www.ijte.net

A Scale Development Study for The Evaluation of Digital Content

Onur Duran
National Defense University, Türkiye

Ayse Oguz Unver
Mugla Sitki Kocman University, Türkiye

To cite this article:

Duran, O. & Oguz Unver, A. (2025). A scale development study for the evaluation of digital content. *International Journal of Technology in Education (IJTE), 8*(2), 482-501. https://doi.org/10.46328/ijte.1074

The International Journal of Technology in Education (IJTE) is a peer-reviewed scholarly online journal. This article may be used for research, teaching, and private study purposes. Authors alone are responsible for the contents of their articles. The journal owns the copyright of the articles. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of the research material. All authors are requested to disclose any actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations regarding the submitted work.

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

International Journal of Technology in Education
2025, Vol. 8, No. 2, 482-501 https://doi.org/10.46328/ijte.1074

A Scale Development Study for The Evaluation of Digital Content

Onur Duran, Ayse Oguz Unver

Article Info

Article History

Received:
13 October 2024
Accepted:
27 March 2025

Keywords

Web 2.0 tools
Rubric
Assessment
Digital content
Validity and reliability

Abstract

This study aimed to develop a rubric for evaluating digital contents specifically Web 2.0 tools in education to improve the quality of digital content, as well as to assess the rubric's validity and reliability. Data were collected from pre-service science teachers in their third year. In the development process, expert opinions were sought, and the rubric's items, criteria, and proficiency levels were created based on relevant literature. The draft rubric was used to evaluate digital contents prepared by these pre-service teachers, followed by analyzes to ensure validity and reliability. Exploratory Factor Analysis (EFA), Confirmatory Factor Analysis (CFA), Cronbach's Alpha, and Kappa analyzes were performed on the collected data. EFA revealed that the 25-item rubric explained 52.70% of the total variance, with factor loadings between .498 and .862 across three factors. CFA results showed a $\chi 2/sd$ value of 1.437 and an RMSEA value of .071. The rubric's internal consistency was measured with a Cronbach's Alpha coefficient of .785. To further assess reliability, kappa analysis indicated a moderate inter-rater agreement with a kappa value of .64. Based on these results, the EWES rubric is considered both a valid and reliable evaluation tool.

Introduction

Rapid technological developments have driven significant changes and adaptations in education. Educational technology is defined as processes that enhance teaching quality and improve performance (Ahmadigol, 2016). It offers advantages such as facilitating learning, enhancing performance, and promoting more effective teaching, thus improving educational environments (UNESCO, 2008; Richey et al., 2008; Januszewski & Molenda, 2013). With the rise of the Internet, digital content has become integral to education, offering benefits like low cost, flexibility, accessibility, and personalized learning (OECD, 2023). The International Society for Technology in Education (ISTE) established standards to ensure equitable and quality learning opportunities through technology (ISTE, 2017). UNESCO's 2023 report stresses that teachers should integrate technology to provide more student-centered teaching (UNESCO, 2023). As digital content's role grows, effectively evaluating and integrating Web 2.0 tools in education has become crucial, making them indispensable in modern teaching.

Web 2.0 Tools in Educational Environments

There are various types of digital content in education, but Web 2.0 tools are the most prominent due to their ease

of use and accessibility. Web 2.0, known as the second era of the Web, enables individuals to produce and participate in content creation (O'Reilly, 2007). These tools play a crucial role in integrating technology into education and fostering collaboration (Murugesan, 2007). They also help minimize individual differences and strengthen teacher-student communication. The COVID-19 pandemic further highlighted the importance of Web 2.0 tools during the shift to mandatory distance learning, addressing challenges like accessibility and communication (Mishra et al., 2020; Nkansah et al., 2020). Web 2.0 tools facilitate information exchange, interactive content, and ease of use in education (O'Reilly, 2005). For students to be active participants in learning, these tools enable the creation, organization, and presentation of interactive content. Teachers, as key integrators of these tools, must develop digital skills to thrive in their professional roles.

In the literature, there are different classification options according to the types of Web 2.0 tools that can be used in educational environments (e.g. Hussein et al. 2009; Majhi & Mahara 2010; Berg, 2011; Baxter & Connolly, 2014). However, in this study, Web 2.0 tools were classified into 11 categories. The classification of Web 2.0 tools is presented in Figure 1.

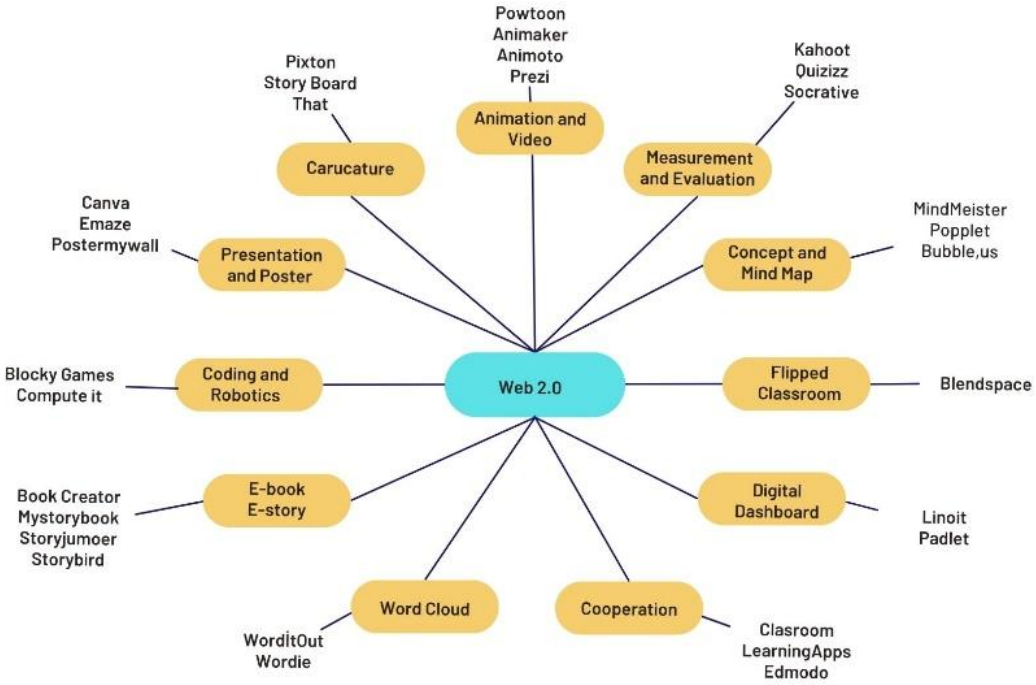

Figure 1. Classification of Web 2.0 Tools

When Figure 1 is analyzed, 30 Web 2.0 tools are presented in 11 different categories. Web 2.0 tools continue to be widely used in education. When the studies on pre-service teachers and Web 2.0 tools in the literature are examined, it is seen that the majority of them focus on participants' views on the use of Web 2.0 tools (e.g. Ahmed et al, Ahmed et al. 2016; Davison et al. 2013), self-efficacy (Pan & Franklin, 2011; Alhassan, 2017), perceptions (Alajmi, 2011; Leow & Neo, 2015; Ozcinar et al. 2020; Venkatesh et al. 2014; Zelick, 2013), attitudes (Leh et al. 2021) and its effect on participants' academic achievement (Hursen, 2021; Newland & Byles, 2014).

Evaluation of Digital Content

A culture of creating quality content can be achieved when the digital content used in educational activities has certain characteristics. In addition, the production of these contents according to certain criteria will ensure that the contents used are of higher quality and more effective. The evaluation of digital content in education focuses on increasing the quality and effectiveness of digital content. In the literature, different criteria and measurement tools have been developed for the evaluation of digital content (e.g. Baya'a et al., 2009; Cochrane & Bateman, 2010; Stefani & Kameas, 2013; Al-Alwani, 2014; Yuan & Recker, 2015; Ferrer & Ramírez, 2016; Romero et al. 2017). Looking at the rubrics developed in the literature, a number of rubrics have been developed, each with its own strengths and weaknesses. Yuan and Recker (2015) underlined the diversity of rubrics developed and the problems in standardization. While Baya'a et al. (2009) focused on variables such as content, usability, and educational value especially for web-based educational environments, Ferrer and Ramírez (2016) presented a rubric emphasizing structure, evidence, language, and technology use to evaluate e-portfolios created using Web 2.0 tools. Looking at the rubrics developed in the literature, it is clear that they all focus on different points and clearly emphasize the importance of evaluation to improve the quality of the content.

Therefore, there is a need to develop an evaluation tool that aligns with established criteria to enhance user experiences and improve the quality of educational content. The main purpose of this research is to develop a comprehensive rubric for the evaluation of digital educational content created using Web 2.0 tools and to ensure that these materials meet high quality and content standards. In the context of science education, the use of Web 2.0 tools offers several opportunities to create interactive and engaging learning experiences. These tools enable educators to design content that not only covers basic scientific concepts but also promotes critical thinking, problem solving and collaborative learning. However, the effectiveness of such digital content is highly dependent on its conformity to certain quality standards. To meet this need, the evaluation tool should consider various criteria such as design, content accuracy, usability, accessibility and technological integration.

In summary, this research aims to develop a detailed and multidimensional rubric that can be applied to the evaluation of digital content created with Web 2.0 tools in the field of education. Thus, it aims to ensure that these digital contents are not only of high quality and effective in achieving educational outcomes, but also accessible, usable and ethically sound.

Method

In this study, which is based on the development of a rubric for the evaluation of Web 2.0 tools prepared in the educational field and the validity and reliability of the developed rubric, descriptive research method was used. Fraenkel and Wallen (2011) defined descriptive research as the description of a situation. In this context, the reason why descriptive research was preferred in the study was to reveal the situation obtained from the evaluation of Web 2.0 tools. Following the development of the relevant rubric, Web 2.0 tools created by pre-service teachers will be evaluated and the validity and reliability analysis of the rubric was carried out.

Working Group

The study group of the research consists of pre-service teachers studying in the 3rd year of science teaching (N=28). Convenient sampling method was used to determine the study group in order to ensure the selection of easily accessible people (Cohen et al., 2007; Fraenkel & Wallen, 2008; Creswell, 2012).

The participant group of the study consisted of N=21 women and N=7 men. In addition, it is considered that it would be appropriate to carry out the study with this group, considering that pre-service teachers graduate within one year, since the national curriculum declares that teachers should actively use digital content.

Research Application Process

The implementation process of this research was carried out with third year science teaching students studying at the Faculty of Education of a Western Anatolian University in Turkey in the 2022-2023 academic year. The implementation was carried out in the fall semester. It continued for a total of 12 weeks starting from September to December, with a total of 48 lesson hours, 4 lesson hours per week. Two branches were selected as the participant group. The pre-service teachers were given information about how to use the Web 2.0 tools selected during the implementation process for a total of five weeks, two of them each week, and how to produce content. At the end of five weeks, topics were selected by the researcher. Among the Web 2.0 tools used in the research, Bubbl.us, Canva, Google Classroom, Kahoot, Padlet, Popplet, StoryboardThat and Storyjumper are presented under the headings respectively.

Which subject and which Web 2.0 tool the prospective teachers would choose was determined by lottery. Two different draw boxes were created, one with the topics and the other with the names of Web 2.0 tools. The prospective teachers drew one paper from the box with the topic and two papers from the box with the Web 2.0 tool. Thus, the pre-service teachers chose a different topic and two different Web 2.0 tools. In line with the choices, pre-service teachers were guided in creating content using Web 2.0 tools during the remaining weeks and meetings were held at certain intervals. At the end of the design process, pre-service teachers introduced the content they prepared in five minutes each. Web 2.0 tools were evaluated during the presentation.

Data Collection Tool

In the research, a rubric was developed that enables the evaluation of Web 2.0 tools prepared in the educational field in different dimensions. Basically, this research focuses on Web 2.0 tools. However, the relevant scale can also be integrated into the evaluation of digital content. Following the development of the rubric, the validity and reliability study of the scale was carried out by evaluating the contents prepared by the relevant working group. The development stages of the relevant rubric are presented in detail below.

Development of the Rubric

Rubrics may appear in the literature with different names such as rubric, scoring key, scoring instruction (Goodrich, 1997). However, in this study, it was used as rubric. There are different definitions of rubrics in the literature (e.g. Brookhart, 1999; Goodrich, 2001). In the most general definition, rubrics are evaluation lists that allow the evaluation of the created products according to certain criteria (Brookhart, 1999). The rubric was developed by Haladyna (1997) based on the stages to be followed in rubric development.

These stages are;
1. The purpose for which the rubric to be developed will be used is determined.
2. The type of rubric to be developed is determined.
3. The evaluation criteria are determined and the issues to be followed in the evaluation are explained.
4. Criteria about the situation to be measured are determined and the meaning of these criteria is stated.
5. The adequacy levels of the rubric are determined.
6. A draft rubric is created.
7. A sample assessment is made with the creation of a draft rubric.
8. The validity and reliability of the scoring using the rubric is determined.
is in the form of. The process of developing the rubric is presented in detail below.

The rubric was developed to evaluate the contents prepared with the help of educational Web 2.0 tools prepared in science education. The type of the related rubric was determined as analytical rubric in order to focus on many sub-dimensions and perform evaluation.

The criteria to be followed in the evaluation were determined by reviewing the relevant literature. A series of rubrics were developed to evaluate the content created with educational Web 2.0 tools. These rubrics collectively emphasize the importance of considering content, usability, educational value and technological aspects in the evaluation of educational content developed with Web 2.0 tools. In line with the criteria determined in the related literature, three criteria were determined for the rubric in this study: design, content and ease of use.

A total of 25 items, eight items under the design criterion, ten items under the content criterion, and seven items under the ease of use criterion, were formed. The items under these criteria are as follows.

A) Design

"Pictures and graphics are functional, not ornamental." In the criterion, it is aimed that the pictures and graphics support the content and are not ornamental. People learn 83% of what they learn by seeing, 11% by hearing and the rest by other sense organs (Chowdhury, 1995). In other words, visuality is very effective in conveying your message. However, visual elements should be used when necessary and for a purpose. Visuals should be related to and support the text in the content (Marsh, 2002).

As can be seen in Figure 2, in the presentation on the left side, the content and the picture overlap, that is, the picture has the feature of supporting the written text. However, in the presentation on the right side, the pictures

do not support the written text, they are only decorative, independent of the subject.

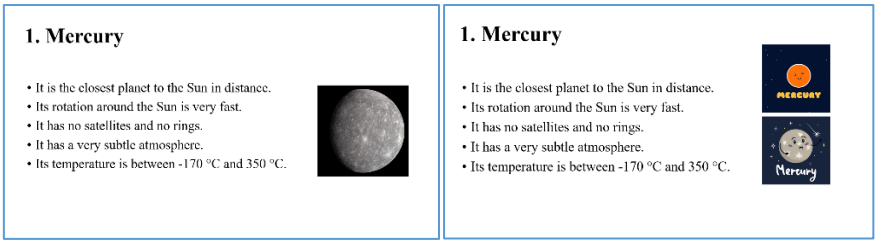

Figure 2. Correct Use (left) and Incorrect Use (right) of Pictures etc.

"The background Color is compatible with the Color of the foreground elements." In the criterion, it is aimed to prevent Color confusion, not to tire the eyes and to facilitate readability. Contrast (contrast) should be created between foreground and background Colors.

Figure 3. Color Circle

Color Wheel: It shows the order of Colors among themselves and their relationship with each other. Colors that are directly opposite each other on the Color wheel are called "complementary Colors". The Colors opposite each other create contrast (Vader, 2001). For example; Red and Green. The group of Colors adjacent to each other are called "harmonious Colors". For example; blue and purple Colors are similar. A good and bad example showing the harmony between background and foreground elements is presented in Figure 4.

Figure 4. Good (right) and Bad (left) Example Showing the Harmony between Background and Foreground Elements

When Figure 4 is analyzed, light blue and dark blue elements are used on the blue background on the left. Since Colors compatible with each other are used here, it is difficult to see the elements in the foreground. On the right side, red and yellow, the opposite Colors of blue, are used. Since contrasting Colors are used here, harmony (Color harmony) is achieved.

"An appropriate Color combination is used between the background Color and the text." With the criterion, it is aimed that the background Colors and text Colors are compatible and readable. Care should be taken to use contrasting Colors due to the shape-ground relationship. Here, we will again use the Color wheel to determine the Color selection. Writing dark text on light Colors and light text on dark Colors is suitable in terms of readability (Reichenstein, 2012).

Figure 5. Use of Non-contrasting (left) and Contrasting (right) Colors

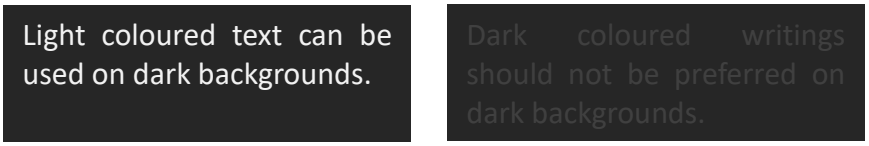

Figure 6. Use of Light Color (left) and Dark Color (right) on Dark Backgrounds

Figure 7. Use of Dark Colors (left) and Light Colors (right) on Light Backgrounds

"The Colors used add attractiveness to the product visuality." In the criterion, it is aimed that the Colors appeal to the eye properly and the Colors are used appropriately. Colors should visually appeal to the eye, there should be no Color confusion.

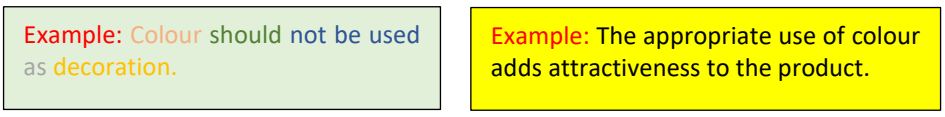

Figure 8. Random Color Use (left) and Attractive Color Use (right)

"The font type and size are appropriate for the user level." and "Appropriate font size is used for text, headings and subheadings." It is aimed that the font type and size used in the criteria are appropriate. For web texts, 18-

point or larger fonts positively affect readability and comprehension (Rello et al., 2016). According to the presentation handbook determined by TUBİTAK (2009), headings should be 32 pt. The maximum font size for first level items is 28, for second level items is 24, and for third level items is 20. The bibliography should be in 11-12 font size. Headings should be clear and precise, preferably not exceeding 5 words. The same font should be used throughout the content (TUBİTAK, 2009).

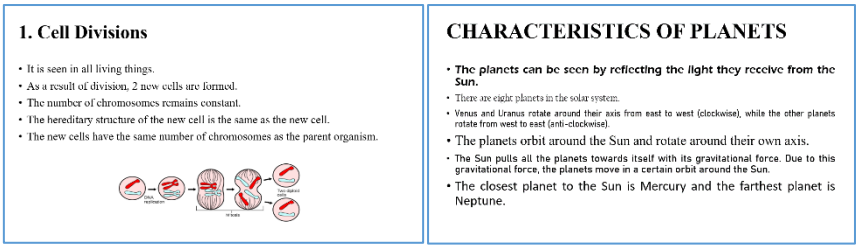

Figure 9. Appropriate Font Size Usage in Text, Headings and Subheadings (left) and Inappropriate Font Size Usage (right)

"The placement of my components is suitable for the aesthetics of the page." In the criterion, the balance and proportions of the elements with each other are emphasised. According to McClurg-Genevese (2005), balance is the symmetrical or asymmetrical position of the elements on the page. Balance refers to the equal distribution of Colors, text and visuals. Again, according to McClurg-Genevese (2005), proportion refers to the relationship of one element to another or where something big will stand. In design, the balance between shapes and writings is called ratio. The ratio provides a stylish appearance.

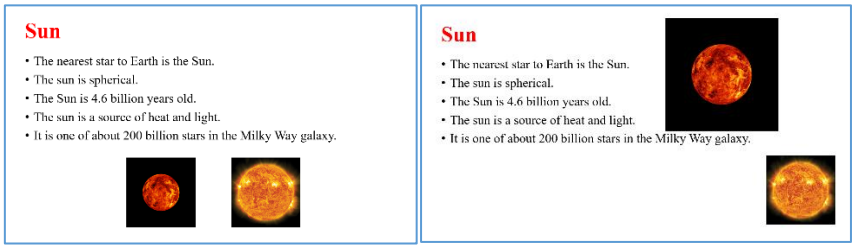

Figure 10. Aesthetically Appropriate (left) and Aesthetically Inappropriate (right) Layout

As can be seen in Figure 11, the content that is aesthetically appropriate (left) complies with the proportion and balance features. This balance is provided as follows; the pictures are symmetrical on both sides of the content. However, this is not the case in the content on the right. There is no balance in the position of the images. As for the ratio, the dimensions of the pictures and the picture-text relationship were considered. In the content on the right, the dimensions of the pictures are the same size, but this is not the case in the content on the right.

"Web 2.0 tool is remarkable and arouses curiosity." In the criterion, it is aimed that the Web 2.0 tool will attract the attention of the users and affect them. In general, the Web 2.0 tool should use semantic information in a balanced way, use appropriate vocabulary, include multisensory aspects whenever possible, and aim to emotionally surprise and affect the audience for maximum impact.

B) Content

"The subject is related to the acquisitions in the curriculum.", "The subject has adequate explanations.", "The subject is suitable for the target user level.", "It can be processed within the specified time.", "The subject is in a logical order." In the criteria, it is aimed that the targeted subject is compatible with the achievements determined in the science curriculum. The content should be compatible with the science curriculum and reflect the subject achievements well. The content should be prepared in accordance with the level of the achievement specified in the Science Curriculum.

"The message intended to be given in the texts is clear and understandable.", "A clear and simple language is used.", "Unnecessary information density is avoided." In the criteria, it is aimed to comply with the principles of clarity, simplicity and clarity. A language appropriate to the principles of expression (clarity, simplicity, clarity) should be used. Expression should be simple and understandable without unnecessary expressions (Reimold, 1981).

"It is up-to-date and scientifically appropriate." In the criterion, the content should be prepared according to the latest Science Curriculum and information consistent with the learning outcomes in the Science Curriculum should be conveyed. The content should be valid, reliable, up-to-date and should not contain information that may lead to misconceptions, inconsistencies and contradictions.

"The language and expression in the texts are appropriate for the level of the target user." It is aimed to use correct language and expression in the criterion. The words used should be written correctly, sentences should be constructed properly and grammar rules should be followed (Blumenthal, 1994).

C) Ease of Use

"Links (Web 2.0 tool link) work smoothly." The criterion aims to avoid the use of faulty or non-working links. Non-functioning links negatively affect the usability of websites. Links should be checked regularly and updated or replaced when necessary (Bustard et al., 2007).

In the criterion "It is possible to access the Web 2.0 tool from any operating system", it was taken into consideration that users may have different operating systems. Accessibility of a web site is defined as the ability of users to access the site without any environmental difficulties (Mankoff et al., 2005). In order to prevent this, a smooth operation test should be performed with browsers on national open source operating systems.

"It is possible to access the Web 2.0 tool from different browsers (Google Chrome, Internet Explorer, etc.)." In the criterion, it is aimed that users do not have problems when logging in to the application from different browsers. In the past years, when websites were viewed from different browsers, differences in appearance emerged. However, today, the consistency between the current versions of browsers has been increased.

"The Web 2.0 tool is personalized and provides feedback." In the criterion, it is aimed to give feedback to users about missing or incorrect information during data entry.

"The desired information can be easily accessed in the Web 2.0 tool." In the criterion, it is aimed that users can access the content prepared in the Web 2.0 tool directly and without any complexity. In this context, putting titles to the whole and sub-sections of the content in an easily understandable way will enable users to make better sense of the content and navigate between the contents more easily (Farkas, 2002; Stauber, 2005).

"There are instructions in the Web 2.0 tool." In the criterion, it is aimed that users can navigate easily in the Web 2.0 tool. Routers, signs, etc. in the Web 2.0 tool make it easier for users to navigate the page by creating a mental map. Thanks to the guidelines, it will allow the content to be presented without any obstacle between the Web 2.0 tool and the user (Rosenfeld & Morville, 2002).

"The user can continue the Web 2.0 tool at any time." In the criterion, it is aimed to comply with the principle of progression according to individual speed from the principles of programmed teaching. This principle emphasizes the individual learning pace, which enables each student to progress independently from others in a teaching-learning situation, enrichment for fast learners and compensation for slower learners (Mularsih, 2007). The learning speed of each individual is different, so it must be respected.

The adequacy level of the rubric was determined as (1) Poor, (2) Needs Improvement, (3) Fair and (4) Good due to Haladyna's (1997) suggestion of four grading levels. Explanations about these proficiency levels are as follows:

(1) Weak: The assessed content weakly reflects the specified characteristics. There is insufficient evidence to support the presence of the qualities in the items in the content.
(2) Should be improved: The assessed content reflects the stated characteristics to a lesser extent. There is evidence to support the presence of the qualities in the items in the content, but it should be improved.
(3) Medium: The assessed content adequately reflects the stated characteristics. Evidence supporting the presence of the qualities in the items in the content is at an acceptable level.
(4) Good: The assessed content reflects the stated characteristics at a good level. It can be stated that the evidence supporting the presence of the qualities in the items in the content is consistent.

In this context, a draft of the rubric was prepared. The rubric developed for the evaluation of educational Web 2.0 tools prepared in science education consists of 25 items. The rubric has a quadruple proficiency level and the lowest score that can be obtained as a result of the evaluation is 25 and the highest score is 100. Finally, with the scale named Educational Web 2.0 Tools Evaluation Scale in Science Education (EWES), the content prepared by the prospective teachers was evaluated. As a result of the related evaluations, the validity and reliability of the rubric was analyzed.

Results

In order to analyze the validity and reliability of the rubric, which is the eighth stage of the rubric development

stages determined by Haladyna (1997), the contents developed by the pre-service teachers were evaluated and then Exploratory Factor Analysis (EFA) for the validity of the rubric and Confirmatory Factor Analysis (CFA) and Cronbach Alpha tests for the reliability of the rubric were performed with the data obtained through SPSS and AMOS package programs.

Findings Related to Validity

The validity of the measurement tool is related to the fact that the measurement tool serves its purpose and accurately measures the situation to be measured (Fraenkel & Wallen, 2008; Middleton, 2019). In other words, the validity of the scale can be expressed as the ability of the measurement tool to consistently measure the situation it aims to measure.

In this study, expert opinions were first sought for the validity study of EWES. After the expert opinions were received, factor analysis was performed. Firstly, Kaiser Meyer Olkin (KMO) sampling adequacy measurement and Bartlett's sphericity test were performed to determine the suitability of the rubric for EFA. In order to perform factor analysis on the data, the KMO value should be higher than 0.60 and the result of Barlett's test should be $p<0.05$ (Tabachinck & Fidell, 1996). The findings obtained as a result of the tests are presented in Table 1.

Table 1. KMO and Bartlett's Sphericity Test Results of EWES

Kaiser-Meyer-Olkin (KMO)	.85
Significance (p)	.000

When Table 1 is examined, as a result of the analyzes, KMO value was .85 and Barttlett's Sphericity test result was $p<0.05$. Since the KMO value obtained in the present study was .85 and Barlett's test of sphericity was significant, it can be said that the data distribution is suitable for factor analysis (Tabachinck & Fidell, 1996; Field, 2005). In line with these, EFA was applied to the data set. Since EWES consists of three sub-dimensions: design, content and ease of use, EFA was conducted with three factors. As a result of EFA, it was found that the EWES rubric consisting of 25 items explained 52.70% of the total variance and the factor loadings of the rubric ranged between .498 and .862 when analyzed as three factors. According to Tabachnick and Fidell (2011), factor loadings of 0.32 and above are sufficient for the scale items to be acceptable. In this context, since the factor loadings of the relevant rubric ranged between .498 and .862, it can be stated that the items are acceptable.

CFA was conducted to check whether the constructs obtained from the EFA results of the related rubric were confirmed (Tabachnick & Fidell, 2011; Brown, 2006). The correlations between the factors, the loadings of the factors and the degree of fit of the model were determined through the SPSS AMOS package program. By performing the relevant analyzes, it was determined that the coefficients in the structural equation model were above .30. This value is acceptable according to Kline (2005). Chi-Square/Degree of Freedom ($\chi 2$/sd) and Root Mean Square Error of Approximation (RMSEA) values were calculated for the fit of EWES. The findings related to these values are presented in Table 2.

Table 2. Fit Indices for Confirmatory Factor Analysis

Index of Fit	Valuable
Chi-Square/Degree of freedom (χ2/sd)	1.437
RMSEA	.071

When Table 2 is analyzed, it is seen that the Chi-Square/Degree of Freedom value (χ2/sd) is 1.437 and the root mean square of the errors of estimation (RMSEA) is .071. According to these data, it can be concluded that the measurement tool has an acceptable level of fit (Kline, 2011; Hooper et al., 2008). According to the results of the validity analysis, it was determined that EWES is a valid scale with three factors.

Findings Related to Reliability

The reliability of the measurement tool can be defined as the consistency and continuity seen in the measurement results (Fraenkel & Wallen, 2008). Cohen et al. (2007) defined reliability as the consistency of measurements made on groups with common characteristics with each other.

In this study, Cronbach's Alpha internal consistency coefficient was calculated for the reliability of EWES. As a result of the reliability analysis, Cronbach's Alpha internal consistency coefficient of the rubric is presented in Table 3.

Table 3. Cronbach's Alpha Internal Consistency Coefficient Value of EWES

Number of items	Cronbach' Alpha
25	.785

When Table 3 was analyzed, it was found that the Cronbach's Alpha internal consistency coefficient of the rubric was .785. Based on the value of Cronbach's Alpha internal consistency coefficient, it is possible to say that the rubric is reliable (Pallant, 2001; Tabachnick & Fidell, 1996; Tuckman & Harper, 2012).

In addition to all these, it is recommended to look at the inter-rater agreement in the reliability analysis of rubrics (Moskal et al., 2002). In this context, Kappa analyzes, which is a technique for determining the inter-rater agreement, were performed in the study. The findings related to kappa analysis are presented in Table 4.

Table 4. Value of Kappa Analysis of EWES

Number of items	Kappa Coefficient
25	.64

When Table 4 is analyzed, Kappa value is 0.64. Considering this value, it can be said that the agreement between the raters is at a significant level (Landish & Koch, 1977). In other words, since the opinions expressed by the raters are in significant agreement with each other, it can be stated that the reliability between the raters is appropriate.

In general, it can be said that EWES is a valid and reliable scale according to the validity and reliability analyzes. In this context, the final version of the relevant rubric is presented in Appendix 1.

Findings Related to The Content Prepared by Pre-Service Teachers Using Web 2.0 Tools

In the context of determining the evaluation of the Web 2.0 tools prepared in the research process in terms of design, content and ease of use, descriptive statistics such as arithmetic mean, standard deviation, minimum and maximum value for the scores obtained from the subsections of the rubric by the prospective teachers participating in the research are presented in Table 5.

Table 5. Descriptive Statistics Results According to the Sub-Dimensions of the Scores Received by the Preservice Teachers from the Content They Prepared Using Web 2.0 Tools

Dimensions of Educational Web 2.0 Tools Evaluation Scale in Science Education	n	$\bar{X}$	Ss	Min	Max
Design	56	28.16	3.014	19	32
Content	56	37.54	1.695	32	40
Ease of Use	56	20.82	2.405	10	28
Total Score from All Dimensions	56	86.52	4.987	74	99

When Table 5 is analyzed, the mean scores of the pre-service teachers in the study for the content they prepared using Web 2.0 tools differ in each dimension. The arithmetic mean that the pre-service teachers received from the sum of all dimensions is 86.52. This arithmetic mean shows that pre-service teachers' ability to create content using Web 2.0 tools is quite high. In addition, when we look at the scores obtained by the pre-service teachers from the content they prepared using web 2.0 tools, it is seen that the highest arithmetic mean is in the "content" dimension with 37.54 and the lowest arithmetic mean is in the "ease of use" sub-dimension with 20.82.

Discussion and Conclusion

This study developed the Evaluation Scale for Educational Web 2.0 Tools in Science Education (EWES) and conducted validity and reliability analyzes. The rubric, based on Haladyna's (1997) stages, consists of 25 items scored from 1 ('poor') to 4 ('good'), with total scores ranging from 25 to 100. It was tested on content prepared by pre-service teachers. Validity analysis showed a KMO value of .85 and a significant Bartlett’s Sphericity test, indicating suitability for factor analysis. EFA revealed a three-factor structure explaining 52.70% of the total variance, with factor loadings between .498 and .862. CFA confirmed that all items fit the scale, with acceptable fit indices ($\chi 2/sd = 1.437$, RMSEA = .071). The reliability analysis showed a Cronbach's Alpha coefficient of .785, demonstrating that the scale is reliable.

In the literature, Moskal et al. (2002) also suggested looking at the inter-rater agreement to determine the reliability of rubrics. In this context, Kappa analyzes were performed to look at the agreement between the scores obtained as a result of the evaluation of Web 2.0 tools prepared by pre-service teachers with EWES. According to the result

of Kappa Analysis, Kappa coefficient= .64 was determined. Based on the value obtained, it can be said that the reliability between the raters is appropriate. In other words, the scores given by the raters are in harmony with each other.

In the research literature, there is no study on the evaluation of the contents prepared with Web 2.0 tools according to "design", "content" and "ease of use" sub-dimensions. However, the research is partially similar to the studies on educational software evaluation. In order to make a comparison with this study, different studies are encountered when the research literature on educational software evaluation is examined (Iglesias et al. (1997); Beale & Sharples (2002); Herring et al. (2005); Escudeiro et al. 2009; Erensayın & GUler, 2017). For example; Iglesias et al. (1997) developed an evaluation method based on user responses in categories such as content, teaching methodology, design and user opinions in his study for the evaluation of educational software. Herring et al. (2005) developed a multimedia software evaluation scale for teachers and created an evaluation form that emphasized the importance of evaluating format, content and learning processes. Beale and Sharples (2002) provided a design guide for educational software developers emphasizing usability, usefulness and desirability as key aspects. In the study conducted by Erensayın and GUler (2017) in order to develop more qualified materials by evaluating the materials on the Educational Informatics Network (EBA) platform according to the educational software evaluation criteria by teachers, it was emphasized that the educational features of the materials should reveal the prior knowledge of the students, provide adequate feedback and the achievements should be appropriate to the student level. In a study conducted by Kartal et al. (2017), it was concluded that the content and educational features of the educational software evaluated were at a good level. The results of the current study are similar to the mentioned studies.

According to the analyzes, it can be stated that the developed EWES is a valid and reliable rubric. Consisting of 25 items, EWES can be said to be a valid and reliable rubric to be used in the evaluation of Web 2.0 tools prepared in the educational field in terms of design, content and ease of use. The related rubric is not only designed to be used in the evaluation of Web 2.0 tools, but can also be integrated into the evaluation of other digital contents.

Recommendations

Web 2.0 tools offer valuable opportunities in science education, but as technology evolves, their necessity may diminish with the rise of artificial intelligence (AI). Although AI applications are not yet widespread in education, both Web 2.0 and AI-generated content need to be evaluated for integration and quality. This study introduces pre-service teachers to digital technologies and evaluates the content they produce in terms of design, content, and usability. It is expected that the findings will enhance teachers' digital competencies and ability to produce quality content, benefiting their long-term professional development.

The relevant rubric was created for the evaluation of Web 2.0 tools prepared in the educational field and can be applied to the evaluation of digital content. New measurement tools can be developed in line with the features that digital content should have. The related rubric can be used in the evaluation of Web 2.0 tools in order to guide the creation of quality content and to encourage the creation of quality content. In this study, a new rubric was

developed and the validity and reliability study of the rubric was conducted with pre-service teachers. Therefore, it is considered that it would be more useful to analyze the relevant rubric by applying it to a larger study group.

Statements and Declarations

Acknowledgements

This manuscript is based on research conducted as part of the author Onur Duran's Master's thesis.

Authors contributions

All authors participated in the research design and writing of the manuscript. The final version of the paper received approval from all the authors.

Funding

This study is based upon work supported by The Scientific and Technological Research Council of Turkey, 1001-Scientific Technological Research Project Support Program under Grant 220K080 entitled "Designing and Evaluation of the Effectiveness of Scientific Inquiry Supported Online Mentoring (e-scaffolding) in In-service Teacher Training"

Availability of data and materials

The data that support the findings of this study are available on request from the corresponding author.

Competing interests

The authors have no potential conflict of interest to report.

References

Ahmadigol, J. (2016). *New definition of educational technology. The Annual Convention of the Association for Educational. 1*, 1–3.

Aksaya, H. (2014). *Examination of open course materials platforms in the perspective of lifelong learning*. (Unpublished master's thesis). Marmara University, Department of Journalism, Istanbul.

Al-Alwani, A. (2014). Evaluation criterion for quality assessment of E-learning content. *E-Learning and Digital Media, 11*(6), 532-542. https://doi.org/10.2304/elea.2014.11.6.532

Alhassan, R. (2017). Exploring the relationship between web 2.0 tools self-efficacy and teachers' use of these tools in their teaching. *Journal of Education and Learning*, *6*(4), 217-228. https://doi.org/10.5539/jel.v6n4p217

Ates, A. (2011). Educational software evaluation scale: validity and reliability study. *Journal of Educational*

Technology Research, 2(1).

Banerjee, J., & Bhattacharyya, M. (2011). Selection of the optimum font type and size interface for on screen continuous reading by young adults: an ergonomic approach. *Journal of human ergology*, *40*(1-2), 47-62. https://doi.org/10.11183/jhe.40.47

Baxter, G. J., & Connolly, T. M. (2014). Using web 2.0 tools to support the theoretical constructs of organisational learning. In *International Joint Conference SOCO'13-CISIS'13-ICEUTE'13: Salamanca, Spain,* September *11th-13th, 2013 Proceedings* (pp. 679-688). Springer International Publishing. https://doi.org/10.1007/978-3-319-01854-6_69

Baya'a, N.F., Shehade, H., & Baya'a, A. (2009). A rubric for evaluating web-based learning environments. *British Journal of Educational Technology*, 40, 761-763. https://doi.org/10.1111/j.1467-8535.2008.00864.x

Beale, R., & Sharples, M. (2002). Design guide for developers of educational software. *British Educational Communications and Technology Agency.* 1-34.

Berg, S. (2011). Web 2.0 technologies in higher education teaching: A practical introduction. *Kentucky Journal of Excellence in College Teaching and Learning*, *8*(1), 2-10.

Blumenthal, J. C. (1994). *English 2200: with writing applications: a programmed course in grammar and usage*, Harcourt.

Brookhart, S. M. (1999). The Art and Science of Classroom Assessment. The Missing Part of Pedagogy. *ASHE-ERIC Higher Education Report, 27*(1).

Brown, T.A. (2006). *Confirmatory factor analysis for applied research*, New York: GuilfordPress.

Bustard, D., Moore, A., Higgins, D., & Ayre, D. (2007, March). Towards self-managing web sites: The link integrity problem. In *Fourth IEEE International Workshop on Engineering of Autonomic and Autonomous Systems,* (pp. 61-67). IEEE. https://doi.org/10.1109/EASE.2007.26

Chowdhury, A. A. (1995). Interactive learning in a beginning CS course. *3C ONLINE, 2*(3), 7-9. https://doi.org/10.1145/213615.213614

Cochrane, T.D., & Bateman, R. (2010). Smartphones give you wings: pedagogical affordances of mobile web 2.0. *Australasian Journal of Educational Technology*, 26, 1-14. https://doi.org/10.14742/ajet.1098

Cohen, L., Manıon, L., & Morrison, K. (2007). *Research methods in education.* London And New York: Routledge/Taylor & Francis Group.

Creswell, J. W. (2012). *Educational research: Planning, conducting and evaluating quantitative and qualitative research*. Pearson.

Creswell, J. W. (2014). *A concise introduction to mixed methods research.* SAGE publications.

Erensayın, E., & Guler, C. (2018). Adaptation of multimedia software evaluation form into Turkish. *Yuzuncu Yıl University Journal of Faculty of Education, 15*(1), 1332-1354. http://doi.org/10.23891/efdyyu.2018.107

Farkas, D. K. (2002, September). Managing headings in print and online documents. In *Proceedings. IEEE International Professional Communication Conference* (pp. 283-296). IEEE. https://doi.org/10.1109/IPCC.2002.1049111

Ferrer, G.T., & Ramírez, S.U. (2016). Rúbrica para la evaluación de portafolios electrónicos en el entorno de la web social. *Píxel-Bit. Revista de Medios y Educación, 48*(2016), 83-96. https://doi.org/10.12795/pixelbit.2016.i48.06

Field, A. (2005). *Discovering Statistics Using SPSS for Windows.* London. Sage Paplications.

Fraenkel, J. R., & Wallen, N. E. (2008). *How to Design and Evaluate Research in Education.* New York. McGraw-Hill Higher Education.

Goodrich, H. (1997). Understanding rubrics: The dictionary may define" rubric," but these models provide more clarity. *Educational Leadership, 54*(4), 14-17.

Gul, S., & Yesilyurt, S. (2011). Preparation and evaluation of a courseware based on constructivist learning approach. *Cukurova University Journal of Faculty of Education, 1*(40), 19-36.

Haladyna, T. M. (1997). *Writing test ıtem toevaluate higher order thinking.* USA: Allyn & Bacon.

Herring, D., Notar, CE, & Wilson, JD (2005). Multimedia software evaluation form for teachers. *Education, 126*(1), 100-112.

Hichang, C., Geri, G., Barry, D.,& Anthony, I. (2007). Social networks, communication styles, and learning performance in a cscl community. *Computers and Education. 49*(2), 309-329. https://doi.org/10.1016/j.compedu.2005.07.003

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods, 6*(1), 53-60.

Hussein, D., Ghada, A., & Hamad, A. (2009). Classifying Web 2.0 supported applications by pattern of usage: Functional & technical ISSUES. The 4th Mediterranean Conference on Information Systems, 25-27 September 2009, Athens University, Athens, Greece.

Iglesias, OA, Paniagua, CN, & Pessacq, R.A. (1997). Evaluation of university educational software. *Computer Applications in Engineering Education*, *5*(3), 181-188. https://doi.org/10.1002/(SICI)1099-0542(1997)5:3<181::AID-CAE5>3.0.CO;2-9

ISTE. (2017). ISTE standarts: Educator. ISTE: https://www.iste.org/standards/iste-standards-for-teachers

Januszewski, A., & Molenda, M. (2013). *Educational technology: A definition with commentary*. Routledge.

Kelleci, O. (2010). *Development and application of an educational software evaluation form.* Unpublished master's thesis). Marmara University, Department of Computer Education and Instructional Technology, Istanbul.

Kline, R. B. (2005). *Principles and practice of structural equation modeling (2nd. Ed.)* New York: The Guilford Pres.

Kline, R. B. (2011). *Principles and practice of structural equation modeling (3rd ed*.). New York: The Guildford Press.

Landis, J. R., & Koch, G. G. (1977). An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics,* 363-374. https://doi.org/10.2307/2529786

Majhi, S., & Maharana, B. (2010). Innovative web 2.0 technologies for integrating the learning process. XI MANLIBNET: Trends and Challenges in Management and Corporate Libraries in Digital Era, 15-17 February 2010, Secundarabad, India.

Marsh, E. E. (2002). *Rhetorical relationships between images and text in web pages*. University of Maryland, College Park.

MEB (2011). *Graphics and photography.* MEB.

Middleton, F. (2019). Reliability vs validity: what's the difference? scribbr. https://www.scribbr.com/methodology/reliability-vs-validity/

Mishra, L., Gupta, T., & Shree, A. (2020). Online teaching-learning in higher education during lockdown period

of COVID-19 pandemic. *International Journal of Educational Research Open, 1*(September). https://doi.org/10.1016/j.ijedro.2020.100012

Moskal, B. M., Leydens, J. A., & Pavelich, M. J. (2002). Validity, reliability and the assessment of engineering education. *Journal of Engineering Education, 91*(3), 351-354. https://doi.org/10.1002/j.2168-9830.2002.tb00714.x

Mularsih, H. (2007). *Pembelajaran* individual dengan menggunakan modul. *Akademika*, *9*(1), 1-17.

Murugesan, S. (2007). Understanding web 2.0. *IT professional, 9*(4), 34-41.

Nkansah, C., Serwaa, D., Adarkwah, L. A., Osei-Boakye, F., Mensah, K., Tetteh, P. & Apodola, A. (2020). Novel coronavirus disease 2019: knowledge, practice and preparedness: a survey of healthcare workers in the Offinso-North District, Ghana. *The Pan African Medical Journal, 35*(Suppl 2).

O'Reilly, T. (2007). What is web 2.0?: Design patterns and business models for the next generation of software. *The Social Media Reader, 65*, 32–52. https://doi.org/10.18574/nyu/9780814763025.003.0008

O'Reilly, T. (2005). What ıs web 2.0 design patterns and business models for the next generation of software. https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html

OECD (2023), *OECD digital education outlook 2023: Towards an Effective Digital Education Ecosystem*, OECD Publishing, Paris, https://doi.org/10.1787/c74f03de-en.

Ozerbas, M.A., & Mart, O.A. (2017). Pre-service English teachers' opinions on the use of web 2.0 and their usage levels. *Journal of Ahi Evran University Kırsehir Faculty of Education, 18*(3), 1152-1167. https://doi.org/10.29299/kefad.2017.18.3.058

Pallant, J. (2001). *SPSS Survival Manuel. Maidenhead,* Open university Press.

Pan, S. C., & Franklin, T. (2011). In-Service teachers' self-efficacy, professional development, and web 2.0 tools for ıntegration. *New Horizons in Education, 59*(3), 28-40.

Reimold, C. (1981). Business writing-clear and simple. *IEEE Transactions on Professional Communication*, (4), 184-185. https://doi.org/10.1109/TPC.1981.6501691

Rello, L., Pielot, M., & Marcos, M. (2016). Make It Big!: The effect of font size and line spacing on online readability. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 3637-3648. https://doi.org/10.1145/2858036.285820

Richey, R. C., Silber, K. H., & Ely, D. P. (2008). Reflections on the 2008 AECT definitions of the field. *TechTrends*, *52*(1), 24-25. https://doi.org/10.1007/s11528-008-0108-2

Romero, O. C., Fernández, I. M. G., & Rodriguez, J. R. (2017). La evaluación de los materiales didácticos. *RELATEC: Revista Latinoamericana de Tecnología Educativa*, *16*(2), 79-95. https://doi.org/10.17398/1695-288X.16.2.79

Rowley, J. (2008). Understanding digital content marketing. *Journal of Marketing Management, 24*(5-6), 517-540. https://doi.org/10.1362/026725708X325977

Stauber, D. M. (2005). Do Mi's second rule or the functions of subheadings. *The Indexer: The International Journal of Indexing, 24*(5), 192-196. https://doi.org/10.3828/indexer.2005.24.4.11

Stefani, A., & Kameas, A. (2013). A Life-Cycle model for the quality evaluation of educational content. *LINQ 2013: Learning Innovations and Quality - The Future of Digital Resources,* 29. 1-8.

Tabachnick, B.G. & Fidell, L.S. (1996). *Using multivariate statistics.* New York: Harper Collins College.

Tabachnick, B. G. & Fidell, L. S. (2007). *Using multivariate statistics (5. Ed.).* Boston: Allyn And Bacon.

Tabachnick, B.G. & Fidell, L.S. (2011). *Using multivariate statistics*. Boston: Allyn and Bacon.

Tuckman, B. W. & Harper, B. E. (2012). *Conducting educational research (6th ed.).* Plymouth: Rowman & Littlefield Publishing.

TUBITAK (2009). *Handbook for effective presentations.* TUBITAK. https://www.tubitak.gov.tr/sites/default/files/content_files/iletisim/sunum_el_kibabi.pdf

UNESCO, (2008). Educational Technology. https://unevoc.unesco.org/home/TVETipedia+Glossary/show=term/term=Educational+technology.

UNESCO, (2018). *UNESCO ICT Competency Framework for Teachers*. https://www.unesco.org/en/digital-competencies-skills/ict-cft.

UNESCO, (2023). *Global Education Monitoring Report 2023: Technology in Education-A Tool on Whose Terms?*. UN. https://www.unesco.org/gem-report/en

Yıldız, H., & Sarıtepeci, M. (2013). A review on educational software in the light of programme evaluation models. *Academic Informatics 2013 - XVth Academic Informatics Conference Proceedings,* 503–508.

Yigit, N., Butuner, S. O., & Dertlioglu, K. (2008). Developing a webbing site evaluation scale for instructional purposes. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education, 2(*2), 38–51.

Yuan, M., & Recker, M. (2015). Not all rubrics are equal: a review of rubrics for evaluating the quality of open educational resources. *The International Review of Research in Open and Distributed Learning*, 16, 16-38. https://doi.org/10.19173/irrodl.v16i5.2389

Author Information

Onur Duran

https://orcid.org/0000-0002-8662-8813

National Defense University, Istanbul

Türkiye

Ayşe Oğuz Ünver

https://orcid.org/0000-0003-2938-5269

Mugla Sitki Kocman University, Science Education Department, Mugla

Türkiye

Contact e-mail: *ayseoguz@mu.edu.tr*

Appendix. The Scale

Criteria	Score Range	Score	Total Score
A. Design			
Pictures and graphics are functional, not decorative.	1-4		/32
The background Color is compatible with the Color of the foreground elements.	1-4		
An appropriate Color combination is used between the background Color and the text.	1-4		
The Colors used add attractiveness to the product visuality	1-4		
The font type and size are appropriate for the user level.	1-4		
Appropriate font size is used for text, headings and subheadings.	1-4		
The placement of my components is suitable for the aesthetics of the page.	1-4		
Web 2.0 tool is remarkable and arouses curiosity.	1-4		
B. Content			
The subject is related to the acquisitions in the curriculum.	1-4		/40
The subject has adequate explanations.	1-4		
The subject is suitable for the target user level.	1-4		
It can be processed within the specified time.	1-4		
The message intended to be given in the texts is clear and understandable.	1-4		
A clear and simple language is used.	1-4		
Unnecessary information density is avoided.	1-4		
It is up-to-date and scientifically appropriate.	1-4		
The language and expression in the texts are appropriate for the level of the target user.	1-4		
The subject is in a logical order.	1-4		
C. Ease of Use			
Links (Web 2.0 tool link) work smoothly.	1-4		/28
It is possible to access the Web 2.0 tool from any operating system.	1-4		
It is possible to access the Web 2.0 tool from different browsers (Google Chrome, Internet Explorer, etc.).	1-4		
The Web 2.0 tool is personalised and provides feedback.	1-4		
The desired information can be easily accessed in the Web 2.0 tool.	1-4		
There are instructions in the Web 2.0 tool.	1-4		
The user can continue the Web 2.0 tool at any time.	1-4		
Total			/100