RESEARCH ON EDUCATION AND PSYCHOLOGY (REP)

Received: October 31, 2021
Accepted: December 4, 2021
http:// dergipark.org.tr/rep

e-ISSN: 2602-3733

December 2021 ◆ 5(2) ◆ 267-277

Research Article

Doi: 10.54535/rep.1017070

Validity and Reliability Study of a Turkish Form of the Machine Learning Attitude Scale

Sinan Hopcan [1]
İstanbul University - Cerrahpaşa

Elif Polat[2]
İstanbul University - Cerrahpaşa

Gamze Türkmen[3]
Manisa Celal Bayar University

Abstract

This study aims to adapt the learners' machine learning technologies attitude scale to Turkish. Participants of the study are 309 university students. Confirmatory factor analysis (CFA) was used on data obtained from Turkish students for construct validity of the scale. Following this, 23 items were excluded. A confirmatory factor analysis was performed again, completing adaptation of the scale to Turkish. Three reasons for excluding the items and factor following the confirmatory factor analysis emerged: item structure, domain self-efficacy, and the cultural adaptation process. This study has enabled the scale of attitudes toward artificial intelligence to be adapted to Turkish specifically for machine learning techniques and technologies. The scale can be used as a resource for further studies.

Key Words

Artificial intelligence • Machine learning • Attitude • Scale adaptation • Teacher candidates

[1] Dr., İstanbul University-Cerrahpaşa, Hasan Ali Yücel Faculty of Education, İstanbul, Turkey. E-mail: sinan.hopcan@iuc.edu.tr **ORCID:** 0000-0002-0255-4600
[2] **Correspondance to:** Dr., İstanbul University-Cerrahpaşa, Hasan Ali Yücel Faculty of Education, İstanbul, Turkey. E-mail: elif.polat@iuc.edu.tr **ORCID:** 0000-0002-6086-9002
[3] Dr., Manisa Celal Bayar University, Faculty of Education, Manisa, Turkey. E-mail: gamze.trkmn@gmail.com **ORCID:** 0000-0002-4695-9159
Citation: Hopcan, S., Polat, E., & Türkmen, G. (2021). Validity and reliability study of a Turkish Form of the Machine Learning Attitude Scale. *Research on Education and Psychology (REP), 5*(2), 267-277.

Attitudes toward artificial intelligence in information technologies has been of interest since the first generation of computers. It has attracted the attention of researchers since human interaction with computers became more widespread (Knezek & Christensen, 2008). In the early 1980s, many researchers agreed that the successful use of computers in the classroom was dependent on positive attitudes toward them. Studies conducted in the past 20 years have shown that teachers' positive attitudes toward information technologies can be passed on to their students (Christensen, 2002). Along with developing techniques and technologies, educational materials including software and hardware involving artificial intelligence (AI) technologies have begun to be used in learning environments. In recent years, machine learning has been used in discourse analyses (Yücel, 2021), interpreting sign language (Öztürk et al., 2021), the prediction of social relationships and delays in cognitive development (Metlek & Kayaalp, 2020), and in research on big data. For the long term development of educational projects in collaboration with field experts, it is important that teachers are aware of such research. In addition, by determining attitudes toward machine learning and technologies, this study can be seen as the first step in determining the kind of relationships and affective arrangements that may be needed in follow-up studies.

Interdisciplinary research into learning facilitates our understanding of learning processes and the nature of the teaching practices that can support these processes. With an understanding of these, machine learning technologies can be designed for use in education. As Luckin and Çukurova (2019) point out, the development of artificial intelligence technologies for use in education requires cooperation between researchers working in the learning sciences. Such collaboration is also needed for AI developers to better understand the training and learning activities involved. The educational challenges of bringing human intelligence to the fore and educating everyone in AI can make the more effective use of AI in educational activities one of the core goals. Although such processes can be investigated using AI on big data, teachers need to develop new skills so that educational processes can progress effectively in the classroom environment and they can regularly integrate AI into their teaching processes (Hampel & Stickler, 2005).

Big data in AI can provide information such as metrics of technical systems or social media filters that can facilitate education management. Subsequent developments can provide applications for students, teachers, researchers, administrators, policy makers, and institutes. Baker et al. (2019) divide “artificial intelligence tools used in education” into three groups oriented toward learning, teaching, and the system. While learning-oriented AI tools are those that students use to learn a subject area, teaching-oriented tools are used by teachers to reduce their workload and ensure effectiveness in certain tasks. System-oriented tools, on the other hand, enable administrators and managers to track information. While these categories of AI tools are used in education and training, understanding the attitudes of current teacher candidates toward them and toward machine learning technologies is important in terms of predicting their long-term use.

In foreign language education, studies have shown how teachers are integrated into educational processes and suggest methods for preparing teachers in the application of AI supported technologies. These studies predict that the application of machine learning technologies in education will make classroom management more effective, as well as providing personalized and flexible teaching. In addition, it is predicted that personalized AI tools will enable teachers to better understand their own development and that of their students, and to organize learning activities more easily (Pokrivcakova, 2019). In order to integrate intelligent computer-assisted language learning tools into ongoing teaching processes, teachers need to develop new skills. Such tools will reduce workload and make repetitive tasks easier, allowing teachers more time to support their students' development (Dodigovic, 2009; Hampel & Stickler, 2005).

As teachers need to be trained in the use of machine learning technologies due to future use of such technologies by embedding into their teaching processes, it is important to understand their attitude toward them. Although attitudes toward computer use has received much attention from researchers, few studies have evaluated attitude dimensions based on machine learning technologies. In addition, where they have been carried out, evaluations have not been done in Turkish, and there is as yet no comprehensive research into attitudes toward machine learning technologies in the context of Turkey. For this reason, traditional methods for measuring attitudes toward computers, the internet, and information technologies are inadequate for applying to the field of machine learning technologies. Although we have insufficient knowledge of the differences between current attitudes, future studies may reveal that individuals' attitudes differ toward machine learning technologies and computer, internet, and information technologies. This could be due to either the different perceptions of developments in machine learning technologies or the differences between individuals' learning styles. In order to understand the core reasons, a reference research on attitudes toward machine learning technologies is needed.

The Learners' Attitude in Artificial Intelligence Scale was developed by Lee (2019) to measure the attitudes of university students studying in different fields toward machine learning technologies. This study aims to adapt that scale to Turkish, specifically for machine learning technologies. As AI technologies have developed, machine learning technologies have become important, and understanding attitudes toward these techniques and technologies within the scope of teacher training practices has become a priority for developing studies and workshops. It is important that teachers develop positive attitudes toward these techniques and technologies in order to prepare the young generation for the machine learning technologies that can develop their ways of thinking, enabling them to evaluate daily events, and improving their job opportunities. In order to prepare this awareness and workshop training, a scale development in Turkish is needed to determine students' attitudes and to clarify what kind of parameters can be added to the training content during the education process. Adapting the scale developed by Lee (2019) to Turkish will help guide content preparation for workshop-based training for university students in Turkey, and will also help to reveal any correlation between attitudes toward machine learning technologies and a variety of variables. In this sense, the scale is a reference resource and will be referred to as such throughout this paper.

Method

Research model

This research is a scale adaptation study. Permissions required for the scale's adaptation were obtained by e-mail. Additionally, required permissions were obtained from the Social Sciences Ethics Committee prior to administering the scale. The researchers first translated the scale into Turkish. After translation, a form was developed that included the original items of the scale and the translation items. The form provided space for experts to comment on each item and make recommendations. Two English language educators, five educational technology specialists, and one measurement and evaluation specialist were consulted for their expert opinions, and revisions were carried out in light of these opinions. Three university students were asked to read the Turkish scale aloud to ensure the language was clear. Following this stage, the original scale and the Turkish version were distributed to 17 students fluent in English and Turkish. The students ranged in age from 18 to 27 years and were either studying English Language Education, studying at an English-medium university, or had mastered English at a preparatory course (and who had a good command of the language). The Pearson Product-Moment Correlation Coefficient was calculated to be 0.90 ($p < 0.001$) between the scores obtained from the original form and its translation. The Turkish and English forms of the scale were therefore determined to be equivalent.

Participants

The participants were 309 university students who volunteered to take part in the study. The participants' demographic information is presented in Table 1. Of the participants in the study, 59.9% were female and 40.1% were male. Almost all participants (95.5%) were students at public universities; half of them were freshmen (50.8%) and approximately a third were seniors (23.6%).

Table 1

Demographic Information

		n (309)	*f*
Gender	Female	185	% 59.9
	Male	124	% 40.1
University Type	Public	295	% 95.5
	Private	14	% 4.5
Grade	1	157	% 50.8
	2	36	% 11.7
	3	27	% 8.7
	4	73	% 23.6
	4+	16	% 5.2

Data Collection Instrument

The original scale consists of seven factors and 62 items. The factors are "interest in technology" (10 items), "gender role of technology" (10 items), "importance and impact of technology" (10 items), "ease of access to technology" (9 items), "technology and courses" (10 items), "technology-related career paths" (9 items), and "technology and creative activities" (4 items). The attitude measurement instrument used in this study was found to be valid and reliable. The Cronbach Alpha (CA) coefficient was found to be 0.89, which was deemed valid based on the face validity measurement. The scale was tested on five elementary school students and found to be appropriate for their age group. The parts of the scale related to artificial intelligence were modified to make them suitable for machine learning technology. The items were rated on a 5-point Likert scale, ranging from "Strongly Agree" (+2 points) to "Strongly Disagree" (-2 points) (Appendix: Turkish Machine Learning Attitude Scale).

Data Analysis

A confirmatory factor analysis (CFA) was conducted on the data collected from 309 Turkish university, obtained during the process of adapting the scale to Turkish. The goal of CFA is to see how well a predetermined structure — a model — fits the obtained data. In this context, the CFA was conducted to test the construct validity of the Machine Learning Attitude Scale. The AMOS 21 software program was used to conduct the CFA. A CA coefficient was calculated for reliability.

Findings

Confirmatory Factor Analysis Results

The construct validity of the scale was evaluated using confirmatory factor analysis (CFA) after the prescribed factor structure of the original scale was found to be suitable as a result of consulting expert opinion. Fit statistics for the model with seven factors specified on the original scale were investigated. Items 2, 4, 5, 9, 11, 14, 16, 18, 24, 26, 28, 34, 41, 45, 46, 47, 49, and 52, which had small factor loads, were removed after the CFA. For reliability, the CA values were examined, and items 51 and 58 were excluded because they reduced the internal consistency of the factor. Because the internal consistency of items 35, 54, and 60 was low (CA = 0.58), they were removed, and thus the factor "Ease of Access to Technology" was removed entirely from the scale. Following these operations, a CFA was again conducted and CA values for each factor were calculated.

For such indices, although it is difficult to establish standards, a value of 0.50 or greater for PNFI (Hu & Bentler, 1999) and PGFI (Meyers et al., 2006), and a value of 0.08 or smaller for SRMR, RMR (Hu & Bentler, 1998) and RMSEA (Browne & Cudeck, 1993) are typically recommended as they correspond to adequate fit.

The p value of the $\chi 2$ (chi-square) statistic was examined by taking into account the fit indices of the model. This value is considered a good fit at 0.00 ($p < 0.05$). As this value is significant in large samples, the $\chi 2$ / df ratio and other indices should be taken into account (Tabachnick & Fidell, 2007). This value is acceptable if it is smaller than five (Wheaton et al., 1977). The $\chi 2$ value was calculated as 1850.512, and the df value was calculated as 681. The $\chi 2$ / df ratio (1850.512 / 681) was calculated as 2.72, which was considered a good fit. Other fit indices are shown in Table 2 and discussed in relation to the literature. All the indices listed in Table 2 were found to fit well. In this way, the model was confirmed to have six factors.

Table 2

Model Fit Measurements

	Model	*Criteria*	*Decision*	*Rationale*
χ2	1850.51	-	-	-
df	681	-	-	-
χ2/df	2.72	<3	Good Fit	Kline (2011)
RMSEA	.07	<.08	Good Fit	Browne and Cudeck (1993)
SRMR	.07	≤ .08	Good Fit	Hu and Bentler (1999)
RMR	.06	≤ .08	Good Fit	Hu and Bentler (1998)
PNFI	.70	≥.50	Good Fit	Hu and Bentler (1999)
PGFI	.65	≥.50	Good Fit	Meyers et al. (2006)

Because the data set had a normal distribution, the Maximum Likelihood Method was used to estimate the parameters, and the covariance matrix method was used to calculate the data matrix. Table 3 shows the factor loads, error variances, and t values. The t values of all items were greater than +1.96. According to the literature, T values should be different from ±1.96 (Kline, 2011). All values were significant ($p < 0.05$). A path diagram is shown in Figure 1. Due to the large number of items, the path diagram only displays the factors. As shown in

Figure 1, the factor with the lowest correlation with the scale results is “Gender Role of Technology”, and that with the highest correlation is “Interest in Technology”.

Table 3

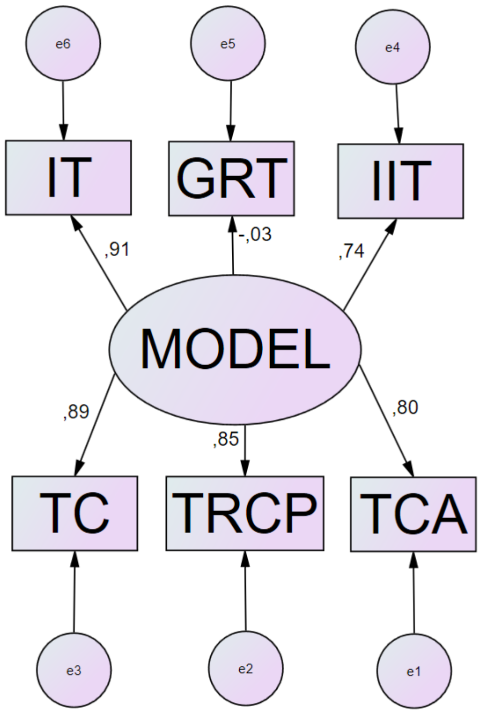
Item Analysis

Factor	*Item Number*	*Factor Loads*	*Error Variances*	*t**
Interest in Technology	25	0.77	0.12	10.63
	13	0.76	0.13	10.53
	50	0.74	0.11	10.37
	38	0.72	0.13	10.13
	44	0.72	0.13	10.15
	57	0.71	0.10	10.10
	7	0.63	0.09	9.28
	31	0.61	0.11	9.02
	1	0.59	-	-
	19	0.57	0.10	8.50
Gender Role of Technology	39	0.83	0.08	13.38
	33	0.80	-	-
	20	0.77	0.07	13.01
Technology and Courses	42	0.76	0.13	8.85
	23	0.74	0.15	8.73
	12	0.70	0.14	8.50
	17	0.69	0.11	8.43
	55	0.56	0.12	7.48
	29	0.56	0.11	7.52
	36	0.54	0.13	7.36
	48	0.53	0.10	10.83
	61	0.50	-	-
Importance and Impact of Technology	15	0.71	0.15	9.68
	10	0.67	0.15	9.33
	40	0.66	0.15	9.22
	59	0.65	0.16	9.13
	27	0.63	0.16	8.93
	3	0.59	-	-
	22	0.55	0.17	7.99
Technology-related Career Paths	30	0.81	0.20	8.55
	43	0.79	0.18	8.46
	56	0.75	0.16	9.36
	37	0.66	0.21	7.82
	6	0.56	0.20	7.15
	62	0.48	-	-
Technology and Creative Activities	21	0.74	0.10	11.07
	8	0.67	0.11	10.16
	32	0.64	0.10	9.83
	53	0.63	-	-

*p<0.001 (for all t values)

Figure 1

Path Diagram of the Factors (Factor Path Diagram)



Note: (it: interest in technology, grt: gender role of technology, iit: importance and impact of technology, tc: technology and courses, trcp: technology-related career paths, tca: technology and creative activities).

The internal consistency coefficient of CA (α) varied from 0.76 to 0.89 (Table 4). A CA value that is greater than 0.6 is acceptable (Cortina, 1993).

Table 4

Reliability

Factor	*Cronbach Alpha*
Interest in Technology	0.89
Gender Role of Technology	0.84
Importance and Impact of Technology	0.82
Technology and Courses	0.86
Technology-related Career Paths	0.83
Technology and Creative Activities	0.76

Conclusion and Discussion

This study aimed to develop a reference scale in Turkish that can determine the attitudes of teacher candidates and newly graduated teachers toward machine learning technologies by adapting the scale created by Lee (2019) to Turkish. Items numbered 2, 4, 5, 9, 11, 14, 16, 18, 24, 26, 28, 34, 41, 45, 46, 47, 49, and 52 were removed because of their low weights according to the results of the confirmatory factor analysis. Also, items numbered 51 and 58 were removed since they affected the internal consistency of the factor, and the “Ease of Access to Technology” factor (items numbered 35, 54, and 50) was removed due to its low internal consistency. Finally, a further confirmatory factor analysis was performed, completing the adaptation of the scale to Turkish.

Three reasons emerged regarding the items and factors removed following the confirmatory factor analysis. These can be classified as item structures, domain self-efficacy, and the cultural adaptation process. The fact that the item structures are directive is thought to be one of the primary reasons for removing the scale items from the original language when they are adapted to Turkish. Statements such as “He is smarter” and “even she can do it”, which are among the items in the original scale, add directiveness to the items. The fact that such items specify the level of intelligence, emphasizes intelligence and the ease of access to technology, which may cause the items to be removed when they are adapted to Turkish. Examination of the removed items shows that the item structures in the original language in the scale are leading. The presence of leading (a bias included in the item sentence) items (questions) in any scale is to be avoided (Colosi, 2006).

The inability to provide domain self-efficacy can be interpreted as another reason for removing the "Ease of Access to Technology" factor from the scale. The distribution of answers given to the items may have been affected because as the current university students' level of self-efficacy was low, they could not make a decision about accessing technology, and also because these students were from different fields. The reason for this recommendation can be investigated with a follow-up study into the relationship between domain self-efficacy and attitude toward machine learning technologies. In addition, different researchers may suggest starting the scale's cultural adaptation process with exploratory factor analysis as one of the solutions. However, the confirmatory factor analysis can be given priority in cases where it is known how many factors exist for the variables (Orcan, 2018). In this study, confirmatory factor analysis was applied twice before and after the items were removed, effectively completing the cultural adaptation process.

As a result, the fact that the item structures are leading and the participants' lack of self-efficacy may be seen as the primary reasons for removing the above-mentioned items and factor from the scale. This scale also opens a road for follow-up researches. For example, studies examining the relationship between attitude toward machine learning technologies, domain self-efficacy, and learning motivation can be considered important.

Recommendations

This study adapted the scale of attitudes toward artificial intelligence into Turkish specifically for machine learning technologies. The results of the study have recommended potential effective items for detecting the attitude towards machine learning technologies among teacher candidates and potential sub-titles for determined machine learning course content. First, by determining attitudes toward machine learning technologies, this study can be seen as the first stage in determining what kind of relationships should be investigated in future research and what kind of emotional regulation may be needed in follow-up studies. Thus, reference studies can be encouraged for the development of positive attitudes toward machine learning technologies. Second, this study can act as a resource scale that facilitates the preparation of future teacher training content. In other words, considering the initial level of teacher candidates before a machine learning training course may encourage the adaptation of the course content.

Limitations

Limitations of this study are three-folded as the number of participant, data collection procedure and further scale testing, and having possible leading items. First of all, this study reached 309 students from 6 universities for the data collection. Second, students filled the scales via an online platform so that the time duration for completing the data may be varied within teacher candidates. Also, although the scale has been shown to be valid, it would be better to have different sample groups for further testing of the scale. Finally, the elimination of one factor ("Ease of Technology Use") and items remind the researchers about the leading items on the originally developed scale so that the final scale should be re-tested for further confirmation with the same group.

Ethic

All procedures in this study involving human participants were carried out in accordance with the ethical standards of Manisa Celal Bayar University Research Ethics Committee with date 12.04.2021 and number E.57768.

Author Contributions

This article was written with the joint contributions of three authors.

Conflict of Interest

The authors declare that they have no conflict of interest.

Funding

No scholarships or payments have been received from any institution for this article.

References

Baker, T., Smith, L., & Anissa, N. (2019). *Educ-AI-tion rebooted? Exploring the future of artificial intelligence in schools and colleges.* Date of access: 31 October 2021, https://media.nesta.org.uk/documents/Future_of_AI_and_education_v5_WEB.pdf.

Browne, M.W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In: Bollen, K.A., & Long, J.S. (Eds.), *Testing Structural Equation Models* (pp. 136-162), Sage, Beverly Hills, CA.

Christensen, R. (2002). Effects of technology integration education on the attitudes of teachers and students. *Journal of Research on Technology in Education*, *34*(4), 411-433.

Colosi L. (2006). *Designing an effective questionnaire.* Date of access: 31 October 2021, https://www.gateshead.gov.uk/DocumentLibrary/council/consultation/Questionnaire-design-guidance-web.pdf.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Spychology*, *78*(1), 98-104.

Dodigovic, M. (2009). *Speech processing technology in second language testing*. In: Proceedings of the Conference on Language & Technology. https://www.cle.org.pk/clt09/download/Papers/Paper17.pdf

Hampel, R., & Stickler, U. (2005). New skills for new classrooms: Training tutors to teach languages online. *Computer Assisted Language Learning*, *18*(4), 311-326.

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*, 424-453.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55.

Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling* (3rd ed.). New York: Guilford.

Knezek, G., & Christensen, R. (2008). The importance of information technology attitudes and competencies in primary and secondary education. *International handbook of information technology in primary and secondary education* (pp. 321-331). Boston, MA: Springer.

Lee, Y. (2019). An analysis of the influence of block-type programming language-based artificial intelligence education on the learner's attitude in artificial intelligence. *Journal of The Korean Association of Information Education*, *23*(2), 189-196.

Luckin, R., & Cukurova, M. (2019). Designing educational technologies in the age of AI: A learning sciences-driven approach. *British Journal of Educational Technology*, *50*(6), 2824-2838.

Metlek, S., & Kayaalp, K. (2020). Otistik spektrum bozukluğunun makine öğrenme algoritmaları ile tespiti [Detection of autistic spectrum disorder with machine learning algorithms]. *Journal of Intelligent Systems: Theory and Applications*, *3*(2), 60-68.

Meyers, L. S., Glenn, G., & Guarino, A. J. (2006). *Applied multivariate research: design and interpretation*. CA: Sage, Thousand Oaks.

Orcan, F. (2018). Exploratory and confirmatory factor analysis: Which one to use first? *Journal of Measurement and Evaluation in Education and Psychology*, *9*(4), 414-421.

Öztürk, A, Karatekin, M, Saylar, İ., & Bardakcı, N. (2021). Recognition of sign language letters using image processing and deep learning methods. *Journal of Intelligent Systems: Theory and Applications*, *4*(1), 17-23.

Pokrivcakova, S. (2019). Preparing teachers for the application of AI-powered technologies in foreign language education. *Journal of Language and Cultural Education*, *7*(3), 135–153.

Tabachnick, B. G., & Fidell, L. S. (2007). *Experimental designs using ANOVA.* CA: Thomson/Brooks/Cole, Belmont.

Wheaton, B., Muthen, B., Alwin, D. F., & Summers, G. F. (1977). Assessing reliability and stability in panel models. *Sociological Methodology*, *8*, 84-136.

Yücel, A. (2021). Tüketici yorumları üzerine bir metin madenciliği ve veri boyutu indirgeme yaklaşımı [A text mining and data size reduction approach on consumer reviews]. *Journal of Intelligent Systems: Theory and Applications*, *4*(1), 8-16.