimensional instrument consisting of 18 items without reverse-coded statements. A five-point Likert scale ranging from "Strongly Disagree (1)" to "Strongly Agree (5)" was used.

To improve data reliability and identify random responses, one control item was added to the scale. Participants identified as responding randomly during any stage of the implementation process were excluded from the

https://doi.org/10.32329/uad.1899455

analyses. Reliability analysis indicated that the overall Cronbach's Alpha coefficient of the scale was .881. Since reliability coefficients above .70 are generally considered acceptable (Büyüköztürk, 2011), the scale demonstrated satisfactory internal consistency.

2.4. Data Collection Procedure

2.4.1. First Stage

First, a comprehensive literature review was conducted during the scale creation and development process, and the scale was prepared according to these rules. When creating the scale items, an item pool consisting of 34 Likert-type items referring to participants' knowledge and opinions about artificial intelligence visual literacy was initially formed. The item pool was submitted to experts and academicians in the field using the Lawshe technique, and the content validity of each item was evaluated. The Lawshe technique is a widely used method to evaluate the content validity of items by utilizing expert opinions. This technique increases the scientific validity of the scale, allowing researchers to develop reliable and valid measurement tools (Baghestania et al., 2017). In this way, developed scales include statements that fully represent the relevant topic, thus preventing off-topic statements (Lawshe, 1975).

Numerous methods of quantifying experts' degree of agreement regarding the content relevance of an instrument have been proposed. These include, for example, averaging experts' ratings of item relevance and using a pre-established criterion of acceptability (Beck & Gable, 2001); using coefficient alpha to quantify agreement of item relevance by three or more experts (Waltz et al., 2005). In this study, 5 academicians working at Gazi University and Ankara Music and Fine Arts University and 3 field experts working in various public institutions in Ankara were selected to obtain expert opinions. The distribution of the titles and genders of the experts participating in the study is given in ▶**Table 3**.

Table 3. Gender and Title Distribution of Experts

Title	Woman	Man	Total (N)
Prof. Dr.	0	2	2
Assoc. Prof. Dr.	1	1	2
Lecturer	0	1	1
AI Specialist	0	1	1
Graphic Designer	1	1	2
Total (N)	2	6	8

Experts evaluated each item as "appropriate," "partially appropriate," or "inappropriate." They were also asked to provide written suggestions for items marked as partially appropriate. The titles of the experts participating in the study and their distribution according to their gender are given in ▶**Table 3**.

The expert form, which was used as a data collection tool in the study, consists of a total of 38 items. Experts were asked to mark one of the options: "appropriate," "partially appropriate," and "not appropriate" for the 34 items in the form. Additionally, if experts marked "partially appropriate," they were asked to provide their suggestions with the statement "If your answer is partially appropriate, please write your suggestion." When creating the item pool, the adequacy of the items in terms of Turkish language was also checked by an academician from the Turkish Language and Literature Department, and necessary corrections were made.

Content Validity Ratio (CVR) is a statistical value used to determine how appropriate each item in a measurement tool is perceived by experts, and it ranges between -1 and +1. If all experts evaluate an item as "appropriate," the CVR value of that item is +1, indicating strong acceptance of the item. If 50% of the experts find an item appropriate, the CVR is zero, meaning the item is subject to a neutral evaluation. However, if less than half of the experts find an item appropriate, the CVR is negative. In this case, the item is removed from the scale (Yurdugül, 2005).

For an item to be included in the scale, its CVR value is compared with the relevant table value (CVO); if the CVR is equal to or greater than the table value, the item is retained in the scale; otherwise, it is removed. This process is critically important to ensure the validity of the scale (Davis, 1992). In the study, 16 items with low CVR, based on the data obtained from expert evaluations, were removed from the item pool. ▶**Figure 1** shows the critical values of the Lawshe technique Content Validity Index (CVI).

Table 1. $CVR_{critical}$ One-Tailed Test (α = .05) Based on Exact Binomial Probabilities.

N (Panel Size)	Proportion Agreeing Essential	$CVR_{Critical}$ Exact Values	One-Sided *p* Value	$N_{critical}$ (Minimum Number of Experts Required to Agree Item Essential)—Ayre and Scally, This Article	$N_{critical}$ Calculated From CRITBINOM Function—Wilson et al. (2012)
5	1	1.00	.031	5	4
6	1	1.00	.016	6	5
7	1	1.00	.008	7	6
8	.875	.750	.035	7	6
9	.889	.778	.020	8	7
10	.900	.800	.011	9	8
11	.818	.636	.033	9	8
12	.833	.667	.019	10	9
13	.769	.538	.046	10	9
14	.786	.571	.029	11	10
15	.800	.600	.018	12	11
16	.750	.500	.038	12	11
17	.765	.529	.025	13	12
18	.722	.444	.048	13	12
19	.737	.474	.032	14	13
20	.750	.500	.021	15	14

Figure 1. CVR Minimum Critical Values Table (Ayre & Scally, 2014).

The formula required for calculating the content validity ratio (CVR) is given below (Lawshe, 1975).

- CVR = $\frac{Ng}{N/2} - 1$
- Ng = Number of experts who rated the item as appropriate
- N = Total number of experts who expressed an opinion on the item

▶**Table 4** below presents the expert evaluations;

Table 4. Expert Evaluations

Items	Suitable	Fixable	Unsuitable	Total Experts
Item 1	8	0	0	8
Item 2	2	0	6	8
Item 3	5	0	3	8
Item 4	6	0	2	8
Item 5	4	0	4	8
Item 6	7	1	0	8
Item 7	8	0	0	8
Item 8	8	0	0	8
Item 9	5	0	3	8
Item 10	8	0	0	8
Item 11	8	0	0	8
Item 12	6	0	2	8
Item 13	7	1	0	8
Item 14	3	0	5	8
Item 15	7	1	0	8
Item 16	8	0	0	8
Item 17	8	0	0	8
Item 18	8	0	0	8
Item 19	5	0	3	8
Item 20	6	0	2	8
Item 21	8	0	0	8
Item 22	4	0	4	8
Item 23	8	0	0	8
Item 24	4	0	4	8
Item 25	6	0	2	8
Item 26	7	1	0	8
Item 27	8	0	0	8
Item 28	4	0	4	8
Item 29	5	0	3	8
Item 30	4	0	4	8
Item 31	8	0	0	8
Item 32	8	0	0	8
Item 33	7	1	0	8
Item 34	5	0	3	8

The calculation of the Content Validity Ratio (CVR) for the scale items for which expert opinions were obtained is shown in ▶**Table 5** below.

Table 5. Content Validity Ratio (CVR) Calculation

Items	Number of Experts Who Rated the Item as Appropriate (Ng)	Half of the Total Number of Experts (N/2)	Content Validity Ratio (CVR)	Verdict
Item 1	8	4	1,00	Accept
Item 6	7	4	0,75	Accept
Item 7	8	4	1,00	Accept
Item 8	8	4	1,00	Accept
Item 10	8	4	1,00	Accept
Item 11	8	4	1,00	Accept
Item 13	7	4	0,75	Accept
Item 15	7	4	0,75	Accept
Item 16	8	4	1,00	Accept
Item 17	8	4	1,00	Accept
Item 18	8	4	1,00	Accept
Item 21	8	4	1,00	Accept
Item 23	8	4	1,00	Accept
Item 26	7	4	0,75	Accept
Item 27	8	4	1,00	Accept
Item 31	8	4	1,00	Accept
Item 32	8	4	1,00	Accept
Item 33	7	4	0,75	Accept
Total Number of Experts: 8				
Calculated Content Validity Index (CVI): 0,930				
Content Validity Criteria (CVC) Table Value: 0,750				

The Content Validity Index (CVI) is calculated by taking the average of the CVR values of all items. This helps to evaluate the overall validity of the scale (Polit & Beck, 2006).

- CVI = $\frac{\sum CVR}{NI}$
- CVR = Sum of content validity ratios of remaining items
- NI = Number of remaining items
- CVI = 0,930

An obtained CVI value greater than the CVO value (CVI > CVR) indicates that the content validity of the remaining items in the scale is statistically significant (Yeşilyurt & Çapraz, 2018). As a result of the calculation, the Content Validity Index (CVI) value (0.930) was found to be higher than the Content Validity Ratio (CVR) value (0.750). Therefore, the content validity of the items was found to be statistically significant. As a result of the content validity analyses, it was observed that experts suggested corrections for five items, and these items were re-examined and corrected. Consequently, a total of 18 items with established content validity were created.

https://doi.org/10.32329/uad.1899455

Before conducting the pilot study of the scale, an attention-check item was added to the scale items to increase the reliability of the scale. Adding a control item to a scale is a method used to check whether participants answer the questions by reading them or randomly. This increases the reliability of the data (Yıldırım Seheryeli & Gören, 2025). In the study, attention items containing very clear instructions, which are a type of control item, were used. For example, the control item "Mark the 'Undecided' option in this item" was randomly added towards the middle of the scale items.

2.4.2. Second Stage

After the scale was finalized, it was sent via e-mail to 63 participants for a pilot (pre-test) application. The pre-test application allows the prepared scale to be reviewed on a small group similar to the intended target audience. The pre-test application is important for determining the readability of the items, the response time of the test, and the parts that respondents do not understand (Crocker & Algina, 1986). After the survey was delivered to the participants, no time limit was applied to their responses.

According to Yıldırım Seheryeli & Gören (2025), at this stage, it is important to first examine and clean the data. Participants who answered all items of the scale in the same way or formed a pattern should not be included in the analysis. Similarly, participants should be excluded from the analysis based on the control item. In the study, as a result of the pre-application, 8 participants who selected the same response for all items or failed the attention-check item were excluded from the analyses, by analyzing the control item added to the scale with other items in the scale, were excluded from the evaluation, and a total of 55 participants' surveys were included in the evaluation. As a result of the data obtained from the pre-application, it was determined that the survey questions were clear and understandable.

2.4.3. Third Stage

In the final stage, the scale was applied to 295 participants in the EFA study group by sending it via e-mail. As in the pilot application and EFA stages, as a result of the CFA application, 23 participants who selected the same response for all items or failed the attention-check item were excluded from the analyses, by analyzing the control item added to the scale with other items in the scale, were excluded from the evaluation, and 272 participants' surveys were included in the evaluation. After that, EFA and reliability analysis were performed on the data obtained from 272 participants. According to Kalaycı (2006), at this stage, the prerequisites for factor analysis, such as sample size and number of items, should be examined, outliers should be removed, and suitability tests for factor analysis should be performed.

In this direction, the measurement tool prepared during the scale development process was distributed to a randomly selected sample group from the research population, and the given responses were scored, and factor analysis was performed. Factor analysis can be used for two purposes in research: exploration and confirmation. Exploratory Factor Analysis (EFA) is used to reveal the common latent structure among variables and to see the possible theoretical structure, while Confirmatory Factor Analysis (CFA) is used to test hypotheses based on previous research/theories about the latent structure and to create empirical evidence as to whether the structure is supported (Goodwin, 1999).

To check whether the study sample size was suitable for factor analysis, reliability tests such as Kaiser Meyer Olkin (KMO) and Bartlett tests were performed. KMO values between 0.80 and 0.89 indicate a "very good" fit for factor analysis (Field, 2009). In the study, the KMO value was .832, indicating suitability for factor analysis. Bartlett's sphericity test, on the other hand, tests the homogeneity and consistency of the factors (Yurdugül, 2005). In the study, the Bartlett's test was statistically significant level of $p<0.01$. After determining the suitability for factor analysis and the principal axis factoring analysis technique was used to reveal the construct validity of the scale. This method was preferred because it reveals the latent structures and common variance among a large number of variables (Büyüköztürk, 2020).

To confirm the 18-item, four-factor derived structure obtained from EFA, CFA was applied to 264 participants from a different sample group by sending it via e-mail. As in the pilot application and EFA stages, as a result of the CFA application, 16 participants who selected the same response for all items or failed the attention-check item were excluded from the analyses, by analyzing the control item added to the scale with other items in the scale, were excluded from the evaluation, and 248 participants' surveys were included in the evaluation. Then, CFA was applied to the data obtained from 248 participants via the SPSS AMOS program. Confirmatory Factor Analysis (CFA) is used when there is prior knowledge about the measured construct and it is known which item measures which dimension, to determine whether this information is correct.

2.5. Data Analysis

Data analysis can be defined as the process of generating scientifically valid conclusions through the application of appropriate statistical techniques to collected data

(Büyüköztürk, 2020). Within the scope of this study, validity and reliability analyses were conducted during the development of the Artificial Intelligence Visual Literacy Scale. Sample adequacy for factor analysis was evaluated and determined to be sufficient (Tabachnick & Fidell, 2013). The scale items were grouped under four factors, and subsequent confirmatory analyses were conducted. SPSS and SPSS AMOS software programs were used for all statistical analyses.

Exploratory Factor Analysis was conducted using Principal Axis Factoring extraction and Direct Oblimin rotation methods to determine the construct validity of the scale. Confirmatory Factor Analysis was subsequently employed to verify the factor structure obtained from EFA. Factor loadings of .30 or higher were considered acceptable (Çokluk, Şekercioğlu, & Büyüköztürk, 2016). Cronbach's Alpha coefficients were calculated for both the overall scale and its subdimensions. Since sample size plays a critical role in the reliability of factor analysis results, careful attention was paid to participant adequacy throughout the study.

A total of 622 participants initially took part in the study. Across all implementation stages (pilot study, EFA, and CFA), 47 participants identified as providing random responses based on the control item were excluded from the analyses. Consequently, the final analyses were conducted using data obtained from 575 participants. This sample size satisfies the minimum sample size recommendations proposed in the literature for scale development studies (Comrey & Lee, 1992).

3. Findings and Interpretations

This section of the research presents the outputs and interpretations obtained from the analyses performed on the data collected from the Artificial Intelligence Visual Literacy Scale, developed by the researchers, using SPSS and SPSS AMOS statistical programs.

3.1. Artificial Intelligence Visual Literacy Scale Reliability Analysis

To determine the reliability of the scale developed in this study, Cronbach's Alpha coefficient was calculated, and item analysis was performed. Reliability analysis is a widely used method to evaluate the consistency of scale items with each other and to determine the internal consistency level of the measurement tool (Tavşancıl, 2006). Generally accepted reliability thresholds in the literature are as follows: 0.90 and above is considered "excellent"; 0.80–0.89 "very good"; 0.70–0.79 "acceptable"; 0.60–0.69 "questionable"; 0.50–0.59 "poor"; and below 0.50 "unacceptable." Generally, values of 0.70 and above are interpreted as indicating that the scale is reliable (Tavakol & Dennick, 2011). For the 18-item Artificial Intelligence Visual Literacy Scale, the prerequisite for factor analysis is a high level of correlation between variables and a KMO value above 0.70 (Pallant, 2001). The reliability analysis results of the study are detailed in ▶**Table 6**.

Table 6. Reliability Analysis Results of the Artificial Intelligence Visual Literacy Scale

Items	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Item 1	65.31	61.367	.570	.634	.873
Item 2	65.53	62.826	.440	.554	.877
Item 3	65.31	62.143	.503	.588	.875
Item 4	65.70	60.772	.495	.727	.876
Item 5	65.54	61.659	.496	.728	.875
Item 6	65.62	60.385	.486	.643	.876
Item 7	65.47	62.472	.486	.706	.876
Item 8	65.42	62.481	.453	.832	.877
Item 9	65.63	61.010	.525	.801	.874
Item 10	65.36	59.397	.576	.487	.872
Item 11	65.25	59.266	.615	.834	.871
Item 12	65.38	60.827	.559	.798	.873
Item 13	65.43	60.335	.449	.430	.878
Item 14	65.37	60.064	.590	.892	.872
Item 15	65.49	59.387	.537	.455	.874
Item 16	65.50	62.443	.484	.745	.876
Item 17	65.38	61.018	.474	.702	.876
Item 18	65.45	63.348	.415	.550	.878

Cronbach's Alpha = ,881

To reveal the relationship of each item with other items in the scale, Item-Total Statistics values were analyzed. This analysis helps to determine the extent to which each item represents the entire scale. When ▶**Table 6** is examined, it is seen that the Corrected Item-Total Correlation values for all items of the scale range from .415 to .615. In the literature, correlations below 0.30 are considered weak, while values of 0.30 and above indicate that the items contribute positively to the scale's integrity (Büyüköztürk, 2021). Therefore, all items being above this threshold indicates that the developed items are consistent with the measured construct and do not disrupt the scale's integrity.

Furthermore, Cronbach's Alpha if Item Deleted values were examined to determine whether there was a significant increase in the Cronbach's Alpha value when an item was removed. According to the analysis results, it was observed that removing any item did not signifi-

https://doi.org/10.32329/uad.1899455

cantly increase the overall reliability of the scale. For example, if Item 2 (.440), Item 13 (.449), and Item 18 (.415), which have relatively lower Corrected Item-Total Correlation values, were removed, the Alpha (α) value would remain at .877 and .878, respectively. This indicates that these items do not harm the integrity of the scale, are consistent with the overall scale, and do not create a structure dependent on any single item. ▶**Table 7** below shows the overall reliability value of the scale, item-total correlation ranges, and Alpha (α) values in case of item deletion.

Table 7. Reliability and Item Analysis Results of the Scale

Scale Feature	Value Range / Result
Cronbach's Alpha (All Scale)	,881
Number of Items	18
Item-Total Correlation Range	,415 – ,615
Alpha Range (When Items is Removed)	,877 – ,878

As a result of the analysis performed in the SPSS program, the Cronbach's Alpha coefficient for the entire 18-item scale was found to be .881. This value indicates that the developed scale is sufficiently reliable and that the items forming the scale work consistently with each other.

In conclusion, both the overall reliability coefficient and the findings related to item analysis indicate that the developed scale has a sufficiently reliable level of consistency statistically and that all items establish significant relationships with the measured construct. In line with these findings, it was not deemed necessary to remove any item from the scale, and it was evaluated that the scale reached an adequate and ideal level of reliability.

3.2. Exploratory Factor Analysis (EFA)

Exploratory Factor Analysis (EFA) was performed to determine the construct validity and reveal the factor structure of the Artificial Intelligence Visual Literacy Scale (AIVLS). EFA is one of the fundamental statistical techniques used to determine the underlying factor structure of a scale and to reveal the loadings of items on these factors. For this purpose, Principal Axis Factoring and Direct Oblimin methods were used. Principal Axis Factoring (PAF) is an EFA extraction method used to reveal latent structures and common variance among a large number of variables. It is particularly preferred when the data do not meet the assumption of normal distribution and when common variances are to be estimated. The Direct Oblimin method is used when there is believed to be a relationship between factors (Büyüköztürk, 2021). Before starting EFA, Kaiser Meyer Olkin (KMO) measure and Bartlett's sphericity test were applied to test the suitability of the data for factor analysis.

Table 8. KMO and Bartlett Test Results of the Artificial Intelligence Visual Literacy Scale

KMO and Bartlett Test Results		
Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy		,832
Bartlett's Test of Sphericity	Approx. Chi-Square	3668,494
	df	153
	Sig.	<,001

When ▶**Table 8** is examined, the Kaiser Meyer Olkin (KMO) sample adequacy value was found to be .832, indicating that the sample size is sufficient for EFA (KMO > .80 = very good level). This value is considered sufficient by Field (2009) when it is above 0.50 and is classified as "very good" between 0.80-0.90. The KMO values calculated for each item were also found to be at least 0.749a, confirming the adequacy of the sample. In addition, as a result of the Bartlett Test, x2=3668.494; df=153; p<0.001 was found, and this result indicates that there are significant correlations between variables suitable for factor analysis.

As a result of EFA, it was determined that the 18-item Artificial Intelligence Visual Literacy Scale consists of a 4-dimensional (factor) structure with an eigenvalue greater than 1, and these 4 factors explain 67.738% of the total variance. This explained total variance is above the minimum limit of 50% accepted in social sciences and indicates a significant structural integrity in the scale (Tabachnick & Fidell, 2013). Accordingly, it is seen that the factor loading values of all items in the 4-factor structure of the scale are high and sufficient (Güler & Günel, 2022). ▶**Table 9** presents the distribution of items according to factors and factor loadings.

When ▶**Table 9** is examined, the variance ratios explained by the factors for the structure are calculated as follows: 31.977% for Factor 1; 14.676% for Factor 2; 12.001% for Factor 3; and 9.083% for Factor 4. In addition, it was concluded that the structure explains 67.738% of the total variance. The first factor consists of 6 items (items 10, 11, 12, 13, 14, 15), the second of 4 items (items 8, 16, 17, 18), the third of 4 items (items 4, 5, 6, 9), and the fourth of 4 items (items 1, 2, 3, 7). The lowest factor loading was observed for Item 13 with a factor loading value of .561, and highest for Item 8 with a factor loading value of .989. Therefore, since factor loadings of 0.40 and above are ideally accepted (Field, 2009), even the item with the lowest factor loading was considered to make a significant contribution.

Table 9. Factor Analysis Findings of the Artificial Intelligence Visual Literacy Scale

Items	Factor 1	Factor 2	Factor 3	Factor 4
Item 14	.949			
Item 11	.945			
Item 12	.844			
Item 10	.597			
Item 15	.574			
Item 13	.561			
Item 8		.989		
Item 16		.864		
Item 17		.833		
Item 18		.732		
Item 9			.930	
Item 4			.866	
Item 5			.822	
Item 6			.743	
Item 7				.960
Item 3				.759
Item 1				.750
Item 2				.716
Total	6.068	2.902	2.413	1.947
Variance Explained	31.977	14.676	12.001	9.083
Total Variance Explained	67.738			

Extraction Method: Principal Axis Factoring.
Rotation Method: Oblimin with Kaiser Normalization.[a]
a. Rotation converged in 7 iterations.

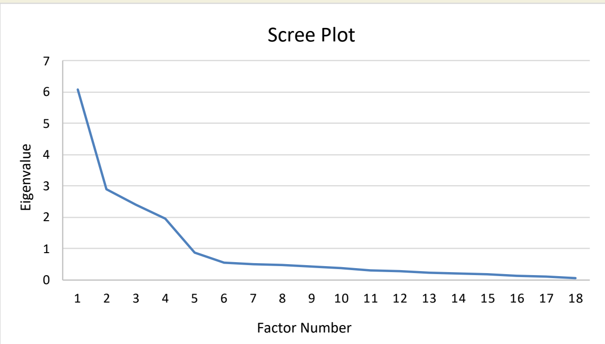

Figure 2. Scree Plot Test Result of the Artificial Intelligence Visual Literacy Scale

When ▶**Figure 2** is examined, it is seen that the eigenvalue of the 18-item scale is above 1, and the common variance of the factors ranges from .561 to .989. In addition, according to the data in the figure, 18 items are grouped into four factors. After the fourth factor, the graph tends towards a horizontal position, indicating that limiting the number of factors to four would be sufficient. ▶**Table 10** below shows the items related to the sub-dimensions of the scale and their overall reliability values.

Table 10. Reliability Results for AIVLS Sub-Dimensions

Sub-Dimensions	Number of Items	Items	Cronbach's Alpha (α)
Interpretation and Analysis	6	10, 11, 12, 13, 14, 15	,882
Design and Creation	4	8, 16, 17, 18	,912
Ethical and Legal Issues	4	4, 5, 6, 9	,909
Identification and Access	4	1, 2, 3, 7	,879
All Scale	18		,881

As seen in ▶**Table 10**, the reliability of the "Interpretation and Analysis" sub-dimension is .882; the reliability of the "Design and Creation" sub-dimension is .912; the reliability of the "Ethical and Legal Issues" sub-dimension is .909; and the reliability of the "Identification and Access" sub-dimension is .879. These findings regarding the sub-dimensions of the scale indicate that the scale consistently represents the concept it aims to measure.

3.3. Confirmatory Factor Analysis (CFA)

Confirmatory Factor Analysis (CFA), also known as a measurement model, is used to test the relationships between observed variables and the construct(s) assumed to be measured through these observed variables (Wetson & Gore, 2006). Like other Structural Equation Models (SEM), CFA models are based on a theoretical foundation (Schreiber, Stage, King, Nora & Barlow, 2006) and thus differ from exploratory factor analysis (Byrne, 2010; Mulaik, 2009).

In Structural Equation Model (SEM) research, although there is a consensus among researchers regarding the reporting of $\chi2/df$ (Mulaik et al., 1989), different researchers have made various suggestions regarding which other fit indices should be reported. McDonald and Ho (2002) suggest CFI, GFI, NFI, and NNFI (TLI); Garver and Mentzer (1999) suggest RMSEA, CFI, and NNFI (TLI); Brown (2006) suggests RMSEA, SRMR, CFI, and NNFI (TLI); and Iacobucci (2010) recommends reporting CFI and SRMR fit indices. Gerbing and Anderson (1985) state that different fit indices can be reported depending on the researcher's objective.

According to Gerbing and Anderson (1985), answering the question "Which fit indices should be reported in research?" is as difficult as answering the question "Which is the best car in a car dealership?" Just as the definition of the best car can vary according to one's purpose, the fit indices that should be reported in research also vary according to the researcher's objective. Just as the best car for one individual might be the fastest car, while for another it might be the safest car, researchers with different objectives may report different fit indices.

https://doi.org/10.32329/uad.1899455

Table 11. Confirmatory Factor Analysis (CFA) Results and Fit Index Values

Fit Indices	Perfect Fit	Acceptable Fit	Fit Indices Obtained
[1]CMIN/DF (χ^2/df)	0 ≤ ^{2}/df ≤ 3	3 ≤ ^{2}/df ≤ 5	3,070
[6]AGFI	0,90 ≤ AGFI ≤ 1	0,85 ≤ AGFI ≤ 0,90	0,876
[2]CFI	0,95 ≤ CFI ≤ 1	0,90 ≤ CFI ≤ 0,95	0,926
[3]GFI	0,95 ≤ GFI ≤ 1	0,90 ≤ GFI ≤ 0,95	0,915
[3]IFI	0,95 ≤ IFI ≤ 1	0,90 ≤ IFI ≤ 0,95	0,927
[3]TLI	0,95 ≤ TLI ≤ 1	0,90 ≤ TLI < 0,95	0,913
[4]RMSEA	0,00 ≤ RMSEA ≤ 0,05	0,05 ≤ RMSEA ≤ 0,08	0,059
[5]SRMR	0,00 ≤ SRMR ≤ 0,05	0,05 ≤ SRMR ≤ 0,10	0,058

1(Kline, 2005), 2(Tabachnick & Fidell, 2013), 3((Baumgartner & Homburg, 1996; Bentler, 1980; Bentler & Bonett, 1980; Marsh vd., 2006), 4(Çokluk, Şekercioğlu & Büyüköztürk, 2025), 5(Browne & Cudeck, 1993), 6(Schermelleh-Engel & Moosbrugger, 2003).

In this study, to confirm the 18-item, 4-factor derived structure obtained from EFA, CFA was applied via the SPSS AMOS program to data obtained from a different sample than the EFA stage, consisting of 264 participants. As in the EFA study group application, as a result of the CFA study group application, 16 participants who selected the same response for all items or failed the attention-check item were excluded from the analyses, by analyzing the control item added to the scale with other items in the scale, were excluded from the evaluation, and 248 participants' surveys were included in the evaluation. This verification process is performed to confirm the stability and consistency of the factor structure across different samples (Karagöz, 2019). When the fit indices obtained from the CFA fit index results for the developed scale were examined, it was determined that the model showed an "acceptable" level of fit. ▶**Table 11** shows the CFA fit index results obtained from the 18-item and 4-factor structure of the scale and some of the critical values that the fit indices should meet.

When ▶**Table 11** is examined, considering the critical values that the fit indices should meet, the CMIN/DF, AGFI, CFI, GFI, IFI, TLI, RMSEA, and SRMR fit index values were determined to indicate an "acceptable" level of fit for the model (Hu & Bentler, 1999). In line with these findings, the model is generally in harmony with the data. Thus, the validity of the 18-item and 4-factor measurement structure revealed by EFA was confirmed on a sample different from the EFA stage sample. The factor loadings for the 18-item and 4-factor model obtained as a result of CFA are shown in ▶**Figure 3**.

As seen in ▶**Figure 3**, factor loadings range from .69 to .81 for the "Interpretation and Analysis" factor, from .68 to .77 for the "Design and Creation" factor, from .66 to .76 for the "Ethical and Legal Issues" factor, and from .65 to .77 for the "Identification and Access" factor. The results indicate that indicate acceptable model fit (Schermelleh-Engel, Moosbrugger & Müller, 2003), and furthermore, the 18-item and 4-factor structure can be said to be a scale that adequately reflects students' artificial intelligence visual literacy levels. All other findings obtained with the CFA result are summarized in ▶**Table 12** below.

Table 12. Summary Table of All Findings Obtained After CFA

		Value Range / Result
Number of Items		18
Number of Participants		622 PILOT (63) + EFA (295) + CFA (264)
Number of Participants Excluded from Evaluation		47
Kaiser Meyer Olkin (KMO)		,832
Bartlett's Test	Approx. Chi-Square	3668,494
	df	153
	Sig.	<,001
Cronbach's Alpha (All Scale)		,881
Item-total Correlation Range		,415 – ,615
Alpha Range (When Items is Removed)		,877 – ,878
Number of Factor		4
Factor Total Values (Cronbach's Alpha)	Factor 1	,882
	Factor 2	,912
	Factor 3	,909
	Factor 4	,879
Total Factor Variance		67,738
KMO Lowest/Highest Item Value Range		,749[a] – ,939[a]

4. Conclusion and Discussion

Measurement tools such as scales contribute to scientifically determining the validity and reliability of a situation. However, developing a measurement tool is a comprehensive and intensive process. The aim is to make evaluations about individuals, events, or objects in terms of the characteristic being measured and to make

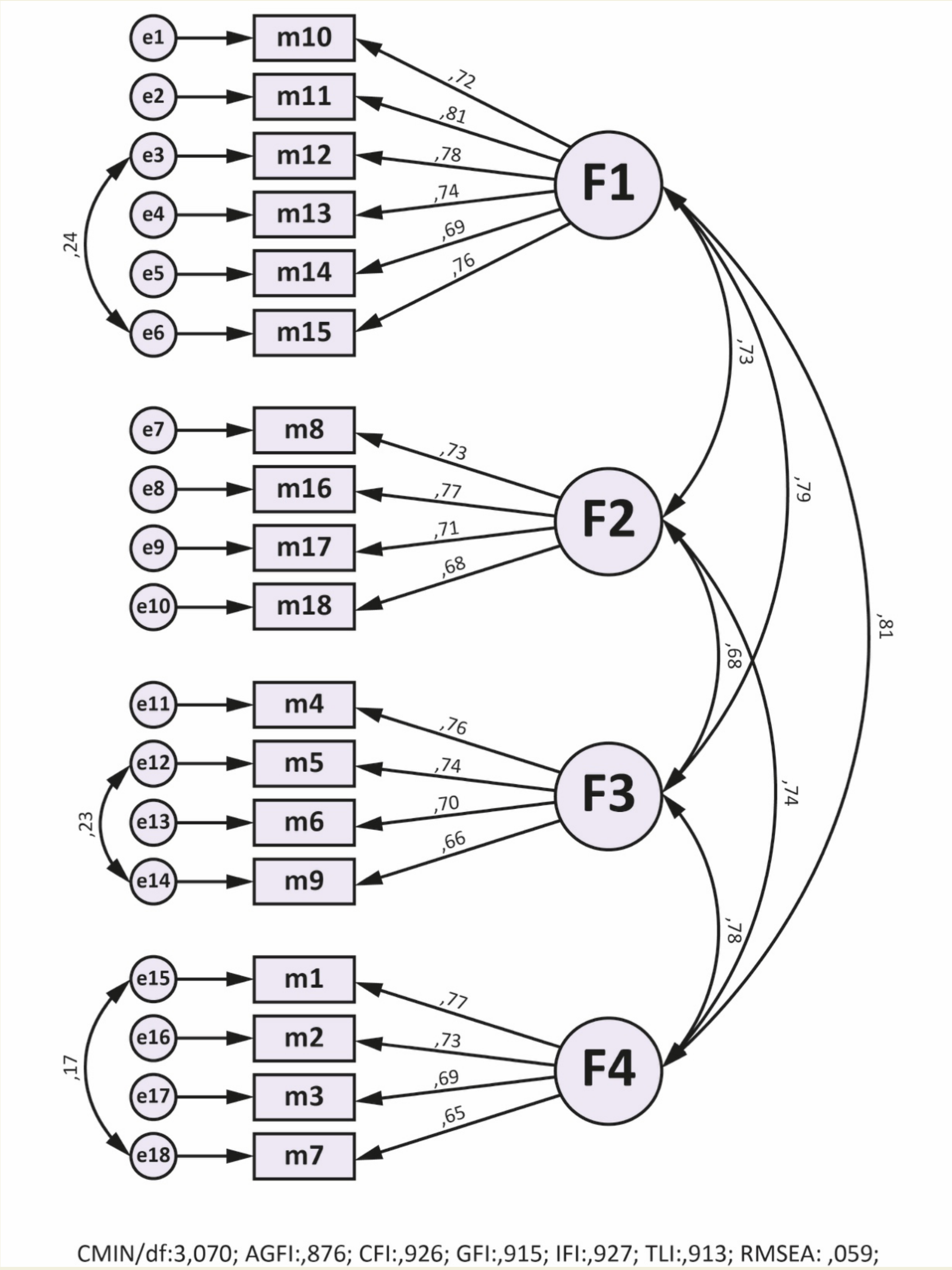

Figure 3. Confirmatory Factor Analysis (CFA) Model

certain decisions based on the evaluation results (Ercan & Kan, 2004). Researchers, in the process of developing a measurement tool, should consider (1) whether they will obtain evidence for the structure of the characteristic they are measuring, (2) whether the items can be defined under a specific structure, and (3) what kind of pattern the correlations between structures will reveal (Karakoç & Dönmez, 2014). The study process should be planned within the framework of these three important points. In the first stage of the research, a comprehensive literature review was conducted, and an item pool was created by utilizing scales developed in the fields of artificial intelligence literacy, visual literacy, media literacy, and digital literacy. It is observed that a large part of the scales developed in the field of artificial intelligence literacy focus on technical knowledge, ethical awareness, evaluation, and usage skills. The four-factor structure of the Artificial Intelligence Literacy Scale developed by Pinski & Benlian (2023) is based on awareness, usage, evaluation, and ethical dimensions. Similarly, in

https://doi.org/10.32329/uad.1899455

the MAILS scale developed by Carolus and colleagues, multi-dimensional structures such as understanding, using, detecting, and ethically evaluating artificial intelligence come to the forefront. The multi-dimensional structure of the scale developed in this research also shows parallelism with current studies in the literature.

In this research, the Artificial Intelligence Visual Literacy Scale (AIVLS) was developed to measure individuals' competencies in understanding, evaluating, and producing AI-supported visual content in today's digital age. The use of artificial intelligence in universities has the potential to create a major transformation in learning processes. However, in this process, justice, access, and ethical values must be prioritized (Slimi & Villarejo Carballido, 2023; cited by Telli & Aydın, 2025). In the study, the scale development process was carried out based on classical measurement theory, and validity and reliability analyses were meticulously performed during this process. The research findings indicate that the developed scale presents a theoretically grounded and psychometrically strong structure.

4.1. Evaluation of Construct Validity and Factor Structure

As a result of the Exploratory Factor Analysis (EFA) conducted to test the construct validity of the scale, it was determined that the AIVLS has a four-dimensional structure. In the analyses regarding the prerequisites for EFA, the Kaiser-Meyer-Olkin (KMO) sample adequacy coefficient was calculated as .832. This value falls within the "very good" level limits defined by Field (2009), indicating that the sample is highly suitable for factor analysis. In addition, the Bartlett's sphericity test result (χ^2=3668.494; df=153; p<.001) was significant, indicating a sufficient level of correlation between variables. These results are considered to provide a statistically sufficient and ideal basis for the application of EFA.

The four factors obtained as a result of the factor analysis explain 67.738% of the total variance. This ratio is well above the 50% limit frequently accepted in social sciences (Tabachnick & Fidell, 2013). The first factor explains 31.977% of the variance, the second factor 14.676%, the third factor 12.001%, and the fourth factor 9.083%. The factor loadings of the items range from .561 to .989. All values above 0.40 are considered ideal (Field, 2009), and the significant conceptual differentiations between factors indicate that the scale presents a structure consistent with its theoretical dimensions. Furthermore, the acceptable level of factor loadings indicates that the items adequately represent the relevant factors.

These results show similarities with other scale development studies conducted in the field of media literacy and artificial intelligence literacy. For example, a four-factor structure was also obtained in the Media Literacy Skills Scale developed by Eristi & Erdem (2017), and the Cronbach's Alpha coefficient was found to be high. In addition, multi-dimensional factor structures are observed in the SNAIL scale developed in the field of AI literacy (Laupichler et al., 2023).

4.2. Interpretation of Reliability and Item Analysis Findings

As a result of the reliability analysis conducted to evaluate the internal consistency of the scale, the overall Cronbach's Alpha coefficient of the scale was determined to be .881. This value indicates that the items in the scale work with a very good level of consistency (George & Mallery, 2003). The Cronbach's Alpha values being above 0.70 for all sub-dimensions indicate that each sub-area of the scale also provides consistent measurements separately. These results appear quite strong when compared to other scales developed in the field of AI literacy. For example, in the Artificial Intelligence Literacy study adapted to Turkish by Çelebi et al. (2023), the Cronbach Alpha coefficients were stated to be at an acceptable level. Similarly, high internal consistency coefficients were obtained in Eristi & Erdem's (2017) Media Literacy Skills Scale study.

The corrected item-total correlations examined within the scope of item analysis range from .415 to .615. These values indicate that the items are consistent with the conceptual structures they measure and contribute to the integrity of the scale. While correlations below 0.30 are considered weak in the literature, all items in this study being above this threshold is a positive finding (Büyüköztürk, 2021).

In addition, Cronbach's Alpha if Item Deleted analyses were also performed, and it was observed that removing any item did not significantly increase the overall reliability of the scale. Specifically, even when item 2, item 13, and item 18, which have lower factor loadings compared to other items, were removed, the Cronbach's Alpha value remained in the average range of .877-.878. This finding indicates that all items make a significant contribution to the scale and that their removal from the scale is not necessary. However, it may also necessitate re-evaluation of these items in future studies. Nevertheless, it is known that scale development studies should not be limited to statistical criteria; the theoretical contributions and content validity of the items should also be taken into consideration. Therefore, it would be more appropriate to re-test the relevant items in different samples.

4.3. Confirmatory Factor Analysis (CFA) Results and Support for the Structure

Confirmatory Factor Analysis (CFA), conducted with the second sample group of the research, was performed to test the validity of the four-dimensional structure revealed by EFA. CFA reveals the extent to which a theoretically defined structure overlaps with observed data (Byrne, 2010). The fit indices obtained from the CFA results are as follows: $\chi^2/df = 3.070$; AGFI = .876; CFI = 0.926; GFI = .915; IFI = 0.927; TLI = 0.913; RMSEA = 0.059 and SRMR = 0.058. These results show similarities with other scales developed in the field of artificial intelligence literacy. For example, in the Conceptualizing AI literacy and AI Literacy Questionnaire studies developed by Ng et al. (2021; 2023), acceptable fit indices were obtained as a result of CFA, and the multi-dimensional structure of the scale was confirmed. Similarly, in the AI Literacy Scale studies developed by Pingmuang, Koraneekij, and Khlaisang (2026) for prospective teachers, model validation was performed with CFA, and multi-dimensional structures were supported.

These values indicate that all fit indices are at an "acceptable" level and that the measurement model adequately fits the data (Hu & Bentler, 1999; Schermelleh-Engel, Moosbrugger & Müller, 2003). In particular, the fact that commonly used indices such as RMSEA, CFI, and SRMR are within "acceptable" limits indicates that the four-factor structure is theoretically supported and that the model works consistently.

4.4. General Evaluation

The Artificial Intelligence Visual Literacy Scale developed within the scope of this study can be said to be a sufficiently reliable measurement tool in terms of both content and construct validity. The scale covers multi-dimensional skills such as establishing a relationship between artificial intelligence and visual content, critically analyzing this content, producing new visual designs, and developing ethical-legal awareness. This approach aims to make the multi-faceted nature of artificial intelligence visual literacy, which is not limited to tool use, visible by ensuring that interpretations are directly based on the study's findings. Each dimension offers sub-scales with high reliability within itself; this indicates that the scale can be easily used in education, communication, media, art, and design research.

One of the important findings of the research is that artificial intelligence visual literacy does not consist solely of technical usage skills. When the factor structures in the scale are examined, it is understood that individuals need to possess not only the ability to use AI-supported visuals but also the skills to analyze, verify, ethically evaluate, and critically interpret them. This situation also overlaps with the media and artificial intelligence literacy approach emphasized by the OECD within the scope of PISA 2029. According to the OECD, the ability of individuals to evaluate the reliability, purpose, and accuracy of digital content has become one of the fundamental components of contemporary literacy (OECD, 2024).

In future research, with more balanced samples covering different universities, the relationship between artificial intelligence visual literacy and more "proximal" educational variables (Yaşlıca & Altay, 2026) such as the use of artificial intelligence tools for production purposes, course and workshop experiences, project-based learning processes, and ethics, beyond demographic variables, can be examined through multivariate models. In addition, examining the test-retest reliability of the scale and investigating its relationships with different variables will contribute to the literature.

In conclusion, the Artificial Intelligence Visual Literacy Scale can be used as an effective, valid, and reliable tool for measuring individuals' abilities to understand and produce complex visual and artificial intelligence content they encounter in the digital age. It is thought that the use of this scale in both academic research and the evaluation of curricula will make significant contributions to monitoring the development of individuals' visual literacy skills for 21st-century technological opportunities and innovations.

5. Recommendations

In line with the research findings, the Artificial Intelligence Visual Literacy Scale is recommended as a valid and reliable measurement tool that can be used in both academic research and educational settings. In this regard, the following recommendations have been developed:

5.1. Recommendations for Practitioners

- The scale can be used for in-class assessment in education faculties or art and design departments. In this way, how students perceive, use, and ethically evaluate artificial intelligence tools can be measured.

- Educators can use this scale to analyze students' critical thinking, digital productivity, and ethical awareness levels and revise their lesson plans accordingly.

https://doi.org/10.32329/uad.1899455

- The scale can be evaluated as a needs analysis tool to understand the level of interaction of students with artificial intelligence-based visual production tools, especially in distance education environments.

5.2. Recommendations for Researchers

- It is recommended to conduct validity-reliability studies of the scale on different age groups (e.g., high school or adult education). This can increase the generalizability of the scale.
- The cross-cultural validity of the scale can be examined by conducting applications in different cultural contexts.
- Relational research can be conducted between the scale and variables such as artificial intelligence education, media literacy, or creativity. Especially correlation analyses with creativity and critical thinking may provide guidance.
- The scale can be used with a pre-test–post-test method to evaluate the effectiveness of artificial intelligence-supported learning environments.
- In addition, a dynamic version of the scale can be created by adding new items in line with developing artificial intelligence technologies.

Research Ethics

This research was conducted with permission granted by the Gazi University Ethics Committee's decision number 06, dated 15/04/2025.

Author Contributions

The authors contributed equally to the research process.

Competing Interests

The authors have no conflict of interest to declare.

Research Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sector.

Appreciation

We would like to thank all the students studying in the fine arts education departments of the faculties, institutes, colleges, and vocational schools of the 45 universities (state/foundation) who contributed to the creation of this research by answering the survey questions during the data collection process.

Orcid

Abdulkerim Turkaya https://orcid.org/0000-0003-0461-0925
Ali Ertuğrul Küpeli https://orcid.org/0000-0003-3485-8489
Adnan Kan https://orcid.org/0000-0002-3610-0033

References

Avgerinou, M. D., & Pettersson, R. (2011). Toward a cohesive theory of visual literacy. *Journal of Visual Literacy*, *30*(2), 1-19. https://doi.org/10.1080/23796529.2011.11674687

Ayçiçek, B. (2025). Attitudes toward artificial intelligence in higher education: A mixed-methods study. *Journal of University Research*, *8*(4), 517-534. https://doi.org/10.32329/uad.1770973

Ayre, C., & Scally, A. J. (2014). Critical values for Lawshe's content validity ratio: revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development*, *47*(1), 79-86.

Baghestania, A., Ahmadi , F., Tanha, A., & Meshkat, M. (2017). Bayesian critical values for lawshe's content validity ratio. *Measurement and Evaluation in Counseling and Development*, *52*(1), 1-5.

Baumgartner, H., & Homburg, C. (1996). Applications of structural equation modeling in marketing and consumer research: A review. *International Journal of Research in Marketing*, *13*(2), 139-161.

Beck, C.T., & Gable, R.K. (2001). Ensuring content validity: An illustration of the process. *Journal of Nursing Measurement*, *9*, 201–215.

Bentler, P.M. (1980). Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology*, *31*, 419-456.

Bentler, P.M., & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588-606.

Bilgili, C., & Şener, G. (2013). *Sosyal medya ve ağ toplumu-1: Kitle iletişiminde yaşanan değişimler*. Grafik Tasarım Yayıncılık. İstanbul, 5, ISBN: 978-605-60892-1-3.

Büyüköztürk, Ş. (2020). *Sosyal bilimler için veri analizi el kitabı. İstatistik, araştırma deseni SPSS uygulamaları ve yorum* (27. Baskı). Ankara: Pegem Akademi.

Büyüköztürk, Ş. (2021). *Sosyal bilimler için veri analizi el kitabı. İstatistik, araştırma deseni SPSS uygulamaları ve yorum* (28. Baskı). Ankara: Pegem Akademi.

Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2024). *Bilimsel araştırma yöntemleri* (36. Baskı). Ankara: Pegem Akademi.

Brandt, D.S. (2001). Information technology literacy: Task knowledge and mental models. *Library Trends*, *50*(1), 73-86. https://www.learntechlib.org/p/93628/

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). NY: The Guilford Press.

Calvani, A., Cartelli, A. & Ranieri, M. (2008). Models and instruments for assessing digital competence at school. *Journal of E-Learning and Knowledge Society* *4*(3), 183-193.

Carolus, A., Koch, M. J., Straka, S., Latoschik, M. E., Wienrich, C., & Gámez Zerban, A. (2023). MAILS – Meta AI literacy scale: Development and testing of an AI literacy questionnaire based on well-founded competency models and psychological change- and meta-competencies. *Computers in Human Behavior*, *1*(2), 1-10. 100014.

Cemaloğlu, N. (2023). *Bilimsel araştırma teknikleri ve etik* (5. Baskı). Ankara: Pegem Akademi.

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, *8*, 75264-75278.

Crocker, L. ve Algina, J. (1986). *Introduction to classical and modern test theory*. New York:

Holt, Rinehart and Winston, Inc.

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Çam, M. B., Çelik, N. C., Turan Güntepe, E., & Durukan, Ü. G. (2021). Öğretmen adaylarının yapay zekâ teknolojileri ile ilgili farkındalıklarının belirlenmesi. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, *18*(48), 263-285. https://izlik.org/JA54YD38JK

Çelebi, C., Demir, U., & Karakuş, F. (2023). Artificial intelligence literacy: A systematic review. *Journal of Necmettin Erbakan University Ereğli Faculty of Education*, *5*(2), 536-560.

Çetin, M., & Aktaş, A. (2021). Yapay zekâ ve eğitimde gelecek senaryoları. *OPUS International Journal of Society Researches*, *18*(Eğitim Bilimleri Özel Sayısı), 4225-4268. https://doi.org/10.26466/opus.911444

Çokluk, O., Şekercioğlu, G., & Büyüköztürk, Ş. (2025). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları* (8. Baskı). Ankara: Pegem Akademi.

Davenport, T. H., & R. Ronanki. (2018). Artificial intelligence for the real world. *Harvard Business Review*, *96*(1), 108-116.

Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research*, *5*, 194-197. https://doi.org/10.1016/s0897-1897(05)80008-4

Ercan, İ., & Kan, İ. (2004). Ölçeklerde güvenirlik ve geçerlik. *Uludağ Üniversitesi Tıp Fakültesi Dergisi*, *30*(3), 211-216.

Eristi, B., & Erdem, C. (2017). Development of a media literacy skills scale. *Contemporary Educational Technology*, *8*(3), 249-267. https://izlik.org/JA68LX82YN

Field, A. (2009). *Discovering statistics using SPSS*, (3rd ed.). Sage Publications Ltd., London.

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference* (4th ed.). Boston: Allyn & Bacon.

Goodwin, L. D. (1999). The role of factor analysis in the estimation of construct validity. *Measurement in Physical Education and Exercise Science*, *3*(2), 85-100.

Güler, A., & Günel, Y. (2022). Kamu kurumlarında sanal kaytarmaya yönelik bir ölçek geliştirme çalışması. İşletme Bilimi Dergisi, *10*(2), 247-274. https://doi.org/10.22139/jobs.1147063

Güney, E., & Yavuz, H. (2022). Mekân ve yaşam kurgusu bağlamında yapay yaşam sanatı. *Uluslararası Kültürel ve Sosyal Araştırmalar Dergisi (UKSAD)*, *8*(2), Kış, 83-94.

Graf, A. (2023). *Yapay zekâ ve gelecek tasarımı*. Editör: Gökhan İnce. 1. Basım. Ankara: Gece Kitaplığı.

Grassian, E. S. & Kaplowitz, J.R. (2001). *Information literacy instruction*. Neal-Schuman, NY.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structural analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Hobbs, R. (2010). *Digital and media literacy. A plan of action*. Washington D.C.: The Aspen Institute.

Hoffman, M., & Blake, J. (2003). Computer literacy: Today and tomorrow. *Journal of Computing Sciences in Colleges*. *18*(5), 221-233.

Hwang, G. J., Xie, H., Wah, B. W., Gasevic, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *1*(2020), 100001.

Hwang, G.J., & Tu, Y.F. (2021). Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review. *Mathematics*, *9*(6), 584.

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons* *61*(4), 577-586.

Kalaycı, Ş. (Ed.) (2006). *SPSS uygulamalı çok değişkenli istatistik teknikleri*. Ankara: Asil Yayın Dağıtım.

Kandlhofer, M., G. Steinbauer, S. Hirschmugl-Gaisch, and P. Huber. (2016). Artificial intelligence and computer science in education: From kindergarten to university. *Paper Presented at the 2016 IEEE Frontiers in Education Conference (FIE)*.

Karagöz, Y. (2019). *SPSS 23 ve AMOS 23 uygulamalı istatistiksel analizler*. 2. Baskı. Ankara: Nobel Akademik Yayıncılık.

Karakoç, F. Y., & Dönmez, L. (2014). Ölçek geliştirme çalışmalarında temel ilkeler. *Tıp Eğitim Dünyası (TED)*, *13*(40), 39-49.

Karpunina, E. K., Dedov, S. V., Kholod, M. V., Ponomarev, S. V., & Gorlova, E. A. (2020). Artificial intelligence and its impact on economic security: trends, estimates and forecasts (pp.213-225). In E. G. Popkova, & B. S. (Eds), *Scientific and Technical Revolution: Yesterday, Today and Tomorrow*.

Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling*, (2nd ed.). New York: The Guilford Press.

Kurudayıcıoğlu, M., & Tüzel, S. (2010). The types of literacy of the 21st century, changing text comprehension and Turkish teaching. *Türklük Bilimi Araştırmaları*, *28*, 281-298.

Laupichler, M. C., Aster, A., Haverkamp, N. & Raupach, T. (2023). Development of the "Scale for the assessment of non-experts' AI literacy" – an exploratory factor analysis. *Computers in Human Behavior Reports*, *Volume 12*, 1-10. https://doi.org/10.1016/j.chbr.2023.100338

Lawshe, C. (1975). A quantitative approach to content validity. *Personnel psychology*, *28*(4), 563-575.

Lee, S., & J. Choi, J. (2017). Enhancing user experience with conversational agent for movie recommendation: Effects of self-disclosure and reciprocity. *International Journal of Human-Computer Studies*, *103*, 95–105.

Lin, T. B. (2021). Artificial intelligence and media literacy: A conceptual framework and research agenda. *Journal of Media Literacy Education*, *13*(2), 57-66. https://doi.org/10.23860/JMLE-2021-13-2-5

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. London: Pearson Education.

Luo, X., Tong, S., Fang, Z., ve Qu, Z. (2019). Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases. *Marketing Science*, *38*(6), 937-947. https://doi.org/10.1287/mksc.2019.1192

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, *13*(4), 410-425.

Long, D., & Magerko, B. (2020). *What is AI literacy? Competencies and design considerations*. CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Honolulu, USA: Association for Computing Machinery.

Marsh, H.W., Hau, K.T., Artelt, C., Baumert, J., & Peschar, J.L. (2006). OECD's brief self-report measure of educational psychology's most useful affective constructs: Cross-cultural, psychometric comparisons across 25 countries. *International Journal of Testing*, *6*(4), 311-360.

Metelskaia, I., Ignatyeva, O., Denef, S., and Samsonowa, T. (2018). A business model template for AI solutions. *Paper Presented at the Proceedings of the International Conference on Intelligent Science and Technology*.

Munro, M. C., Huff, S. L., Marcolin, B. L., and Compeau, D. R. (1997). Understanding and measuring user competence. *Information & Management*, *33*(1), 45-57.

Ng, T. K., Wu, W., Leung, J. K. L., & Chu, S. K. W. (2023). Artificial intelligence (AI) literacy questionnaire with confirmatory factor analysis. *In Proceedings of 2023 IEEE International Conference on Advanced Learning Technologies*, *ICALT 2023* (pp. 233-235). IEEE. https://doi.org/10.1109/ICALT58122.2023.00074

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*, pp. 1-11, 100041. https://doi.org/10.1016/j.caeai.2021.100041

Norman, D. (2013). *The design of everyday things: Revised & expanded edition*. Published by Basic Books.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

OECD. (2024). PISA 2029 media and artificial intelligence literacy (MAIL) assessment framework. OECD Publishing. https://www.oecd.org/en/about/projects/pisa-rdi-programme.html

Özmantar, Z. (2018). Örnekleme yöntemleri ve örnekleme süreci. Beycioğlu, K., Özer, N. ve Kondakçı, Y. (Ed.), *Eğitim yönetiminde araştırma* (ss.88-109). Ankara: Pegem Akademi.

Pallant, J. (2001). The SPSS survival manual: A step-by-step guide to data analysis using SPSS, *St. Leonards Publications*. https://doi.org/10.4324/9781003117452

Pingmuang, P., Koraneekij, P., & Khlaisang, J. (2026). Development and validation of an AI literacy scale for pre-service teachers in Thailand. *Electronic Journal of e-Learning*, *24*(1), 01-18. https://doi.org/10.34190/ejel.24.1.3990

Pinski, M., & Benlian, A. (2023). AI literacy - towards measuring human competency in artificial intelligence. Proceedings of the 56th Hawaii International Conference on System Sciences, pp. 165-174. https://doi.org/10.24251/HICSS.2023.021

Poe, M. T. (2015). İletişim tarihi: Konuşmanın evriminden internete medya ve toplum. (Çev: U. Kara). İstanbul: Islık Yayınları.

Polit, D., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, *29*, 489-497.

Rosemann, A., & Zhang, X. (2022). Exploring the social, ethical, legal, and responsibility dimensions of artificial intelligence for health-a new

https://doi.org/10.32329/uad.1899455

column in Intelligent Medicine. *Intelligent Medicine*, *2*(2), 103-109.

Rosling, A. & Littlemore, K. (2011). Improving student mental models in a new university ınformation setting (pp.89-101). *In Digitisation Perspectives*, Cham, Switzerland: Springer.

Serafini, F. (2013). *Reading the visual: An introduction to teaching multimodal literacy*. NY: Teachers College Press.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research*, *8*(2), 23–74.

Schreiber, B., Stage, F.K., King, J., Nora, A., & Barlow, E.A. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *Journal of Educational Research*, *99*(6), 323-337.

Su, G. (2018). Unemployment in the AI age. *AI Matters 3*(4), 35-43.

Stembert, N., and Harbers, M. (2019). Accounting for the human when designing with AI: challenges identified. *CHI'19-Extended Abstracts*, Glasgow, Scotland Uk. May 04-09, 2019.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson.

Tarafdar, M., C. M. Beath, and J. W. Ross. (2019). Using AI to enhance business operations. *MIT Sloan Management Review 60*(4), 37–44.

Tavakol, M., & Dennick, R. (2011). Making sense of cronbach's alpha. *International Journal of Medical Education*, *2*, 53-55.

Tavşancıl, E. (2006). *Tutumların ölçülmesi ve SPSS ile veri analizi* (2. baskı). Ankara: Nobel Yayın Dağıtım.

Telli, S. G., & Aydın, S. (2025). Üniversitelerde yapay zekânın kullanımı: Dönüşümler, getiriler ve geleceğe hazırlık. *Journal of University Research*, *8*(1), 139-148. https://doi.org/10.32329/uad.1609305

Tiernan, P. D., Costello, E., Donlon, E., Parysz, M., & Scriney, M. (2023). Information and media literacy in the age of AI: Options for the future. *Education Sciences 13*(9), 906, 1-11. https://doi.org/10.3390/educsci13090906

Tobin, C. D. (1983). Developing computer literacy. *Arithmetic Teacher 30*(6), 22-60.

Waltz, C.F., Strickland, O.L., & Lenz, E.R. (2005). *Measurement in nursing and health research* (3rd ed.), New York: Springer Publishing Co.

Wetson, R., & Gore Jr, P.A. (2006). A brief guide to structural equation model. *The Counseling Psychologist*, *34*(5), 719-751.

Yaşlıca, E., & Altay, I. (2026). İletişim fakültesi öğrencilerinin yapay zekâ okuryazarlıklarının farklı değişkenlere göre incelenmesi. *Journal of University Research*, *9*(1), 1-12. https://izlik.org/JA27DM75XN

Yavuz Aksakal, N., ve Ülgen, B. (2021). Artificial intelligence and jobs of the future. *TRT Akademi*, *6*(13), 834-852.

Yıldırım Seheryeli, M., & Gören, S. (2025). R ile Psikometri. Güler, N., Atar, B. ve Atalay Kabasakal, K. (Ed.), Ölçme aracı seçme, geliştirme ve uyarlama içinde (1. Baskı). Ankara: Pegem Akademi.

Yılmaz, R., Yurdugül, H., Karaoğlan Yılmaz, F. G., Sahin, M., Sulak, S., Aydin, F., Tepgec, M., Müftüoğlu, C. T., Ömer, O. R. A. L. (2022). Smart MOOC integrated with intelligent tutoring: A system architecture and framework model proposal. Computers and Education: Artificial Intelligence, *3*(2022), 100092.

Yurdugül, H. (2005). Ölçek geliştirme çalışmalarında kapsam geçerliği için kapsam geçerlik indekslerinin kullanılması (ss.1-6). *XIV. Ulusal Eğitim Bilimleri Kongresi*. Pamukkale Üniversitesi Eğitim Fakültesi.