



OPEN Psychometric validation of the 20-item K-DOCS in Chinese adolescents: a multi-method approach

Cui-Hong Cao¹, Xue-Lian Wang^{2,4}✉, Yin-Ping Ji¹ & I-Hua Chen³✉

Creativity is crucial for addressing new world challenges. The Kaufman Domains of Creativity Scale, a widely used self-report scale, assesses distinct creativity facets but lacks validation in adolescents under 18 and is limited by classical test theory (CTT). This study evaluated the psychometric properties of the 20-item Kaufman Domains of Creativity Scale (20-K-DOCS) in Chinese adolescents aged 15–18. 3,079 adolescents completed the 20-K-DOCS. CTT, Rasch analysis, and Exploratory Structural Equation Modeling (ESEM) were used to assess reliability and various types of validity. Concurrent validity was examined by correlating 20-K-DOCS scores with Creative Self-Efficacy (CSE). The 20-K-DOCS demonstrated excellent internal consistency (McDonald's omega: 0.91–0.95 per subscale). Rasch analysis supported item validity and optimal functioning. The oblique five-factor model and the ESEM five-factor model indicated satisfactory model fit. ESEM revealed convergent validity and addressed high factor correlations, improving discriminant validity. Significant positive correlations were found between 20-K-DOCS and CSE, particularly in Artistic and Scholarly Creativity. The 20-K-DOCS is a reliable and valid measure of domain-specific creativity in Chinese adolescents. Findings emphasize the importance of domain-specific approaches to creativity assessment and suggest the potential for developing a more concise scale version.

Keywords Kaufman domains of creativity scale, Psychometrics, Rasch analysis, Exploratory structural equation modeling, Adolescents, Creativity

The importance of creativity in the 21st century cannot be overstated, as it plays a crucial role in driving innovation, problem-solving, and personal and societal progress¹. According to Plucker et al., the concept of creativity refers to the interaction of abilities, processes, and environment through which an individual or group can produce a perceived product of novelty and usefulness defined in a social context², which develops significantly during adolescence, with distinct trajectories of development for insight, verbal divergent thinking, and visual divergent thinking³. As Barbot and Heuser assert, creativity functions as a vital component of adolescent self-identity, enabling teenagers to cultivate positive self-image through creative expression⁴. Tested in a sample of 146 late adolescents, Giancola et al. supported that adolescents' openness and creativity contribute to their ability to critically think about today's ecological issues and adopt environmentally friendly behaviors⁵.

Historically, the study of creativity has been approached from two main perspectives: domain-general theories that focus on universal aspects of creativity, such as divergent thinking and personality traits, and domain-specific theories that emphasize the role of expertise, knowledge, and skills specific to a particular domain^{6,7}. These distinct approaches, however, have led to a need for a more comprehensive framework that can reconcile both perspectives.

In response to this need, the Amusement Park Theoretical (APT) model of creativity^{8,9} offers a unified framework that integrates both domain-general and domain-specific perspectives, providing a more comprehensive understanding of the nature of creativity. The APT model suggests that creativity operates at different levels of specificity, with domain-general factors (e.g., intelligence, motivation) acting as the foundation for creativity across all domains, while domain-specific factors (e.g., expertise, knowledge) influence creativity within specific areas. In this model, creativity is likened to an amusement park, with general thematic areas

¹School of Foreign Languages, Shandong Women's University, Jinan 250300, China. ²Yancheng Mechatronic Branch, Jiangsu Union Technical Institute, Yancheng 224006, China. ³Chinese Academy of Education Big Data, Qufu Normal University, Qufu, Shandong, China. ⁴International College, Krirk University, Bangkok, Thailand. ✉email: 412901508@qq.com; aholichen@gmail.com

represented by different sections of the park and more specific domains represented by individual rides within each section. The model incorporates domain-general factors as the park's general requirements that apply to all visitors, regardless of the specific rides they choose, while domain-specific factors are represented by the unique requirements and experiences of each individual ride. By integrating both domain-general and domain-specific perspectives, the APT model acknowledges the importance of universal factors that contribute to creativity across domains while also recognizing the role of domain-specific expertise and knowledge in shaping creative outcomes within specific fields, allowing for a more nuanced understanding of the complex nature of creativity and its manifestation in various contexts.

Building upon the APT model, Kaufman¹⁰ developed the Kaufman Domains of Creativity Scale (K-DOCS), a self-report measure that assesses individuals' self-perceived creativity across five distinct domains: Self/Everyday, Scholarly, Performance, Mechanical/Scientific, and Artistic. The scale consists of 50 items that ask respondents to rate their creativity in various areas of life, such as daily problem-solving, academic pursuits, performing arts, engineering, and visual arts⁹. The five domains assessed by the K-DOCS align with the APT model's conceptualization of creativity as a multi-faceted construct, providing a valuable tool for researchers and practitioners to assess creativity across these domains and gain insights into individuals' creative strengths and preferences. Notably, the K-DOCS can effectively assess an individual's level of creativity in a specific domain even if they have not directly experienced the behaviors described in the domain-specific items. This is achieved by asking respondents to imagine themselves in those scenarios¹¹.

The K-DOCS has been widely used in creativity research and has demonstrated good psychometric properties, including internal reliability and factorial validity^{10,12–17}. Among the studies regarding to factorial validity, Kapoor et al., recruiting a large sample of 22,013 American participants, adopting confirmatory factor analyses (CFA), found that the K-DOCS demonstrated good reliability and construct validity, with the five- and nine-factor model of good fit for both men and women while the one-factor model was not recommended¹⁵. In another study, using CFA as well, Kapoor, Zheng, et al. supported both five- and nine- factor structure for the K-DOCS across ethnicity¹⁶. However, in Miroshnik et al.'s study validating the K-DOCS on a Russian-speaking sample, exploratory and confirmatory factor analyses demonstrated that the five-factor structure was the best fit but the nine-factor one was not supported¹³.

As for criterion validity, choosing the Big Five Inventory–2 as criterion measures, Miroshnik et al. reported to find that openness to experience was correlated with the K-DOCS¹³. Moreover, adopting Chinese big five personality inventory brief version (CBF-PI-B) and creative self-efficacy scale as criterion measures, Zhang et al. found that openness to experience and creative self-efficacy had positive correlations with the K-DOCS factors¹⁷.

Although the 50-item K-DOCS has been verified as a reliable and valid scale, it has limitations that have led to the creation of shortened versions. Firstly, some items have been found inappropriate for certain populations, with Tan et al. removing four items and Kandemir and Kaufman removing three items due to low factor loadings in Malaysian¹⁸ and Turkish samples¹⁹, respectively. Additionally, completing 50 items takes participants more time to respond and could be burdensome, resulting in lowered administration efficiency²⁰. To address these issues, researchers have developed various shorter versions to facilitate more practical applications²¹, including a 20-item English version²⁰, a 42-item Turkish version²², and a 16-item Chinese version¹⁷.

The use of shortened creativity scales offers several advantages. They allow for more efficient data collection, reducing participant fatigue and potentially increasing response rates. Moreover, shorter scales provide greater flexibility in research designs, enabling researchers to examine the relationships between creativity and other variables more comprehensively within time-constrained studies. This increased efficiency and flexibility can lead to more diverse and nuanced investigations of creativity across various contexts and populations. Among these shortened versions, the 20-item short form of the K-DOCS (20-K-DOCS) was recently developed by Tan et al.²⁰ by selecting the four items with the highest factor loadings from each of the five dimensions of the original K-DOCS. In their study involving 1,409 Malaysian undergraduates, Tan et al. found that the 20-K-DOCS was a psychometrically equivalent short form of the original K-DOCS, maintaining its strong psychometric properties while offering a more concise assessment of creativity.

To date, the reliability and validity of several shortened versions, including the 16-item Chinese and 42-item Turkish versions, have been established in various studies^{17,22}. However, the 20-K-DOCS is relatively new and has primarily been validated with university student populations. Despite its demonstrated suitability for university students, the applicability of the 20-K-DOCS to mid-adolescents (aged 15–18) remains unknown. This gap in the literature presents an important area for further investigation, as understanding the psychometric properties of the 20-K-DOCS in younger populations could significantly enhance its utility in educational and developmental research. This focus is particularly crucial as adolescence, being a transitional phase from childhood to adulthood, is often seen as a critical period for the development of creativity^{4,23}. The importance of this period stems from two main factors: firstly, it is marked by significant improvements in cognitive abilities thought to be related to creativity, such as abstract reasoning and problem-solving, along with the gradual maturation of emotional and social skills²⁴; secondly, adolescents' creative identities are more susceptible to influence from emotions, relationships, and social experiences²⁵. The various values, criteria, and behaviors they encounter during this period can either support or suppress the development of their creative identity. Given the importance of this developmental stage, special attention needs to be paid to the creativity of this population to support its development effectively. Therefore, testing the suitability of the 20-K-DOCS for mid-adolescents is not only necessary but also potentially valuable for enhancing our understanding of creativity during this crucial developmental period.

Furthermore, another notable gap in the literature on the 20-K-DOCS is the scarcity of studies employing Rasch analysis to evaluate its psychometric properties. Item Response Theory (IRT), of which Rasch analysis is a specific model, and Classical Test Theory (CTT) are two primary approaches used to analyze and validate measurement scales in psychometrics. Embretson and Reise highlighted a key distinction between these

approaches²⁴. While CTT focuses on overall test scores, IRT provides a more nuanced perspective by considering the interaction between individual item characteristics and respondent traits²⁶. Specifically, IRT posits that the probability of a correct response to an item is a function of both the respondent's ability and the item's difficulty. This relationship allows for a more precise estimation of respondent abilities and item properties.

Given that Rasch analysis, a form of IRT, offers a more comprehensive understanding of a scale's functioning at the item level, it provides detailed information on item difficulty and fit, yielding psychometric insights unavailable through CTT alone²⁷. Additionally, it allows for the examination of scale unidimensionality, a crucial assumption in psychological measurement²⁸. Surprisingly, this approach has seldom been used to assess the K-DOCS or its shortened versions. In contrast, other creativity measures, such as the Creative Achievement Questionnaire²⁹ and the Academic Writing Creativity scale³⁰, have benefited from Rasch analysis, providing valuable insights into their psychometric properties. The lack of Rasch analysis in K-DOCS literature represents a significant opportunity for future research. To provide a more comprehensive picture of the 20-K-DOCS' psychometric properties and further establish its validity, it is necessary to extend the analysis beyond CTT and employ Rasch analysis. This approach would allow researchers to gain a more nuanced understanding of the scale's functioning and its suitability for assessing creativity across various domains.

To address the aforementioned knowledge gaps, this study aims to conduct a comprehensive evaluation of the psychometric properties of the 20-K-DOCS utilizing both CTT and Rasch analysis, while concurrently assessing its concurrent and known-groups validity. The investigation of concurrent validity will be pursued through an examination of the scale's relationship with the Creative Self-Efficacy (CSE) scale. This approach is predicated on the understanding that creative self-efficacy—defined as an individual's belief in their creative potential—serves as a critical driver of creative motivation¹². This line of inquiry is supported by Sternberg's⁷ assertion that creativity development is inextricably linked to motivation, and Bandura's³¹ postulation that a robust sense of self-efficacy is fundamental to creative productivity and the discovery of new knowledge. Extant empirical studies on the K-DOCS have consistently revealed significant correlations between creative self-efficacy and creativity, with correlation coefficients ranging from 0.23 to 0.48²⁰, 0.30 to 0.46³², and 0.25 to 0.41¹⁷, thus underscoring the intrinsic relationship between these constructs. Furthermore, the study will investigate the known-groups validity of the K-DOCS scale, a crucial indicator of construct validity given the absence of a gold standard for creativity measurement. This methodological approach is validated when the scale demonstrates the ability to differentiate between groups known to vary on the variable under investigation³³. Through this multifaceted evaluation, incorporating CTT, Rasch analysis, concurrent validity assessment, and known-groups validation, this study endeavors to provide a comprehensive examination of the 20-K-DOCS, thereby contributing to the ongoing refinement and validation of creativity measurement instruments.

Method

Participants and data collection

To investigate the psychometric properties of the 20-K-DOCS among adolescents aged 15–18, a convenience sampling method was employed, beginning in January 2024. Assistance was sought from school administrators at two vocational secondary schools and two vocational high schools in a province of mainland China. To ensure adherence to ethical data collection procedures, the research purpose was initially presented to the classroom teachers of these schools through the administrators' arrangement. The teachers were then encouraged to discuss the purpose with the students' parents, either through face-to-face meetings or online chat groups. Electronic consent was obtained from parents who allowed their children to participate in the survey. Subsequently, participating students completed the online questionnaire in their school's computer lab.

The inclusion criteria for participation in the study were as follows: (i) absence of clinically diagnosed mental illnesses, and (ii) provision of electronic consent forms by parents. The exclusion criterion was based on responding too quickly (less than 400 s) and failing an attention test question, which required respondents to select their favorite color. If a respondent selected a non-relevant response instead of a color, they were excluded from the analysis. Ultimately, 3,079 respondents were included in the study's analysis. There were no missing data due to the platform's functionality, which required all responses to be complete before submission.

It needs to be clarified that rather than offering course credits or other direct incentives to students, our recruitment approach mainly relied on school administrators to encourage teachers to promote student participation in the questionnaire survey. This method aimed to increase voluntary involvement while reducing potential bias from external rewards. To show appreciation for their support, our research team provided educational benefits to the participating schools. Specifically, we arranged for our team members, who are experienced in vocational skills competitions, to deliver two lectures on competition strategies to all students in each school. Additionally, we offered personalized coaching sessions for student teams preparing for national-level competitions. This arrangement not only aided our data collection but also supported the educational goals of the participating schools, creating a mutually beneficial research partnership.

Table 1 presents the demographic information of the participants, revealing an average age of 16.48 years (SD = 0.93), with a majority being male (57.8%, $n = 1780$). Furthermore, a higher proportion were enrolled in vocational high schools (61.1%, $n = 1881$). The distribution of participants across different grades was relatively even, with the exception of a notably smaller number in the third year of vocational high school (12.0%, $n = 371$). Regarding their fields of study, participants were enrolled in a diverse range of programs, with engineering being the most common, representing 33.1% ($n = 1019$) of the participants. This aligns with the characteristics of vocational education. The study was approved by the Jiangxi Psychological Consultant Association (IRB ref: JXSXL-2022-CL15). Informed consent was obtained from all individual participants included in the study.

Variable	Mean (SD) or n (%)
Age	16.48 (0.93)
Sex	
Male	
Female	1780 (57.8)
School type	
Vocational Secondary School	1198 (38.9)
Vocational high school	1881 (61.1)
Grade	
Vocational Secondary School- Grade 1	699 (22.7)
Vocational Secondary School- Grade 2	701 (22.8)
Vocational high school Grade- Grade 1	682 (22.2)
Vocational high school Grade- Grade 2	626 (20.3)
Vocational high school Grade- Grade 3	371 (12.0)
Professional	
Engineering	1019 (33.1)
Science	381 (12.3)
Agriculture	376 (12.2)
Social Sciences (Economics, Management, Education, and Law)	591 (19.2)
Arts and Humanities	356 (11.6)
Others (Medicine, Military, Interdisciplinary)	356 (11.6)

Table 1. Participant characteristics (*N* = 3,079).

Instruments

20-item Kaufman domains of creativity scale

The 20-K-DOCS is a shortened version of the K-DOCS. As mentioned earlier, this version was proposed by Tan et al.²⁰ to select the four highest factor loadings from the original K-DOCS. Therefore, the oblique five-factor structure is consistent between the shortened version and the original version. The scale includes five domains: Self/Everyday, Scholarly, Performance (encompassing writing and music), Mechanical/Scientific, and Artistic, with 4 items per domain, totaling 20 items. The instructions are as follows: “Compared to people of your age and life experience, how would you rate your creativity in the following behaviors? For behaviors you have not specifically engaged in, assess your creative potential based on your performance in similar tasks. The scale employs a 5-point Likert scale, ranging from 1 (very little creativity) to 5 (very creative). In the study by Tan et al.¹⁵, the Cronbach’s alpha coefficients for the five domains of the scale were 0.73 (Self/Everyday), 0.72 (Scholarly), 0.87 (Performance), 0.84 (Mechanical/Scientific), and 0.81 (Artistic), respectively.

Creative self-efficacy scale

In the present study, we utilized the self-belief in creative thinking strategies subscale from the CSE scale to assess participants’ creative self-efficacy. This subscale was chosen not only because of the absence of items requiring reverse scoring, which could potentially lead to confusion in the analysis, but more importantly, the highest Cronbach’s alpha internal consistency coefficient (0.83) of this subscale among the three³⁴. The other two subscales of the CSE scale were included in the online survey; however, they were not analyzed as they contained items that needed to be reversed. The self-belief in creative thinking strategies subscale, developed by Hung³⁵, comprises four items. Representative items include: “When facing novelty problems, I believe I can imagine a lot of solutions very quickly” and “When facing a challenging task, I believe I can imagine a lot of related knowledge.” Responses were recorded using a 4-point Likert scale, ranging from 1 (strongly disagree) to 4 (strongly agree). In Hung’s³⁴ original study, the internal consistency of the four items, as measured by Cronbach’s alpha coefficient, was found to be 0.83 as presented above. In the current study, the McDonald’s omega for the CSE scale was 0.93, demonstrating high reliability.

Data analysis

The descriptive statistics for the 20-K-DOCS, along with the Pearson correlations among its subscales and their association with CSE scale were presented. In addition to basic analyses, this study conducted a multivariate analysis of variance (MANOVA) to compare 20-K-DOCS subscale levels among participants based on their experience in Vocational Schools Skills Competitions, testing the scale’s known-groups validity. The MANOVA was chosen to simultaneously analyze all dependent variables (K-DOCS subscales), accounting for potential interrelationships among them. Participants were categorized by their highest level of competition: national, provincial, municipal, school, or no participation. This competition, a significant event in China’s vocational education system with nearly 90 categories, progressively evaluates participants’ work from school to national levels.

The known-groups validity test is based on the hypothesis that participants with higher-level competition experience, particularly at the national level, should demonstrate higher 20-K-DOCS scores compared to those

with less or no experience. This expectation stems from creativity being a key factor in competition success, the progressive selection process favoring creative individuals, and the potential for competitions to foster creative skill development and enhance creative self-efficacy. If confirmed, this would provide evidence for the 20-K-DOCS scale's ability to differentiate between groups theoretically expected to have varying levels of creativity, thus partially supporting its construct validity. This approach is particularly valuable given the absence of a gold standard measure for creativity.

Subsequent analyses employed multidimensional Rasch modeling to examine the 20 K-DOCS items, accounting for the correlations among the five subscales. This approach allows for the simultaneous calibration of all subscales, utilizing the inter-subscale correlations to enhance measurement precision. The initial phase of analysis compares unidimensional and multidimensional Rasch models. Within this framework, we assess whether the Partial Credit Model (PCM) or the Andrich Rating Scale Model (RSM) is more appropriate for the data. Following this, the unidimensionality of each subscale and the monotonicity of response categories were evaluated. To assess unidimensionality, we applied the criterion that the eigenvalue of the first component in the principal component analysis (PCA) of the residuals should be less than 2.00³⁶. Monotonicity was examined by calculating the category boundary parameters for each of the five subscales.

Having met the prerequisites for Rasch analysis, item validity was evaluated using information-weighted (INFIT) and outlier-sensitive (OUTFIT) fit statistics, with acceptable ranges for valid item responses set between 0.50 and 1.50³⁷. The multidimensional Wright map was utilized to assess item targeting, offering a comprehensive visual representation of items from all five subscales in a single display. A person separation reliability threshold of 0.50, as suggested by Boone et al.³⁸, was applied to examine the scale's ability to differentiate participants with high and low creativity traits.

Following item diagnostics, model fit assessment focused on factorial, convergent, and discriminant validity using a CFA approach. This included testing several models: a unidimensional model, an oblique five-factor model, a nine-factor model (based on the original K-DOCS structure), a second-order model, a bifactor model (comprising one general factor and five specific factors), an Exploratory Structural Equation Modeling (ESEM) five-factor model, and an ESEM bifactor model. It should be noted that while the nine-factor model was included for comparison with the original K-DOCS structure, it was expected to be psychometrically problematic for the 20-item version due to having factors with too few items.

In addition to the common structures of second-order model and oblique five- and nine-factor models in K-DOCS, the rationale for incorporating both bifactor and ESEM approaches was to provide a comprehensive assessment of the 20-K-DOCS' psychometric properties and underlying structure. The bifactor model, represented by a general factor and specific factors corresponding to the proportional structure of the 20-K-DOCS, was employed to address high factor correlations and determine the unique contributions of general and specific factors³⁹. Simultaneously, ESEM was utilized to overcome limitations of traditional CFA, particularly in avoiding inflated factor correlations⁴⁰. Unlike CFA, which constrains cross-loadings to zero, ESEM allows estimation of all possible cross-loadings while maintaining a confirmatory framework. This approach provides a more realistic representation of complex constructs, often yielding reduced and more accurate factor correlations, improved model fit, and a balance between exploratory flexibility and confirmatory hypothesis-testing⁴⁰. By allowing items to load on multiple factors, ESEM mitigates issues of misspecification and biased parameter estimates common in CFA, especially for multidimensional constructs like creativity. Both bifactor and ESEM methods are well-suited for the 20-K-DOCS, enabling a nuanced examination of its factor structure while avoiding pitfalls of overly restrictive modeling, thus allowing for a more comprehensive understanding of the scale's underlying structure.

Model fit was assessed using established criteria: Comparative Fit Index (CFI) and Non-Normed Fit Index (NNFI) values above 0.90, Root Mean Square Error of Approximation (RMSEA) below 0.06, and Standardized Root Mean Square Residual (SRMR) below 0.08, following Brown⁴¹. The mean- and variance-adjusted weighted least squares (WLSMV) estimation was employed, given the ordinal nature of the 20-K-DOCS, as it provides robust estimates for non-normal distributions typical in Likert scales⁴².

The selection of the most suitable model considered the enhanced model fit potentially afforded by bifactor, ESEM five-factor and ESEM bifactor models, as evidenced by notable improvements in RMSEA/SRMR and TLI/CFI indices surpassing 0.015 and 0.01, respectively^{43,44}. After identifying the model that adequately represented the factor structure, the internal reliability of the 20-K-DOCS subscales was assessed using McDonald's omega coefficients.

Upon selecting the best-fit model, convergent and discriminant validity were further examined based on standardized factor loadings and Average Variance Extracted (AVE). An item's loading significance and magnitude were used as indicators of convergent validity for the construct, deemed satisfactory with an AVE above 0.50, in line with Fornell and Larcker⁴⁵. Additionally, the 20-K-DOCS's multidimensional factor discriminant validity was evaluated, with sufficient discriminant validity established when the AVE for two factors exceeded the squared correlation between those factors⁴⁶.

In the subsequent phase of the CFA, we broadened our investigation to assess the concurrent validity of the 20-K-DOCS in relation to CSE, utilizing the superior model identified in the CFA phase. Employing SEM, we examined the path coefficients stemming from the latent constructs within the 20-K-DOCS towards CSE. The evaluative criteria applied to this model remained consistent with those established in the initial CFA, adhering to the same estimation procedures.

Results

Descriptive statistics, Pearson correlations, and internal reliability

Among the five subscales of the 20-K-DOCS, Everyday Creativity had the highest mean score ($M=13.81$, $SD=3.87$), followed by Scholarly Creativity ($M=13.16$, $SD=4.01$), Artistic Creativity ($M=13.03$, $SD=4.15$),

Performance Creativity ($M = 12.12$, $SD = 4.57$), and Mechanical/Scientific Creativity ($M = 12.07$, $SD = 4.46$). All subscales showed significant positive correlations with creativity self-efficacy ($r = .45$ to 0.69 , $p < .001$) (refer to Table 2).

A MANOVA was conducted to examine the difference in previous experience of attending the Vocational Schools Skills Competition on five subscales within the 20-K-DOCS. The overall MANOVA revealed a significant multivariate main effect for previous experience, Wilks' $\lambda = 0.99$, $F(20, 10183) = 1.76$, $p = .020$. Given the significance of the overall test, the univariate main effects were examined. Significant univariate main effects were found for Performance Creativity ($F = 4.29$, $p = .002$), Mechanical/Scientific ($F = 4.81$, $p < .001$), and Artistic categories ($F = 4.82$, $p < .001$), but not for Self/Everyday Creativity and Scholarly Creativity. Post hoc analyses using Games-Howell indicated that participants who had experience attending the national level scored higher than those who did not participate in the competition. For Performance Creativity, national level participants ($n = 78$, $M = 13.74$, $SD = 5.45$) scored higher than non-participants ($n = 2664$, $M = 12.00$, $SD = 4.57$). Similarly, for the Mechanical/Scientific category, national level participants ($n = 78$, $M = 13.77$, $SD = 5.48$) outperformed non-participants ($n = 2664$, $M = 11.95$, $SD = 4.44$). In the Artistic category, national level participants ($n = 78$, $M = 14.72$, $SD = 4.97$) also scored higher than non-participants ($n = 2664$, $M = 12.93$, $SD = 4.13$). These findings provide partial support for the known-groups validity of the 20-K-DOCS, a component of construct validity. The results indicate that individuals who demonstrated exceptional creativity by reaching the national level of competition scored higher in specific categories than those who did not participate, aligning with the theoretical expectation that groups with different levels of creative achievement should be distinguishable by the scale.

Multidimensional Rasch analysis

Table S1 in the supplementary materials compares unidimensional and multidimensional Rasch models using RSM and PCM. The results indicate that PCM demonstrated superior fit in both unidimensional and multidimensional contexts, with lower AIC and BIC values. Likelihood ratio tests confirmed significantly lower deviance for PCM compared to RSM in both cases ($p < .001$). Within the PCM framework, the multidimensional Rasch model outperformed its unidimensional counterpart, exhibiting lower AIC and BIC values. A further likelihood ratio test revealed significantly lower deviance for the multidimensional PCM ($p < .001$). These findings suggest that the multidimensional Rasch model using PCM provides the most appropriate fit to the data, supporting its use in subsequent analyses.

Unidimensionality of each subscale was confirmed by PCA of the residuals. First component eigenvalues ranged from 1.51 to 1.81 across the five subscales, all below the 2.00 threshold, indicating no significant secondary dimensions. Table 3 presents the thresholds for the response categories and reveals that several items did not demonstrate monotonic category increases. Specifically, items 1, 2, and 3 in the Subscale of Self/Everyday Creativity, as well as all items in the Subscales of Performance Creativity and Artistic Creativity, exhibited violations of monotonicity in steps 3 and 4. To address this issue, we adopted the approach of collapsing response categories 4 and 5 into a single category. The revised items demonstrated satisfactory monotonicity, and consequently, we utilized these modified items to conduct further analyses of factorial and concurrent validity.

Moreover, Table 3 also shows that the infit/outfit of all items fell within the acceptable range of 0.5 to 1.5, suggesting that each item possessed validity and provided optimal information for the trait being measured. Within each subscale, the items identified as the most challenging were item 2 for the Self/Everyday Creativity subscale, item 5 for the Scholarly Creativity subscale, and items 12, 13, and 19 for the Performance Creativity, Mechanical/Scientific Creativity, and Artistic Creativity subscales, respectively.

The multidimensional Wright Map analysis, as depicted in Fig. 1, reveals a distribution of items across different difficulty levels for each subscale. While there is a concentration of items near the zero point (the mean difficulty level), the map also shows some spread of item difficulties. This distribution indicates that the test items cover a range of ability levels, with many items appropriately targeted to correspond with the average ability of the participant group. However, the map also highlights measurement gaps, particularly at the extreme ends of the latent ability spectrum. This suggests that while the scale is generally well-calibrated for the cohort to which it was administered, there may be limitations in assessing individuals with very high or very low levels of creativity in each dimension. The presence of these gaps indicates potential areas where additional items might be needed to fully capture the breadth of the constructs being measured, especially at the extremes of the ability range.

The person separation reliability, quantified by the Expected A Posteriori (EAP) reliability coefficient, exceeded 0.85 for each subscale within the multidimensional Rasch model. This indicates that the scales demonstrate adequate precision to differentiate between varying levels of the latent traits for most of the study

	1	2	3	4	5	6	7
1. Total score of 20- K-DOCS	—						
2. Self/Everyday Creativity	0.85	—					
3. Scholarly creativity	0.92	0.86	—				
4. Performance creativity	0.92	0.67	0.80	—			
5. Mechanical/scientific creativity	0.91	0.66	0.76	0.84	—		
6. Artistic creativity	0.89	0.69	0.75	0.77	0.79	—	
7. CSE	0.73	0.63	0.67	0.63	0.65	0.69	—

Table 2. Pearson correlations between 20-K-DOCS and CSE scale ($N = 3079$). CSE = Creative self-efficacy; All $p < .001$.

	Difficulty	Infit MnSq	Outfit MnSq	Step 1	Step 2	Step 3	Step 4
Subscale of self/everyday creativity							
Item 1: Helping other people cope with a difficult situation	0.07	0.94	0.72	−1.15	−1.96	1.79	1.33
Item 2: Teaching someone how to do something	0.19	0.93	0.79	−1.52	−1.66	1.83	1.35
Item 3: Maintaining a good balance between my work and my personal life	−0.05	0.87	0.86	−1.43	−1.91	1.73	1.60
Item 4: Understanding how to make myself happy	−0.21	0.86	0.79	−1.13	−1.80	1.46	1.48
Subscale of scholarly creativity							
Item 5: Writing a nonfiction article for a newspaper, newsletter, or magazine	0.45	0.94	0.73	−2.19	−1.47	1.60	1.76
Item 6: Responding to an issue in a context-appropriate way	−0.12	0.92	0.83	−2.34	−2.06	2.13	2.27
Item 7: Arguing a side in a debate that I do not personally agree with	−0.28	0.86	0.85	−2.39	−2.00	2.02	2.38
Item 8: Being able to offer constructive feedback based on my own reading of a paper	−0.05	0.86	0.76	−2.45	−1.76	2.03	2.17
Subscale of performance creativity							
Item 9: Making up lyrics to a funny song	−0.34	0.94	0.67	−1.96	−1.57	1.94	1.59
Item 10: Composing an original song	0.04	0.93	0.78	−1.66	−1.51	2.00	1.17
Item 11: Spontaneously creating lyrics to a rap song	0.08	0.87	0.86	−1.65	−1.54	1.93	1.25
Item 12: Playing music in public	0.22	0.86	0.78	−1.66	−1.50	1.93	1.22
Subscale of mechanical/scientific creativity							
Item 13: Writing a computer program	0.23	0.94	0.62	−2.63	−1.82	1.90	2.55
Item 14: Solving math puzzles	−0.12	0.92	0.78	−2.91	−1.88	1.91	2.87
Item 15: Taking apart machines and figuring out how they work	−0.08	0.89	0.88	−2.76	−1.74	1.88	2.62
Item 16: Building something mechanical (like a robot)	−0.03	0.87	0.83	−2.62	−1.71	1.94	2.38
Subscale of artistic creativity							
Item 17: Sketching a person or object	0.16	0.88	0.67	−1.64	−1.46	1.64	1.46
Item 18: Doodling/drawing random or geometric designs	−0.16	0.87	0.79	−1.72	−1.56	1.71	1.57
Item 19: Making a scrapbook page out of my photographs	0.18	0.75	0.69	−1.73	−1.46	1.80	1.39
Item 20: Appreciating a beautiful painting	−0.18	0.73	0.56	−1.68	−1.63	1.69	1.61

Table 3. Item difficulties and response category boundary parameters for the five factors of K-DOCS-20. Expected A Posteriori (EAP) reliability for each subscale was as follows: Self/Everyday Creativity (0.93), Scholarly Creativity (0.93), Performance Creativity (0.85), Mechanical/Scientific Creativity (0.92), and Artistic Creativity (0.87).

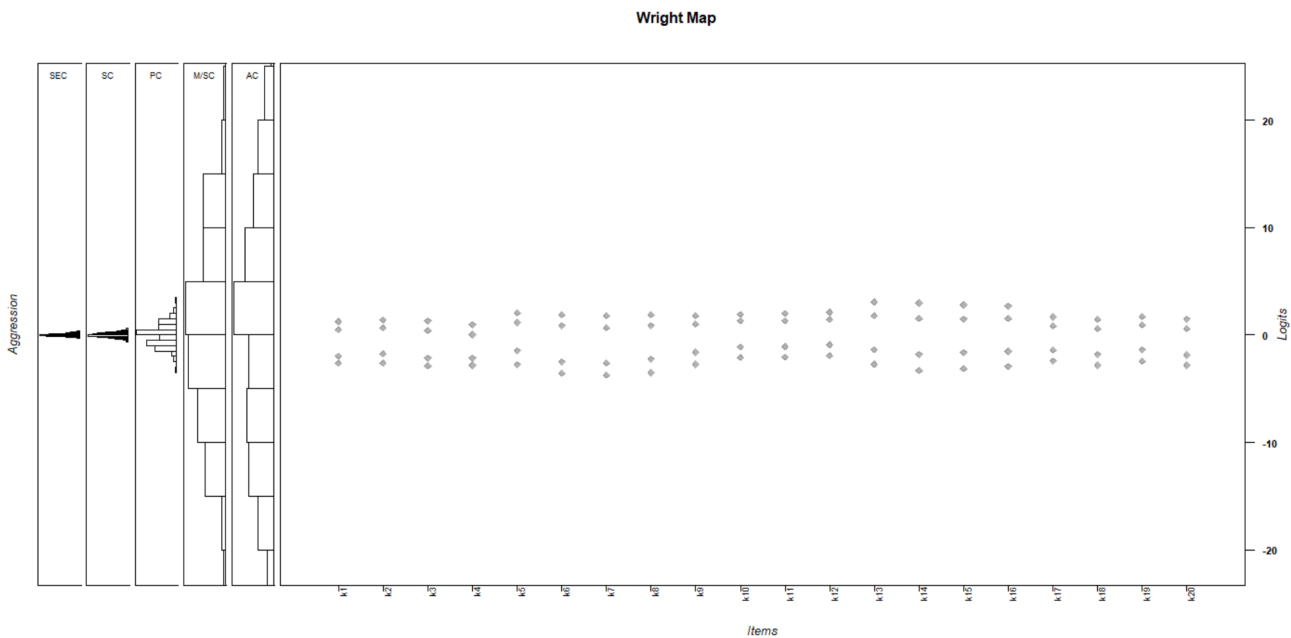


Fig. 1. Wright map of the five subscales for the 20-item K-DOCS. SEC, Self/everyday creativity; SC, Scholarly creativity; PC, Performance creativity; M/SC, Mechanical/scientific creativity; AC, Artistic creativity.

participants. However, in light of the Wright Map analysis, it is important to note that this high reliability may not extend to participants at the extreme ends of the ability spectrum.

Factorial, convergent, and discriminant validity

Table 4 provides a detailed summary of the model fit indices for various structural models using the revised items (those with category collapsing). These models include the one-factor model, the oblique five-factor model, nine-factor model, the second-order model, the bifactor model, ESEM five-factor model, and the ESEM bifactor model. The findings reveal that the oblique five-factor, the nine-factor, ESEM five-factor, and ESEM bifactor models all exhibited acceptable levels of model fit. In contrast, the one-factor, second-order, and bifactor models did not fit well with the data. It is particularly noteworthy that the improvement in model fit from the ESEM five-factor and ESEM bifactor models, as the differences between these models did not meet the predefined significance thresholds. Given the principle of parsimony, which prefers simpler models when the fit is similarly good, ESEM five-factor model appears to be a better choice than ESEM bifactor model. Most importantly, the factor correlations in the oblique five-factor model, which ranged from 0.71 to 0.93, were notably reduced in the ESEM five-factor model to a range of 0.34 to 0.69, highlighting the efficacy of the ESEM in addressing overly high factor correlations. Besides, although the nine-factor model fits better than the five-factor model, due to the limitation of having only a single item in two dimensions, there is no way to truly test their psychometric properties. Based on the above, the oblique five-factor model and ESEM five-factor model were identified as adequately representing the factor structure for further analysis. Moreover, the subscales demonstrated excellent internal consistency reliability in both the oblique five-factor and ESEM five-factor models. The McDonald’s omega coefficients ranged from 0.91 to 0.95 for the oblique five-factor model and from 0.80 to 0.83 for the ESEM five-factor model, indicating strong reliability across both approaches.

In the oblique five-factor structure of the 20-K-DOCS, the factor loadings for each item within the subscales surpassed 0.80, indicating robust item-specific contributions to their respective factors. Subsequent calculations of AVE produced values ranging from 0.74 to 0.82, affirming a commendable level of convergent validity for the five-factor model of the 20-K-DOCS. Nevertheless, the pronounced correlations among the latent variables challenge the discriminant validity within this structural framework.

Conversely, the ESEM outcomes, as detailed in Table 5, revealed that the majority of item-specific target loadings exceeded 0.50, with cross-loading issues being generally negligible, predominantly under 0.30, although a few exceptions were noted. Specifically, Items 4, 9, and 13 exhibited cross-loadings exceeding 0.30, which, while notable, were observed in only a small number of cases. This observation largely stipulated that convergent validity is substantiated by the magnitude of an item’s factor loadings on its designated factor surpassing those on non-associated factors, with the aforementioned exceptions warranting further consideration.

Concurrent validity

In the exploration of the concurrent validity of the 20-K-DOCS in relation to CSE, the SEM analysis, based on the oblique five-factor model, yielded the following results: χ^2 (df) = 3427.05 (237), CFI = 0.986, NNFI = 0.984, RMSEA = 0.066, and SRMR = 0.026. Notably, the factors of Artistic Creativity ($\beta = 0.39, t = 9.51, p < .001$), Scholarly Creativity ($\beta = 0.20, t = 2.13, p = .033$), and Mechanical/ Scientific Creativity ($\beta = 0.15, t = 3.29, p < .001$) demonstrated significant positive associations with CSE, while Self/Everyday Creativity and Performance Creativity showed no significant relationship (Refer to Figure S1 in the supplementary file).

The ESEM five-factor model (Refer to Figure S2 in the supplementary file) also revealed an exemplary model fit: χ^2 (df) = 835.29 (177), CFI = 0.997, NNFI = 0.996, RMSEA = 0.035, SRMR = 0.011. Within this model, all path coefficients were significantly positively correlated with CSE, in descending order of magnitude: Artistic Creativity ($\beta = 0.34, t = 13.43, p < .001$), Scholarly Creativity ($\beta = 0.21, t = 6.23, p < .001$), Mechanical/ Scientific Creativity ($\beta = 0.18, t = 7.85, p < .001$), Self/Everyday Creativity ($\beta = 0.17, t = 5.90, p < .001$), and Performance Creativity ($\beta = 0.09, t = 4.20, p < .001$).

Discussion

This study comprehensively assessed the psychometric properties of the 20-K-DOCS on a large sample of Chinese adolescents aged 15–18. The results demonstrated excellent internal consistency for each subscale, and satisfactory infit and outfit statistics supported the notion that the items within each subscale effectively

	χ^2 (df)	p-value	CFI	NNFI	RMSEA (90% Confidence Interval)	SRMR	AIC	BIC	Range of the factor correlations
One- factor CFA	11997.64 (170)	<0.001	0.934	0.927	0.150 (0.145–0.156)	0.068	12117.64	12375.87	Not applicable
Five-factor CFA	2990.80 (160)	<0.001	0.984	0.981	0.076 (0.070–0.082)	0.029	3130.80	3429.59	0.71 to 0.93
Nine-factor CFA	2646.959 (136)	<0.001	0.986	0.981	0.077 (0.070–0.085)	0.026	2834.96	3040.92	0.62 to 0.96
Second-order CFA	6238.94 (165)	<0.001	0.966	0.961	0.109 (0.104–0.115)	0.046	6368.94	6648.17	Not applicable
Bifactor CFA	5937.21 (150)	<0.001	0.968	0.959	0.112 (0.106–0.118)	0.044	6097.21	6456.00	Not applicable
ESEM Five-factor	600.02 (100)	<0.001	0.997	0.995	0.040 (0.029–0.052)	0.009	860.02	1408.48	0.34 to 0.69
ESEM Bifactor	369.02 (85)	<0.001	0.998	0.996	0.033 (0.020–0.047)	0.007	659.02	1268.59	Not applicable

Table 4. Model fit across various factor structures in the 20-K-DOCS. CFI, comparative fit index; NNFI, non-normed fit index; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual; BIC, Bayesian information criterion; AIC, Akaike information criterion.

	Self/Everyday creativity	Scholarly creativity	Performance creativity	Mechanical/scientific creativity	Artistic creativity
Self/everyday creativity					
Item 1: Helping other people cope with a difficult situation	0.71	0.08	0.08	0.06	0.08
Item 2: Teaching someone how to do something	0.60	0.15	0.12	0.07	0.07
Item 3: Maintaining a good balance between my work and my personal life	0.68	0.20	0.01	0.07	0.09
Item 4: Understanding how to make myself happy	0.49	0.31	−0.07	0.05	0.14
Scholarly creativity					
Item 5: Writing a nonfiction article for a newspaper, newsletter, or magazine	0.21	0.42	0.26	0.18	−0.01
Item 6: Responding to an issue in a context-appropriate way	0.28	0.53	0.06	0.09	0.09
Item 7: Arguing a side in a debate that I do not personally agree with	0.22	0.56	0.04	0.05	0.15
Item 8: Being able to offer constructive feedback based on my own reading of a paper	0.10	0.62	0.13	0.12	0.12
Performance creativity					
Item 9: Making up lyrics to a funny song	0.08	0.37	0.45	0.06	0.14
Item 10: Composing an original song	0.03	0.19	0.69	0.09	0.10
Item 11: Spontaneously creating lyrics to a rap song	0.07	0.10	0.71	0.13	0.09
Item 12: Playing music in public	0.09	0.02	0.58	0.28	0.13
Mechanical/scientific creativity					
Item 13: Writing a computer program	0.07	0.05	0.39	0.49	0.08
Item 14: Solving math puzzles	0.15	0.07	0.18	0.49	0.14
Item 15: Taking apart machines and figuring out how they work	0.06	0.10	0.07	0.75	0.09
Item 16: Building something mechanical (like a robot)	0.03	0.10	0.05	0.73	0.14
Artistic creativity					
Item 17: Sketching a person or object	0.03	0.09	0.09	0.29	0.56
Item 18: Doodling/drawing random or geometric designs	0.07	0.11	−0.01	0.12	0.72
Item 19: Making a scrapbook page out of my photographs	0.07	0.07	0.25	0.12	0.56
Item 20: Appreciating a beautiful painting	0.13	0.08	0.08	0.01	0.72

Table 5. Standardized factor loadings derived from ESEM models.

measured their respective factors. Furthermore, the person separation reliability for all subscales was greater than 0.80, suggesting that the scale can effectively differentiate between different levels of creativity facets. The five-factor structure was confirmed by both CFA and ESEM. Concurrent validity was also established, with positive associations found between the five facets of creativity in the K-DOCS and creative self-efficacy⁴⁷.

Notably, this study utilized multidimensional Rasch analysis to diagnose each item of the 20-item K-DOCS, a method rarely employed in previous psychometric studies of the K-DOCS. To our knowledge, only Susanto et al.⁴⁸ have applied Rasch analysis to the original K-DOCS using Indonesian students, though their sample size of less than 100 participants limited the reliability of their results. In our study, with a larger and more suitable sample size, we employed multidimensional Rasch analysis to evaluate the 20-item K-DOCS. While we confirmed the general validity of each item, our analysis also revealed nuanced insights into the scale's performance across different levels of the measured constructs. Our findings indicate that while the 20-K-DOCS demonstrates overall validity, it functions most effectively for individuals within the average to moderately high or low range of creative abilities. For participants at the extremes of the creativity spectrum in each dimension, the measurement precision may be reduced due to the lack of items targeting these exceptional ability levels. This insight, rather than contradicting the scale's validity, highlights an area for potential refinement. Further development of the 20-K-DOCS could focus on enhancing its ability to discriminate across the full spectrum of creative abilities in each dimension, particularly by adding items that better target very high and very low levels of creativity. Such improvements could increase the scale's overall measurement precision and utility, especially for assessing individuals with exceptional creative abilities.

Moreover, the unidimensional structure of the 20-item K-DOCS did not fit the data well, providing initial support for the multifaceted nature of creativity rather than a singular construct. Poor fit of the one-factor structure has also been found in previous studies of the K-DOCS (50-item English version^{15,16}; 16-item Chinese version¹⁷). Further supporting this perspective, our analyses using bifactor and ESEM bifactor models indicated that the addition of a general factor did not significantly improve model fit. Specifically, the improvements in model fit indices were minimal when a general factor was included, suggesting that the general factor lacks robustness. To our knowledge, this is the first time the bifactor approach has been used to assess the factor structure of the K-DOCS. Our findings echo those of Kapoor et al.¹⁶, who suggested that it is not recommended to summate K-DOCS scores on all items to obtain a unitary “creativity” score. In this vein, our findings favor the domain-specific perspective of creativity, highlighting the importance of considering distinct facets rather than an overall construct (i.e., domain-general). This finding is consistent with previous research, which has shown that creativity is better understood as a set of domain-specific factors rather than a single, overarching

construct^{9,10}. This aligns with the arguments of Kaufman and Baer⁴⁹, who proposed that creativity manifests differently across various domains and that a domain-specific approach is more appropriate.

However, it is important to note that the correlations between the creativity facets shown in the ESEM five-factor model in our study still had moderate effects, suggesting potential overlap among these distinct factors. This finding aligns with the recent work of Acar et al.⁵⁰, who tested the K-DOCS using a meta-analytic structural equation modeling approach. Their results supported both a general domain of creativity and a two-factor solution consisting of Arts and Sciences factors using a second-order CFA. Although our study diverges from Acar et al.'s findings in that the second-order CFA model did not fit well in our analysis, we still advocate for a conservative conclusion that considers both domain-general and domain-specificity. This perspective aligns more closely with the APT model of creativity⁸. Given that bifactor models and second-order CFAs are rarely reported in creativity research, we acknowledge that further evidence is needed to definitively establish the most appropriate structural model for creativity as measured by the 20-K-DOCS. This cautious approach recognizes the complexity of creativity as a construct and the need for continued investigation into its underlying structure.

Surprisingly, the factor correlations demonstrated excessively high correlations among the five creativity facets of the 20-item K-DOCS, contrary to most studies of the K-DOCS^{17,20,51}. It is worth considering whether the unique characteristics of the adolescent participants (including the age period and their enrollment in vocational education) in this study led to this exception, and future research should investigate this further, given that most studies of the K-DOCS have been conducted with university students or general adult populations⁵⁰. Despite the excessive correlations found in the zero-order correlations and traditional CFA, the application of ESEM in this study provides a novel approach. Previous research on the K-DOCS encountered cross-loading issues with items¹³ and resolved this by removing the problematic items. However, using ESEM, a more effective solution may be to retain the original items. The use of ESEM was particularly effective in addressing the issue of excessive correlations between factors, thereby enhancing discriminant validity. This aligns with the recommendations of Alamer and Marsh⁴⁴ and van Zyl and ten Klooster³⁹, who highlighted that ESEM can effectively mitigate high factor correlations that often compromise the discriminant validity of measurement models. In our study, the factor correlations in the oblique five-factor model ranged from 0.70 to 0.92 but were significantly reduced in the ESEM five-factor model to a range of 0.33 to 0.68. This reduction underscores the utility of ESEM in providing a clearer differentiation between the distinct facets of creativity. Furthermore, echoing previous studies^{15,19}, our analysis found that the nine-factor model provided an acceptable fit. However, due to the presence of single-item subscales in this structure, a comprehensive assessment was not feasible in the current study. This limitation presents an opportunity for future research to further investigate and validate the nine-factor structure using a version of the K-DOCS that includes multiple items for each subscale.

Lastly, the five creativity facets of the 20-item K-DOCS demonstrated significant associations with creative self-efficacy, with artistic creativity exhibiting the strongest correlation. This finding aligns with the results reported by Tan et al.²⁰, suggesting that the relationships between different domains of creativity and other creativity-related variables may vary. Moreover, our study provided evidence for known-groups validity. Students who had participated in national-level competitions demonstrated significantly higher scores in both mathematical and artistic creativity compared to those who had not engaged in such competitions. This observation further underscores the differential impact of creativity dimensions on individuals and their roles in creativity-related tasks. To further advance our understanding of this complex construct, future research should continue to explore and validate the domain-specific factors of creativity across diverse contexts, employing rigorous methodologies and considering potential moderating variables.

This study has important implications for future research on creativity both in theoretical and practical perspectives. Theoretically, through thoroughly evaluating the psychometric properties of the K-DOCS, it contributes significantly to the growing body of literature on this scale and provides valuable insights into the nature of creativity in Chinese adolescents in addition to university students. Particularly, the application of multidimensional Rasch analysis and ESEM offers a nuanced understanding of the scale's structure and item functioning. While our findings largely support the use of the 20-item K-DOCS as a reliable and valid measure of domain-specific creativity and underscore the importance of considering both domain-general and domain-specificity in creativity research, they also highlight areas that require further investigation. Specifically, the multidimensional Rasch analysis identified some items with non-monotonically increasing category thresholds, and the ESEM analysis revealed cross-loading issues for certain items. These findings suggest that while the 20-K-DOCS is a promising tool, there is room for refinement. Future research should address these issues, potentially through item revision or the development of additional items to enhance the scale's psychometric properties. Furthermore, continued exploration of the interplay between domain-general and domain-specific aspects of creativity could lead to a more comprehensive understanding of this complex construct.

Practically speaking, this research has important implications for practitioners using the 20-K-DOCS: First, the assessment tool is most reliable when evaluating individuals at the higher or lower ends of the creativity spectrum. Second, to obtain meaningful insights, practitioners should avoid using composite scores and instead analyze each creative domain independently, as this better captures domain-specific creative abilities and can better inform targeted interventions or development strategies.

Conclusions and limitations

This study provides robust evidence for the psychometric properties of the 20-item K-DOCS in a large sample of Chinese adolescents, supporting the multifaceted and domain-specific nature of creativity. Our findings demonstrate the scale's overall validity and reliability for measuring domain-specific creativity. The differential associations between creativity facets and creative self-efficacy, along with higher mathematical and artistic creativity scores among national-level competition participants, underscore the importance of domain-specific considerations in creativity research. However, our analysis also identified areas for potential improvement,

particularly in measuring extreme levels of creativity and in the functioning of some items (i.e., non-monotonic thresholds and cross-loading issues). These results suggest directions for future scale development and refinement to improve measurement across the full spectrum of creative abilities.

This study has some limitations that should be considered. Firstly, the sample consisted of Chinese adolescents aged 15–18, which may limit the generalizability of the findings to other age groups and cultural contexts. Particularly, the exceptionally high correlations among the creativity facets raise the question of whether the vocational school setting or the developmental stage of the participants contributed to this finding. Future studies should investigate the psychometric properties of the 20-item K-DOCS in diverse populations to determine if these high correlations persist across different contexts and age groups. Secondly, the criterion variable in this study was limited to creative self-efficacy, without extending to other relevant constructs. To establish a more comprehensive understanding of the 20-item K-DOCS's criterion validity, future research should examine its relationships with a broader range of creativity-related variables, such as creative achievement, divergent thinking, and personality traits associated with creativity. Thirdly, the 20-item version adopted in the current study has limited sensitivity at the extremes of the creativity spectrum. Future research could explore alternative item selection strategies or develop supplementary items to better capture exceptional levels of creativity, expanding the scale's applicability across a broader range of creative capacities. Lastly, this study relied on self-report measures, which may be subject to response biases, despite the inclusion of an attention check. Self-report assessments are inherently vulnerable to social desirability bias and other forms of response distortion, which could potentially influence the results. Future research could incorporate other assessment methods, such as performance-based tasks or observer ratings, to provide a more comprehensive and objective evaluation of creativity.

Data availability

The data generated for the present study are available from the corresponding authors on reasonable request.

Received: 22 June 2024; Accepted: 3 February 2025

Published online: 10 February 2025

References

- Nakano, T. D. C. & Wechsler, S. M. Creativity and innovation: Skills for the 21st Century. *Estud Psicol.* **35**, 237–246 (2018).
- Plucker, J. A., Beghetto, R. A. & Dow, G. T. Why isn't creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. *Educ. Psychol.* **39**(2), 83–96 (2004).
- Baer, J. *Domain Specificity of Creativity* (Academic, 2015).
- Kleibeuker, S. W., De Dreu, C. K. & Crone, E. A. Creativity development in adolescence: Insight from behavior, brain, and training studies. *New Dir. Child Adolesc. Dev.* **2016**(151), 73–84 (2016).
- Barbot, B. & Heuser, B. Creativity and identity formation in adolescence: A developmental perspective. In *The Creative Self* 87–98 (Academic, (2017).
- Giancola, M., Palmiero, M. & D'Amico, S. The green adolescent: The joint contribution of personality and divergent thinking in shaping pro-environmental behaviours. *J. Clean. Prod.* **417**, 138083 (2023).
- Sternberg, R. J. The nature of creativity. *Creat. Res.* **18**(1), 87–98 (2006).
- Baer, J. & Kaufman, J. C. Bridging generality and specificity: The amusement park theoretical (APT) model of creativity. *Roeper Rev.* **27**(3), 158–163 (2005).
- Baer, J. & Kaufman, J. C. The amusement park theoretical model of creativity: An attempt to bridge the domain-specificity/generality gap. In *The Cambridge Handbook of Creativity Across Domains* (eds. Kaufman, J. C., Glăveanu, V. P. & Baer, J.) 15–30 (Cambridge University Press, 2017).
- Kaufman, J. C. Counting the muses: Development of the Kaufman domains of creativity scale (K-DOCS). *Psychol. Aesthet. Creat. Arts* **6**(4), 298–308 (2012).
- Kaufman, J. C., Cole, J. C. & Baer, J. The construct of creativity: A structural model for self-reported creativity ratings. *J. Creat. Behav.* **43**(2), 119–134 (2009).
- McKay, A. S., Karwowski, M. & Kaufman, J. C. Measuring the muses: Validating the Kaufman domains of creativity scale (K-DOCS). *Psychol. Aesthet. Creat. Arts* **11**(2), 216–230 (2017).
- Miroshnik, K. G., Shcherbakova, O. V. & Kaufman, J. C. Kaufman domains of creativity scale: Relationship to occupation and measurement invariance across gender. *Creat. Res.* **34**(2), 159–177 (2021).
- Silvia, P. J., Wigert, B., Reiter-Palmon, R. & Kaufman, J. C. Assessing creativity with self-report scales: A review and empirical evaluation. *Psychol. Aesthet. Creat. Arts* **6**(1), 19–34 (2012).
- Kapoor, H., Reiter-Palmon, R. & Kaufman, J. C. Norming the muses: Establishing the psychometric properties of the Kaufman domains of creativity scale. *J. Psychoeduc. Assess.* **39**(6), 680–693 (2021).
- Kapoor, H., Zheng, M., Reiter-Palmon, R. & Kaufman, J. C. Toward equitable creativity self-assessment: Measurement invariance and mean differences across ethnicity in the Kaufman domains of creativity scale. *Psychol. Aesthet. Creat. Arts* **17**(4), 487–494 (2023).
- Zhang, S., Zhong, L., Zhang, Y., Ren, J. & Zhang, J. Development and validation of the Kaufman domains of creativity scale-chinese short version (K-DOCS-CSV). *Think. Skills Creat.* **48**, 101268 (2023).
- Tan, C. S., Tan, S. A., Cheng, S. M. & Hashim, I. H. M. Factor structure and psychometric qualities of the Kaufman domains of creativity scale. In *International Conference on Education and Psychology (ICEduPsy16)*, Sabah, Malaysia 389–400 (2016).
- Kandemir, M. A. & Kaufman, J. C. The Kaufman domains of creativity scale: Turkish validation and relationship to academic major. *J. Creat. Behav.* **54**(4), 1002–1012 (2020).
- Tan, C. S., Tan, S. A., Cheng, S. M., Hashim, I. H. M. & Ong, A. W. H. Development and preliminary validation of the 20-item Kaufman domains of creativity scale for use with Malaysian populations. *Curr. Psychol.* **40**, 1946–1957 (2021).
- Sen, S. & Yörük, S. A reliability generalization meta-analysis of the Kaufman domains of creativity scale. *J. Creat. Behav.* <https://doi.org/10.1002/jocb.620> (2023).
- Şahin, F. Kaufman Alanları Yaratıcılık Ölçeği'nin Türkçeye Uyarlanması ve Psikometrik Özelliklerinin İncelenmesi. *Ilkogretim Online* **15**(3), 1100–1115 (2016).
- Rothenberg, A. Creativity, mental health, and alcoholism. *Creat. Res.* **3**(3), 179–201 (1990).
- Crone, E. A. & Dahl, R. E. Understanding adolescence as a period of social-affective engagement and goal flexibility. *Nat. Rev. Neurosci.* **13**(9), 636–650 (2012).

25. van der Zanden, P. J., Meijer, P. C. & Beghetto, R. A. A review study about creativity in adolescence: Where is the social context? *Think. Skills Creat.* **38**, 100702 (2020).
26. Embretson, S. E. & Reise, S. P. *Item Response Theory for Psychologists* (Lawrence Erlbaum Associates, 2000).
27. Sbeglia, G. C. & Nehm, R. H. Do you see what I-SEA? A Rasch analysis of the psychometric properties of the inventory of student evolution acceptance. *Sci. Educ.* **103**(2), 287–316 (2019).
28. van der Wal, M. B. et al. Rasch analysis of the patient and observer scar assessment scale (POSAS) in burn scars. *Qual. Life Res.* **21**, 13–23 (2012).
29. Wang, C. C., Ho, H. C., Cheng, C. L. & Cheng, Y. Y. Application of the Rasch model to the measurement of creativity: The creative achievement questionnaire. *Creat. Res.* **26**(1), 62–71 (2014).
30. Mulyono, H., Ningsih, S. K. & Raufi, B. S. Validating the academic writing creativity and self-efficacy scale: A rasch model analysis. *Indones. Res. J. Educ. (IRJE)* **7**(1), 121–135 (2023).
31. Bandura, A. *Self-efficacy: The Exercise of Control* (Freeman, 1997).
32. Tu, C. P. & Fan, F. M. Verification and preliminary application of Chinese version of Kaufman domains of creativity scale. *Stud. Psychol. Behav.* **13**(6), 811–816 (2015).
33. Davidson, M. Known-groups validity. In *Encyclopedia of Quality of Life and Well-Being Research* (ed Michalos, A. C.) (Springer, 2014).
34. Hung, S. P. Validating the creative self-efficacy student scale with a Taiwanese sample: An item response theory-based investigation. *Think. Skills Creat.* **27**, 190–203 (2018).
35. Hung, S. P. *The Development and Examining the Psychometric Properties of Creative self-efficacy Student Scale* (2009).
36. Linacre, J. M. *A User's Guide to WINSTEPS: Rasch-Model Computer Programs* (Winsteps.com, 2016).
37. Jafari, P., Bagheri, Z. & Safe, M. Item and response-category functioning of the Persian version of the KIDSCREEN-27: Rasch partial credit model. *Health Qual. Life Outcomes* **10**(1), 127 (2012).
38. Boone, W. J., Staver, J. R. & Yale, M. S. *Rasch Analysis in the Human Sciences* (Springer, 2013).
39. van Zyl, L. E. & ten Klooster, P. M. Exploratory structural equation modeling: Practical guidelines and tutorial with a convenient online tool for Mplus. *Front. Psychiatry* **12**, 795672 (2022).
40. Marsh, H. & Alamer, A. When and how to use set-exploratory structural equation modelling to test structural models: A tutorial using the R package lavaan. *Br. J. Math. Stat. Psychol.* 1–18 (2024).
41. Brown, T. A. *Confirmatory Factor Analysis for Applied Research* (Guilford, 2015).
42. Brauer, K., Ranger, J. & Ziegler, M. Confirmatory factor analyses in psychological test adaptation and development. *Psychol. Test. Adapt. Dev.* **4**, 4–12 (2023).
43. Alamer, A. Exploratory structural equation modeling (ESEM) and bifactor ESEM for construct validation purposes: Guidelines and applied example. *Res. Methods Appl. Linguist.* **1**(1), 100005 (2022).
44. Alamer, A. & Marsh, H. Exploratory structural equation modeling in second language research: An applied example using the dualistic model of passion. *Stud. Second Lang. Acquis.* **44**(5), 1477–1500 (2022).
45. Fornell, C. & Larcker, D. F. Evaluating structural equation models with unobservable variables and measurement error. *J. Mark. Res.* **18**(1), 39–50 (1981).
46. Hair, J. F., Risher, J. J., Sarstedt, M. & Ringle, C. M. When to use and how to report the results of PLS-SEM. *Eur. Bus. Rev.* **31**(1), 2–24 (2019).
47. Beghetto, R. A. Creative self-efficacy: Correlates in middle and secondary students. *Creat. Res.* **18**(4), 447–457 (2006).
48. Susanto, E., Novitasari, Y., Rakhmat, C., Hidayat, A. & Wibowo, S. B. Rasch model analysis of Kaufman domains of creativity scale (K-DOCS) to Indonesian students. *J. Phys. Conf. Ser.* **1114**(1), 012027 (2018).
49. Kaufman, J. C. & Baer, J. Sure, I'm creative—but not in mathematics! Self-reported creativity in diverse domains. *Empir. Stud. Arts.* **22**(2), 143–155 (2004).
50. Acar, S., Cevik, E., Fesli, E., Bozkurt, R. N. & Kaufman, J. C. Testing the domain specificity of creativity with Kaufman domains of creativity scale: A meta-analytic confirmatory factor analysis. *J. Creat. Behav.* **58**(1), 171–189 (2024).
51. Taylor, C. L. & Kaufman, J. C. Values across creative domains. *J. Creat. Behav.* **55**(2), 501–516 (2021).

Acknowledgements

We thank all the participants to take part in the study.

Author contributions

C.H.C: Conceptualization, Writing - original draft, Project administration. X.L.W: Data curation, Formal analysis. Writing - review & editing. Y.P.J.: Resources, Validation; Visualization. I.H.C: Resources, Investigation, Methodology, Data curation, Formal analysis, Supervision. All authors reviewed the manuscript.

Funding

This study was supported by a special grant from the Taishan Scholars Project [Project No. tsqn202211130].

Declarations

Competing interests

The authors declare no competing interests.

Ethics approval

The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of the Jiangxi Psychological Consultant Association (JXSL-2022-CL15).

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-88980-8>.

Correspondence and requests for materials should be addressed to X.-L.W. or I.-H.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025