Development of Analytical Rubric for Evaluating the Effect of Mobile Design Applications on Artistic Creativity: Validity-Reliability Study

Sevda CEYLAN DADAKOGLU
Ministry of National Education, Ankara-Turkey

Atilla OZDEMIR
Suleyman Demirel University, Isparta-Turkey

Abstract

Purpose: Creativity and technology are essential components of education and the future around the world. Developing students' creative thinking skills and their relationship with technology is among the main objectives of both 21st century skills and education. This study, which is based on the relationship between digital technologies and artistic creativity, aims to develop an analytical rubric to evaluate 11th grade students' artistic creativity that emerges with their use of mobile design applications.

Design & Methodology: This research is a descriptive research aiming to develop an analytical rubric with high validity, reliability and usefulness to evaluate the effect of mobile design applications on artistic creativity.

Findings: There are five criteria in the analytical rubric developed to evaluate the effect of mobile design applications on artistic creativity. Expert opinion was sought in determining the content validity of the developed analytical rubric, and principal component analysis was used in determining the construct validity. In classical test theory, Cronbach's alpha coefficient was calculated for internal consistency, and Kendall's coefficient of concordance (W) was calculated for inter-rater reliability, and in generalizability (G) theory, G and Phi coefficients were calculated. Fully crossed design was preferred in G theory analyzes. The findings obtained show that the developed analytical rubric gives reliable results in evaluating the artistic creativity of the students.

Implications & Suggestions: An analytical rubric was developed to evaluate the 11th grade students' artistic creativity that emerge through the designs made with mobile design applications and the validity and reliability proofs of the rubric were revealed. As a result of the analysis, it was concluded that the developed Creativity Level Analytical Rubric (CLAR) is a reliable and valid measurement tool. It is recommended to increase the number of studies to develop tools that can evaluate artistic creativity and the relationship between creativity and technology at all levels of education

Mobil Tasarım Uygulamalarının Sanatsal Yaratıcılığa Etkisini Değerlendirmeye Yönelik Analitik Rubrik Geliştirme: Geçerlik-Güvenirlik Çalışması

Sevda CEYLAN DADAKOĞLU
Milli Eğitim Bakanlığı, Ankara-Turkey

Atilla ÖZDEMİR
Süleyman Demirel Üniversitesi, Isparta-Turkey

Öz

Amaç: Dünya genelinde yaratıcılık ve teknoloji, eğitimin ve geleceğin temel bileşenlerindendir. Öğrencilerin yaratıcı düşünme becerilerini ve teknoloji ile olan ilişkilerini geliştirmek, hem 21. yy. becerileri hem de eğitimin temel amaçları içerisinde yer almaktadır. Dijital teknolojiler ve sanatsal yaratıcılık arasındaki ilişki üzerine temellendirilen bu çalışmanın amacı, ortaöğretim 11. sınıf düzeyindeki öğrencilerin, mobil tasarım uygulamalarını kullanmalarıyla ortaya çıkan sanatsal yaratıcılıklarını değerlendirmek üzere analitik dereceli puanlama anahtarı geliştirmektir.

Yöntem: Bu araştırma, mobil tasarım uygulamalarının sanatsal yaratıcılığa etkisini değerlendirmeye yönelik geçerliği, güvenirliği ve kullanışlılığı yüksek bir analitik rubrik geliştirilmesini amaçlayan betimsel araştırma niteliğindedir.

Bulgular: Mobil tasarım uygulamalarının sanatsal yaratıcılığa etkisini değerlendirmeye yönelik geliştirilen analitik rubrikte beş ölçüt yer almaktadır. Geliştirilen analitik rubriğin kapsam geçerliğini belirlemede uzman görüşü, yapı geçerliği belirlemede ise temel bileşenler analizi yöntemi kullanılmıştır. Klasik test kuramın'da, Cronbach alfa iç tutarlılık güvenirlik katsayısı; puanlayıcılar arası güvenirlik için Kendall'ın uyuşum katsayısı (W) belirlenirken, genellenebilirlik (G) kuramında ise G ve Phi katsayıları kestirilmiştir. G kuramı analizlerinde tümüyle çaprazlanmış desen uygulanmıştır. Elde edilen bulgular, geliştirilen analitik rubriğin, öğrencilerin sanatsal yaratıcılığını belirlemede güvenilir sonuçlar verdiğini göstermektedir.

Sonuçlar ve Öneriler: Ortaöğretim 11. sınıf öğrencilerinin mobil tasarım uygulamalarını kullanarak elde ettikleri tasarımlar aracılığıyla ortaya çıkan sanatsal yaratıcılığı değerlendirebilmek amacıyla analitik dereceli puanlama anahtarı geliştirilmiş ve bu puanlama anahtarının geçerlik ve güvenirlik kanıtları ortaya konulmuştur. Yapılan analizler sonucunda geliştirilen YDADPA'nın (yaratıcılık düzeyi analitik dereceli puanlama anahtarının) güvenilir ve geçerli bir ölçme aracı olduğu sonucuna varılmıştır. Sanatsal yaratıcılığı ve yaratıcılık-teknoloji ilişkisini eğitimin tüm kademelerinde ölçebilecek araçlar geliştirilmesi için çalışmaların artırılması önerilmektedir

INTRODUCTION

The rapid development of technology has integrated the use of technological tools and applications in all areas of our lives, creating a digital age. There have also been changes in educational approaches with the development of digital technologies. The methods and techniques used in visual arts education are not independent of these changes. Therefore, educators need to keep up with digital technologies and adapt these to their lessons. With the integration of the technological developments to education and the technology competence developed in individuals, the phenomenon of creativity is included in a series of skills called 21st-century skills, and these skills are expected from today's students. The World Economic Forum ranked creativity third among the ten most important skills for the fourth industrial revolution and creativity is also listed in the section of 21st century skills (Kanlı, 2020). Moreover, in the past two decades, creativity has been one of the key learning goals of the main learning goals of the 21st century (Bolden et al., 2020). Creative thinking is a high-level mental process (Kutlu et al., 2017). Therefore, special measurement tools should be developed and used to evaluate such skills. Considering these, it is evident that using digital tools in visual arts education is important and they impact visual arts students' creativity. In this context, the problem statement of the research is the use of mobile design applications in art education and the development of a DPA in order to evaluate this situation in terms of students' creativity. Therefore, a review will be made on the concept of creativity and the use of digital tools in art education.

From a general perspective, the concept of "creativity", which we can define as "the state of being creative, the ability to create, the hypothetical predisposition that is accepted to exist in every individual, that drives something to create" (Turkish Language Association [TLA], 2020) has been seen as a feature belonging to God in history and the origin of creativity has been associated with God. Creativity gained its present meaning in the age of enlightenment. In the 18th century, creativity began to be associated with imagination rather than inspiration. In the 19th century, Galton analyzed creativity in terms of hereditary with his research (Bacanlı et al., 2011, p. 539). The definitions of the term 'creativity' vary according to researchers. For example, Lowenfeld, in his work titled "Creative and Mental Growth", described creativity as a basic instinct that all human beings are born with, that makes human beings human. He defined it as an instinct primarily used to solve and express life problems (Lowenfeld & Brittain, 1964). Guilford (1967), who conducted valuable studies on creativity, perceives creative thinking as "unconventional thinking" and defines this way of thinking as "flexible, original and fluent". He stated that the core of creativity resides in divergent thinking (Arieti, 2016).

Known for his research on assessing creativity, Torrance considers creative thinking as a process of intuition and defines creativity as detecting gaps, disturbing or missing items, generating thoughts or assumptions about them, testing them, comparing results, and possibly changing and retesting these assumptions (San, 1979). Wallach and Kogan (1965, as cited in Jersild, 1983), who also conducted research on the assessment of creativity, defines creativity as being able to produce a large number of unique associations; however, while doing these, not leaving the question altogether and not deviating to the quirks. According to Csikszentmihalyi (1996, as cited in Özaşkın & Bacanak, 2016), one of the prominent researchers in the field of creativity, creativity is ideas or actions that are new and valuable. With a general and broad definition, creativity can be defined as extracting something brand new from known things, reaching a unique synthesis, finding new solutions to some problems (San, 2004). It is known that there are many definitions, methods and theories in the field of creativity. For example, Treffinger has presented more than 100 different definitions of creativity (Treffinger, 1996 as cited in Kanlı, 2020). When the definitions are examined, it is seen that concepts such as innovation, invention and problem solving are frequently emphasized. Many contemporary psychologists and educators agree

66

that creativity is an interactive, complex process in which the relationships between persons, processes, products and social and cultural contexts are of great importance (Zimmerman, 2009).

When the concept of creativity is viewed in terms of artistic creativity, art is either an imitation of a divine work or the result of a divine inspiration. Therefore, the origin of artistic creativity is God. It is thought that the researcher who gave creativity its present meaning is J.P. Guilford (Bacanlı et al., 2011). According to American author Conrad (1990, as cited in San, 2004), artistic creativity is the emergence of an effective and harmonious metaphor in the process of finding research that includes sensation, perception, emotion, imagination. The artistic creative process includes thinking with images. Eisner describes the artistic creativity types that he identified in 1972 on the basis of pushing the boundaries, making inventions, breaking the boundaries and organizing them aesthetically (Kırışoğlu, 2005). Haladyna (1997, as cited in Gürkan & Dolapçıoğlu, 2020) divides creativity into cognitive creativity in the mental context and aesthetic creativity in the affective context. He identifies aesthetic creativity as creating a unique and original product that is not expected to fully coincide with scientific facts. Kırışoğlu (2005) defined the artistic creation process as thinking with values and problem-solving process. He described artistic problem solving as organizing by considering the relationships between artistic values such as color, line, texture and space in order to achieve the desired quality product. Thus, Kırışoğlu concluded that achieving a qualified result is solving the artistic problem and when the artistic problem is solved, artistic creation is successful.

In studies on creativity, researchers mentioned different dimensions of creativity. For example, Stein and Heinze (1960, as cited in Yavuzer, 1996) examine creativity in terms of efficiency, process, measurement and personality. Rhodes (1961) classified the dimensions of creativity and developed the 4P model. The content of the 4P Model consists of person (individual characteristics), process, product and press (environment). Later, Simonton added Persuasion dimension in 1990 and Runco added Potential dimension in 2003 to these four dimensions (Özaşkın & Bacanak, 2016). Mooney (1962, as cited in Yavuzer, 1996) defines the concept of creativity within the framework of product/efficiency, process, personality and the environment in which the product is created. He also stated that examining creativity in terms of process is seen in individuals who are interested in art. According to the process approach, the psychological structure and environment at the time of creation are considered the best criteria. May (2007) defined creativity as the process of revealing existence, an expression, the creation of something new. Amabile (1996, as cited in Onur & Zorlu, 2017) defined creativity as intelligence in recognizing the paths to be taken and maneuverability in the process of creating a product. He also argued that the dimensions of creativity are expertise, creative thinking skills and motivation. On the other hand, Higgins and Morgan (2000) described creativity as the discovery of ideas and reproduction of knowledge in an original way and, as in the other two views, tapped on the person, process and product dimensions of the concept of creativity. It is noteworthy that views on the dimensions of creativity are generally shaped around person, process, product and environmental factors. In general, these four main factors come to the fore when it comes to evaluating creativity (Pelowski et al., 2017).

The problem of evaluating creativity is as old as the concept of creativity itself. Researchers from various cultures and disciplines have tried to define the concept of creativity and offer a valid way to assess it (Kanlı, 2020). Although assessing is considered as a situation that hinders creativity, generally, researches support assessment of creativity, therefore, giving feedback (Bolden et al., 2020). "What are the mental processes of creative thinking? Which personality traits are associated with creativity?, How can a product be understood as creative? and What are the external forces that affect creativity?". The answers to these questions constitute the criteria of the most used creativity assessment tools (Kanlı, 2020). The evaluation of creativity varies according to the content of the creativity definitions. For example, Guilford evaluated creativity according to four attributes: Flexibility (trying different ways for a problem), Originality (being able to generate different thoughts), Efficiency (being able to offer multiple solutions), Elaborating (being able to examine the thought in detail). Torrence developed the Torrance

Tests of Creative Thinking, which allows the assessment of creativity in the fields of psychology and education, including three of the qualities presented by Guilford (Bacanlı et al. 2011). Lindstrom stated in 1998 that they used seven criteria in the assessment of creative performance, three of which included completed products/works, four of which were related to the research and study process. He stated that he determined these criteria according to the qualities accepted in the art world and researches about the creative process (Lindstrom, 2006). Fox and Schirrmacher (2014) stated that the creation process can be used as the basis for evaluation. The authors' evaluation criteria for artistic creativity are as follows;
- Willingness to experiment, explore and examine,
- Demonstrating methods of creatively combined tools, materials and artistic waste,
- Detail, decoration and careful use,
- Showing originality, imagination and creativity,
- Making unique individual and personal art expressions.
Many overlapping skills such as perception, memory, motor control, language, spatial reasoning, imagination are needed in order to produce products in the field of visual arts. This situation makes art one of the most complex human activities. Evaluating creativity, which is an important component of artworks, is also very difficult due to the nature of art (Pelowski et al., 2017). It is stated that rating scales for measuring some difficult cognitive skills such as creativity or group work, collaboration, leadership, communication, performance evaluations and simulations, skills and talent history, portfolios, tools containing different item types (multiple choice, computer-assisted and open-ended items) can be used (Yalçın, 2018). Kaygın and Çetinkaya (2015) stated that unconventional alternative evaluations such as expert opinion and teacher observations, multiple assessment methods, dynamic evaluations and performance evaluations are also utilized in the evaluation of creativity. Additionally, Kutlu et al. (2017) stated that rubric and self-peer-group evaluation forms can be used in measuring high-level mental processes and monitoring student achievement. It points out that rubrics that are scored by more than one rater are important in terms of providing reliability.

The term creativity is defined as thinking with values, problem-solving process (Kırışoğlu, 2005), problem-solving with the use of technology, visual reasoning, exploring and expressing creative thinking (Black & Browning, 2011). Therefore, creativity is related to "problem-finding, problem-solving, divergent and convergent thinking, self-expression and adaptation to new situations" (Zimmerman, 2009). In this research, creativity is defined as a phenomenon that emerges with the combination of person, process, product and environment. Therefore, creativity is analyzed in terms of person, process, product and environment. While developing the analytical rubric and determining the criteria, these definitions and dimensions were used.

Various researches have been done on the effects of art education on creativity. Regarding this issue, San (2004) argues that the most appropriate field for developing creativity of individuals from an early age is the artistic field and that art education should be widely included in general education. Zimmerman (2010) stated that teachers and students should take risks and that student works should be allowed to come into being through self-learning over time. According to him; art classes are places where true creativity and self-expression can be promoted and valued. Teachers play a leading role in their students' creative performances and the work they produce. Thus, creativity can be thought of as a form of social reasoning translated between an art educator and students in the cultural context of the art class (Thomas, 2009, as cited in Zimmerman, 2010). Technology is also included in the context of the art class today. Interest in creativity has increased with the possibilities provided by digital technologies. In the future, digital technologies, contemporary teaching practices and education are seen as the core of creativity. Technologies, mixed creative applications, and innovative montages can offer many options for creativity (Henriksen et al., 2021).

The Purpose and Importance of the Research

Education is a very important factor in raising creative individuals. Advances in technology are reflected in training programs and diversify training methods. Today, educators should be able to include new learning tools, applications, or new learning environments into the education favorably to the characteristics of the new generation called "digital natives". Digital natives, who were born into a completely digital world, spend most of their time in a virtual world and frequently use digital/mobile applications. Mobile design applications used in this research are included in this framework. Consequently, in this study, it is aimed to develop an analytical rubric in order to evaluate the 11th grade students' creativity that arise from using mobile design applications. For this purpose, answer to the question "How should the rubric that assesses students' artistic creativity through designs that emerge as a result of the use of mobile design applications in Visual Arts lessons be?" was sought.

Around the world, creativity and technology stand out as the basic components of education and the future. The importance of creativity in 21st century skills is indisputable, especially given the need for innovation to survive (Henriksen et al., 2021). Therefore, it is necessary to include all kinds of tools and materials that will support creativity in educational environments. In this context, it is important to use mobile design applications in art lessons and to associate the results with artistic creativity. Developing a rubric that can evaluate the obtained results makes the research important. In addition, the use of mobile design applications is considered to serve the purposes of education and to be interesting for the participants. This research can be considered important and up-to-date in today's world where technology is present in all areas of life.

METHOD

This research is a descriptive research aiming to develop an analytical rubric with high validity, reliability and usefulness to evaluate the effect of mobile design applications on artistic creativity. In descriptive research, the aim is to reveal a situation or phenomenon in a topic (Johnson & Christensen, 2004).

Participants

In the study, the study group was determined according to the convenience sampling, considering the accessibility. Study group consisted of 38 students studying in a public school in the city center of Ankara. Table 1 displays the distribution of students according to gender.

Table 1.
Distribution of students according to gender

Gender	Frequency (n)	%
Female	25	66
Male	13	34
Total	38	100

As shown in Table 1, a total of 38 students participated in the study. The research started by determining the problem statement and continued with instruction of art history subjects in the 11th grade Visual Arts course curriculum. After the subjects in the curriculum were completed, the participants conducted a research to explore mobile design applications. Participants tried the discovered mobile design applications and decided which applications to use. Then, the participants were asked to take the

artworks out of their context and reinterpret them using mobile design applications (Appendix 2). Table 2 shows the applications used by the participants and their frequency of use.

Table 2.
Mobile design applications used in the study and their frequency of use

Application used	Frequency of use
PicsArt	20
Caps yap	5
You Cut	5
CS6	4
Photo editör	2
Pixlr	2
PS6	2
Fotoğraf Birleştir	2
Photo Studio	2
Snapchat	2
Paint 3D	1
Photo Grid	1
Pizzap	1
Adobe Photoshop Express	1

Participating students stated that they used 14 different programs and applications at different frequencies of use. Most of the students used the PicsArt program. PicsArt is known as a photo editing application designed for mobile devices. The application, which offers convenience such as cutting and pasting photos, changing colors, adding texture, making simple drawings, also has a Turkish language option. For this reason, it has been the most used program by students. The mobile design applications presented in the table have almost similar features. These applications can help the students freely shape their design and reveal their creativity within their own means. Each participant was free to use the application they wanted, no restrictions were made, so no application was suggested or taught. The applications used can only make changes on student work with the commands given by the student.

The Creativity Level Analytical Rubric (CLAR) was developed in order for the studies designed by the participants to be evaluated by researchers and expert academicians (Appendix 1). In this research, the evaluation was carried out by three experts. One of the researchers is the course teacher of the participants and one of the raters. The other two raters consist of academicians who are experts in art and design.

Data Collection Tool (Development of Rubric)

CLAR was developed by the researchers to obtain the opinions of experts on artistic creativity developed in students through designs that emerge as a result of the use of mobile design applications in art lessons and to score and evaluate this developed artistic creativity. Rubric is one of the measurement tools used especially in performance-based evaluations (Goodrich, 1997). Thus, it not only gives students more

reliable feedback about their performance but also minimizes the amount of error and bias with its multilevel structure (Parlak & Doğan, 2014).

Determining the rubric type

Rubric types are divided into two as holistic and analytical. While process skills are mostly evaluated in analytical rubrics, outcome evaluation is emphasized in holistic rubrics (Campell, 2005; Taylor, 2003). Therefore, for the purpose of this study, it was aimed to develop an analytical rubric.

Determining the criteria

During the development process of the rubric, the relevant literature was reviewed and rubrics used in this field were examined. Creativity is defined differently from various perspectives by many researchers. For some, creativity is a process, for some it is a product, for some it is the environment and for some it is a whole that covers them all. The definition of creativity used in this study is a multidimensional concept that unites these definitions. Therefore, it was taken into consideration that creativity should be examined from many angles while developing CLAR. Basically, based on the researches of Rhodes (1961), Kırışoğlu (2005) and Zimmerman (2009) on creativity, the criteria of CLAR were determined. Views of Rhodes (1961), which allows to assess the creativity level of the students in terms of person, product, process and environment, formed the basis for the behaviors that should be evaluated in applied fields; Kırışoğlu (2005), who grouped thinking with values, problem-solving, technical skills, aesthetic, expressive and creative elements in the product, and Zimmerman (2009), who defined creativity as problem-finding, problem-solving, divergent and convergent thinking, self-expression and adaptation to new situations) formed the basis of CLAR. The draft rubrics prepared were presented to the field experts and the evaluation and measurement specialist, and after the necessary corrections were made, a consensus was achieved on the final form of the rubric. The criteria defined as a result of the researches were determined to evaluate the process, product, person and environment. In this context, CLAR has been prepared to evaluate student studies in terms of "Knowledge and Perception Capacity", "Technical Skill", "Design", "Research and Working Approach" and "Self-Evaluation Capacity". While creating CLAR, considering the very different levels of the students, basic areas related to the level of creativity were determined. For this, attention has been paid to ensure that the limits of the criteria determined are clear and the features to be observed increase the usefulness. For this, 5 different criteria were determined and scored in 5 different degrees of success. The degrees are defined with expressions appropriate to each criterion.

Defining the behavior levels

The content of the five criteria in the CLAR is as follows:

1- Knowledge and Perception Capacity, student's ability to know and use methods and techniques related to mobile design applications; (Rhodes, 1961; Zimmerman, 2009; Black & Browning, 2011)
2- Technical Skills, student's ability to use digital tools, analyze and effectively use materials and techniques, produce solutions to technical problems, take risks when necessary, and use mobile design applications; (Rhodes, 1961; Kırışoğlu, 2005; Zimmerman,2009; Black & Browning, 2011)
3- Design, student's ability to use visual design elements and principles as well as affective qualities unusually and effectively, to develop their ideas in a way that reflects them to the other side, to create an original visual layout, to make continuous experiments to ensure visual integrity and to create original

and extraordinary designs with mobile design applications (Rhodes, 1961; Kırışoğlu, 2005; Black & Browning, 2011)

4- Research and Working Approach, student's outward-looking, curious and investigative attitude in acquiring information, searching for different solutions and developing alternative ideas, combining ideas with new methods, fulfilling their responsibilities completely within the given time, not being afraid of difficulties and contributing to the work of their friends (Rhodes, 1961; Zimmerman, 2009)

5- Self-Assessment Capacity, student's multi-faceted critical attitude towards his/her design and technical skills and awareness of his/her needs within the framework of learning deficiencies (Rhodes, 1961; Zimmerman, 2009) were evaluated.

Although the criteria are interrelated, they are not overlapping. Thus, it was aimed to evaluate the artistic creativity levels of the students from different angles. After the criteria were determined, the definitions and explanations of each criterion were examined. Later, these characteristics were properly defined, performance criteria and levels of performance expected from the student were determined, and each criterion was graded with scores between 0 and 5. The criteria are expressed as 0: Poor, 1: Minimal, 2: Sufficient, 3: Above Average and 4: Excellent (Appendix 1). The scores the students got from the rubric were obtained by scoring student products separately in each criterion and summing up the scores. While determining these qualifications, rubric qualifications of "The Harriet W. Sheridan Center for Teaching and Learning" were taken as the basis (Brown University, 2020). Students were informed about the implementation of CLAR.

Analysis and Interpretation of Data

In the analysis of the data, for the validity, necessary calculations were made in accordance with the formula of Miles and Huberman (1994). The percentage of agreement as a result of the calculation of above 70% is considered as a critical value for the reliability of the study (Miles-Huberman, 1994; Tavşancıl & Aslan, 2001). In order to test the reliability of the scores obtained from the rubric, the results obtained from Kendall's W and Generalizability (G) theory, in which the classical test theory (CTT) was used, was examined to determine the consistency between raters. Inter-rater reliability is the degree of agreement and consistency between two or more raters (Crocker & Algina, 1986). Calculations regarding the Kendall's W coefficient, which is based on CTT were made using SPSS. Although there are different methods to determine the reliability between raters in the classical test theory, the Kendall W fit coefficient was preferred because it can examine the level of harmony between more than 2 raters and syntax is not used. On the other hand, G Theory has emerged from the limitations of CTT such as not being able to distinguish the error sources in the measurement process and to determine the interaction between errors. In G theory, it is possible to separately calculate the errors and their quantities that are involved in measurement from various error sources with a single analysis. In accordance with the G theory, the variance components for the main and common effects (interaction) were estimated, and the G and Phi coefficients were calculated for the reliability of the scores. EduG program was used in these calculations and decision study (D study). Both methods are effective in determining the reliability between more than two raters. All raters who took part in the study were informed about the purpose of the study, student products, and the scoring key. All the students participating in the study completed their products, and since all the products obtained were scored independently by three raters, a fully crossed (individualxitemxrater) pattern was used for the sources of variability. In this way, it was aimed to compare the findings obtained from different methods. Thus, it was aimed to reveal the quality of the study by comparing the data obtained on reliability with two different methods.

FINDINGS

Findings Regarding the Validity of Rubric

The validity of a measurement tool is related to the extent the tool measures the desired property. In order to determine the validity of the relevant rubric, the opinions of field experts were consulted. Therefore, three experts from the Visual Arts field, one expert from the measurement and evaluation field and one expert from the Turkish Language and Literature field were asked to examine the developed draft CRAL in terms of content validity (content, structure and criteria).

For that reason, expert opinion was consulted for the developed draft CRAL. Opinions of the experts about the content, structure and criteria were obtained in accordance with the purpose and scope of the study. For expert opinions, the following equation recommended by Miles-Huberman (1994) was taken into account:

$$\text{Reliability} = \frac{\text{Agreement}}{\text{Agreement} + \text{Disagreement}}$$

It is considered reliable if the percentage of agreement obtained as a result of this calculation exceeds 70% (Miles-Huberman, 1994; Tavşancıl & Aslan, 2001). The opinions of three experts (two in the field of visual arts, one in the field of measurement and evaluation) were taken on the usefulness, the comprehensibility of the criteria and the appropriate expression of the levels of the developed draft CRAL and it was observed that the percentage of agreement between expert opinions was 90% in the first round and 100% in the second round after making necessary corrections. Thus, the form was finalized in line with expert opinions (Appendix 1).

The researchers took part as "participant observers" during the development of CLAR. In participant observation, the researcher is involved in the life of a group, community or organization for a long time to understand the habits and thoughts of people and to solve the social structure that holds them together (Punch, 2014).

Findings Regarding the Construct Validity of Rubric

Exploratory factor analysis (EFA) was performed for the scores obtained from three raters regarding the construct validity of the rubric. Analysis results showed that there is only one factor with eigenvalue greater than 1 for each rater. In other words, the rubric was collected under a single factor in the scores given by the raters to each sub-dimension of the rubric. These findings reveal that each sub-dimension of the rubric measures a single common structure that is intended to be measured. Table 3 displays the EFA results.

Table 3.
Exploratory factor analysis results

Dimensions	Factor Loads			
	Rater 1	Rater 2	Rater 3	Average Score
Knowledge and Perception Capacity	0,82	0,89	0,78	0,86
Ability to use digital tools	0,83	0,90	0,84	0,83
Design	0,79	0,77	0,86	0,85
Research and Working Approach	0,84	0,87	0,83	0,85
Self-Assessment	0,82	0,84	0,83	0,83
Eigenvalue	3,40	3,69	3,46	3,59
Variance explained	68,14	73,92	69,28	71,82

Descriptive Statistics

In this section, firstly descriptive statistics of the scores obtained from CLAR are presented. Table 4 presents the results of the raters for each sub-dimension in the rubric.

Table 4.
Descriptive statistics regarding the scores obtained from CLAR

Dimensions	Rater	$\bar{x}$	Sx	Min	Max	Skewness	Kurtosis
Knowledge and Perception Capacity	P1	3,21	0,77	2	4	-0,393	-1,21
	P2	3,10	0,79	2	4	-0,196	-1,38
	P3	3,18	0,76	2	4	-0,332	-1,19
Ability to use digital tools	P1	3,07	0,78	2	4	-0,142	-1,33
	P2	3,02	0,75	2	4	-0,44	-1,18
	P3	3,02	0,71	2	4	-0,03	-0,96
Design	P1	3,10	0,83	2	4	-0,20	-1,52
	P2	3,00	0,80	2	4	0,00	-1,446
	P3	3,07	0,78	2	4	-0,14	-1,33
Research and Working Approach	P1	3,39	0,63	2	4	-0,57	-0,54
	P2	3,39	0,63	2	4	-0,57	-0,54
	P3	3,36	0,67	2	4	-0,60	-0,62
Self-Assessment	P1	3,36	0,63	2	4	-0,48	-0,58
	P2	3,36	0,63	2	4	-0,48	-0,58
	P3	3,34	0,66	2	4	-0,52	-0,65
Total	P1	16,15	3,02	10	20	-0,33	-0,90
	P2	15,89	3,11	10	20	-0,21	-1,18
	P3	16,00	3,00	11	20	-0,05	-1,38

While the lowest score that can be obtained from each sub-dimension of the rubric is "0", the highest score is "4". In general, it is observed that students got high scores from both each sub-dimension and the whole rubric. When the averages of the scores given by the raters in each sub-dimension are examined, it can be said that they are close to each other. When the skewness-kurtosis values of the scores are analyzed, it is seen that the scores show a slightly skewed distribution to the left. Table 5 presents the data regarding the correlation between raters.

Table 5.
Correlation Coefficient Between the Scores Given by the Three Raters

	Rater 1	Rater 2	Rater 3
Rater 1	-	0,96*	0,92*
Rater 2		-	0,90*
Rater 3			-

*p<0,01

Table 5 ilustrates that the correlation coefficients between the scores given by the three raters for the five sub-dimensions are between 0.90 and 0.96. There is a high and significant relationship between all raters.

Findings Regarding the Reliability of Rubric

In line with the purpose of the research, firstly the reliability analysis findings related to CTT, then the analysis results related to the G Theory are presented below.

Classical test theory

In order to estimate the reliability of the scores obtained from CLAR, Cronbach's alpha internal consistency coefficient, which is one of the methods based on KTK, and Kendall's W correspondence coefficient were calculated. The Cronbach alpha coefficient was calculated by taking each rater's scores for the students separately and on the average scores. Table 6 provides the obtained Cronbach alpha coefficients.

Table 6.
Cronbach Alpha Reliability Coefficients

Rater	Cronbach alfa
P1	0,87
P2	0,90
P3	0,88
Average score	0,90

As can be seen from Table 6, the internal consistency values calculated for each rater vary between 0.87 and 0.90. The Cronbach alpha reliability coefficient, calculated by taking the average of the points given by the raters, was found to be 0.90. It can be said that the values obtained are in high reliability range (Özdamar, 1999). Then, the consistency between raters was analyzed with Kendall's W fit coefficient, which is a non-parametric statistical technique. Table 7 shows the obtained results.

Table 7.
Kendall's W coefficient between raters calculated for total scores

Coefficient of concordance	CLAR Total Score
Kendall's W	0,95*

*p<0,01

As Table 7 shows, according to the total CLAR scores, the coefficient of agreement between the three raters was calculated as 0.95, and this value was found to be significant. This value shows that the agreement between raters is high (Von Eye & Mun, 2005). This finding can be interpreted as raters show high similarity in the ranking of individuals. The results obtained suggest that CLAR is sufficiently reliable, but the Kendall's W coefficient alone was not considered sufficient for reliability. Therefore, findings are supported by generalizability theory.

Generalizability theory

In coefficient of concordance approach, situations other than those caused by the difference between rater scores as a source of error do not affect the result. In G theory, on the other hand, individual (in), item (it), rater (r), individual-item interaction (inxit), individual-rater interaction (inxr), item-rater interaction (itxr) and individual-item-rater interaction (inxitxr) are considered as sources of error. Table 8 provides the variance values obtained regarding the surface and conditions in the G theory analysis.

Based on the data presented in the table, the reliability is interpreted by obtaining the G and Phi coefficients.

Table 8.
Estimated Variance Components and Percentages (ANOVA table)

Source of Variance	Degree of Freedom (df)	Mean Square (ms)	Variance components estimation	Total variance percentage estimation
İndividual (in)	37	5,32	0,32	56,20
Item (it)	4	0,11	-0,00	0,00
Rater (r)	2	3,88	0,02	2,70
inxit	148	0,60	0,10	17,00
inxr	74	0,12	-0,00	0,00
itxr	8	0,54	0,01	1,90
inxitxr, e	296	0,12	0,12	22,2

In Table 8, it is seen that the individual variance, within the variance components estimated for the first three variables (individual, item, rater) explains 56.2% of the total variance. This is the highest value among the main components, and having a high difference among students is a desired and expected situation. A variance of 0% for the items indicates that the difficulty levels are the same. Likewise, the variance value between raters is close to zero with 2.7%, and it is seen that it explains a very small part of the total variance. This is also a desirable situation because a value close to zero can be interpreted as having a very small variation between rater scores. The percentage of the variance related to the individual-item (inxit) interaction within the total variance is 17%. This situation can be explained as the difficulty level of the items varies according to the individuals. Another component is the variance regarding the individual-rater (inxr) interaction. It is desirable to have this value by 0%. It shows that the scores of the individuals do not change from rater to rater. The rate of change between raters in the scores given to the items is 1.9%. Finally, the (inxitxr,e) individual-item-rater joint effect value is 22.2%. This value being higher than expected suggests that there is an error source between the individual-item-rater that is not systematic, which does not systematically affect the scores of the raters. This error can be expressed as an error that cannot be controlled in research. As a result of the analyses, it was found that the values of coefficients are quite high considering that G coefficient value calculated for 5 items, 4 raters, and 38 students being 0.89, and the value of the Phi coefficient being 0.88.

Table 9 shows the values for the G ($E\rho^2$) and Phi (Φ) coefficients estimated for the decision D study scenarios arranged according to the cases where the numbers of raters are 2, 3, 4, 5 and 6 created by increasing or decreasing the number of three raters.

Table 9.
G and Phi Coefficients According to the Scenarios Made by Increasing and Decreasing the Number of Items and Raters with D Studies of the inxitxr Pattern

	Number of Items							
Number of Raters	2		3		5		7	
	$E\rho^2$	Φ	$E\rho^2$	Φ	$E\rho^2$	Φ	$E\rho^2$	Φ
2	0,80	0,78	0,82	0,80	0,84	0,82	0,85	0,83
3	0,86	0,84	0,87	0,86	**0,89**	**0,88**	0,89	0,88
5	0,91	0,90	0,92	0,91	0,93	0,92	0,93	0,92
7	0,93	0,93	0,94	0,93	0,95	0,94	0,95	0,94

Table 9 presents the G and Phi coefficients obtained by taking the number of individuals as 38 and changing the number of items and raters. When the data are analyzed, it is seen that when the number of items is kept constant and the number of raters is increased, the amount of increase in the coefficients obtained is more than the amount of increase in the coefficients obtained when the number of raters is kept constant and the number of items is increased. In the inxitxr design, it was found that the G coefficient obtained by scoring 38 individuals in 5 items by three raters was 0.89 and the Phi coefficient was 0.88. It is possible to increase the coefficients by increasing the number of raters, but considering the usefulness of a measurement tool, it can be said that increasing the number of raters is not economical in terms of labor and time.

DISCUSSION, CONCLUSION and SUGGESTIONS

Many of the most popular methods of assessing creativity rely heavily on human judgment. Generally, professionals (for example, artists) or other experts are asked to judge creative work (Türkman, 2018). In assessing creativity, expert assessments are assumed to be the best possible assessment (Kaufman et al., 2008). In this study, creativity was assessed by experts. This assessment benefited from the possibilities of rubrics that were not made randomly.

Rubrics, one of the effective measurement tools, have been used in performance-based evaluations since 1990. When the studies on this subject are examined, it can be seen that rubrics are used in many stages of the education (instructional design, evaluation, eliminating learning deficiencies, etc.) (Aktaş & Alıcı, 2018; Arter, 2002; Brooks ve Miller, 2006; Çelik, et al, 2014; Çıralı Sarıca & Koçak Usluel Dunbar, 2016; Hall ve Salmon, 2003; Moskal ve Leydens, 2000; Oaklef, 2009; Öztürk & Güdek, 2016; Parlak & Doğan, 2014; Şen & Karagül, 2020; Yılmaz ve İnceağaç, 2017; Wolf ve Steven, 2007). Additionally, creativity is a high-level mental process that includes skills such as flexibility and originality. Rubrics are also used to evaluate high-level mental processes (Doğan et al., 2017). In this study, a multidimensional phenomenon such as creativity is discussed within the scope of performance-based evaluation. Therefore, an analytical rubric was developed in order to evaluate the artistic creativity that emerged through the designs obtained by secondary school students using mobile design applications and the validity and reliability evidence of the scoring key were presented. Developing students' creative thinking skills and their relationship with technology is among the main objectives of both 21st century skills and education. The CLAR we have developed is intended to measure these skills. As a result of the analysis, it was concluded that CLAR (Creativity level analytical rubric) is a reliable and valid measurement tool.

In CLAR, there are five criteria to measure the artistic creativity that emerges through the products obtained by students using mobile design applications. Findings based on expert opinion have shown that the criteria and the explanations regarding the criteria in the CLAR are sufficient and appropriate to evaluate the artistic creativity of the students. The consistency between the opinions of the experts has formed the proof of the content validity of the developed CLAR. Inter-rater agreement and reliability coefficient were calculated through the Kendall's W coefficient. The values obtained by scoring the CRAL showed that the inter-rater reliability in the criteria of the developed CRAL was high. According to these findings, it can be said that CRAL is not sufficient enough to completely eliminate the disagreements between raters. In respect to this, Bıkmaz Bilgen and Doğan (2017), stated that analytical rubric gives more objective results among raters, enables more consistent scoring, therefore it is more reliable.

According to the results of the research conducted by Türkman (2018), having more referees ensures higher inter-rater reliability values. Regarding this issue, Arslan Mancar (2019) stated in her study that by increasing the number of sub-dimensions and raters in the analytical rubric, the reliability of the

77

ratings made will increase. Increasing the number of experts is critical because researchers working with few experts to assess the creativity of products have found that using a small number of experts results in low inter-rater reliability values (Kaufman & Sternberg, 2010 as cited in Türkman, 2018). This study also found a similar result in D study.

Researchers studying on creativity share the same opinion that this issue is a complex phenomenon that has no clear definition and boundaries. Regardless of its scope, the type of creativity studied has its own limitations in assessment of creativity (Kanlı, 2020). When the literature was reviewed, creativity was generally considered as a psychological construct and evaluation scales were developed in this context. An accepted scale for evaluating artistic creativity could not be found. It is accepted in the literature that the best way to measure artistic creativity is the opinion of the experts (Kaufman et al., 2008). It is a well-known fact that there is a need for generally accepted measurement tools or methods in this field. For future research, it is recommended to increase the number of studies on developing tools that can measure artistic creativity at all levels of education. CLAR, developed in this research, was prepared to reflect the technical features of mobile design applications. In case of using another digital tool or application, it is recommended for users to use CLAR by rearranging it according to the characteristics of the tools to be used and testing it. Additionally, more reliable results can be obtained by increasing the number of experts evaluating CLAR. In this study, a total of 16 behaviors expected from students were expressed under 5 main headings. In a different study, these behaviors can also be scored separately and compared with the results of this study.

REFERENCES

Aktaş, M. & Alıcı, D. (2018). Analytical rubric development for story writing: validity and reliability study. mersin university journal of the faculty of education, 2018; 14(2): 597-610" Research Article / DOI: 10.17860/mersinefd.424198

Arieti, S. (2016). *Büyülü bireşim yaratıcılık.* [Magical synthesis of creativity]. Kurgu Kültür Merkezi Yayınları.

Arter, J. (2002). Rubrics, scoring guides, and performance criteria. C. Boston (Yay. haz.). Understanding scoring rubrics a guide for teachers. Washington: Office of Educational Research and Improvement.

Arslan Mancar, S. (2019). *Performansa dayalı durum belirlemede puanlayıcılar arası güvenirlik tekniklerinin karşılaştırılması*. [The comparison of inter rater reliability estimating techniques in performance based assessment] [Master's thesis, Instıtute of Educatıon Scıences]. Master's Thesis. Educational Sciences Institute, Ankara University.

Bacanlı H., Dombaycı M. A., Demir M., Tarhan S. (2011). Quadruple thinking: creative thinking. *Procedia-Social and Behavioral Sciences, 12, 536-544.*

Bıkmaz Bilgen, Ö. & Doğan, N. (2017). Puanlayıcılar arası güvenirlik belirleme tekniklerinin karşılaştırılması [The comparison of interrater reliability estimating techniques]. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi.8(1), 63-78.*

Bolden, B., DeLuca, C., Kekkonen, T., Roy S., & Wearing J. (2020). Assessment of creativity in k-12 education: a scoping review. *Review of Education. 8(2), 343-376.*

Black, J. & Browning, K (2011). Creativity in digital art education teaching practices. *Art Education, 64(5), 19-34.*

Brown University (2020). *Rubric scales*, Retrieved from https://www.brown.edu/sheridan/teaching-learning-resources/teaching-resources/course-design/classroom-assessment/grading-criteria/rubrics-scales at 12.12.2020.

Campell, A., (2005). Application of ICT and rubrics to the assessment process where professional judgement is involved: the features of an e-marking tool. *Assessment and Evaluation in Higher Education. 529-537.*

Crocker, L. M. & Algina, L. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston.

Conrad, S. D. (1990). Toward a phenomenological analysis of artistic creativity. *Journal of Phenomenological Psychology, 21(2), 103-120.*

Çelik, T., Demirgüneş, S. & Baştuğ, M. (2014). Metin oluşturma ve değerlendirme çalışmalarına yönelik "okur dostu" metin değerlendirme rubriği geliştirme çalışması. *Uluslararası Avrasya Sosyal Bilimler Dergisi, Cilt: 5, Sayı: 14, ss: (65-82).*

Çıralı Sarıca, H. & Koçak Usluel, Y. (2016). Eğitsel bağlamda dijital hikâye anlatımı: bir rubrik geliştirme çalışması . *Eğitim Teknolojisi Kuram ve Uygulama , 6 (2) , 65-84 . DOI: 10.17943/etku.12600*

Dunbar, N. E., Brooks, C. F. & Miller, T. K. (2006). Oral communication skills in higher education: using a performance-based evaluation rubric to assess communication skills. *Innovative Higher Education, 31(2), 2006, 115-128.*

Fox, J. E. & Schirrmacher, R. (2014). *Çocuklarda sanat ve yaratıcılığın gelişimi [The development of art and creativity in children]*. (N. Aral, G. Duman, Çev.). Ankara:Nobel.

Goodrich, H., (1997). Understanding rubrics. *educational leadership 54(4), 14-17.*

Guilford, J. P. (1967). Creativity: yesterday, today and tomorrow. *The Journal of Creative Behavior, 1 (1), 3-14.*

Gürkan, B. & Dolapçıoğlu, S. (2020). Sosyal bilgiler dersinde estetik yaratıcılık öğretim etkinlikleriyle yaratıcı düşünme becerilerinin geliştirilmesi [Development of creative thinking skills with aesthetic creativity teaching activities in social studies course]. *Eğitim ve Bilim, 45 (202), 51-77.*

Hall, E. K. & Salmon, S. J. (2003). Chocolate chip cookies and rubrics helping students understand rubrics in inclusive settings. *Teaching Exceptional Children, 35(4), 8-11.*

Henriksen, D., Creely, E., Henderson, M. & Mishra, P. (2021). *Creativity and technology in teaching and learning: a literature review of the uneasy space of implementation*. Educational Technology Research and Development.

Higgins, M. & Morgan, J. (2000). The role of creativity in planning: the 'creative practitioner'. *Planning Practice & Research, 15, (1-2), 117-127.*

Jersild, A. (1983). *Çocuk psikolojisi [Child psychology]*. (G. Günce, Çev.). Ankara Üniversitesi.

Johnson, B., & Christensen, L. (2004). Educational research: Quantitative, qualitative, and mixed approaches. NY: Pearson/Allyn & Bacon.

Kanlı, E. (2020). Assessment of creativity: theories and methods [Online First], *IntechOpen,* DOI: 10.5772/intechopen.93971. Retrieved from https://www.intechopen.com/online-first/assessment-of-creativity-theories-and-methods at 12.12.2020.

Kaufman, J. C., Baer, J., Cole, J. C., & Sexton, J. D. (2008). A comparison of expert and nonexpert raters using the consensual assessment technique. *Creativity Research Journal, 20(2), 171-178.*

Kayakçı Danışmaz, Z., & Kurnaz, Adıbatmaz, B. F. (2020). Çocuk gelişimi alanında dereceli puanlama hazırlama: deneysel bir uygulama [Creating rubric on the field of child development: an experimantal practice]. *Çocuk ve Gelişim Dergisi 3(5), 12-28.*

Kaygın, B. & Çetinkaya, Ç. (2015). Yaratıcılığın değerlendirmesinde yeni yaklaşımlar [new approaches in the evaluation of creativity]. *Üstün Zekâlılar Eğitimi ve Yaratıcılık Dergisi 2(1), 1-11.*

Kırışoğlu, O. T. (2005). *Sanatta eğitim görmek öğrenmek yaratmak [Education in art, seeing, learning, creating]*. Ankara: Pegem A.

Kutlu, Ö., Doğan, C. D. & Karakaya, İ. (2017). *Öğrenci başarısın belirlenmesi (performansa ve portfolyoya dayalı durum belirleme) [Determining of student success (performance and portfolio based assessment]*. Ankara: Pegem A.

Lindström, L. (2006). "Creativity: what is it? can you assess it? can it be taught? *The International Journal of Art Design Education, 25(1), 53-66.*

Lowenfeld, V. & Brittain, L. (1964). *Creative and mental growth.* McMillan

May, R. (2007). *Yaratma cesareti [Courage to create].* Metis.

Miles, M. B. & Huberman, A. M. (1994). *Qualitative data analysis*. Thousand.

Moskal, B. M. & Leydens, J. A. (2000). Scoring rubric development: validity and reliability. *Practical Assessment, Research & Evaluation, 7(10).*

Oaklef, M. (2009). Using rubrics to assess information literacy: An examination of methodology and ınterrater reliability. *Journal of the American Society for Information Science and Technology, 60(5), 969-983.*

Onur, D. & Zorlu, T. (2017). Yaratıcılık kavramı ile ilişkili kuramsal yaklaşımlar [Theoretical approaches towards the concept of creativity]. *İnsan ve Toplum Bilimleri Araştırmaları Dergisi, 6(3), 1535-1552.*

Özaşkın, A. G. & Bacanak, A. (2016). Eğitimde yaratıcılık çalışmaları: neler biliyoruz? [Creativity studies in education: what do we know?]. *Eğitim ve Öğretim Araştırmaları Dergisi Journal of Research in Education and Teaching 2016 Cilt:5 Özel Sayı, pp.212-226.*

Özdamar, K. (1999). *Paket programlar ile istatistiksel veri analizi 1*. Kaan Kitabevi, Eskişehir

Öztürk, D. & Güdek, B. (2016). Viyolonsel performans değerlendirmesine yönelik dereceli puanlama anahtarının (rubrik) geliştirilmesi [An essay concerning improving the graded scoring method (rubric) for rating the violoncello performance]. *Akademik Müzik Araştırmaları Dergisi, 2 (3),1-20.*

Parlak, B. & Doğan, N., (2014). Dereceli puanlama anahtarı ve puanlama anahtarından elde edilen puanların uyum düzeyleri [Comparison of answer key and scoring rubric for the evaluation of the student performances]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 29(2), 189-197.*

Pelowski, M., Leder, H., & Tinio, P. (2017). *Creativity in the Visual Arts. In J. Kaufman, V. Glăveanu, & J. Baer (Eds.), The cambridge handbook of creativity across domains (Cambridge handbooks in psychology, pp. 80-109).* Cambridge University Press.

Punch, F. K. (2014). *Sosyal araştırmalara giriş, nicel ve nitel yaklaşımlar [Introduction to social research, quantitative and qualitative approach].* Siyasal.

Rhodes, M. (1961). An analysis of creativity. *The Phi Delta Kappan, 42 (7), 305-310.*

San, İ. (1979). *Sanatta yaratma ve çocukta yaratıcılık [Creation in art and creativity in children*]. Ankara: İş Bankası Kültür.

San, İ. (2004). *Sanat ve eğitim [Art and education*]. İstanbul: Ütopya.

Şen, E. & Karagül, S. (2020). Kurmaca metinleri değerlendirmeye yönelik rubrik geliştirilmesi: geçerlik ve güvenirlik çalışması, *Turkish Studies - Education, 15(3), 1951-1961.* https://dx.doi.org/10.29228/TurkishStudies.41821

Tavşancıl, E. & Aslan E. (2001). *İçerik analizi ve uygulama örnekler [Content analysis and application examples].* Ankara: Epsilon Yayınları.

Taylor, G. R., (2003). *Informal classroom assessment strategies for teachers*. The Scarecrow Pres. Lanham, Maryland and Oxford.

Taylor, L. & Parsons, J. (2011). Improving student engagement. *Current Issues in Education, 14(1), 1-33.*

Von Eye, A., & Mun, E. Y. (2005). *Analyzing rater agreement: manifest variable methods*. Lawrence Erlbaum Associates.

Türk Dil Kurumu (2020). *Yaratıcılık*. Retriewed from https://sozluk.gov.tr/ at 10.12.2020

Türkman, B. (2018). Açık yönergeler ve yaratıcılığın değerlendirilmesi. [Explicit Instructions & Creativity Judgment]. *Turkish Journal of Giftedness and Education. 8 (1), 33-46.*

Wolf, K. & Stevens, E. (2007). The role of rubrics in advancing and assessing student learning. *The Journal of Effective Teaching, 7(1), 3-14.*

Yalçın, S. (2018). *21. yüzyıl becerileri ve bu becerilerin ölçülmesinde kullanılan araçlar ve yaklaşımlar. [21st century skills and tools and approaches that are used to measure these skills].* Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi. 51(1), 183-201.

Yavuzer, H.S. (1996). *Yaratıcılık [Creativity].* Boğaziçi Üniversitesi Yayınları.

Yılmaz, M. & İnceağaç, M. (2017). Görsel sanatlar eğitimi uygulamalarında edebiyat ürünlerinin etkisini belirlemeye yönelik dereceli puanlama anahtarının (rubrik) geliştirilmesi. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi, 19 (3) , 1-16 . DOI: 10.17556/erziefd.341456.*

Zimmerman, E. (2009). Reconceptualizing the role of creativity in art education theory and practice. *Studies in Art Education A Journal of Issues and Research, 50(4), 382-399.*

Zimmerman, E. (2010). Creativity and art education: a personal journey in four acts. *Art Education, 63(5), 84-9.*

APPENDIX 1. Creativity Level Analytical Rubric (CLAR)

Ölçütler	Zayıf (Poor) 0	Minimal (Minimal) 1	Yeterli (Sufficient) 2	Ortalamanın üzerinde (Above Average) 3	Mükemmel (Excellent) 4	Puan
Bilgi ve Algılama Kapasitesi	Mobil tasarım uygulamaları konusunda hiç bilgisi yoktur. Mobil tasarım uygulamaları kullanarak tasarım oluşturmaya yönelik teknik ve yöntemleri uygulamaz.	Mobil tasarım uygulamaları konusunda yeterli bilgiye sahip değildir. Mobil tasarım uygulamaları kullanarak tasarım oluşturmaya yönelik teknik ve yöntemleri yeterli derecede uygulayamaz.	Mobil tasarım uygulamalarından çok azını bilir. Mobil tasarım uygulamalarını kullanarak tasarım oluşturmaya yönelik teknik ve yöntemlerden çok azını bilir.	Mobil tasarım uygulamalarından bazılarını bilir. Mobil tasarım uygulamalarını kullanarak tasarım oluşturmaya yönelik teknik ve yöntemlerden birazını bilir.	Mobil tasarım uygulamalarını bilir. Mobil tasarım uygulamalarını kullanarak tasarım oluşturmaya yönelik teknik ve yöntemleri bilir.	
Dijital araçları kullanabilme	Tasarımın anlamını güçlendirmek için materyal ve tekniği analiz etmez. Tekniğe ilişkin problemlere çözüm üretmez ve çözüme yönelik risk almaz. Mobil tasarım uygulamalarını kullanmaz. Mobil tasarım uygulamaları ile özgün tasarımlar yapmaz.	Tasarımın anlamını güçlendirmek için materyal ve tekniği analiz etmede yeterli değildir. Tekniğe ilişkin problemlere çözüm üretmede zorlanmakta ve çözüme yönelik risk almamaktadır. Mobil tasarım uygulamaları kullanamaz ve uygulamada sıkıntılar yaşamaktadır. Mobil tasarım uygulamaları ile özgün tasarımlar yapma konusunda yeterli değildir.	Tasarımın anlamını güçlendirmek için materyal ve tekniği nadiren analiz edip kullanır. Tekniğe ilişkin problemlere zorlukla çözüm üretir, nadiren risk alır. Mobil tasarım uygulamalarını nadiren kullanır ve uygulamada sıkıntılar yaşamaktadır. Mobil tasarım uygulamaları ile kısmen özgün tasarımlar yapar.	Tasarımın anlamını güçlendirmek için materyal ve tekniği analiz edip kullanır. Tekniğe ilişkin problemlere bazen çözüm üretir, bazen risk alır. Mobil tasarım uygulamalarını kullanır ve uygulamada nadiren sıkıntılar yaşamaktadır. Mobil tasarım uygulamaları ile özgün tasarımlar yapar.	Tasarımın anlamını güçlendirmek için materyal ve tekniği analiz edip etkili bir biçimde kullanır. Tekniğe ilişkin problemlere çözüm üretir, gerektiğinde risk alır. Mobil tasarım uygulamalarını kullanır ve uygulamada sorun yaşamaz. Mobil tasarım uygulamaları ile özgün tasarımlar yapar.	
Tasarım	Görsel tasarım öge ve ilkeleri ile duyuşsal nitelikleri bir arada kullanamaz. Fikirlerini karşı tarafa yansıtacak şekilde geliştirip, görsel bir düzen oluşturamaz. Görsel bütünlüğü sağlamak için	Görsel tasarım öge ve ilkeleri ile duyuşsal nitelikleri bir arada kullanma konusunda sıkıntı yaşar. Fikirlerini karşı tarafa yansıtacak şekilde geliştirip, görsel bir düzen oluşturma konusunda yeterli değildir.	Görsel tasarım öge ve ilkeleri ile duyuşsal nitelikleri bir arada kullanır. Fikirlerini karşı tarafa çok az yansıtacak şekilde geliştirip, görsel bir düzen oluşturur. Görsel bütünlüğü sağlamak için çok az deneme yapar.	Görsel tasarım öge ve ilkeleri ile duyuşsal nitelikleri etkili bir biçimde kullanır. Fikirlerini karşı tarafa yansıtacak şekilde geliştirip, görsel bir düzen oluşturur. Görsel bütünlüğü sağlamak için	Görsel tasarım öge ve ilkeleri ile duyuşsal nitelikleri alışılmadık ve etkili bir biçimde kullanır. Fikirlerini karşı tarafa yansıtacak şekilde geliştirip, özgün görsel bir düzen oluşturur.	

82

|						
	deneme yapmaz ve mobil tasarım uygulamaları ile tasarım oluşturamaz.	Görsel bütünlüğü sağlamak için deneme yapmada ve mobil tasarım uygulamaları ile tasarım oluşturma yeterli değildir.	Mobil tasarım uygulamaları ile orta düzeyde tasarımlar oluşturur.	denemeler yapar. Mobil tasarım uygulamaları ile tasarımlar oluşturur.	Görsel bütünlüğü sağlamak için sürekli denemeler yapar. Mobil tasarım uygulamaları ile özgün, sıra dışı tasarımlar oluşturur.	
Araştırma çalışma yaklaşımı	Bilgi edinmede dışa dönük, meraklı ve araştırmacı tavır sergilemez. Kendisine verilen konu üzerine ayrıntılı düşünmez, farklı çözümler aramaz ve alternatif fikirler geliştirmez. Fikirleri yeni yöntemlerle birleştirmez. Alışılmışın dışında çözümler bulmak için risk almaz. Öğrenme sürecine ilişkin sorumluluklarını kendisine verilen süre içerisinde eksiksiz yerine getirmez. Arkadaşlarının çalışmalarına katkıda bulunmaz.	Bilgi edinmede dışa dönük, meraklı ve araştırmacı tavır sergilemeye çalışmaz. Kendisine verilen konu üzerine ayrıntılı düşünme, farklı çözümler arama ve alternatif fikirler geliştirmede istekli değildir. Fikirleri yeni yöntemlerle birleştiremez. Alışılmışın dışında çözümler bulmak için gerektiğinde risk alamaz. Öğrenme sürecine ilişkin sorumluluklarını kendisine verilen süre içerisinde eksiksiz yerine getiremez. Zorluklar karşında çekimser kalır. Arkadaşlarının çalışmalarına katkıda bulunamaz.	Bilgi edinmede nadiren meraklı ve araştırmacı tavır sergiler. Kendisine verilen konu üzerine ayrıntılı düşünme, sorunlara farklı çözümler arama, alternatif fikirler geliştirmede nadiren isteklidir. Fikirleri yeni yöntemlerle birleştirme, alışılmışın dışında çözümler bulmada nadiren risk alır. Öğrenme sürecine ilişkin sorumluluklarını kendisine verilen süre içerisinde bazen yerine getirir. Zorluklar karşında sıklıkla tereddüt edebilir. Arkadaşlarının çalışmalarına nadiren katkıda bulunur.	Bilgi edinmede bazen dışa dönük, meraklı ve araştırmacı tavır sergiler. Kendisine verilen konu üzerine düşünür. Bazen farklı çözümler arar ve alternatif fikirler geliştirir. Fikirleri bazen yeni yöntemlerle birleştirir. Alışılmışın dışında çözümler bulmak için bazen risk alır. Öğrenme sürecine ilişkin sorumluluklarını kendisine verilen süre içerisinde yerine getirir. Zorluklar karşında tereddüt edebilir. Arkadaşlarının çalışmalarına bazen katkıda bulunur.	Bilgi edinmede dışa dönük, meraklı ve araştırmacı tavır sergiler. Kendisine verilen konu üzerine ayrıntılı düşünür. Farklı çözümler arar, alternatif fikirler geliştirir. Fikirleri yeni yöntemlerle birleştirir. Alışılmışın dışında çözümler bulmak için gerektiğinde risk alır. Öğrenme sürecine ilişkin sorumluluklarını kendisine verilen süre içerisinde eksiksiz yerine getirir. Zorluklar karşında yılmaz. Arkadaşlarının çalışmalarına katkıda bulunur.	
Öz değerlendirme	Kendi tasarımı ve teknik becerisi ile ilgili çok yönlü eleştirel bir tavır sergilemez. Öğrenme eksikliklerini çözümlemez.	Kendi tasarımı ve teknik becerisi ile ilgili çok yönlü eleştirel bir tavır sergileyemez. Öğrenme eksikliklerini yeterince çözümleyemez.	Kendi tasarımını /çalışmasını ve teknik becerisini birkaç ölçüt çerçevesinde değerlendirebilir. Öğrenme eksikliklerini çözümlemekte zorlanır.	Kendi tasarımını /çalışmasını ve teknik becerisini belli ölçütler çerçevesinde değerlendirebilir. Öğrenme eksiklikleri çerçevesinde bazı ihtiyaçlarının farkındadır.	Kendi tasarımı ve teknik becerisi ile ilgili çok yönlü eleştirel bir tavır sergiler. Öğrenme eksiklikleri çerçevesinde ihtiyaçlarının farkındadır.	

APPENDIX 2. Examples of student work
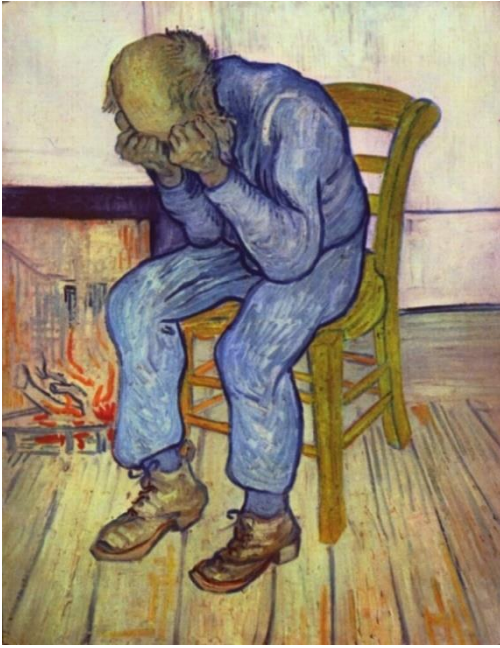
Image 1_A. *"At Eternity's Gate"* Vincent van Gogh

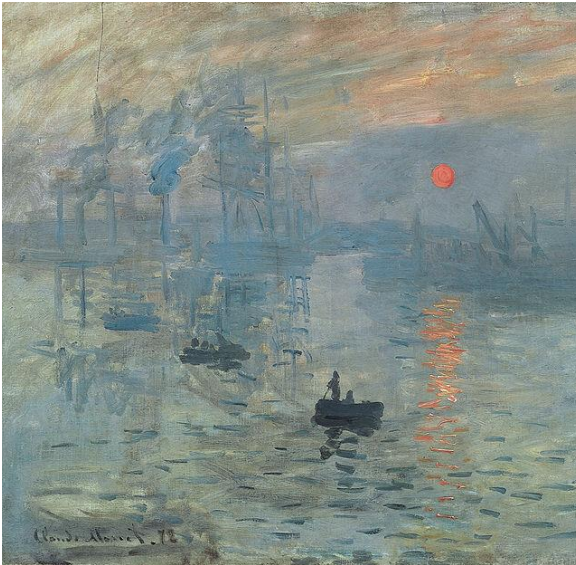**Image 1_B**. Work of student number 17

Image 2_A. *"Impression, soleil levant"* Claude Monet

Image 2_B. Work of student number 30