INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN BUSINESS & SOCIAL SCIENCES

Vol 16, Issue 2, (2026) E-ISSN: 2222-6990

Factorial Invariance of the Kaufman Domains of Creativity Scale (K-DOCS) among Post-Basic Education Students in the Sultanate of Oman

Ahmed Bin Abdullah Bin Bakhit Al-Subaie, Muhammad Nasir Muhammad Saleh, Ahmad Bin Yussuf

Department of Islamic History, Civilization and Education, Academy of Islamic Studies, University of Malaya

Email: amdysf@um.edu.my, Ahmedum2023@gmail.com

Corresponding Authors Email: nasir1990@um.edu.my

DOI Link: http://dx.doi.org/10.6007/IJARBSS/v16-i2/27646

Published Date: 10 February 2026

Abstract

This study aimed to examine the factorial invariance of the Kaufman Domains of Creativity Scale (Kaufman Domains of Creativity Scale – K-DOCS) among post-basic education students in the Sultanate of Oman by testing the four levels of measurement invariance: configural, metric, residual, and scalar invariance. The study sample consisted of Grade 11 and Grade 12 students, and the Arabic version of the scale was administered after ensuring its linguistic and cultural appropriateness. Multi-Group Confirmatory Factor Analysis (MG-CFA) was employed to examine the stability of the scale's factorial structure across selected demographic variables. The results indicated that configural and metric invariance were supported, suggesting equivalence in the factorial structure and the meanings of the latent constructs across groups. However, partial non-invariance was observed at the residual and scalar levels, reflecting potential differences in students' responses to certain items without compromising the overall conceptual structure of the scale. The findings support the use of the K-DOCS for research and group comparisons within the Omani educational context, while recommending caution when interpreting mean-level comparisons. Overall, the results provide empirical support for the validity of the K-DOCS in Arabic contexts and contribute to the literature on measuring creativity from a multidimensional perspective.

Keywords: Measurement Invariance, Factorial Structure, Multidimensional Creativity, Kaufman Domains of Creativity Scale (K-DOCS), Post-Basic Education Students, Sultanate of Oman

Introduction

Creativity is receiving increasing attention in contemporary educational and psychological studies as a fundamental pillar for preparing learners for the demands of knowledge societies. In this context, multidisciplinary models of creativity have emerged, viewing it as a complex

construct manifested across diverse domains, rather than a single, overarching characteristic. The Kaufman Creativity Domains Scale (K-DOCS) is one of the most prominent tools developed to measure this perspective, assessing creativity across relatively independent domains.

Measurement invariance is a cornerstone of psychological and educational measurement, ensuring that the instrument used measures the same psychological structure in the same way across different groups, without bias resulting from cultural, linguistic, or demographic characteristics. The importance of this concept lies in its being a methodological prerequisite for any meaningful comparison between groups, whether this comparison concerns the averages of underlying traits or the relationships between them (Meredith, 1993; Vandenberg & Lance, 2000).

The lack of measurement equivalence indicates that observed differences in scores may not reflect genuine differences in the measured trait but rather may result from differences in understanding the items or in how individuals respond to them. This leads to inaccurate and potentially misleading conclusions (Byrne, 2016). Recent Arab studies have confirmed this problem, showing that applying psychological measurement tools developed in different cultural contexts without verifying their equivalence can lead to interpretive biases that affect the credibility of the results, especially in Arab educational settings (Al-Sibai, 2020).

Measurement equivalence is often examined within the framework of multi-group confirmatory factor analysis, through a hierarchy of nested models. This sequence begins with formal equivalence testing, which is achieved when the number of factors and the pattern of item associations with them remain constant across groups and is considered the minimum threshold for structural comparison. This is followed by metric equivalence, which reflects the uniformity of factor item weights and indicates that the underlying factors have the same meaning across groups (Cheung & Rensvold, 2002).

Residual equivalence, on the other hand, relates to the uniformity of measurement errors associated with each item and reflects the degree of accuracy with which the items measure the underlying construct in each group. This level of equivalence is among the most rigorous and is often difficult to achieve in multicultural applied studies. At the top of this hierarchy is strong equivalence, which allows for fair comparisons of the means of underlying factors and is a prerequisite for interpreting differences between groups as genuine differences in psychological trait (Meredith & Teresi, 2006; Putnick & Bornstein, 2016).

In this context, Al-Siba'i (2020) pointed out that verifying the equivalence of measurements is a necessary methodological step in contemporary Arab psychological research, especially when studying gender differences or educational environments. He emphasized that neglecting this procedure could lead to inaccurate generalizations, even if traditional validity and reliability indicators are met. This view aligns with modern trends in psychometrics, which affirm that measurement equivalence is not a secondary statistical procedure, but rather a fundamental condition for the soundness of scientific inference (Brown, 2015).

With the widespread use of translated psychological scales, the concept of measurement equivalence has emerged as a fundamental methodological requirement to ensure that the

instrument measures the same psychological structure across different groups. Meredith (1993) asserts that any comparison between groups loses its scientific meaning if the stability of the factor structure and the meanings of the underlying factors are not verified. Vandenberg and Lance (2000) also point out that neglecting measurement equivalence is one of the most significant sources of methodological error in comparative research.

In this context, systematic studies have shown that relying solely on traditional validity and reliability indicators may mask underlying standard biases, as item weights or their psychological significance can differ between groups despite internal consistency (Byrne, 2016). Cheung and Rensvold (2002) demonstrated that multigroup confirmatory factor analysis provides a precise framework for testing these differences using graded equivalence models.

The literature indicates that measurement equivalence is examined through a hierarchy that begins with formal equivalence, which is achieved when the number of factors and the pattern of item association are constant, then metric equivalence, which reflects the uniformity of item weights and underlying factor meanings, followed by residual equivalence, which relates to the uniformity of measurement errors, and finally strong equivalence, which allows for fair comparisons of factor means (Meredith & Teresi, 2006; Putnick & Bornstein, 2016).

In the Arab world, serious research efforts have begun to apply the concept of measurement equivalence in psychological and educational studies. Al-Siba'i (2020) pointed out that many Arab studies compare groups without verifying measurement equivalence, which weakens the validity of the conclusions, especially when studying differences between genders or educational stages. His findings confirmed that achieving validity and reliability does not necessarily guarantee measurement fairness.

In another study, Al-Zahrani and Al-Ghamdi (2019) investigated the measurement equivalence of a psychological trait scale among secondary school students. The results showed that formal and metric equivalence was achieved, but strong equivalence was not fully established. This suggests that some differences in the means may reflect differences in response patterns rather than true differences in the measured trait. These findings are consistent with global trends highlighting the difficulty of achieving advanced levels of equivalence in multicultural contexts.

Al-Otaibi's 2021 study, which examined the psychometric properties of a psychological scale used in the Gulf region, found that while items may retain their general factor structure, they differ in error coefficients or constants, thus affecting the accuracy of comparisons between groups. The study emphasized the necessity of using measurement equivalence models as a fundamental procedure before interpreting differences between means.

Regarding the K-DOCS scale specifically, several foreign studies have demonstrated the stability of its factor structure across different cultures. However, some have indicated only partial achievement of advanced measurement equivalence levels, particularly when comparing genders or age groups, reflecting the sensitivity of certain areas of creativity to cultural and educational context (Kaufman et al., 2016). These findings highlight the

importance of verifying measurement equivalence before employing the scale in the Omani context, especially among post-basic education students, whose developmental and cognitive characteristics are diverse.

With the increasing use of psychological scales in non-Western settings, the need arises to verify the equivalence of the measurements to ensure fair comparison and accurate interpretation, especially when using the scale with culturally or demographically diverse groups. This is particularly important in the Omani context, given the limited number of studies examining the psychometric properties of multi-domain creativity scales among post-basic education students.

Based on the above, integrating multi-domain creativity models with measurement equivalence tests represents an integrated methodological approach to studying creativity in different educational contexts, and enhances the reliability, interpretability, and generalizability of the results.

The Problem of the Study

The problem of the study is to question the extent to which the factor structure of the K-DOCS scale is equivalent to the performance of post-basic education students in the Sultanate of Oman across different measurement levels.

Despite the growing interest in measuring creativity as a multidisciplinary construct, and the accompanying development of advanced measurement tools such as the Kaufman Creativity Domains Scale (K-DOCS), the application of these tools in Arab educational settings still faces methodological challenges related to the validity and fairness of the measurement. Many Arab studies rely on translated scales without adequately verifying the reliability of their factor structures across different groups, raising fundamental questions about the accuracy of interpreting differences in students' levels of creativity.

This issue is particularly significant in post-basic education in the Sultanate of Oman, a transitional phase characterized by diverse developmental and cognitive characteristics among students, as well as varying educational and social experiences that may influence their understanding of and response to psychological assessment items. Furthermore, using the K-DOCS scale in this context without verifying its equivalence may lead to results reflecting underlying standard biases rather than genuine differences in areas of creativity.

Although foreign studies have addressed the psychometric properties of the K-DOCS scale, Arabic literature—and Omani literature in particular—still suffers from a clear lack of studies that verify the equivalence of measurement at its various levels, especially metric equivalence, residual equivalence, and strong equivalence, which are essential conditions for fair comparison between groups. Therefore, the problem addressed in this study is the need to verify the extent to which the factor structure of the K-DOCS scale is equivalent among post-basic education students in the Sultanate of Oman across different measurement levels, thus ensuring the soundness of interpretation and the fairness of scientific inference.

Research Questions

In light of the above, the current study seeks to answer the following questions:

1. What is the factor structure of the Kaufman Creativity Domains Scale (K-DOCS) using exploratory and confirmatory factor analysis in the Omani environment?
2. How well does the factor structure of the Kaufman Scale of Creativity Domains (K-DOCS) match student data by gender and grade level in the Omani environment?
3. What are the indicators of reliability and internal consistency of the Kaufman Scale of Creativity Domains (K-DOCS) for data from post-basic education students in the Omani environment?

Importance of the Study

The importance of the current study stems from its theoretical and applied significance in the field of psychological and educational measurement, especially in the area of measuring multi-domain creativity among students in post-basic education in the Sultanate of Oman.

Theoretically, this study contributes to enriching the Arabic literature on measurement equivalence and factor structure of creativity scales by testing the different levels of measurement equivalence of the Kaufman Creativity Domains Scale (K-DOCS) within an Arab cultural and educational context. Furthermore, the study enhances the scientific understanding of creativity as a multidimensional construct and supports modern trends that emphasize the necessity of verifying measurement fairness before interpreting differences between groups.

On a practical level, the study's significance lies in providing a measurement tool with psychometric properties suitable for use in the Omani educational environment. This allows researchers and educational practitioners to rely on more accurate and objective results when studying students' levels of creativity. Furthermore, the study's findings contribute to supporting educational decision-makers in designing educational and developmental programs that consider genuine differences in areas of creativity, moving beyond standardized biases.

The study also gains importance from its focus on the post-basic education stage, which is a pivotal stage in shaping students' creative abilities and in guiding their future academic and professional inclinations, making it extremely important to verify the validity of the measurement tools used at this stage.

Research Objectives

The current study aims to achieve the following:

Verifying the extent to which the measurement equivalence of the Kaufman Scale of Creativity Domains (K-DOCS) is achieved among post-basic education students in the Sultanate of Oman across different equivalence levels.

Determining the level of creativity areas among post-basic education students in the Sultanate of Oman as measured by the Kaufman Creativity Areas Scale (K-DOCS).

Verifying the suitability and reliability of the Kaufman Scale of Creativity Domains (K-DOCS) factor structure among post-basic education students in the Sultanate of Oman.

Research Terms

Measurement Invariance

Measurement equivalence refers to the stability of the relationship between observed and underlying variables across different groups, ensuring that the instrument measures the same psychological construct in the same way without bias resulting from cultural or demographic differences (Meredith, 1993; Vandenberg & Lance, 2000).

The researcher defines it operationally as the extent to which the factor structure of the Kaufman Scale of Creativity Domains (K-DOCS) is stable among post-basic education students in the Sultanate of Oman, and it is verified using multigroup confirmatory factor analysis across levels of formal equivalence, metric, residual equivalence, and strong equivalence.

Factorial Structure

Factorial structure is defined as the latent organization of items into a set of factors that represent the dimensions of the measured psychological structure, and reflect the nature of the relationships between observed variables and latent variables (Brown, 2015).
The researcher defines it operationally in this study as the extent to which the factorial model of the K-DOCS scale is suitable as tested using confirmatory factor analysis, through approved statistical fit quality indicators.

Kaufman Domains of Creativity Scale (K-DOCS)

The K-DOCS scale is a self-reporting tool developed to measure creativity across several qualitative domains, reflecting different patterns of creative behavior in everyday life (Kaufman, 2012).

The researcher defines operationally the K-DOCS scale in this study, the Arabic version applied to post-basic education students in the Sultanate of Oman, which is used to measure the level of creativity in its various fields according to the students' responses to the scale items.

Methodology

The current study adopted a descriptive-analytical approach to verify the psychometric properties of the Kaufman Creativity Domains Scale (K-DOCS), test the suitability of its factor structure and its measurement equivalence within the Omani educational environment, and describe the levels of creativity domains among post-basic education students. This design is suitable for such studies.

Which aims to analyze psychological and educational phenomena as they are in reality, while employing advanced statistical methods to verify the validity of measurement tools and the accuracy of their results.

Population of the Study

The study population comprises all post-basic education students in the Sultanate of Oman enrolled in grades 11 and 12 in government schools during the academic year in which the study was conducted. This stage is a pivotal educational phase characterized by the diversity of students' developmental and cognitive characteristics, as well as the variety of their educational and social experiences, making it a suitable environment for studying multidisciplinary creativity and verifying the validity of its measurement tools.

Study Sample

The study sample was selected using stratified random sampling to ensure adequate representation of the study population. The sample comprised male and female eleventh and twelfth-grade students from schools representing diverse educational environments within the Sultanate of Oman. The sample size was carefully chosen to meet the requirements of confirmatory factor analysis and equivalence analysis, thus ensuring the accuracy of the statistical estimates and the reliability of the results.

The sample size was 930 students, a sufficient size to conduct the advanced statistical analyses used in the study, particularly multigroup confirmatory factor analysis. Participants were also confirmed to have met the study requirements and provided complete answers to the scale items before data entry into the statistical analysis.

Description of the research instrument (scale)

The current study used the Kaufman Domains of Creativity Scale (K-DOCS) after the researcher translated it into Arabic according to the methodological principles adopted in educational and psychological research, and in accordance with the guidelines indicated by Hamilton (1998) in translating measurement tools across cultures, as care was taken to maintain the semantic and conceptual equivalence of the scale items, and to take into account the linguistic and cultural characteristics of the Omani educational environment, in order to ensure the integrity of the measurement and the validity of its results.

The scale, in its final form, contained two main parts:

- **Part One:**This is preliminary data about the study sample, consisting of: Demographic information includes (grade, type of education, gender, educational district).
- **Part Two:**It consists of the dimensions of the scale, which are made up of (5) dimensions, as follows:

- **The first dimension: the subjective field,** it consists of phrase number (1) to phrase number (11).
- **The second dimension: the practical field,** it consists of phrase number (12) to phrase number (22).
- **The third dimension: the performance domain,** it consists of phrase number (23) to phrase number (32).
- **The fourth dimension: the mechanical/scientific field,** it consists of phrase number (33) to phrase number (41).
- **The fifth dimension: the artistic field,** it consists of phrase number (42) to phrase number (50).

The five-point Likert scale (always, often, sometimes, rarely, never) was used to correct the study instrument, where the response is given as never (1), rarely (2), sometimes (3), often (4), always (5).

Study Procedures

The Arabic version of the Kaufman Creativity Domains Scale (K-DOCS) was administered to the study sample after ensuring its linguistic and cultural suitability, following organized procedures that adhered to the ethical and methodological standards of scientific research. These procedures included obtaining the necessary official approvals, clarifying the study

objectives to the participants, and emphasizing the confidentiality of the data and its use solely for scientific research purposes.

The data was collected over a specific time through a questionnaire, ensuring the completeness and validity of student responses for statistical analysis. The data was then coded and entered specialized statistical software, with preliminary checks performed to verify its integrity and the absence of outliers or missing values that could affect the analysis results.

The Validity and Reliability of the Scale

To verify the validity of the Kaufman Creativity Domains Scale (K-DOCS), the Arabic version of the scale was presented to a panel of expert reviewers specializing in psychometrics and educational measurement. These reviewers assessed the suitability of the items in terms of linguistic formulation, clarity, and their relevance to the measured construct. The review results yielded a consensus rate of [percentage missing in original text]. **89%**Among the judges, with their recommendation to delete one paragraph due to its inappropriateness, these comments were considered in the final version of the scale, which supports its content validity.

As for the reliability of the scale, it was verified using a coefficient **Cronbach's Alpha** For internal consistency, where the overall reliability coefficient reached**(α = 0.88)**This value indicates a high level of reliability and demonstrates good internal consistency of the scale items, making it suitable for application in the Omani educational environment.

Statistical Analysis

The data were analyzed using a range of integrated statistical methods. Confirmatory factor analysis was used to verify the suitability of the K-DOCS factor structure by assessing the consistency of the proposed factor model with the experimental data based on established statistical fit quality indices.

Multigroup confirmatory factor analysis was also used to test the measurement equivalence of the scale across different groups, by examining the levels of formal, metric, residual, and strong equivalence, in a hierarchical order, relying on the change in conformity indices to judge whether each level was achieved.

In addition, descriptive statistics, represented by arithmetic means and standard deviations, were employed to describe the levels of creativity among the sample members. All analyses were conducted using specialized structural modeling and factor analysis software (AMOS), adhering to established statistical standards in psychometric studies.

Study Results and Discussion

Question 1: What is the factor structure of the Kaufman Creativity Domains Scale (K-DOCS) using exploratory and confirmatory factor analysis?

First: The exploratory factor structure of the Kaufman Creativity Domains Scale (K-DOCS)

This part of the study sought to verify the underlying factor structure of the Kaufman Creativity Domains Scale (K-DOCS), in its modified form, using Exploratory Factor Analysis to verify the number of basic factors that explain the variation in the responses of the sample

members, and to determine the degree of consistency with the theoretical construction of the scale in its original form.

Checking the suitability of the data for factor analysis

The suitability of the data for factor analysis was examined using Bartlett's test of sphericity and the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMA). The results showed a chi-squared value of 13180.85 (df = 1176, $p < 0.001$), indicating high statistical significance and suggesting that the correlation matrix is suitable for factor analysis. The overall KMO value was 0.91, a high value indicating strong suitability for factor analysis, as all sub-values for the items exceeded the minimum acceptable threshold of 0.60.

Determine the number of extracted factors

The study determined the number of factors based on the eigenvalue ≥ 1, the screen plot test, and consideration of the theoretical interpretation of the factors and the number of saturated items for each factor. The analysis revealed that only five factors exceeded the eigenvalue of 1, with values of 9.69, 3.55, 2.43, 1.81, and 1.49, respectively. This five-factor model explained 31.5% of the total variance in responses.

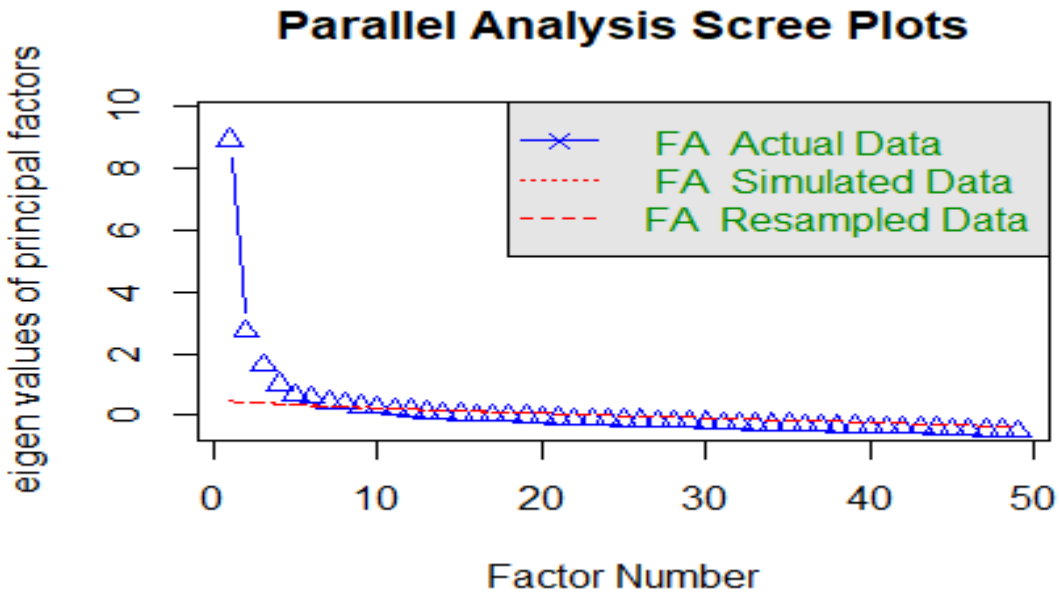

Figure (1): Graphical representation of the latent root values of students' responses on the Kaufman Scale of Creativity domains

Results of the exploratory factor analysis

Principal Axis Factoring (PAF) with Promax rotation was used to assess the correlation between factors. The researcher employed Kaiser's criterion to test the factor loadings of the items, considering loadings of 0.3 or higher to be statistically significant. The analysis revealed that most scale items were statistically significantly distributed across five principal factors representing the domains of creativity in the original Kaufman Scale: the personal/daily domain (E), the academic domain (S), the performance domain (P), the mechanical/scientific domain (MS), and the artistic domain (A). The following table illustrates the factor loadings of the scale items, their subjective values, and the percentage of variance explained for each extracted factor.

Table (1)

Factor saturations, eigenvalues, and explanatory variance for the five factors

The fifth factor	The fourth factor	The third factor	The second factor	First factor	Phrase number
					E1
		0.62			E2
		0.51			E3
		0.42			E4
		0.44			E5
		0.54			E6
		0.53			E7
		0.34			E8
		0.42			E9
		0.40			E10
		0.55			E11
					S1
				0.32	S2
					S3
	0.37				S4
		0.35			S5
	0.46				S6
	0.44				S7
	0.66				S8
	0.53				S9
	0.36				S10
	0.36				S11
				0.68	P1
				0.55	P2
				0.75	P3
				0.84	P4
				0.49	P5
					P6
				0.49	P7
				0.61	P8
				0.37	P9
0.36					MS1
0.65					MS2
0.30					MS3
	0.34				MS4
0.70					MS5
0.77					MS6
0.44					MS7
	0.52				MS8
					MS9
			0.80		A1
			0.80		A2
			0.75		A3
			0.51		A4
			0.55		A5

The fifth factor	The fourth factor	The third factor	The second factor	First factor	Phrase number
			0.40		A6
			0.50		A7
			0.36	0.32	A8
			0.40		A9
1.49	1.81	2.43	3.55	9.69	latent root
4.7%	5.2%	6.1%	7.4%	8.1%	Explained variance ratio
31.5%	26.8%	21.6%	15.5%	8.1%	Total percentage of variance explained

The results in Table (1) indicate a high degree of agreement between the extracted factors and the five original theoretical domains of the scale, as follows:

1- The first factor was related to the dimension of performance creativity, and included the statements: P1, P2, P3, P4, P5, P7, P8, P9, where the saturations of these statements came within a range between 0.37 and 0.84, which reflects a behavioral dimension related to applied and practical achievements.
2- The second factor was the dimension of artistic creativity, and it included statements A1 to A9, with saturations ranging between 0.36 and 0.80, which clearly reflects the connection of this factor to creative productions of an aesthetic and artistic nature.
3- The third factor is the opposite of the Everyday Creativity dimension, and it includes the statements: E2 to E11, where their saturations ranged between 0.34 and 0.62, which confirms the association of these statements with creative behaviors in daily life and ordinary situations.
4- The fourth factor was related to the dimension of academic creativity, and included the statements: S4, S5, S6, S7, S8, S9, S10, S11, MS4 and MS8, with saturations ranging between 0.34 and 0.66, which reflects the dimension of critical thinking and cognitive achievement in the academic context.
5- The fifth factor was represented by the dimension of Mechanical/Scientific Creativity, and included the statements: MS1, MS2, MS3, MS5, MS6, MS7, with saturations ranging between 0.30 and 0.77, which shows a focus on analytical and technical skills.

Furthermore, the results indicated that some statements—E1, S1, S3, P6, and MS9—did not show significant saturations for any of the extracted factors, and their factor saturations did not exceed the minimum dependent threshold (0.30) for any of the factors, suggesting a weak structural correlation with the extracted factors. Statement A8 also showed weak double saturation for the first and second factors (0.32 and 0.36, respectively), which may indicate ambiguity in the wording or content overlap. These results suggest the need to reconsider the unsaturated statements. Nevertheless, the researcher decided to retain these statements and study them considering the results of the confirmatory factor analysis and differential performance analysis at a later date.

Based on the results of the confirmatory factor analysis (CFA) of the five-structured model of the Kaufman Creativity Domains Scale (K-DOCS), the model's fit to the actual data was tested using several fit quality indices. The results showed a chi-squared value (χ^2 = 3225.38, df =

1080, p < 0.001), indicating a statistically significant difference between the predicted and observed variance matrix. However, this statistic is affected by the large sample size. Therefore, relative and absolute fit quality indices were used to evaluate the model.

Table (2)
Indicators of conformity quality to the Kaufman scale for areas of creativity

Index value	The ideal value of the index	Conformity Quality Index
3225.38	To be non-functional	chi-square χ^2
2.98	Less than 3 good matches, Fewer than 5 matches are acceptable	Standardized chi-square / df χ^2
0.047	Less than 0.05 is a good match. Less than 0.08 matches are acceptable	The square root of the mean error RMSEA
0.965	Greater than 0.95 is a good match. Greater than 0.90 matches are acceptable	Comparative Conformity Index (CFI)
0.962	Greater than 0.95 is a good match. Greater than 0.90 matches are acceptable	Lewis Tocker Index (TLI)
0.055	Less than 0.08 is a good match. Less than 0.1 matches are acceptable	Root Mean Square Residue Index (SRMR)

According to the results in Table (2), the Comparative Fit Index (CFI) was 0.965, and the Tucker Lewis Index (TLI) was 0.962, both within the optimal range (greater than 0.95), indicating a good fit for the model. The RMSEA value was 0.047, also within the optimal range (less than 0.05), and the 90% confidence intervals were between 0.045 and 0.049, further supporting the quality of the fit. The Root Mean Square Residuals Index (SRMR) was 0.055, also within the optimal range (less than 0.08), reinforcing the model's validity.

Table (3)
Standard saturations of Kaufman scale items for the domains of creativity

Artistic creativity	Mechanical/Scientific Creativity	Creativity in performance	Academic Creativity	Self/Daily Creativity
0.599	0.721	0.561	0.547	0.348
0.617	0.448	0.679	0.616	0.477
0.618	0.435	0.747	0.547	0.534
0.562	0.487	0.645	0.543	0.417
0.494	0.555	0.509	0.463	0.355
0.709	0.554	0.552	0.478	0.542
0.706	0.700	0.599	0.454	0.564
0.700	0.479	0.660	0.522	0.542
0.489	0.673	0.682	0.616	0.588
			0.61	0.494
			0.575	0.449

Upon examining the standardized factorization coefficients for all items (Table 13), they were found to range from 0.348 to 0.588 in the first domain (everyday creativity), from 0.454 to 0.616 in the second domain (academic creativity), from 0.509 to 0.747 in the third domain (performance creativity), from 0.435 to 0.721 in the fourth domain (mechanical/scientific creativity), and from 0.489 to 0.709 in the fifth domain (artistic creativity). All these values are high and statistically significant ($p < 0.001$), indicating a strong correlation between the items and their assumed dimensions. Therefore, the results suggest that the proposed five-dimensional model has a good fit, and the factorization coefficients reflect statistical significance and understandable content for each dimension of the scale, thus supporting the theoretical factorization structure of the scale after cultural and procedural adjustment (Figure 2).

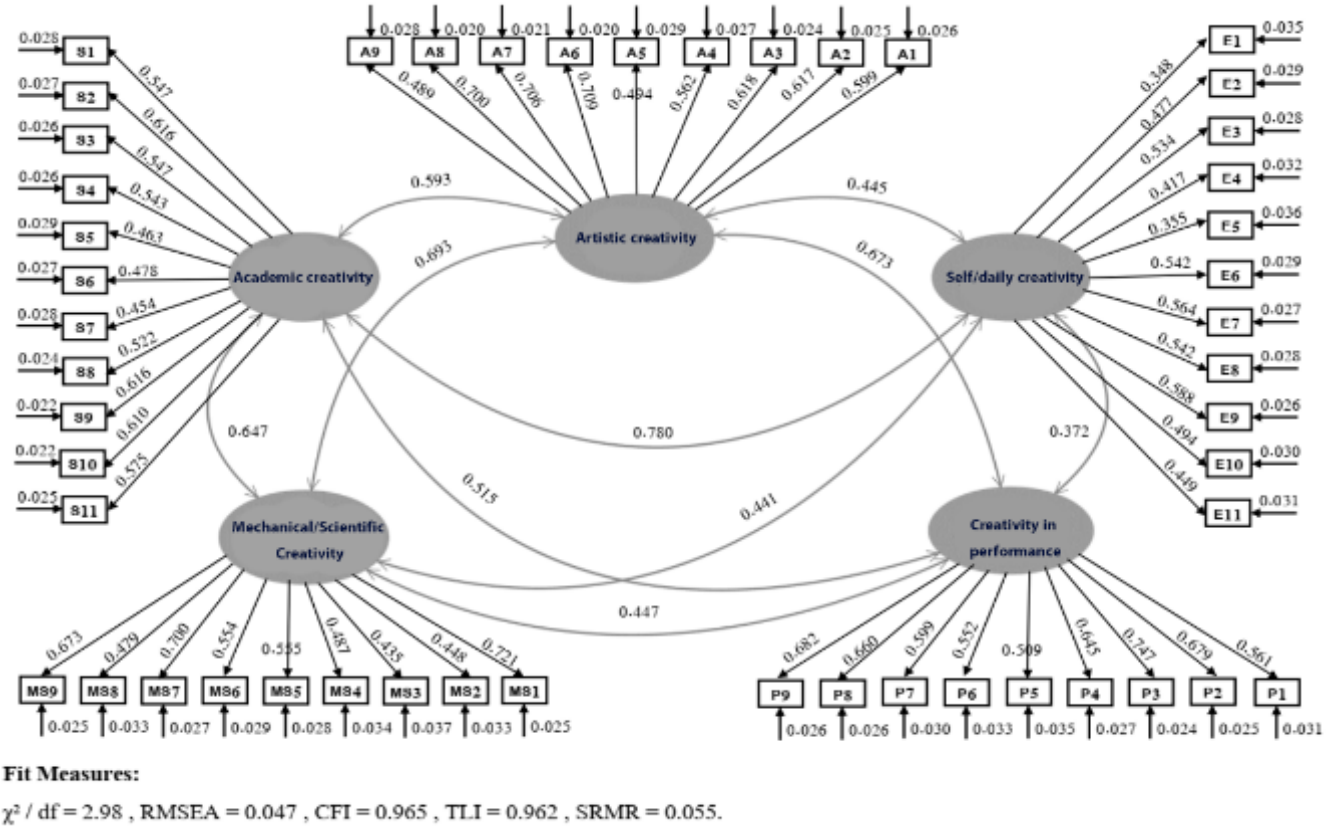

Figure (2): The confirmatory factor model of the Kaufman scale of creative domains

The results of the current study are consistent with the majority of previous studies that used confirmatory factor analysis (CFA) to verify the construct validity of the Kaufman Scale of Creativity Domains.

The results of the current study supported the good fit of the original five-factor model of the scale, a finding corroborated by several previous studies in different cultural and educational contexts. For example, ÅupaniÄ's study (2021), conducted on a sample of early childhood and preschool students in Croatia, found that the five-factor model exhibited good fit, reflecting the consistency of its factor structure across different age groups. Similarly, Fatade and Awofala's study (2015) in Nigeria reached similar conclusions when the scale was administered to a sample of teachers, with the five-factor model demonstrating good fit to the data and stable psychological properties.

Similarly, the study by McKay et al. (2017), which tested the validity of the scale on a multinational student sample, confirmed that the five-factor model maintains high structural quality, reflecting its applicability in diverse cultural settings. Likewise, the study by Metwaly et al. (2024), conducted on a sample of university students in Egypt, supported the good fit of the five-factor model to the data, indicating that the Kaufman scale retains its original factor structure.

Question 2: To what extent does the factor structure of the Kaufman Creativity Domains Scale (K-DOCS) match student data according to gender and grade level?

To answer the second question, we investigated the structural consistency of the Kaufman Scale of Creativity among students by gender (male, female) and by grade level (10th and 11th). This was achieved through multigroup confirmatory factor analysis (MG-CFA) using a hierarchical sequence of models: the Configural, Metric, Scalar, and Strict models. The Chi-Squared Difference Test was used, along with comparisons of the Quality of Fit Indices (CFI and RMSEA) between the successive models.

Table (4)
Indicators of good conformity to the Kaufman scale of creativity domains by gender

ΔRMSEA	ΔCFI	RMSEA	CFI	Δx^2	df	x^2	The model
-	-	0.037	0.963	-	2160	3527.0	formal
0.004	0.008	0.041	0.955	96.24*	2204	3910.2	Metric
0.001	0.002	0.042	0.953	190.87*	2248	4059.1	Standard
0.003	0.007	0.045	0.946	414.98*	2297	4448.9	The rest

* Statistically significant at the 0.001 level.

The results of the morphological model showed good data concordance across genders (CFI = 0.963, RMSEA = 0.037), indicating that the scale's basic factor structure is met across genders. When comparing the metric model to the morphological model, the results showed a statistically significant difference in chi-squared ($\Delta\chi^2 = 96.24$, df = 44, $p < 0.001$), with a decrease in the CFI value (ΔCFI = 0.008) and a decrease in the RMSEA value (ΔRMSEA = 0.004). Since the difference in CFI and RMSEA did not exceed the acceptable limit (0.01), this indicates that the Kaufman Scale of Creativity domains is metrically equivalent across genders.

Similarly, the comparison between the metric and standard models showed a statistically significant difference ($\Delta\chi^2 = 109.87$, df = 44, $p < 0.001$). However, the ΔCFI (0.002) and ΔRMSEA (0.001) indices remained below the acceptable limit (less than 0.01), indicating that standard equivalence of the Kaufman Scale of Creativity Domains by Gender was not achieved.

Finally, the residual model showed significant differences in chi-squared when compared to the standard model ($\Delta\chi^2 = 414.98$, df = 49, $p < 0.001$), with a decrease in the value of the CFI indices (ΔCFI = 0.007) and a decrease in the value of the RMSEA (ΔRMSEA = 0.003). Since the difference in CFI and RMSEA is still within the acceptable limit (0.01), this indicates that the residuals of the Kaufman scale of creativity domains by gender have been achieved.

Based on the above, it can be said that Kaufman's scale of creative domains showed good levels of structural consistency across genders (males and females), and this indicates that the scale works in the same way in both genders, allowing for objective comparisons between male and female students (Figure 3).

Similarly, the structural consistency of the Kaufman Scale of Creativity domains among students was confirmed by grade level (10th and 11th grade). Multigroup confirmatory factor analysis (MG-CFA) was applied across four sequential models: the Configural model, the Metric model, the Scalar model, and the Strict model. The chi-squared difference test ($\Delta\chi^2$) was used, along with quality of fit indices (CFI and RMSEA). Differences between models were measured using ΔCFI and ΔRMSEA. Non-significant differences were considered evidence of equivalence if ΔCFI and ΔRMSEA were less than 0.01, even if $\Delta\chi^2$ was statistically significant, as it is greatly affected by sample size. The results are shown in Table (5).

Table (5)
Indicators of good conformity to the Kaufman scale of creativity domains by grade level

ΔRMSEA	ΔCFI	RMSEA	CFI	Δx^2	df	x^2	The model
-	-	0.036	0.968	-	2160	3408.9	formal
0.003	0.006	0.039	0.962	71.29*	2204	3702.4	Metric
0.001	0.001	0.0.38	0.961	77.07*	2248	3758.8	Standard
0.000	0.000	0.038	0.961	78.28*	2297	3830.3	The rest

* Statistically significant at the 0.001 level.

The results indicated that the scale was metrically and formally equivalent across grade levels, since although the difference in chi-squared between the metric and formal models was statistically significant ($\Delta\chi^2 = 71.29$, df = 44, $p < 0.01$), the differences in the congruence indices were limited (ΔCFI = 0.006, ΔRMSEA = 0.003), indicating that the Kaufman Scale of Creativity Domains was metrically equivalent across grade levels.

The results also indicated standard and residual equivalence, as the chi-squared differences showed statistical significance when switching to the standard model ($\Delta\chi^2 = 77.07$, df = 44, $p < 0.01$) and then to the residual model ($\Delta\chi^2 = 78.28$, df = 49, $p < 0.01$). Despite this, the CFI and RMSEA differences remained within the acceptable limit (0.01), indicating standard and residual equivalence for the scale across grade levels. These results suggest that the scale maintains all levels of measurement equivalence (formal, metric, standard, and residual) across grade levels (Figure 3).

The results of the current study are also consistent with those of the study by Kapoor et al. (2021).The study aimed to verify the psychometric properties of the Kaufman Creativity Domains Scale (K-DOCS) and the consistency of its factor structure across the variables of gender and age. The study was conducted on a sample of university students in the United States, and the researchers used multigroup confirmatory factor analysis (MG-CFA) to verify the equivalence of the measurements across groups. The results showed that all levels of the measurement equivalence gradient were met, includingConfigurative invariance, andMetric invariance, andscalar invarianceIn additionResidual invarianceAmong male and female students, as well as among different age groups according to grade level.

These results indicate that the Kaufman scale has a high degree of structural reliability across gender and grade level. This enhances its ability to measure creativity fairly and reliably without bias towards any particular group. The results of this study support the findings of the current study regarding the stability of the scale's factor structure across multiple demographic variables, such as gender, type of education, and educational region, thus strengthening the legitimacy of using the scale for comparisons between different groups. (Figure4)

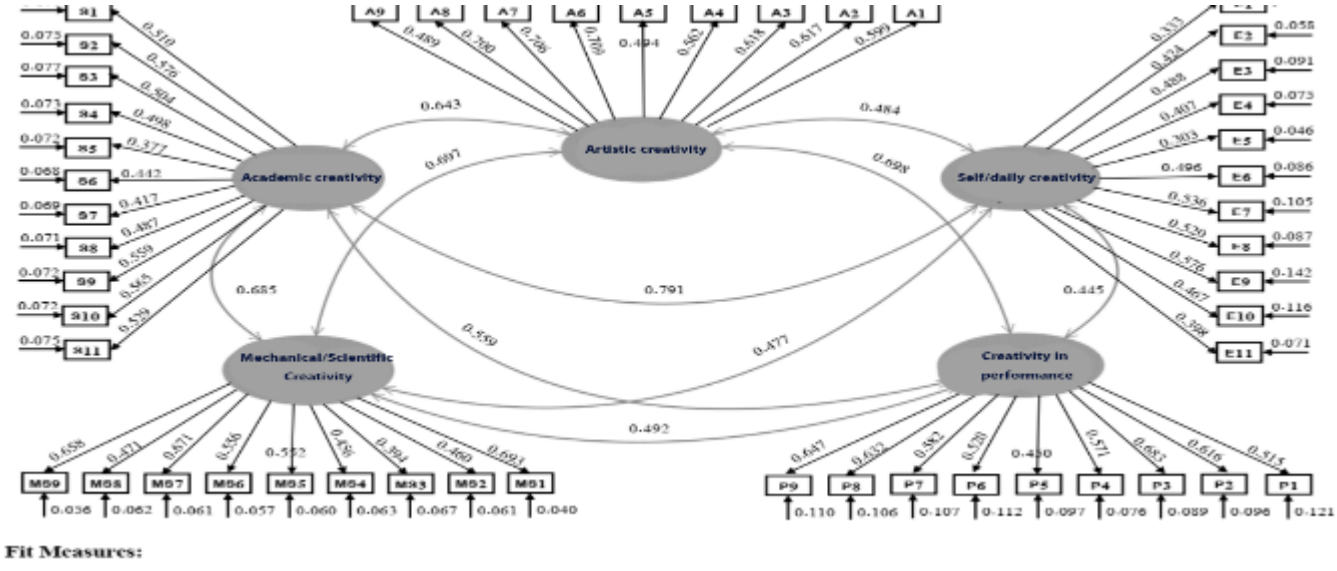

Figure (3): Measurement equivalence model for Kaufman's scale of creativity domains according to gender.

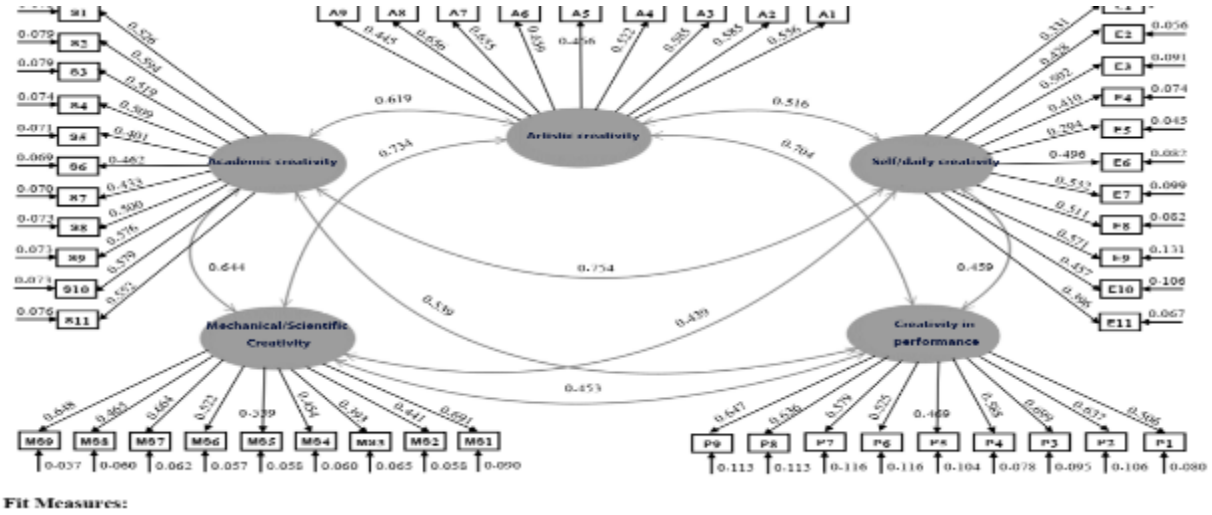

Figure (4): Measurement equivalence model for the Kaufman scale of creative domainsAccording to the grade level.Question 3:

What are the indicators of reliability and internal consistency of the Kaufman Scale of Creativity Domains (K-DOCS) for post-basic education student data?

To verify the psychometric properties of the Kaufman Creativity Domains Scale (K-DOCS), the reliability and internal consistency indices of the scale's five dimensions were estimated using both Cronbach's Alpha and McDonald's Omega coefficients. The results showed that all five

dimensions of the scale exhibited good reliability, with Alpha coefficients ranging from 0.747 to 0.828 and Omega values ranging from 0.772 to 0.861, both within the scientifically acceptable range, indicating a good degree of reliability. Specifically, the reliability indices are presented in Table 6:

Table (6)
Stability using Cronbach's alpha coefficient and omega coefficient

Omega Factor	Cronbach's alpha coefficient	Number of paragraphs	Field
0.772	0.747	11	Self/Daily Creativity
0.817	0.787	11	Academic Creativity
0.857	0.828	9	Creativity in performance
0.837	0.783	9	Mechanical/Scientific Creativity
0.861	0.821	9	Artistic creativity

Furthermore, the internal consistency of the scale across its dimensions was verified by measuring the correlation coefficient between each item and the total score for its respective domain. The results showed that all correlation coefficients were statistically significant at the significance level ($p < 0.001$), indicating that the items measure the same underlying concept within each domain. The correlation coefficients ranged between 0.25 and 0.64, further supporting the internal consistency validity of the Kaufman Scale of Creativity Domains (Table 7).

Table (7)
Internal consistency of items in the Kaufman Scale for Creativity Domains

Artistic creativity	Mechanical/Scientific Creativity	Creativity in performance	Academic Creativity	Self/Daily Creativity
0.585*	0.430*	0.500*	0.315*	0.246*
0.607*	0.480*	0.528*	0.386*	0.452*
0.591*	0.390*	0.618*	0.423*	0.421*
0.447*	0.360*	0.639*	0.417*	0.378*
0.432*	0.560*	0.462*	0.395*	0.354*
0.552*	0.520*	0.400*	0.462*	0.469*
0.566*	0.570*	0.534*	0.437*	0.454*
0.511*	0.410*	0.615*	0.487*	0.394*
0.404*	0.450*	0.482*	0.519*	0.403*
			0.510*	0.340*
			0.480*	0.390*

* The correlation coefficient is statistically significant at the 0.001 level.

The results of the current study regarding the reliability and internal consistency indicators of the Kaufman Scale of Creativity domains are consistent with what many previous studies have shown, such as the study by ÅupaniÄ (2021), the study by Fatade & Awofala (2015), the study by McKay et al. (2017), and the study by Metwaly et al. (2024), as these studies confirmed that the five-point model of the scale has a high degree of reliability, whether through Cronbach's alpha coefficients or the omega coefficient, which indicates good internal consistency between the items within each dimension of the scale.

These studies demonstrated that the items within each of the five creativity domains (personal/daily creativity, academic, performance, artistic, and mechanical/scientific) measure the same underlying structure, and that the correlation coefficients between the items and the overall domain were statistically significant, supporting internal consistency. This aligns with the findings of the current study, which showed that alpha and omega coefficients ranged from acceptable to high levels, in addition to the significance of the inter-item correlation coefficients. These results thus strengthen the legitimacy of using the scale in the Omani context and confirm its consistency with the original structure of the scale as applied in different cultural and educational settings.

Conclusion and Recommendations

The study concluded that the K-DOCS scale has an acceptable degree of measurement equivalence among post-basic education students in the Sultanate of Oman, while recommending its use with caution when making direct comparisons of mean scores. The study recommends further analyses using Item Response Theory models and expanding the scale's application to other educational groups to enhance the generalizability of the findings.

Research Recommendations

1. Reapply the scale to larger and more diverse samples in terms of gender, grade levels, and educational districts to verify the reliability of the factor structure and the validity of the instrument in different educational environments.
2. Expand the scope of future studies to include additional variables such as academic specialization, socioeconomic level, or urban and rural environment; to examine the fairness of the scale across multiple groups.
3. Combining Confirmatory Factor Analysis (CFA) and Item Response Theory (IRT) to thoroughly verify construct validity and measurement reliability across different groups.

References

Al-Zahrani, M., & Al-Ghamdi, S. (2019). Verifying the equivalence of measurement for a psychological scale among high school students. The Arab Journal of science Psychology، 11(3) ، 45–72.

Al-Saba'i, Ahmed bin Abdullah. (2020). Equivalence of Measurement in Psychological and Educational Research: Concept and Methodological Applications. Journal of Educational and Psychological Sciences، 4(2) ، 115–138.

Al-Otaibi, F. Bin Hamad. (2021). Psychometric properties and measurement equivalence of a psychological scale in the Gulf environment. Journal of Educational Measurement and Evaluation، 5(1) ، 89–112.

ÅupaniÄ, M. (2021). Translation and validation of the Kaufman domains of creativity scale on a Croatian sample of early childhood and preschool education students. Center for Educational Policy Studies Journal, 11(3), 163-179

Awofala, A. O., & Fatade, A. O. (2015). Validation of the domains of creativity scale for Nigerian preservice science, technology, and mathematics teachers. Electronic Journal of Research in Educational Psychology, 13(1), 131-150

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

Byrne, B. M. (2016). Structural equation modeling with AMOS: Basic concepts, applications, and programming (3rd ed.). Routledge.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. Structural Equation Modeling, 9(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Hamilton, M. (1998). Translation and cross-cultural adaptation of instruments. In J. P. Guillemin, C. Bombardier, & D. Beaton (Eds.), Cross-cultural adaptation of health-related quality of life measures (pp. 1–12). Oxford University Press.

Kapoor, H., Reiter-Palmon, R., & Kaufman, J. C. (2021). Norming the muses: Establishing the psychometric properties of the Kaufman Domains of Creativity Scale. Journal of psychoeducational assessment, 39(6), 680-693.

Kaufman, J. C. (2012). Counting the muses: Development of the Kaufman Domains of Creativity Scale (K-DOCS). Psychology of Aesthetics, Creativity, and the Arts, 6(4), 298–308. https://doi.org/10.1037/a0029751

Kaufman, J. C., et al. (2016). Openness to experience and intellect differentially predict creative achievement. Journal of Personality, 84(2), 248–258.

McKay, A. S., Karwowski, M., & Kaufman, J. C. (2017). Measuring the muses: validating the Kaufman domains of creativity scale (K-DOCS). Psychology of Aesthetics, Creativity, and the Arts, 11(2), 216.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. Psychometrika, 58(4), 525–543. https://doi.org/10.1007/BF02294825

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. Psychometrika, 58(4), 525–543. https://doi.org/10.1007/BF02294825

Meredith, W., & Teresi, J. A. (2006). An essay on measurement and factorial invariance. Medical Care, 44(11 Suppl 3), S69–S77. https://doi.org/10.1097/01.mlr.0000245438.73837.89

Metwaly, S. S., Abdalmageid, A. F., Chen, C., & Van den Noortgate, W. (2024). Validation and Network Structure of the Kaufman Domains of Creativity Scale: A Psychometric Investigation in an Arabic Context. Thinking Skills and Creativity

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. Developmental Review, 41, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. Organizational Research Methods, 3(1), 4–70. https://doi.org/10.1177/109442810031002