

Test of Figural Creativity – version for children and adults: an investigation of reliability by different methods

Test de creatividad figural – versión para adolescentes y adultos: investigación de la fiabilidad mediante diferentes métodos

Tatiana de Cassia Nakano , Isabella Wonsik Cano , Ana Luiza Pelaquim Silva ,
Isabel Cristina Camelo de Abreu , Laura Gandia Romancini 
Pontifical Catholic University of Campinas

Abstract

Creative thinking has been valued in different contexts and is considered an essential skill in the 21st century. Despite this emphasis on the construct, the instruments for its assessment remain restricted in Brazil. Due to this scenario, this article presents the results of two studies aimed at investigating the reliability of the Figural Creativity Test for adolescents and adults. In study 1, the temporal stability is estimated in a sample composed by 179 Brazilian participants aged 14 to 80 years. Correlation between test and retest was $r = .665$. In study 2, evaluator reliability, two judges evaluated the responses of 97 participants and the correlation was $r = .88$. Favorable results motivate the continuation of the studies until the instrument can be available for professional use.

Keywords: reliability, psychometric qualities, psychological assessment

Resumen

La creatividad es una característica que ha sido valorada en diferentes contextos y considerada una de las habilidades esenciales en el siglo XXI. A pesar de este énfasis en el constructo, los instrumentos para su evaluación aún son escasos en Brasil. Ante este escenario, este artículo presenta los resultados de dos estudios orientados a investigar las cualidades psicométricas de un test en proceso de desarrollo: el Test de Creatividad Figural para adolescentes y adultos. En el estudio 1, se estimó la estabilidad temporal en una muestra compuesta por 179 participantes brasileños de 14 a 80 años. La correlación entre el test y el retest fue de $r = .665$. En el estudio 2, confiabilidad del evaluador, dos jueces evaluaron las respuestas de 97 participantes y la correlación fue de $r = .88$. Los resultados favorables motivan la continuación de los estudios hasta que el instrumento pueda estar disponible para uso profesional.

Palabras clave: confiabilidad, cualidades psicométricas, evaluación psicológica

Cómo citar: Nakano, T. C., Cano, I.W., Pelaquim Silva, de Abreu, I. C. C & Romancini, L.G. (2025). Test of Figural Creativity – version for children and adults: an investigation of reliability by different methods. *Escritos de Psicología – Psychological Writings*, 18(1), 14-25. <https://doi.org/10.24310/escpsi.18.1.2025.20111>

Corresponding author: Tatiana de Cassia Nakano. Pontifícia Universidade Católica de Campinas - Programa de pós-graduação em Psicologia. Avenida John Boyd Dunlop, s/n. Jardim Ipaussurama. Campinas, SP. Brasil. E-mail: tatiananakano@hotmail.com

Editado por:

· Rafael Alarcón. Universidad de Málaga. España

Revisado por:

· Calixto Tapullima-Mori. Universidad Peruana Unión. Perú
· Anónimo



FACULTAD DE
PSICOLOGÍA Y LOGOPEDIA
Universidad de Málaga

Escritos de Psicología – Psychological Writings es una revista *Diamond Open Access*
publicada por la Facultad de Psicología y Logopedia de la Universidad de Málaga
ISSN: 1989-3809 | @PsyWritings | revistas.uma.es/index.php/escpsi



Introduction

Scientific literature has identified creativity as an essential skill for the 21st century (Kupers et al., 2019; Runco, 2016). As a result, it has become a focal point for research, education, arts, industry, and public policy (Green et al., 2022). With the current context, which is characterized by complexity and uncertainty, as well as the rapid rate at which things and ideas become obsolete (Lassig, 2019; Ziegler, 2009), creativity has emerged as an important tool. In response to this, efforts have been made to develop creativity in various contexts (Zhu et al., 2019).

Several benefits of creative thinking are demonstrated by Yahn and Kaufman (2016): at a global level, it can facilitate countries' economic success; at work, it can improve employee satisfaction; and at an individual level, it can lead to improvements in physical health, positive moods, and the ability to cope with trauma more effectively. Besides promoting economic growth, Plucker et al. (2018) also highlight the importance of creativity as a sign of success in life and as a means of maintaining mental health (Oliveira et al., 2016).

Several benefits of creative thinking are demonstrated by Yahn and Kaufman (2016): at a global level, it can facilitate countries' economic success; at work, it can improve employee satisfaction; and at an individual level, it can lead to improvements in physical health, positive moods, and the ability to cope with trauma more effectively. Besides promoting economic growth, Plucker et al. (2018) also highlight the importance of creativity as a sign of success in life and as a means of maintaining mental health (Oliveira et al., 2016).

As a multidimensional and complex construct, creativity is understood as a potential inherent in everyone (Berg et al., 2020). It is characterized by an interaction between aptitude, process, and environment that produces a socially useful outcome (Plucker et al., 2004). This characteristic may manifest in an intense manner because of social and personal factors. There is an inherent creative ability present in every individual, which represents the type of creativity required to solve everyday problems requiring a combination of imaginative solutions (Helfand et al., 2017), value, authenticity, and intentionality (Runco, 2022). Therefore, creative ideas are often evaluated in terms of their novelty and usefulness (O'Tolle & Horvát, 2024).

The scientific literature provides an overview of the main benefits of creativity assessment. The Table 1 below summarizes the findings, and it is important to note that they are limited to certain areas, even though creativity may be beneficial in other areas, including social ones.

Table 1
Creativity Assessment Benefits

Aspect	Main Benefits
Psychological Assessment	Contribute to the understanding of creativity's nature and development Identifying whether it is a unidimensional or multidimensional construct Determine whether creativity is a general or specific domain characteristic Investigate the effectiveness of creative techniques in situations before and after participation in a creative development program Assist in dispelling myths that are still prevalent in common sense Develop a broader understanding of human potential Recognize other important skills in addition to intelligence
Education	Enhance teachers' understanding of creative students' characteristics Promote a creative environment in the classroom Identification of students with high creative potential Contribute to the inclusion of creativity in school curricula Incorporate creative practices into the classroom Increasing parents' and teachers' awareness of the importance of creativity in the classroom Provide information that can be used to plan programs to develop students' creativity
Clinic	Identify different levels of creativity Identify the individual, environmental, educational, and social factors that influence creativity Understanding how creativity develops throughout a person's lifetime Assist in selecting and implementing interventions that will enhance creativity Support the development and recognition of talent
Workplace	Predicting and understanding the productivity of individuals at work Analyze the characteristics of a creative product Provide guidance to creative individuals to help them realize their full potential Enhance company competitiveness Facilitate the development of new products

Source: Long and Wang (2022), Puccio and Murdock (1999), Nakano and Wechsler (2006), Treffinger (1995)

Due to its importance, the interest in measuring and identifying creativity has grown exponentially across a wide range of fields (Plucker, 2022). An assessment of creativity can assist researchers in discovering which domains or areas a person has the greatest potential in, what degree of creativity he or she exhibits, what their strengths and weaknesses are, and how to stimulate creativity in individuals with this profile (Treffinger et al., 2013). Several methods, instruments, and techniques have been proposed to assess creativity (Nakano, 2018).

The use of divergent thinking tasks, product-based evaluations, expert judgments, and self-report measures (Barbot et al., 2019), along with observations, biographical information, interviews, and peer feedback, may help to determine which personal characteristics are associated with creativity and which environments promote or hinder it. Such techniques enable a subject's performance to be compared with that of other individuals to estimate their creative potential (Lubart et al., 2013). As a part of a positive psychology perspective, researchers have found that the study of individual creativity impacts the guidance of children and adolescents with respect to educational and vocational decisions (Krumm et al., 2024).

Additionally, special education has been interested in recognizing creativity as a component of giftedness (Lee et al., 2021) or a characteristic of gifted students (Sierra et al., 2024), incorporating it into most theoretical models concerning this phenomenon (Acar et al., 2022). It is possible to mention, for example, several theoretical models of giftedness that incorporate creativity as part of their conceptions. These include Renzulli's Three Rings Model (Renzulli, 1986), Sternberg's Triarchic Theory of Intelligence (Sternberg, 1997), Wisdom, Intelligence, and Creativity Synthesized Model of Gifted Leadership (Sternberg, 2005), Tannenbaum's Star Model (1986), and Gagné's Differentiated Model of Giftedness and Talent (Gagné, 2004).

Individuals who demonstrate exceptional abilities and accomplishments in one or more domains, like intellectual, creative, social and physical, are commonly referred to as "gifted" (Jung, 2022). A high degree of creative potential is commonly referred to as a creative productive giftedness (Renzulli, 2005) that emphasize "the use and application of information (content) and thinking processes in an integrated, inductive, and real-problem-oriented manner, which allows students to be self-determined firsthand inquirers" (p. 63).

As part of the process of identifying this phenomenon, creativity measures are recommended as one of the criteria to be considered. This increases the possibility of identifying other types of giftedness, as well as cognitive and academic (Lee et al., 2024). Most states in the United States, for example, use creativity tests as part of their assessment of giftedness, with their importance emphasized at the same rate as performances and intelligence tests (Alabbasi et al., 2022). According to the authors, underrepresented populations, such as ethnic minorities or students with low academic achievement, can benefit from the adoption of creativity tests, so that their creative potential can be recognized.

A problem in Brazil involves the fact that, despite the literature review indicating the existence of different instruments for research on creativity (Nakano, 2018), only one, for the assessment of figural creativity in children, was available to professional use. Currently, there are gaps in tests that evaluate adolescents, adults, and the elderly. As a result, a significant amount of potential could be wasted in the country, notably because it is difficult to measure these individuals (Nakano et al., 2023). Consequently, several contexts have used tests that have not been subjected to studies evaluating their psychometric properties for use in Brazil (Rocha & Wechsler, 2018). In this sense, an accurate and reliable assessment is needed to support the development of individual creativity (Cropley et al., 2024).

It's important to recognize that creativity is a domain-specific concept and cannot be categorized into a single type (Sternberg et al., 2022). Thus, it can be expressed in various ways in music, drawing, language, body, visual, scientific, artistic (Nakano et al., 2023). Different aspects of creativity can be captured within these different areas of creative expression (Kim, 2011). This research examines creativity in drawings. This type of creativity was chosen due to a variety of advantages that have been demonstrated in the scientific literature.

Several advantages have been noted by De La Torre (1991) for using a figurative test to assess creativity: (1) there is no age limit for this test, so it can be applied to adults and children alike, ensuring that drawing is a common activity for most, (2) reduction of economic cost and time considering that the test can extract more than ten indicators of creativity using just one sheet, whereas few tests are able to extract so much information with so little material, (3) The test is short in duration, (4) respondents generally find the activity enjoyable, (5) offers the possibility of assessing a greater number of creative indicators than are usually included in other types of creativity assessment instruments (fluency, flexibility, elaboration, and originality), and (6) is independent of language and culture, and may even be used by people who are illiterate.

In face of this gap, the process of investigating the psychometric properties of an existing test, the Children's Figural Creativity Test (Nakano et al., 2011), was initiated with the intention of expanding the test's use to adolescents, adults, and the elderly. This process requires the conduct of a series of studies including validity evidence, reliability, and standardization.

Based on the observation of different advantages figural creativity presents over other forms of creative expression, it was chosen to evaluate this form of creativity (Kasirer et al., 2020). The main advantage the language-independent task (Becker & Cabeza, 2023), the same test can be used for a wide range of ages (including young children, those with verbal difficulties, reduced linguistic abilities, illiterates, and uneducated populations), a reduction in economic costs since different measures of creativity can be determined in a few activities, and the short duration of the test (20 minutes in the case presented here).

The Test for Figural Creativity for adolescents and adults (TCF-AA) is based on the Torrance Test of Creativity Thinking - Figural, one of the most used instruments for evaluating this construct due its applicability to a variety of genders, cultures, socioeconomic statuses, and linguistics (Bart et al., 2017). The test has been translated into more than 35 languages and it is the most widely used and referenced of all creativity tests (Kim, 2006). It's also used to identify gifted students (Acar et al., 2021), being the most frequently test employed to identify giftedness (Bart et al., 2017) as well their eligibility to participate in gifted programs (Acar et al., 2022).

Since reliability is a requirement to be evaluated during the test development process, the studies presented here investigate this psychometric quality as part of the test development process. In all fields of science, reliability is a critical issue, which can be measured in a variety of ways, but it has the same objective: to permit inference based on the measurement of individuals or groups (Revelle & Condon, 2019) and the consistency of scores in relation to measurement error (Peixoto & Ferreira-Rodrigues, 2019). It is important to note that the reliability value can vary from one sample to another depending on several factors, including the variability of scores in the sample (Sánchez-Meca et al., 2008).

Considering that creativity assessments are often rated by raters, the reliability of ratings is crucial (Terai et al., 2024). A review of the type of precision most used in studies on creativity assessment showed that different methods were found in the 84 studies reviewed by Long and Wang (2022): internal consistency (50%), correlations between different evaluators (27%), index of composite reliability (14%), and test-retest (4%). Throughout the studies presented here, this psychometric quality was examined by using two different methods: temporal stability and inter-scorer reliability. Results of these studies are presented separately below.

Study 1: Test-and-retest reliability

Test and retest reliability, also known as temporal stability, refers to how similar a test's results are when applied at different times. The researcher can use this procedure to estimate the correlation between scores derived from assessments conducted at different times to estimate the true score and the measurement error (Peixoto & Ferreira-Rodrigues, 2019).

In this type of precision, a sample is selected, and the same instrument is applied at different times, with a sufficient period to minimize memory effects. This method is appropriate for constructs that do not undergo substantial changes in a short period of time (Andrade & Valentini, 2018). In this way, we can indicate the extent to which results may be generalized over time (Nunes & Primi, 2010).

Method

Participants

The sample consisted of 179 Brazilian participants aged 14 to 80 years ($M = 40.22$; $SD = 19.27$). According to gender, 59 were men (32.94%) and 120 were women (67.04%). There were 25.00% of participants who had completed or were continuing their primary education, 25.00% who had secondary education, and 48.25% who had a higher education. Other participants did not provide this information.

The inclusion criteria used were: (1) present the age selected for the study, (2) obtain parental/guardian authorization for participation (by signing the free and informed consent form) and sign the consent form for those under 18 years of age. Participants over the age of 18 are required to voluntarily agree to participate, and they are required to respond to the instrument during two times. Exclusion criteria included (1) giving up participation in the research, (2) withdrawal of parental authorization, (3) withdrawal of consent by the minor under the age of 18, and (4) response to only one of the two application moments.

Instrument

Test of Figural Creativity for adolescents and adults – TCF-AA

It consists of two activities in a set of incomplete stimuli. The participant is instructed to complete a drawing using these stimuli. The first activity requires you to complete ten different stimuli, while the second requires you to make 30 drawings based on the same stimulus. Based on the drawings, trained raters evaluated 12 objective criteria that include: fluency (number of relevant ideas presented), flexibility (various types of ideas), elaboration (adding details to the basic drawing), originality (unique ideas), expression of emotion (expression of feelings in drawings or titles), fantasy (presence of imaginary beings from fairy tales or science fiction), movement (expressing movement in drawings or titles), unusual perspective (drawings made from unusual angles), internal perspective (internal view of objects in the form of transparency), context (creating an environment in which the drawing takes place), extension of limits (extending the stimulus before to close them) and expressive titles (expressing the drawing in ways that go beyond its obvious description).

Some studies have been conducted using the TCF-AA. According to the validity evidence derived from the internal structure, the 12 creative characteristics, assessed by the instrument, can be grouped into four factors (Nakano et al., 2023): Factor 1 (F1) - elaboration: includes the characteristics of the elaboration, the use of context, the movement, the internal perspective, and the expressive titles. The ability to see things from multiple perspectives, taking account of the details of an idea; Factor two (F2) - external aspects: includes the characteristics of unusual perspective, fan-

tasy, and the extension of limits. The ability to consider external aspects of the problem, especially the environment, to formulate a solution; Factor 3 (F3) - cognitive aspects - involves fluency, flexibility, and originality. It combines the characteristics considered cognitive of creativity to enable original and diverse responses; Factor 4 (F4) - Emotional aspects: comprised of the characteristics of expression of emotion. It is characterized by emotionally charged reactions that involve feelings, skills that are essential to the creation process, facilitating it. By identifying these factors, it is possible to identify the areas where an individual has the most creative potential and those where they must improve. In addition, the test offers the possibility of obtaining an overall creative factor (based on the total performance of the test). The reliability of the test is $\omega = .776$.

A second study aimed to generate a list of common responses to each stimulus and task, to guide the scoring of the originality characteristic in this new population (Nakano, Fusaro, et al., 2023). Finally, evidence of validity based on external criteria of the convergent type was investigated using the Torrance Test of Creativity Thinking - Figural as a comparison measure, and the results indicated positive and significant correlation ($r = .665$; $p < 0,001$) between the measures (Nakano, Batagin, et al., 2023).

Procedures

The study has a cross-sectional, descriptive, correlational and comparative design, based on convenience sample. Research Ethics Committee analyzed and approved this project. The tests was administered individually and collectively, and each participant was required to respond to the same instrument twice. A second application was conducted 15 days following the first application (retest).

It was ensured that the same evaluator corrected both the first and second test for each participant, thus eliminating the possibility of the "evaluator" variable having an impact on the results.

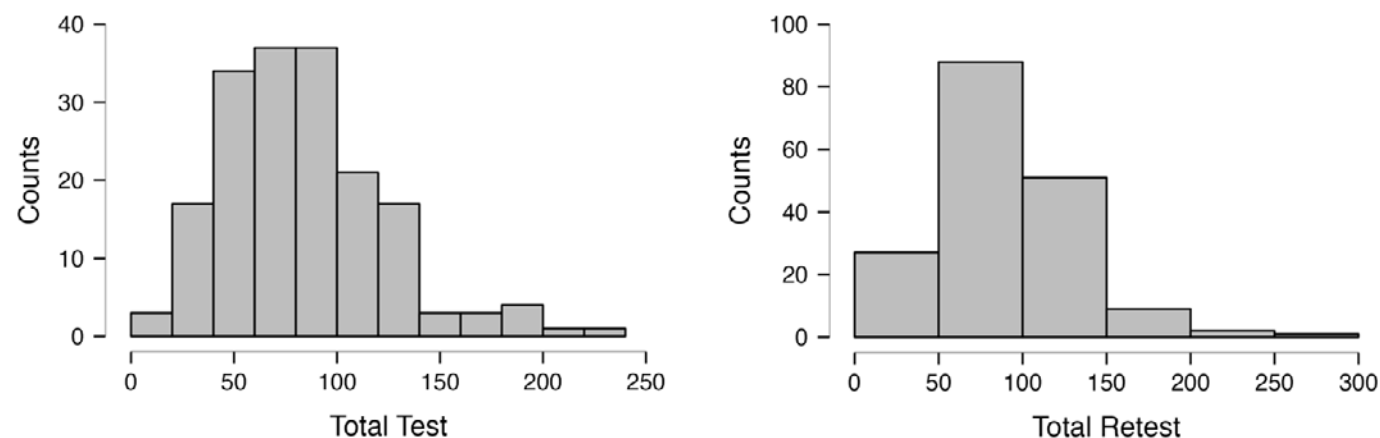
Data Analysis

Data were entered into a spreadsheet and analyzed using JASP. The participants' responses were corrected based on the presence or absence of each of the 12 creative characteristics as determined by the instrument, as described in the section entitled 'instrument'. Moreover, the scores were added together to yield a raw total across the four factors as well as a total score.

For verifying the normality of the sample, the Shapiro-Wilk test was applied. Some measurements showed a normal distribution, while others did not. This led to the decision to use non-parametric analyses in the present analysis (Spearman correlation). A significance level of 0.05 was adopted. A positive skew asymmetry has been observed in the Gaussian normal curve (Figure 1).

Figure 1

Distribution for total score in test and retest



A statistical analysis was conducted on the results obtained in test and retest total score and four factors. It was expected that the correlation value be equal or greater than $r = .60$, to meet the minimum requirements to assert the instrument's reliability.

To interpret the Spearman correlation, the proposal of Valentini and Iglesias (2021) was used: values between .10 to .30 are considered weak, between .31 to .50 are considered moderate, .51 to .99 is considered strong, and a perfect score of 1.00 is considered excellent. The significance level used in all analysis is $p \leq 0.05$.

Results

Detailed statistics in terms of four factors and the total score are presented in Table 2 for each measurement and test moment (test and retest). It can be observed that the retest presented higher means in factor 1, 2, 3 and the total score. In contrast, the means at the time of the test were higher in factor 4.

Table 2
Descriptive statistics for each TCF-AA Measurement at the Time of Test and Retest

Moment	Measurement	Mean	SD	Minimum	Maximum
Test	F1	36.72	28.02	0	184
	F2	4.35	5.94	0	59
	F3	41.84	18.54	0	100
	F4	0.81	4.15	0	4
	Total	83.64	39.29	14	223
Retest	F1	38.47	27.39	1	179
	F2	4.88	4.83	0	25
	F3	44.95	18.25	0	103
	F4	0.28	0.78	0	7
	Total	88.65	40.32	18	284

The Spearman correlations and effect sizes were calculated (Table 3). There was a significant, positive, and strong correlation between the total scores ($r = .77$).

Table 3
Spearman Correlations and Effect Sizes for TCF-AA Factor and Total Scores at Test and Retest

Variable		F1_Test		F2_Test		F3_Test		F4_Test		Total Test	
F1_Retest	Spearman's r	0.627	***	0.350	***	0.191	*	0.270	***	0.536	***
	p	< .001		< .001		.011		< .001		< .001	
	Fisher's Z	0.736		0.505		0.290		0.246		0.677	
F2_Retest	Spearman's r	0.466	***	0.668	***	0.426	***	0.192	*	0.589	***
	p	< .001		< .001		< .001		.010		< .001	
	Fisher's Z	0.366		0.808		0.436		0.059		0.567	
F3_Retest	Spearman's r	0.282	***	0.411	***	0.628	***	0.070		0.502	***
	p	< .001		< .001		< .001		.353		< .001	
	Fisher's Z	0.193		0.455		0.738		-0.027		0.514	
F4_Retest	Spearman's r	0.241	**	0.059		-.027		0.640	***	0.123	
	p	.001		.431		.719		.001		0.103	
	Fisher's Z	0.277		0.195		.070		0.347		0.229	
Total Retest	Spearman's r	0.590	***	0.513	***	0.473	***	0.225	**	0.774	***
	p	< .001		< .001		< .001		.003		< .001	
	Fisher's Z	0.599		0.676		0.551		0.123		0.818	

* $p < .05$; ** $p < .01$; *** $p < .001$

With respect to the factors, F1 showed a significant and positive correlation ($r = .62$), as well F2 ($r = .66$), F3 ($r = .62$) and F4 ($r = .64$). The magnitude can be interpreted as strong in these cases. Results confirm the test's reliability through temporal stability, obtaining the minimum value expected for this type of study ($r \geq .60$).

Study 2: Inter-score reliability

The purpose of the present study was to explore the reliability of the instrument using inter-scores judgement. During this type of reliability, the degree of agreement obtained between independent raters is estimated. In this method, two different raters score the test protocol separately for the same subjects, then the results of both corrections are compared (Anastasi & Urbina, 2000). In this type of study, the authors suggest that error sources can be related to the subjectivity of the evaluator or to the inconsistency of correction standards presented in manuals.

Inter-score reliability has been utilized to verify the accuracy of instruments that require some form of subjectivity when raters evaluate the responses (Ambiel & Carvalho, 2017). Specifically, the accuracy of evaluators was measured by the correlation between the scores given by independent judges. To confirm this type of reliability are expected correlations equal or higher than .60.

Method

Participants

Two female judges, 3rd year undergraduate psychology students, research assistants, independently assessed the responses of 97 participants. The sample, whose results were analyzed, ranged in age from 14 to 56 years old ($M = 20.63$ years; $SD = 7.41$ years), with 73.20% of them being female. Regarding education levels, 35.00% attended or completed secondary education, and 65.00% higher education, with 62.92% coming from the Brazilian southeast region and 37.10% from the northeast.

The inclusion criteria used for judges were: (1) sign the free and informed consent form, (2) having experience in correcting the test used, and (3) being a student in scientific initiation.

Exclusion criteria included (1) giving up participation in the research, and (2) not being able to correct all the tests requested.

Instruments

The same used in Study 1.

Procedures

The study has a cross-sectional, descriptive, correlational and comparative design, based on convenience samples. Each protocol was corrected independently by the two evaluators. Evaluator 1 was not aware, at any time, of the correction made by Evaluator 2. The inverse situation also was conducted.

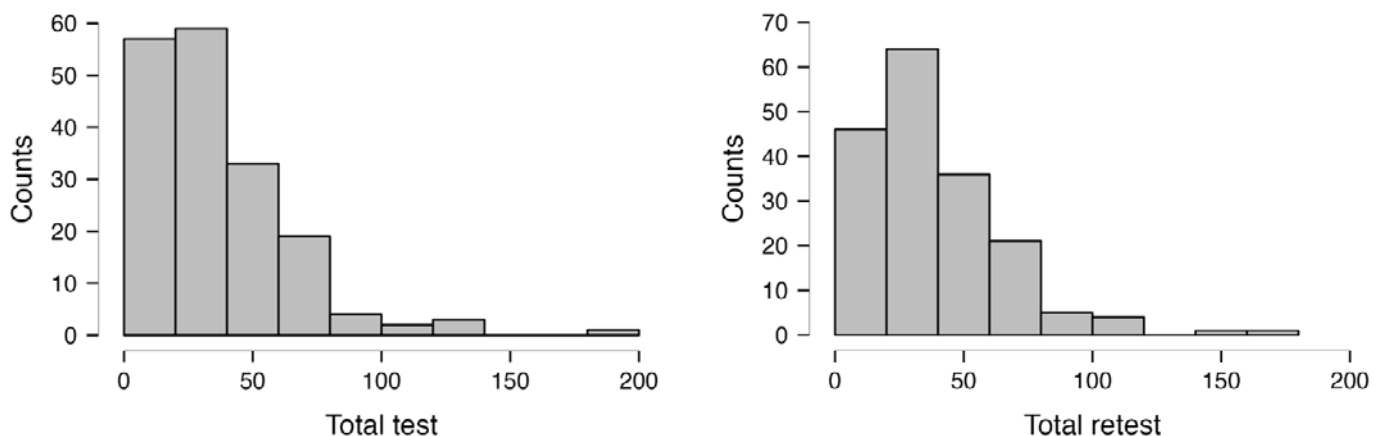
The correction process considers the presence or absence of 12 creative characteristics assessed by the test. The raw data of the four factors and the total score were calculated for each judge and produced the descriptive statistics table. The Spearman correlation coefficients between the two sets of scores was estimated.

Data Analysis

For verifying the normality of the sample, the Shapiro-Wilk test was applied. Some measurements showed a normal distribution, while others did not. This led to the decision to use non-parametric analyses in the present analysis (Spearman correlation). A positive skew asymmetry has been observed in the Gaussian normal curve (Figure 2). A significance level of 0.05 was adopted.

Figure 2

Distribution for total score in test and retest



Based on Valentini and Iglesias (2021) proposal, correlations between .10 and .30 are considered weak, between .31 and .50 are considered moderate, above .51 are considered strong, and 1.00 is considered perfect.

Results

In Table 4 we present the descriptive statistics for each judge, considering each of the factors and the total score. Afterwards, Spearman correlations were conducted. Table 5 presents the results, including the effect size.

Table 4

Descriptive Statistics for Judge and TCF-AA Factor

Measurement	Judge 1		Judge 2	
	M	SD	M	SD
Factor 1	35.20	18.53	39.51	20.03
Factor 2	6.29	3.78	5.06	3.30
Factor 3	43.42	13.67	43.53	13.93
Factor 4	0.47	1.01	1.42	6.74
Total	85.33	29.53	88.69	30.61

Table 5
Correlations Between Judges' Score

Variable		Factor1 J1		Factor2 J1		Factor3 J1		Factor4 J1		Total J1	
Factor1 J2	Spearman's <i>r</i>	0.834	***	0.450	***	0.367	***	0.268	**	0.771	***
	<i>p</i>	< .001		< .001		< .001		.008		< .001	
Factor2 J2	Spearman's <i>r</i>	0.434	***	0.744	***	0.517	***	0.092		0.570	***
	<i>p</i>	< .001		< .001		< .001		.370		< .001	
Factor3 J2	Spearman's <i>r</i>	0.366	***	0.495	***	0.896	***	-0.035		0.657	***
	<i>p</i>	< .001		< .001		< .001		.734		< .001	
Factor4 J2	Spearman's <i>r</i>	0.391	***	0.308	**	0.183		0.664	***	0.386	***
	<i>p</i>	< .001		.002		.073		< .001		< .001	
Total J2	Spearman's <i>r</i>	0.759	***	0.591	***	0.707	***	0.191		0.887	***
	<i>p</i>	< .001		< .001		< .001		.061		< .001	

Legend: J1 = judge 1; J2 = judge 2; **p* < .05; ***p* < .01; ****p* < .001.

According to the results, there was a strong correlation between the corrections made by the two judges in the total score of the instrument of $r = .88$ ($p \leq 0.001$). This value meets the minimum requirement, confirming appropriate reliability by this method. There were also significant correlations between the four factors of the instrument: factor 1 ($r = .83$; $p \leq 0.001$), factor 2 ($r = .74$; $p \leq 0.001$), factor 3 ($r = .89$; $p \leq 0.001$), and factor 4 ($r = .66$; $p \leq 0.001$). These results indicated that the scoring procedure demonstrate satisfactory agreement.

Discussion

The steps that were presented in the present study provide positive evidence for the expansion of the audience target to the TCF-AA. Although other sources of validity evidence have already been investigated, it is necessary to continually investigate the psychometric properties of the test (Sato & Seabra, 2022). To investigate the reliability of the TCF-AA, two types were investigated.

The empirical study indicated satisfactory reliability of the TCF-AA considering test and retest method. A positive and significant value ($r = .774$) confirmed the type of reliability investigated. The inter-rated reliability indicated a significant and positive correlation ($r = .887$). When the factors are analyzed separately, it can be verified that all correlations were positive and significant, with only factor 4 showing a less value than expected. It is important to point that participants generally score lower in this factor, presenting infrequent occurrence of emotional characteristics in their responses (average less than 1). There is a possibility that this situation may have influenced the results of the analysis conducted here. The method used (two judges), the correlation value found, and the interval between the two applications are consistent with the findings reported by Long and Wang (2022) during their review of studies on creativity assessment. According to the authors, approximately 30% of the studies used two to five judges, 60% of whom were undergraduate students. The most common interval was two weeks between the test and the retest.

The results confirm the reliability of the test in this new version for use with adolescents and adults. The results are higher than those typically found when investigating the psychometric qualities of creativity tests. Long and Wang (2022) reported on average, a value of $r = .52$, when analyzing reliability using correlations, especially if considerer that scores can be affected by labor-intensiveness and subjectivity, resulting in diminished reliability (Patterson et al., 2023). Although we recognize that every instrument is subject to some measurement error (Ambiel & Carvalho, 2017) and that in the instrument under study, this error value can be known.

In this case, figural creativity refers to the presentation of an unusual task, which requires the individual to provide the greatest number of ideas possible within a limited and standardized response period (Chu & Lin, 2013). To maximize the accuracy of the measurement, it is important to standardize test application and correction procedures to minimize errors produced by the measurement. The reliability results can be added to those previously obtained. By this way, we can say that the instrument permits valid and reliability comparisons between individual performances aiming the identification and estimation of the creative potential (Lubart et al., 2013).

Although the results are positive, they must be interpreted within the context in which they were obtained, especially since culture can influence the development and expression of creativity (Barth & Stadtmann, 2024). By norms, values, and experiences, culture can offer favorable or unfavorable conditions for creativity, resulting in different levels of individual across contexts (Niu & Sternberg, 2006). Since the test relies on creative drawing, the influence of language is limited, being assessed only in one characteristic referred to as expressive titles.

Final considerations

In this study, we sought to address the need for Brazilian instruments for assessing creativity to be the subject of research studies designed to investigate their psychometric properties. It is possible to achieve growth in the area by ensuring that tests present evidence of validity and precision, thus minimizing the existing gap.

In this context, the research aims to expand the range of methods available for evaluating creativity. It is recommended that, given the need for choice, it be made by considering, mainly, the amount of information about the psychometric qualities of the test. Information of this nature may provide evidence that the instrument can accurately and adequately measure the target construct within the target population. The adolescent and adult version of creativity test can assist educators and researchers in evaluating the creativity using a psychometric test, extending the age range in which the test is currently available (only children).

Furthermore, it is important to recognize that no single measure can be used to assess all aspects and forms of creative expression. Thus, it is necessary to restrict the use of the instrument investigated here to a specific type of creativity, figural creativity. In terms of creativity, one aspect to consider is its domain-specific nature, which means that each creative domain requires specific skills and knowledge across different fields (Xu et al., 2024) and, consequently, specific measures.

A limitation of the study was the fact that only two judges judged the protocols. This criterion meets the requirements for such a study; however, more commonly, the accuracy of evaluators has relied upon a larger number of judges, who evaluate a smaller number of response protocols. Additionally, it cannot be overlooked that the results of the test and retest method may have been affected by memory and motivation effects.

In both studies, it is important to note that the data presented low levels of creativity due to the absence of a normal distribution of data. Since the results did not include a wide range of performance and higher strata of creativity, the results should be interpreted with caution. Consequently, little is known about the accuracy of the instrument for assessing, for example, creative giftedness. It is also possible that methodological choices may have influenced the results, so that other studies may be conducted to evaluate the reliability of the instrument in different groups (individuals of different ages, genders, and levels of education). In general, the results confirm the reliability of the expanded version of the Figural Creativity Test, adding yet another positive finding to its psychometric qualities.

Author Notes:

Funding: The studies were financed by CNPq as a research productivity grant for the first author (process 302273/2019-0), by Fapesp, through a scientific initiation scholarship for the second author (process 2022/12020-5) and the fifth author (process 2022/15154-2), by Fapic for the third author, and by CNPq for the fourth author.

Conflict of interest: The authors declare that they have no conflict of interest in publishing the article.

References

- Acar, S., Berthiaume, K., Grajzel, K., Dumas, D., Flemister, C. T., & Organisciak, P. (2023). Applying automated originality score to the verbal form of Torrance Tests of Creative Thinking. *Gifted Child Quarterly*, 67(1), 3-27. <https://doi.org/10.1177/00169862211061874>
- Acar, S., Branch, M. J., Burnett, C., & Cabra, J. F. (2021). Assessing the Universality of the Zero Originality Lists of the Torrance Tests of Creative Thinking (TTCT)-Figural: An Examination with African American College Students. *Gifted Child Quarterly*, 65(4), 354-369. <https://doi.org/10.1177/00169862211012964>
- Acar, S., Lee, L. E., & Hodges, J. (2023). Assessing the Robustness of the Factor Structure of TTCT-Figural: A Meta-CFA Replication-Extension. *Creativity Research Journal*, 35(4), 547-567. <https://doi.org/10.1080/10400419.2023.2209393>
- Alabbasi, A. M. A., Paek, S. H., Kim, D., & Cramond, B. (2022). What do educators need to know about the Torrance Test of Creative Thinking: a comprehensive view. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.1000385>
- Ambiel, R. A. M., & Carvalho, L. F. (2017). Validade e precisão de instrumentos de avaliação psicológica [Validity and reliability of psychological tests]. In M. R. C. Lins & J. C. Borsa (Orgs.), *Avaliação psicológica: Aspectos teóricos e práticos* [Psychological assessment: theoretical and practices aspects] (pp. 115-125). Vozes.
- Anastasi, A. & Urbina, S. (2000). *Testagem Psicológica* [Psychological testing]. Artes Médicas.
- Andrade, J.M. & Valentini, F. (2018). Diretrizes para a construção de testes psicológicos: a Resolução CFP n.009/2018 em destaque [Guidelines for the Construction of Psychological Tests: Regulation CFP No: 009/2018 in Highlight]. *Psicologia Ciência e Profissão*, 38 (n.spe.), 28-39. <http://doi.org/10.1590/1982-3703000208890>
- Barbot, B., Hass, R. W., & Reiter-Palmon, R. (2019). Creativity assessment in psychological research: (Re)setting the standards. *The Psychology of Aesthetics, Creativity, and the Arts*, 13(2), 233-240. <https://doi.org/10.1037/aca0000233>
- Bart, W. M., Hokanson, B., & Can, I. (2017). An investigation of the factor structure of the Torrance Tests of Creative Thinking. *Educational Sciences: theory and practice*, 17(2), 515-528. <https://doi.org/10.12738/estp/2017.2.0051>
- Barth, P., & Stadtmann, G. (2024). Creativity in the West and the East: A Meta-Analysis of Cross-Cultural Differences. *Creativity Research Journal*, 1-47. <https://doi.org/10.1080/10400419.2024.2369442>
- Becker, M., & Cabeza, R. (2023). Assessing creativity independently of language: A language-independent remote associate task (LI-RAT). *Behavior Research Methods*, 55, 85-102. <https://doi.org/10.3758/s13428-021-01773-5>
- Berg, J., Vestena, C. L. B., & Costa-Lobo, C. (2020). Creativity in Brazilian education: review of a decade of literature. *Creative Education*, 11(3), 420-433. <https://doi.org/10.4236/ce.2020.113030>

- Chu, T., & Lin, W. (2013). Uniqueness, integration, or separation? Exploring the nature of creativity through creative writing by elementary school students in Taiwan. *International Journal of Experimental Educational Psychology*, 33(5), 582-595. <https://doi.org/10.1080/01443410.2013.821459>
- Cropley, D. H., Theurer, C., Mathijssen, A. C. S., & Marrone, R. L. (2024). Fit-For-Purpose Creativity Assessment: Automatic Scoring of the Test of Creative Thinking – Drawing Production (TCT-DP). *Creativity Research Journal*, 1–16. <https://doi.org/10.1080/10400419.2024.2339667>
- De La Torre, S. (1991). *Evaluación de la creatividad*. Editorial Escuela Española.
- Gagné, F. (2004). Transforming gifts into talents: the DMGT as a developmental theory. *High Ability Studies*, 15(2), 119–147. <https://doi.org/10.1080/1359813042000314682>
- Green, A. E., Beaty, R. E., Kenett, Y. N., & Kaufman, J. C. (2023). The Process Definition of Creativity. *Creativity Research Journal*, 36(3), 544–572. <https://doi.org/10.1080/10400419.2023.2254573>
- Helfand, M., Kaufman, J. C., & Beghetto, R. A. (2017). The Four C Model of Creativity: Culture and context. In V. P. Glăveanu (Ed.), *Palgrave handbook of creativity and culture research* (pp. 15-360). Palgrave.
- Jung, J. Y. (2022). Physical giftedness/talent: A systematic review of the literature on identification and development. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.961624>
- Kasirer, A., Adi-Japha, E., & Mashal, N. (2020). Verbal and figural creativity in children with autism spectrum disorder and Typical Development. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.559238>
- Kim, K. H. (2006). Can we trust creativity tests? A review of the Torrance Tests of Creative Thinking (TTCT). *Creativity Research Journal*, 18(1), 3-14. https://doi.org/10.1207/s15326934crj1801_2
- Kim, K. H. (2011). The APA 2009 division 10 debate: are the Torrance Tests of Creative Thinking still relevant in the 21st century? *Psychology of Aesthetics, Creativity, and the Arts*, 5(4), 302-308. <https://doi.org/10.1037/a0021917>
- Krumm, G., Arán-Filippetti, V., & López, M. B. (2024). Evaluación de la creatividad: dos estudios normativos del Test de Pensamiento Creativo de Torrance (TTCT), Figuras, Formas A y B en niños argentinos. *Psicogente* 27(51), 1-24. <https://doi.org/10.17081/psico.27.51.6419>
- Kupers, E., Lehmann-Wermser, A., McPherson, G., & van Geert, P. (2019). Children's Creativity: A Theoretical Framework and Systematic Review. *Review of Educational Research*, 89(1), 93–124. <https://doi.org/10.3102/0034654318815707>
- Lassig, C. (2019). Creativity talent development: fostering creativity in schools. In S.R. Smith (Ed.), *Handbook of giftedness and talent development in the Asia-Pacific* (pp. 1-25). Springer.
- Lee, L. E., Meyer, M. S., & Crutchfield, K. (2021). Gifted Classroom Environments and the Creative Process: A Systematic Review. *Journal for the Education of the Gifted*, 44(2), 107-148. <https://doi.org/10.1177/01623532211001450>
- Lee, L. E., Rinn, A. N., & Rambo-Hernandez, K. E. (2024). What Happens After Nomination? Evaluating the Probability of Gifted Identification with the Torrance Test of Creative Thinking. *Gifted Child Quarterly*, 68(2), 119-136. <https://doi.org/10.1177/00169862231222886>
- Long, H., & Wang, J. (2022). Dissecting Reliability and Validity Evidence of Subjective Creativity Assessment: A Literature Review. *Educational Psychology Review*, 34, 1399–1443. <https://doi.org/10.1007/s10648-022-09679-0>
- Lubart, T., Zenasni, F., & Barbot, B. (2013). Creative potential and its measurement. *International Journal for Talent Development and Creativity*, 1(2), 41-50. <https://files.eric.ed.gov/fulltext/EJ1301375.pdf>
- Nakano, T. C., & Wechsler, S. M. (2006). Teste Brasileiro de Criatividade Figural: proposta de instrumento [Brazilian test of figural creativity.: instrumental proposal] *Interamerican Journal of Psychology*, 40(1), 103-110. <https://www.redalyc.org/pdf/284/28440111.pdf>
- Nakano, T. C., & Wechsler, S. M. (2018). Creativity and innovation: Skills for the 21 Century. *Estudos de Psicologia (Campinas)*, 35(3), 237-246. <https://doi.org/10.1590/1982-02752018000300002>
- Nakano, T. C., Batagin, L. R., Cano, I. W., & Fusaro, G. J. (2023). Teste de Criatividade Figural versão adolescentes e adultos: evidências de validade convergente [The Figural Creativity Test: Teen and Adult Version: Evaluation of the Converging Type's Validity]. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 69(3), 5-14. <https://doi.org/10.21865/RIDEP69.3.01>
- Nakano, T. C., Fusaro, G. J., Cano, I. W., Batagin, L. R., & de Abreu, I. C. C. (2023). A influência da idade na cotação de originalidade em um teste de criatividade figural [The effect of age on originality quotations in a figural creativity test]. *Revista Ibero-Americana de Criatividade e Inovação-RECRIA*, 4. <https://recriai.emnuvens.com.br/revista/article/view/92>
- Nakano, T. C., Primi, R., & Zaia, P. (2023). Revisão dos modelos de criatividade [Review of creativity models]. In Instituto Ayrton Senna, R. Primi, T. C. Nakano, K. McGrew, & J. Schneider (Orgs.), *Educação no século XXI: inteligência, pensamento crítico e criatividade* (pp. 185-291). Hogrefe e Instituto Ayrton Senna.
- Nakano, T. C. (2018). A criatividade pode ser medida? Reflexões sobre métodos utilizados e questões envolvidas [Can creativity be measured? Reflections on methods used and issues involved]. *Arquivos Brasileiros de Psicologia*, 70(1), 128-145. <https://doi.org/10.1590/1413-82712018230313>
- Niu, W., & Sternberg, R. J. (2006). The philosophical roots of Western and Eastern conceptions of creativity. *Journal of Theoretical and Philosophical Psychology*, 26(1–2), 18–38. <https://doi.org/10.1037/h0091265>

- Nunes, C. H. S. S., & Primi, R. (2010). Aspectos técnicos e conceituais da ficha de avaliação dos testes psicológicos [Technical and conceptual aspects of the psychological test evaluation form]. In Conselho Federal de Psicologia (CFP), *Avaliação Psicológica: diretrizes na regulamentação da profissão* (pp. 101-128). CFP.
- O'Tolle, K., & Horvát, E. (2024). Extending human creativity with AI. *Journal of Creativity*, 34(2), 100080. <https://doi.org/10.1016/j.yjoc.2024.100080>
- Oliveira, K. S., Nakano, T. C., & Wechsler, S. M. (2016). Criatividade e saúde mental: uma revisão da produção científica na última década [Creativity and Mental Health: A Review of Scientific Production in the Last Decade]. *Temas em Psicologia*, 24(4), 1493-1506. <http://doi.org/10.9788/TP2016.4-16>
- Patterson, J. D., Merseal, H. M., Johnson, D. R., Agnoli, S., Baas, M., Baker, B. S., Barbot, B., Benedek, M., Borhani, K., Chen, Q., Christensen, J. F., Corazza, G. E., Forthmann, B., Karwowski, M., Kazemian, N., Kreisberg-Nitzav, A., Kenett, Y. N., Link, A., Lubart, T.,... Beaty, R. E. (2023). Multilingual semantic distance: Automatic verbal creativity assessment in many languages. *Psychology of Aesthetics, Creativity, and the Arts*, 17(4), 495–507. <https://doi.org/10.1037/aca0000618>
- Peixoto, E. M., & Ferreira-Rodrigues, C. F. (2019). Propriedades psicométricas dos testes psicológicos [Psychometrics Properties of psychological tests]. In M. N. Baptista et al. (Orgs.), *Compêndio de avaliação psicológica* (pp. 29-39). Vozes.
- Plucker, J. A. (2022). The Patient is Thriving! Current Issues, Recent Advances, and Future Directions in Creativity Assessment. *Creativity Research Journal*, 35(3), 291–303. <https://doi.org/10.1080/10400419.2022.2110415>
- Plucker, J. A., & Beghetto, R. (2004). Why creativity is domain general, why it looks domain specific, and why the distinction does not matter. In R. J. Sternberg, E. L. Grigorenko, & J. L. Singer (Eds.), *Creativity: from potential to realization* (pp. 153-167). American Psychological Association. <https://doi.org/10.1037/10692-009>
- Plucker, J. A., Guo, J., & Makel, M. C. (2018). Creativity. In S. I. Pfeiffer (Ed.), *Handbook of giftedness in children* (2nd ed., pp. 81-99). Springer.
- Puccio, G. J., & Murdock, M. C. (1999). *Creativity assessment: Readings and resources*. Creative Education Foundation Press.
- Renzulli, J. S. (2005). The Three-Ring Conception of Giftedness: A Developmental Model for Promoting Creative Productivity. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 246–279). Cambridge University Press. <https://doi.org/10.1017/CBO9780511610455.015>
- Renzulli, J. S. (2016). The three-ring conception of giftedness: A developmental model for promoting creative productivity. In S. M. Reis (Ed.), *Reflections on gifted education* (pp. 55–86). Prufrock Press.
- Revelle, W., & Condon, D. M. (2019). Reliability from α to ω : A tutorial. *Psychological Assessment*, 31(12), 1395–1411. <https://doi.org/10.1037/pas0000754>
- Rocha, K. N. & Wechsler, S. M. (2018). A avaliação da criatividade nas organizações: estudo comparativo entre técnicas usuais e estandardizadas de avaliação psicológica [The Evaluation of Creativity in Organizations: Comparative Study of Usual and Standardized Psychological Evaluation Techniques]. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 48(3), 139-149. <https://www.redalyc.org/journal/4596/459657523013/>
- Runco, M. A. (2023). Updating the Standard Definition of Creativity to Account for the Artificial Creativity of AI. *Creativity Research Journal*, 37(1), 1–5. <https://doi.org/10.1080/10400419.2023.2257977>
- Runco, M.A. (2016). We must prepare for the unforeseeable future. In D. Ambrose & R.J. Sternberg (Eds.), *Creative intelligence in the 21st century: grappling with enormous problems and huge opportunities* (pp. 65-73). Sense Publishers.
- Sánchez-Meca, J., López-Pina, J. A., & López-López, J. A. (2008). A review of the meta-analytic studies of reliability generalization. *Escritos de Psicología*, 2(1), 110-121. <https://revistas.uma.es/index.php/espsi/article/view/13365/13660>
- Sato, D. T. B., & Seabra A., G. (2022). *Avaliação na Neuropsicopedagogia clínica* [Assessment in Clinical Neuropsychology]. Wak Editora.
- Sierra, M. D. V., Cervantes, C. J. R., Verche, E., & Espinoza, B. V. P. (2024). Creativity and Decision Making in Giftedness. *Education Sciences*, 14(3), 251. <https://doi.org/10.3390/educsci14030251>
- Sternberg, R. J. (1997). The triarchic theory of intelligence. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 92–104). The Guilford Press.
- Sternberg, R. J., Glăveanu, V., & Kaufman, J. C. (2022). In Quest of Creativity: Three Paths toward an Elusive Grail. *Creativity Research Journal*, 36(1), 155–175. <https://doi.org/10.1080/10400419.2022.2107299>
- Sternberg, R.J. (2005). The WICS model of giftedness. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of Giftedness* (pp. 243-327). Cambridge University Press.
- Tannenbaum, A. J. (1986). Giftedness: A psychosocial approach. In R. Sternberg, y J. Davidson, J. (Eds.). *Conceptions of giftedness* (pp. 21-52). Cambridge University Press.
- Terai, H., Miwa, K., & Mizuno, S. (2024). Stability of Assessment Using Creativity-Related Indices. *Japanese Psychological Research*, 1-14. <https://doi.org/10.1111/jpr.12519>
- Treffinger, D. J. (1995). School development, talent development, and creativity. *Roeper Review*, 18, 93-97. <https://doi.org/10.1080/02783199509553707>
- Treffinger, D. J., Schoonover, P. F., & Selby, E. C. (2013). *Educating for creativity & innovation*. Prufrock.

- Xu, S., Reiss, M. J., & Lodge, W. (2024). Comprehensive Scientific Creativity Assessment (C-SCA): A New Approach for Measuring Scientific Creativity in Secondary School Students. *International Journal of Science and Mathematics Education*, <https://doi.org/10.1007/s10763-024-10469-z>
- Yahn, L., & Kaufman, J. C. (2016). Asking the wrong question: Why shouldn't people dislike creativity? In D. Ambrose & R. J. Sternberg (Eds.), *Creative intelligence in the 21st century: Grappling with enormous problems and huge opportunities* (pp. 75-87). Sense.
- Zhu, W., Shang, S., Jiang, W., Pei, M., & Su, Y. (2019). Convergent thinking moderates the relationship between divergent thinking and scientific creativity. *Creativity Research Journal*, 31(3), 320-328. <http://doi.org/10.1080/10400419.2019.1641685>
- Ziegler, A. (2009). Research on giftedness in the 21st century. In L.V. Shavinina (Ed.), *International Handbook on Giftedness* (pp.1509-1524). Springer. <https://doi.org/10.1007/978-1-4020-6162-228>

RECEIVED: 10th June 2024
MODIFIED: 17th March 2024
ACCEPTED: 9th April 2025