Iranian Journal of Psychoeducational Assessment
University of Zanjan

Spring & Summer 2025. Vol 1. Issue 1

Research Paper

Unidimensional or Bidimensional Creativity? Examining the Bidimensional Model in the Iranian Context

Saeed Akbari Zardkhaneh[1*], Amir Hossein Daneshmand Kafteroudi[2], Jalil Fathabadi[3], Majid Salimi[4]

1. Assistant Professor, Department of Applied Psychology, psychology and Educational Science Faculty, Shahid Beheshti University, Tehran, Iran
2. MA student, Educational Psychology, Faculty of Psychology and Educational Science, Shahid Beheshti University, Tehran, Iran
3. Associate Professor, Department of Applied Psychology, psychology and Educational Science Faculty, Shahid Beheshti University, Tehran, Iran
4. PhD in Measurement and Measurement of Allameh Tabatabaei University and Head of assessment center in Nation Iran Oil Company, Tehran, Iran

Article info:
Received: 02.03.2025
Revised: 09.05.2025
Accepted: 21.05.2025

Keywords:
psychometrics, creativity assessment, measurement invariance, torrance tests of creative thinking

CC BY

Publisher: University of Zanjan

Abstract

This fundamental descriptive-survey study aimed to standardize the latest figural version of the Torrance Tests of Creative Thinking (TTCT) for the Iranian adult population and to evaluate the validity and reliability of the two-dimensional creativity structure (innovation and adaptation) in the Iranian cultural context. The sample comprised 241 employees from an industrial company, selected via simple random sampling, who completed Forms A and B of the figural TTCT. Mplus software was utilized for confirmatory factor analysis (CFA) and measurement invariance testing between the two forms, while R programming language was employed to calculate internal consistency coefficients (Alpha and Omega). CFA results indicated that the two-dimensional creativity model exhibited a better fit than the unidimensional model, with the 'resistance to premature closure' component loading solely on the adaptation factor. The overall creativity index (Omega: 0.80–0.84) demonstrated higher reliability than individual factors. Furthermore, Form B showed better reliability and fit compared to Form A. Measurement invariance testing confirmed configural and metric invariance between the two forms, but scalar invariance was rejected. Overall, the findings validate the two-dimensional structure of creativity in the Iranian population and indicate high reliability for the overall creativity index. Form B is recommended for scientific research and sensitive assessments due to its superior performance. Employing a two-dimensional scoring framework, instead of a unidimensional index, allows for a more precise analysis of creativity profiles and can be effective in educational and organizational policies for fostering creativity.

Use your device to scan and read the article online

Citation: Akbari Zardkhaneh, S., Daneshmand Kafteroudi, A. H., Fathabadi, J., & Salimi, M. (2025). Unidimensional or bidimensional creativity? Examining the bidimensional model in the Iranian context. *Iranian Journal of Psychoeducational Assessment*, 1(1), 38–55.
https://doi.org/10.30470/ijpa.2025.726645

***Corresponding Author:** Saeed Akbari Zardkhaneh
Address: Department of Applied Psychology, Faculty of Psychology and Educational Science, Shahid Beheshti University, Tehran, Iran
Email: s_akbari@sbu.ac.ir

Extended Abstract

Introduction

Despite decades of research to develop reliable assessment tools, measuring creativity remains a persistent challenge (Baer & McKool, 2009; Batey & Yoon, 2017; Shao et al., 2019). The Torrance Tests of Creative Thinking (TTCT; Torrance, 1998) are among the most popular and extensively used creativity measures (Kaufman et al., 2007; Plucker et al., 2011). The figural version is most widely used due to its lower dependence on language abilities and greater comprehensiveness and validity (Kim, 2006a; Yoon, 2017; Kim, 2017). TTCT scoring has evolved from its original four components (Fluency, Flexibility, Originality, Elaboration) to the current version, which adds Abstractness of Titles and Resistance to Premature Closure, and removes Flexibility due to its high correlation with Fluency (Hebert et al., 2002). While the general Creativity Index remains widely used, reliance solely on this unidimensional score has been cautioned against (Torrance, 1966, 1974; Treffinger, 2004; Said-Metwaly et al., 2018; Acar et al., 2024). Cultural context may influence the prominence of these dimensions, with innovation more valued in individualistic (Western) and adaptation in collectivistic (Eastern) societies (Nijstad et al., 2010; Xie & Paik, 2019; Adair & Xiong, 2018). Therefore, validating TTCT's factor structure within specific cultures is essential (Kim, 2006a). In Iran, Studies using the figural TTCT often relied on outdated manuals (Pirkhaefi, 1373/1994; Karami, 1378/1999) and did not examine its factor structure. Recent efforts (Saeedi & Shakiba, 1400/2021; Tavana et al., 1401/2022; Samadi et al., 1402/2023) have begun to address norming, but none have validated the latest scoring criteria or bidimensional model. This study aimed (1) to standardize the latest figural TTCT (Forms A and B) for Iranian adults and (2) to investigate the construct validity of creativity in Iran, specifically testing the bidimensional model.

Methods

The population consisted of all employees of an industrial company assessed in 1403 (2024–2025 CE). From all completed answer sheets, 241 were selected via simple random sampling (127 for Form A, 114 for Form B). The sample included 14% women, ages 30–55 (M = 43.41, SD = 6.20). The figural TTCT Forms A and B (Torrance, 1974) were administered, each with three activities (10 minutes per activity). Scoring followed the latest TTCT manual (Torrance, 2018), adapted for Iran by Daneshmand Kafteroudi et al. (1403/2024). Components scored were Fluency, Originality, Elaboration, Abstractness of Titles, and Resistance to Premature Closure. Originality was scored using an Iranian list of non-creative responses (Akbari-Zardkhaneh et al., under review), following the criterion that responses in 5% or more of the sample (Activities 1 & 2) or 20% (Activity 3A)/10% (Activity 3B) are non-original (Zimmerman, 2018, as cited in Acar et al., 2021). Flexibility was also scored for comparison. Confirmatory factor analysis (CFA) was conducted in Mplus 8.0 (Muthén & Muthén, 1998–2021; Asparouhov & Muthén, 2009), testing four models: unidimensional (Torrance, 1998) and three bidimensional models varying the placement of Resistance to Premature Closure (Said-Metwaly et al., 2018). Measurement invariance (configural, metric, scalar) was tested (Putnick & Bornstein, 2016). Due to multivariate non-normality, robust maximum likelihood estimation (MLR; Satorra & Bentler, 1994) was used. Reliability was assessed via Cronbach's Alpha (Cronbach, 1951) and McDonald's Omega (McDonald, 1999) in R ('psych' package; Revelle, 2023).

Results

Descriptive statistics indicated all components except Abstractness of Titles had skewness and kurtosis values within ±2.0 (Gravetter & Wallnau, 2014). Fluency and Flexibility were highly correlated ($r = .91$, $p < .001$). CFA results showed the best fit for Model 3: a bidimensional model with Fluency and Originality on Innovation, and Elaboration, Abstractness of Titles, and Resistance to Premature Closure on Adaptation (CFI = .98, TLI = .95, RMSEA = .08, SRMR = .04). For Form B, Model 3 showed exceptionally good fit (CFI = 1.00, TLI = 1.00, RMSEA = .000, SRMR = .021). Standardized factor loadings were: Fluency (0.97) and Originality (0.61) on Innovation; Elaboration (0.85), Abstractness of Titles (0.32), and Resistance to Premature Closure (0.39) on

Adaptation. The two factors had a strong covariance (0.76).
Measurement invariance analysis supported configural and metric invariance (ΔCFI = -.007, ΔRMSEA = .007), but not scalar invariance (ΔCFI = .002). Reliability coefficients for the combined sample were: Innovation (ω = .75), Adaptation (ω = .58), Overall Index (ω = .82). Form B showed higher reliability: Innovation (ω = .85), Adaptation (ω = .64), Overall Index (ω = .84). For Form A: Innovation (ω = .61), Adaptation (ω = .52), Overall Index (ω = .80).

Discussion

The study supports the bidimensional model of TTCT (Said-Metwaly et al., 2018; Acar et al., 2024) in Iranian adults, with strong evidence for separating Innovation and Adaptation. The strong correlation between Fluency and Flexibility supports their redundancy (Torrance & Ball, 1984). The loading of Resistance to Premature Closure on Adaptation aligns with meta-analyses (Said-Metwaly et al., 2018; Acar et al., 2023), though some studies differ (Humble et al., 2018; Yoon, 2017). The Innovation factor (Fluency, Originality) reflects rapid novel idea generation (Heausler & Thompson, 1988; Kirton, 1976), while Adaptation (Elaboration, Abstractness of Titles, Resistance to Premature Closure) may reflect persistence and task commitment (Lubart & Zenasni, 2010).

This study provides robust evidence for a two-dimensional structure (Innovation and Adaptation) of the figural TTCT in Iranian adults. The Overall Creativity Index and Form B demonstrated high reliability, supporting Form B's use for scientific and high-stakes purposes in Iran. Bidimensional scoring offers more precise creative profiling and can inform educational and organizational strategies to foster creativity.

Ethical Considerations

All individuals who took part in this research provided informed consent by completing a designated form. The researchers ensured that all data collected throughout the study would be treated with strict confidentiality and used solely for academic purposes.

Funding

All financial resources and costs for research and publication of the article have been paid by the authors and no financial support has been received.

Authors' Contributions

This article is extracted from a Master's thesis. The corresponding author served as the thesis supervisor. The second author is the Master's student and main executor of the thesis. The third author was the thesis advisor, and the fourth author, as the organization's director, coordinated the data collection process.

Conflicts of interest

The authors declare that they have no conflicts of interest.

مقاله پژوهشی

خلاقیت تک‌بعدی یا دوبعدی؟ بررسی مدل دوبعدی در زمینه ایرانی

سعید اکبری‌زردخانه[۱*]، امیرحسین دانشمند کفترودی[۲]، جلیل فتح‌آبادی[۳]، مجید سلیمی[۴]

۱. استادیار گروه روانشناسی کاربردی، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه شهید بهشتی، تهران، ایران
۲. دانشجوی کارشناسی ارشد روان‌شناسی تربیتی، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه شهید بهشتی، تهران، ایران
۳. دانشیار گروه روانشناسی کاربردی، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه شهید بهشتی، تهران، ایران
۴. دکتری سنجش و اندازه‌گیری، دانشگاه علامه طباطبایی و رئیس مرکز ارزیابی و سنجش کارکنان شرکت ملی نفت ایران، تهران، ایران

اطلاعات مقاله:
تاریخ دریافت: ۱۴۰۳/۱۲/۱۲
تاریخ داوری: ۱۴۰۴/۰۲/۱۹
تاریخ پذیرش: ۱۴۰۴/۰۲/۳۱

واژه‌های کلیدی:
روانسنجی، سنجش خلاقیت، تغییرناپذیری اندازه‌گیری، آزمون تفکر خلاق تورنس

CC BY

ناشر: دانشگاه زنجان

چکیده

این پژوهش بنیادی با هدف استانداردسازی جدیدترین نسخه تصویری آزمون تفکر خلاق تورنس برای جمعیت بزرگسال ایرانی و ارزیابی روایی و اعتبار ساختار دوبعدی خلاقیت (شامل نوآوری و انطباق) در بافت فرهنگی ایران، به شیوه توصیفی-پیمایشی اجرا شد. نمونه آماری شامل ۲۴۱ نفر از کارکنان یک شرکت صنعتی بود که به روش نمونه‌گیری تصادفی ساده انتخاب شدند و فرم‌های A و B نسخه تصویری آزمون تورنس را تکمیل کردند. به منظور اجرای تحلیل عاملی تأییدی و بررسی هم‌ارزی اندازه‌گیری میان دو فرم از نرم‌افزار امپلاس و برای محاسبه ضرایب همگونی درونی آلفا و امگا از زبان برنامه‌نویسی R استفاده شد. نتایج تحلیل عاملی تأییدی نشان داد که مدل دوبعدی خلاقیت نسبت به مدل تک‌بعدی از برازش بهتری برخوردار است و مؤلفه عدم نتیجه‌گیری زودهنگام تنها بر عامل انطباق بارگذاری شد. شاخص کلی خلاقیت (امگا: ۰/۸۰ تا ۰/۸۴) از اعتبار بالاتری نسبت به عوامل جداگانه برخوردار بود. همچنین، فرم B در مقایسه با فرم A اعتبار و برازش بهتری نشان داد. بررسی هم‌ارزی اندازه‌گیری، هم‌ارزی پیکربندی و متریک بین دو فرم را تأیید کرد، اما هم‌ارزی اسکلار رد شد. در مجموع، یافته‌ها تأییدکننده ساختار دوبعدی خلاقیت در جمعیت ایرانی است و نشان می‌دهد شاخص کلی خلاقیت از اعتبار بالایی برخوردار می‌باشد. فرم B به دلیل عملکرد مطلوب‌تر برای پژوهش‌های علمی و ارزیابی‌های حساس توصیه می‌شود. استفاده از چارچوب امتیازدهی دوبعدی به‌جای شاخص تک‌بعدی، امکان تحلیل دقیق‌تر نیمرخ خلاقیت را فراهم کرده و می‌تواند در سیاست‌گذاری‌های آموزشی و سازمانی برای پرورش خلاقیت مؤثر باشد.

استناد: اکبری زردخانه، س.، دانشمند کفترودی، ا. ح.، فتح‌آبادی، ج و سلیمی، م. (۱۴۰۴). خلاقیت تک‌بعدی یا دوبعدی؟ بررسی مدل دوبعدی در زمینه ایرانی. *مجله ایرانی سنجش روانی-تربیتی*، ۱(۱)، ۳۸-۵۵.
ttps://doi.org/10.30470/ijpa.2025.726645

* **نویسنده مسئول:** سعید اکبری‌زردخانه
نشانی: گروه روانشناسی کاربردی، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه شهید بهشتی، تهران، ایران
پست الکترونیکی: s_akbari@sbu.ac.ir

مقدمه

از خلاقیت، به عنوان عاملی کلیدی در تسهیل هماهنگی اجتماعی، توسعه پایدار انسانی، اختراعات فنّاورانه و انقلاب علمی، در سطوح مختلف فعالیت‌های انسانی نام برده می‌شود (شاو[1] و همکاران، ۲۰۱۹). با این‌که در طول چند دهه گذشته شاهد گسترش مطالعات حوزه خلاقیت برای جستجوی یک ابزار اندازه‌گیری قابل‌اعتماد بوده‌ایم، چالشی بودن اندازه‌گیری خلاقیت همواره مورد تأیید محققان بوده است (بائر و مک کول[2]، ۲۰۰۹؛ باتی، یون[3]، ۲۰۱۷؛ شاو و همکاران، ۲۰۱۹). در این میان، آزمون تفکر خلاق تورنس[4] (تورنس، ۱۹۹۸) یکی از محبوب‌ترین و پرکاربردترین ابزارهای سنجش خلاقیت است (کافمن[5] و همکاران، ۲۰۰۷؛ پلاکر[6] و همکاران، ۲۰۱۱) که به بیش از ۳۲ زبان ترجمه شده است (فریزر[7]، ۱۹۹۰).

محبوبیت تاریخی آزمون تورنس تا حد زیادی ناشی از دهه‌ها پژوهش در مورد ویژگی‌های روانسنجی آن نظیر روایی پیش‌بین (تورنس ۱۹۷۴؛ هبرت[8] و همکاران، ۲۰۰۲؛ تورنس، ۱۹۸۱) و روایی سازه (کلاپام[9]، ۱۹۹۸؛ همبل[10] و همکاران، ۲۰۱۸؛ کیم[11]، ۲۰۰۶ ب؛ کیم و همکاران، ۲۰۰۶؛ کروم[12] و همکاران، ۲۰۱۴، ۲۰۱۶؛ یون[13]، ۲۰۱۷) بوده است. سال‌ها پژوهش در زمینه آزمون تفکر خلاق تورنس منجر به توسعه سه نسخه مختلف شده است: نسخه تصویری[14]، نسخه کلامی[15] و نسخه مختصر[16]. از میان این سه، نسخه تصویری بیشترین استفاده را در پژوهش، آموزش و کاربرد داشته است (لی[17] و همکاران، ۲۰۲۴). نسخه تصویری آزمون تفکر خلاق تورنس نسبت به نسخه کلامی[18] کم‌تر تحت تاثیر توانایی زبانی شرکت‌کننده قرار می‌گیرد (کیم، ۲۰۰۶ الف؛ یون، ۲۰۱۷). همچنین با اینکه نسخه تصویری و کلامی کاملاً مرتبط هستند؛ اما نسخه تصویری، جامع‌تر، معتبرتر و رواتر عمل کرده است (کیم، ۲۰۱۷). هر دو فرم موازی A و B نسخه تصویری از سه فعالیت تشکیل شده است: ۱. ساخت تصویر[19]، ۲. تکمیل تصویر[20] و ۳. تصاویر (خطوط یا دایره) تکرارشده[21] (تورنس و هنزلی[22]، ۲۰۰۳).

اصل آزمون نسخه‌ی تصویری تورنس تاکنون بدون تغییر باقی‌مانده است (کیم، ۲۰۰۶ الف). اولین نسخه آزمون تورنس (۱۹۶۶) دارای چهار مؤلفه سیالی[23]، انعطاف‌پذیری[24]، اصالت[25] و بسط[26] بود؛ اما در آخرین نسخه دو مؤلفه انتزاع عنوان‌ها[27] و عدم نتیجه‌گیری زودهنگام[28] نیز به آن اضافه شده و مولفه‌ی انعطاف‌پذیری به دلیل ارتباط قوی با سیالی حذف شده است (هبرت و همکاران، ۲۰۰۲)؛ اما شرکت خدمات آزمون‌های آموزشگاهی[29] (۲۰۲۳) به عنوان صاحب‌امتیاز آن، کماکان در گزارش‌ها انعطاف را درج می‌کند.

حاصل ترکیب مؤلفه‌های آزمون تورنس شاخص خلاقیت[30] است که معمولاً از آن در شناسایی افراد خلاق استفاده می‌شود. با این حال، پژوهشگران از جمله تورنس (۱۹۶۶؛ ۱۹۷۴)، نسبت به اتکای صرف به این شاخص هشدار داده‌اند (ترِفینگر[31]، ۲۰۰۴). علاوه بر این، فرضیه ساختار تک‌بعدی به‌عنوان زیربنای این شاخص کلی، به چالش کشیده شده است. از سوی دیگر فراتحلیل‌ها (سیدمتوالی[32] و همکاران، ۲۰۱۸؛ آکار[33] و همکاران، ۲۰۲۴) حاکی از تطابق بهتر مدلی دوبعدی و بارگذاری سیالی و اصالت در عامل نوآوری[34] و انتزاع عنوان‌ها و بسط در عامل انطباق[35] دارند. با این حال، در مورد مؤلفه عدم نتیجه‌گیری زودهنگام تناقض‌هایی وجود دارد (کیم، ۲۰۰۶ ب). در این راستا شواهد قابل توجهی نشان می‌دهد که در فرهنگ‌های غربی یا فردگرایانه[36] (زی و پیک[37]، ۲۰۱۹) عامل نوآوری و در مقابل در فرهنگ شرقی یا جمع‌گرا عامل انطباق دارای اهمیت بیشتری است (نیستاد[38] و همکاران، ۲۰۱۰؛ آدایر و شیونگ[39]، ۲۰۱۸). همچنین آرانگورن[40] (۲۰۱۴) در گروه جوانان آرژانتینی نشان داد که مولفه‌ی عدم نتیجه‌گیری زودهنگام در ذیل هر دو عامل انطباق و نوآوری قرار می‌گیرد؛ درحالی‌که کروم و همکاران (۲۰۱۶) در گروه کودکان اسپانیایی‌زبان، قرارگیری این مؤلفه در ذیل عامل انطباق را نتیجه‌گیری کرده و بیان کرده‌اند قبل از هر مطالعه خلاقیت در هر نمونه خاص، بررسی روایی سازه و ساختار عاملی آن برای سن و فرهنگ خاص ضروری است (کیم، ۲۰۰۶ الف).

---

1 Shao
2 Baer & McKool
3 Yoon
4 Torrance Tests of Creative Thinking (TTCT)
5 Kaufman
6 Plucker
7 Frasier
8 Hebert
9 Clapham
10 Humble
11 Kim
12 Krumm
13 Yoon
14 Figural (TTCT-F)
15 Verbal (TTCT-V)
16 Abbreviated (TTCT-A)
17 Lee
18 TTCT Verbal
19 Picture construction
20 Incomplete figures
21 Repeated figures
22 Haensly
23 Fluency
24 Flexibility
25 Originality
26 Elaboration
27 Abstractness of titles
28 Resistance to premature closure
29 Scholastic Testing Service
30 Creativity index (CI)
31 Treffinger
32 Said-Metwaly
33 Acar
34 Innovative
35 Adaptive
36 Individualistic
37 Xie & Paik
38 Nijstad
39 Adair & Xiong
40 Aranguren

تاریخچه به‌کارگیری آزمون تورنس در ایران با ابهام‌های جدّی (دانشمند کفترودی و همکاران، ۱۴۰۳) همراه است. به‌عنوان نمونه شمار بسیاری از مقالات منتشرشده که در بخش ابزار یا حتی عنوان خود به آزمون خلاقیت تورنس اشاره کرده‌اند، درواقع ابزار خلاقیت عابدی (۱۳۷۲) را به کار گرفته‌اند (برای مثال: رضایی و منوچهری، ۱۳۸۷؛ حمیدی و همکاران، ۱۳۹۱؛ بابایی و همکاران، ۱۳۹۴؛ رضوانی و برداران توکلی، ۱۴۰۳). پرسشنامه عابدی یک ابزار خودگزارشی مشتمل بر ۶۰ سؤال است که بر اساس شاخص‌های چهارگانه تورنس و در تهران ساخته شده است (دائمی و مقیمی‌بارفروش، ۱۳۸۳). طبیعتاً با توجه به تفاوت جدی در ماهیت آزمون نمی‌توان این دو آزمون را یکی دانست؛ ولو اگر بر اساس گروه مشابه‌ای از شاخص‌ها نمره‌گذاری شوند.

دو نسخه فارسی از آزمون تفکر خلاق تورنس در ایران وجود دارد. نخست، پیرخائفی (۱۳۷۳، به نقل از پیرخائفی و همکاران، ۱۳۸۸) اعتبار بازآزمایی فرم B نسخه تصویری را در گروه دانش‌آموزان و دانشجویان ۰/۸۰ گزارش کرده است؛ لذا در ایران معدود پژوهش‌هایی که از نسخه تصویری آزمون تورنس استفاده کرده‌اند، به این یافته‌ها اشاره داشته‌اند (برای مثال: اوبالاسی و حسینی‌نصب، ۱۳۹۳؛ ذاکری و برقی،۱۳۹۴؛ عباس‌پور و همکاران، ۱۴۰۰). در این پژوهش‌ها نمره‌گذاری بر اساس چهار شاخص اصالت، سیالی، انعطاف و بسط (براساس نسخه تورنس (۱۹۶۶) و قدیمی‌ترین نسخه) صورت گرفته است. در هیچ یک از این مطالعات نیز به ساختار عاملی این ابزار در جمعیت ایرانی اشاره‌ای نشده است. در تلاشی دیگر فرم A و دستورالعمل نمره‌دهی آن توسط ابوالفضل کرمی (۱۳۷۸) ترجمه و در برخی از پژوهش‌ها به کار گرفته شده است (برای مثال: سعیدی و شکیبا، ۱۴۰۰). این دستورالعمل نیز مربوط به ۱۹۷۰ است که در آن همان چهار شاخص نسخه اولیه مورد نظر بوده‌اند. از سوی دیگر در این ترجمه به منظور نمره‌گذاری اصالت از ۵ نقطه برش استفاده شده است که به نظر می‌رسد این شیوه نمره‌گذاری توسط خود تورنس (۱۹۶۶ یا ۱۹۷۴) معرفی نشده است. روشن است به دلیل ماهیت این اقدام، شواهد مبنی بر اعتبار و روایی آن در جمعیت ایرانی گزارش نشده است. با این وصف نمی‌توان بر دقت و صحت کلیه یافته‌های حاصل از پژوهش‌های انجام شده با این ابزارها مطمئن بود.

در مجموع می‌توان نتیجه گرفت که برخلاف ماهیت سازه خلاقیت که به نوعی بر مفهوم ابداع، نوآوری و به‌روز بودن دلالت دارد، در ایران ابزارهای سنجش خلاقیت از نوآوری و به‌روزرسانی به دور بوده‌اند؛ گواه آنکه هیچ یک از دستورالعمل‌های نمره‌گذاری نسخه تصویری آزمون تورنس که در پژوهش‌های داخلی مورد استفاده قرارگرفته‌اند، نسخه‌های اخیر این ابزار (۲۰۰۸، ۲۰۱۷ یا ۲۰۲۳) را مورد استفاده قرار نداده‌اند. گرچه در سال‌های اخیر پژوهشگرانی همچون سعیدی و شکیبا (۱۴۰۰) در جمعیت دانش‌آموزان، به لزوم تجدید نظر در طبقات پاسخ جمعیت ایرانی تأکید داشتند و تشکیل طبقات پاسخ بومی را در دستور کار قراردادند؛ با این حال هیچ‌کدام شاخص‌ها و شیوه‌های جدید نمره‌گذاری این آزمون را مورد مطالعه قرار ندادند. از سوی دیگر توانا و همکاران (۱۴۰۱) و صمدی و همکاران (۱۴۰۲) در اقدامی خلاقانه با هدف ایجاد نسخه‌ای بومی از آزمون تفکر خلاق تورنس، دو فرم A و B را ترکیب کرده و ابزاری شامل پنج فعالیت و ۲۳ محرک تدوین کردند و سپس به استانداردسازی آن در جمعیت دانش‌آموزان متوسطه و ابتدایی (به ترتیب) پرداختند؛ با این حال، در هیچ یک از این تلاش‌ها، شاخص‌های اخیر و سازه دوبعدی خلاقیت در جمعیت ایرانی مورد بررسی قرار نگرفته است.

با توجه به گسترش مطالعات حوزه خلاقیت و ضرورت نیاز به ابزاری معتبر، روا و استاندارد برای جمعیت ایرانی، این پژوهش با دو هدف انجام شد. نخست، استانداردسازی جدیدترین نسخه تصویری آزمون تفکر خلاق تورنس (فرم‌های A و B) برای جمعیت بزرگسال ایرانی؛ دوم، بررسی سازه خلاقیت در ایران، با تأکید بر پاسخ به این پرسش کلیدی که آیا ساختار دوبعدی خلاقیت مورد پذیرش در جوامع دیگر، در جمعیت ایرانی نیز معتبر است یا خیر.

روش پژوهش

این پژوهش از نظر هدف بنیادی و از نظر روش گردآوری داده توصیفی پیمایشی شمرده می‌شود (سرمد و همکاران، ۱۳۸۰). همچنین از حیث نوع داده‌ها، در دسته پژوهش‌های کمّی قرار می‌گیرد (بازرگان، ۱۴۰۰).

مشارکت‌کنندگان

جامعه آماری این پژوهش را کلیه کارکنان یک شرکت صنعتی تشکیل می‌دادند که در سال ۱۴۰۳ به مرکز ارزیابی آن شرکت مراجعه کرده و آزمون تفکر خلاق تورنس بر روی آن‌ها اجرا شده بود. پس از اجرای آزمون توسط مرکز ارزیابی بر روی تمامی مراجعین واجد شرایط در بازه زمانی مشخص، کلیه پاسخ‌برگ‌های تکمیل‌شده جمع‌آوری گردید. از میان این مجموعه کلی از پاسخ‌برگ‌ها، کارمندان مرکز ارزیابی با استفاده از روش نمونه‌گیری تصادفی ساده، تعداد ۲۴۱ پاسخ‌برگ را برای ورود به مطالعه حاضردر اختیار پژوهشگران قرار دادند. این فرآیند انتخاب تصادفی به این صورت انجام شد که ابتدا به هر پاسخ‌برگ یک کد منحصربه‌فرد اختصاص داده شد و سپس با استفاده از نرم‌افزار اکسل، نمونه نهایی به گونه‌ای انتخاب گردید که هر پاسخ‌برگ شانس برابری برای انتخاب شدن داشته باشد. در نهایت، از این تعداد، ۱۲۷ نفر فرم A و ۱۱۴ نفر فرم B را تکمیل کردند. ۱۴ درصد حجم نمونه را زنان تشکیل دادند و سن مشارکت‌کنندگان بین ۳۰ تا ۵۵ سال متغیر بود (M =۴۳/۴۱؛ sd =۶/۲۰).

ابزار پژوهش

در این قسمت ابزار جمع‌آوری داده‌ها معرفی شده است.

آزمون تصویری تفکر خلاق تورنس ـ نسخه جوانان ایرانی: مشارکت‌کنندگان به‌عنوان بخشی از مراحل ارزیابی سازمان، هر دو فرم A و B نسخه تصویری آزمون تفکر خلاق تورنس (۱۹۷۴) را تکمیل کردند. هر دو فرم این آزمون از سه فعالیت تشکیل شده‌اند که پاسخ‌دهندگان می‌بایست در

یک بازه زمانی ده دقیقه‌ای برای هر تکلیف، بر اساس محرک‌های بصری ارائه‌شده، نقاشی‌هایی خلق کنند. فعالیت ۱ (ساخت تصویر) شامل یک محرک بصری است، درحالی‌که فعالیت ۲ (تکمیل تصویر) شامل ده محرک و فعالیت ۳ (خطوط یا دایره‌ها در فرم A و B) شامل ۳۶ محرک تکراری هستند. دستورالعمل نمره‌دهی تورنس (۲۰۱۸) توسط دانشمند کفترودی و همکاران (۱۴۰۳) ترجمه و مبنای نمره‌گذاری قرار گرفت؛ لذا نمره‌گذاری بر اساس شاخص‌های زیر بود:

سیالی: تعداد کل پاسخ‌های مرتبط و معنادار در فعالیت‌های ۲ و ۳ را اندازه‌گیری می‌کند. این شاخص به‌صورت دو ارزشی عمل می‌کند (یک برای پاسخ‌های مرتبط و معنادار، صفر برای پاسخ‌های نامرتبط یا غیرمعنادار). پاسخ‌هایی که نمره صفر می‌گیرند، برای سایر شاخص‌ها در نظر گرفته نمی‌شوند.

اصالت: در ابتدا، تورنس (۱۹۶۶) به پاسخ‌هایی که در کمتر از ۲ درصد نمونه ظاهر می‌شدند، دو امتیاز و به پاسخ‌هایی که بین دو درصد تا پنج درصد نمونه ظاهر می‌شدند، یک امتیاز اختصاص می‌داد. در نسخه فعلی آزمون تفکر خلاق تورنس، نمره‌گذاری اصالت بر اساس نادر بودن و غیرمعمول بودن آماری پاسخ انجام می‌شود. هر فعالیت دارای فهرستی به نام فهرست پاسخ‌های غیرخلاق[1] است که در دفترچه راهنمای آزمون (تورنس، ۲۰۱۸) ارائه‌شده و شامل پاسخ‌هایی است که با فراوانی زیادی تولید می‌شوند. در این شیوه، برای فعالیت ۱ و ۲ در هر دو فرم، پاسخ‌هایی که در ۵ درصد یا بیشتر از نمونه‌ها مشاهده می‌شوند، غیر اصیل تلقی می‌شوند و در فهرست پاسخ‌های غیرخلاق قرار می‌گیرند. در فعالیت ۳ فرم A، پاسخ‌هایی که در بیست درصد یا بیشتر از نمونه مشاهده می‌شوند، غیراصیل هستند، درحالی‌که در فرم B این حد آستانه ده درصد است. در مجموع برای پاسخ‌هایی که در فهرست پاسخ‌های غیرخلاق قرار دارند صفر امتیاز و باقی پاسخ‌ها یک امتیاز دریافت می‌کنند (زیمرمن، ۲۰۱۸ به نقل از آکار و همکاران، ۲۰۲۱). علاوه بر این، شرکت‌کنندگان می‌توانند برای اصالت با ترکیب اشکال در فعالیت‌های ۲ و ۳ امتیاز اضافی دریافت کنند (تورنس، ۲۰۰۸). در ایران اکبری‌زردخانه و همکاران (زیرچاپ) در پژوهش خود نسخه ایرانی فهرست پاسخ‌های غیرخلاق را تدوین کرده و بر لزوم استفاده از فهرست‌های بومی تاکید کردند؛ بنابراین نمره‌گذاری اصالت در این پژوهش بر اساس فهرست پاسخ‌های غیرخلاق ایرانی صورت گرفت.

بسط: سطح جزئیات و زیبایی در نقاشی‌های تولیدشده در هر سه فعالیت را ارزیابی می‌کند. امتیازها برای ویژگی‌های فراتر از شکل پایه اعطا می‌شوند.

انتزاع عنوان‌ها: سطح انتزاعی بودن عناوین برای پاسخ‌ها در فعالیت ۱ و ۲ را ارزیابی می‌کند. امتیازها بر اساس ماهیت توصیفی (صفر تا یک امتیاز) و تخیلی (دو تا سه امتیاز) عناوین اعطا می‌شوند.

عدم نتیجه‌گیری زودهنگام: سطح باز بودن و حفظ قضاوت در فعالیت ۲ را ارزیابی می‌کند. امتیازها بر اساس اینکه آیا محرک (مثلاً یک شکل منحنی) زود بسته شده است (بدون امتیاز) یا نه (دو امتیاز)، یا اگر اضافاتی پس از بسته شدن انجام شده باشد (یک امتیاز) اعطا می‌شود.

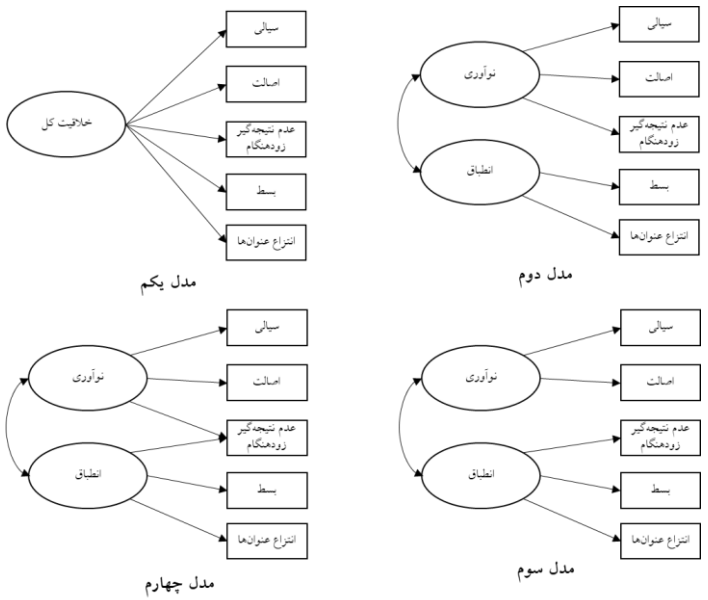
روش اجرای پژوهش

گردآوری داده: داده‌ها در درون مرکز ارزیابی اختصاصی یک سازمان صنعتی و طی فرایند ارزیابی کارکنان در بازه بهار و تابستان ۱۴۰۳ صورت گرفته است. در طی این فرایند ارزیابی و آزمون به صورت گروه‌های ده تا پانزده نفری اجرا شده است.

استخراج داده‌ها: در این مرحله برای هر یک از مشارکت‌کنندگان نمرات خام پنج شاخص سیالی، اصالت، بسط، انتزاع عنوان و عدم نتیجه‌گیری زودهنگام محاسبه گردید. علاوه‌برآن، نمرات انعطاف‌پذیری نیز که به دلیل رابطه بالا با سیالی کنار گذاشته شده بود، با هدف بررسی این فرضیه در جمعیت ایرانی برای تمام مشارکت‌کنندگان محاسبه گردید.

بررسی روایی و اعتبار ابزار نهایی: در مرحله پایانی ابتدا روایی ابزار از طریق تحلیل عاملی تائیدی و سپس اعتبار آن به روش ضرایب همگونی درونی مورد ارزیابی قرار گرفت. روایی از طریق بررسی روایی سازه[2] و انجام تحلیل عاملی تأییدی[3] (آسپاروهوف و موتن[4]، ۲۰۰۹) مورد بررسی قرار گرفت. در این مطالعه، چهار مدل نظری که در ادبیات مربوط به آزمون تفکر خلاق تورنس مورد بحث و اختلاف‌نظر قرار دارند، به‌عنوان مدل‌های رقیب انتخاب شدند (شکل ۲) که عبارت‌اند از یک مدل تک‌عاملی که توسط تورنس (۱۹۹۸) پیشنهاد شده است و سه مدل دوعاملی که به دلیل عدم توافق بر سر جای‌گیری شاخص عدم نتیجه‌گیری زودهنگام توسعه یافته‌اند (سید متوالی و همکاران، ۲۰۱۸). علاوه بر این، برای اطمینان از عملکرد یکسان آزمون در بین گروه‌های مختلف، هم‌ارزی اندازه‌گیری[5] بر اساس فرم آزمون (A یا B) بررسی شد. این فرایند شامل بررسی هم‌ارزی پیکربندی[6]، هم‌ارزی متریک[7] و هم‌ارزی اسکلار[8] بود (پوتنیک و بورنشتاین[9]، ۲۰۱۶).

[1] zero originality list (ZOL)
[2] Construct validity
[3] Confirmatory Factor Analysis (CFA)
[4] Asparouhov & Muthén
[5] Measurement invariance
[6] Configural invariance
[7] Metric invariance
[8] Scalar invariance
[9] Putnick & Bornstein



تصویر ۱
مدل‌های رقیب تحلیل

روش تجزیه و تحلیل

تحلیل‌ها با نرم‌افزار امپلاس[1] (موتن و موتن[2]، ۱۹۹۸–۲۰۲۱) انجام شد. به‌منظور ارزیابی اعتبار، از ضریب همگونی درونی[3] استفاده شد. به‌طور خاص، از دو شاخص ضریب آلفای کرونباخ[4] و ضریب امگا مک‌دونالد[5] بهره گرفته شد. ضریب آلفای کرونباخ (کرونباخ، ۱۹۵۱) به‌عنوان معیار پیش‌فرض برای سنجش همگونی درونی شناخته می‌شود، اما به دلیل فرضیات زیربنایی آن، مانند تک‌بعدی بودن، تائو–معادل‌بودن[6] و توزیع نرمال، مورد انتقاد قرار گرفته است (کراتزن و پیترز[7]، ۲۰۱۵؛ دان[8] و همکاران، ۲۰۱۴؛ گلدهاف[9] و همکاران، ۲۰۱۴؛ رایکوف[10]، ۱۹۹۷)؛ به همین دلیل، ضریب آلفا اغلب به‌عنوان تخمین حد پایین همگونی درونی مورد استفاده قرار می‌گیرد (بندلس[11]، ۲۰۱۸). در مقابل، ضریب امگا مک‌دونالد (مک‌دونالد، ۱۹۹۹) به‌عنوان جایگزینی مناسب و بدون چنین فرضیات سخت‌گیرانه‌ای ارائه شده است و تخمین‌های دقیق‌تر و گاهی بالاتری از آلفا ارائه می‌دهد (هیز و کوتس[12]، ۲۰۲۰؛ مک‌نیش[13]، ۲۰۱۷). در این پژوهش، ضرایب همگونی درونی برای مولفه‌های مدل تأییدشده از تحلیل عاملی تأییدی و همچنین برای نمره کل خلاقیت محاسبه شد. این بخش از محاسبات با استفاده از زبان برنامه‌نویسی R و بسته نرم‌افزاری سایک[14] انجام گرفت (رول[15]، ۲۰۲۳).

یافته‌های پژوهش

ضرایب همبستگی پیرسون و شاخص‌های توصیفی هر شش مؤلفه نسخه تصویری تورنس در جدول ۱ گزارش شده است. مقادیر کشیدگی[16] و کجی[17] برای تمام مؤلفه‌ها جز مؤلفه انتزاع عنوان‌ها کم‌تر از ۲/۰۰ بوده است؛ لذا می‌توان توزیع تقریباً نرمال برای این مؤلفه‌ها متصور بود (گراوتر و والنائو[18]، ۲۰۱۴). با این حال عدم نرمال بودن توزیع مؤلفه انتزاع عنوان‌ها ممکن است بر نرمال بودن توزیع چندمتغیره که مفروضه مدل‌سازی معادلات ساختاری است، اثر گذار باشد. ازاین‌رو، نرمال بودن چندمتغیره[19] با استفاده از آزمون ماردیا[20] ارزیابی شد که عدم نرمال بودن چند متغیره را نشان داد (کجی ماردیا: ۳۸۸/۲۵، $p < ۰/۰۰۱$؛ کشیدگی ماردیا: ۱۸/۳۱، $p < ۰/۰۰۱$). در مواجه با این چالش، در اجرای مدل‌سازی معادلات ساختاری، برآوردگر شباهت بیشینه خدشه‌ناپذیر[21] مورد استفاده قرار گرفت (ساتورا و بنتلر[22]، ۱۹۹۴).

1 Mplus
2 Muthén & Muthén
3 Internal consistency
4 Cronbach's Alpha coefficient
5 McDonald's Omega coefficient
6 Tau-equivalence
7 Crutzen & Peters
8 Dunn
9 Geldhof
10 Raykov
11 Bandalos
12 Hayes & Coutts
13 McNeish
14 Psych
15 Revelle
16 Kurtosis
17 Skewness
18 Gravetter and Wallnau
19 Multivariate normality
20 Mardia's test
21 Robust Maximum Likelihood (MLR) estimator
22 Satorra & Bentler

جدول ۱

شاخص‌های توصیفی و ضرایب همبستگی مؤلفه‌ها

متغیر	۱	۲	۳	۴	۵	۶
۱. سیالی	۱					
۲. اصالت	۰/۶۰**	۱				
۳. بسط	۰/۶۳**	۰/۳۷**	۱			
۴. انتزاع عنوان‌ها	۰/۲۷**	۰/۲۶**	۰/۲۷**	۱		
۵. عدم نتیجه‌گیری زودهنگام	۰/۲۷**	۰/۲۵**	۰/۳۴**	۰/۰۷	۱	
۶. انعطاف	۰/۹۱**	۰/۵۹**	۰/۶۳**	۰/۲۸**	۰/۲۶**	۱
میانگین	۲۰/۰۵	۲۲/۵۹	۸۰/۱۴	۱۱/۱۰	۱۲/۵۹	۱۷/۲۹
انحراف استاندارد	۵/۷۸	۷/۰۸	۲۹/۳۹	۲/۵۷	۲/۵۷	۴/۷۹
کجی	−۰/۲۳	۰/۱۵	۰/۳۸	۲/۱۲	−۰/۵۱	−۰/۳۴
کشیدگی	۰/۲۲	۰/۶۲	۰/۱۵	۱۶/۳۳	−۰/۱۲	۰/۲۹

** معنادار در سطح کم‌تر از ۰/۰۰۱

طبق نتایج جدول ۱ تمام ضرایب همبستگی در سطح کم‌تر از ۰/۰۰۱ معنادار بودند، جز همبستگی میان عدم نتیجه‌گیری زودهنگام و انتزاع عنوان‌ها. ازسوی‌دیگر همبستگی میان دو شاخص سیالی و انعطاف بسیار بالا (۰/۹۱) شد که در اجرای یک تحلیل رگرسیون خطی[1] میان این دو شاخص، نتایج نشان داد که مدل حاصل میان دو مؤلفه قادر است تا ۸۳ درصد از واریانس را تبیین کند (R^2=۰/۸۳؛ P<۰/۰۰۱، ۵۵۱۱/۵۰ = (۱،۲۳۹) =F).

به منظور مطالعه ساختار مکنون[2] سازه خلاقیت در جمعیت جوانان ایرانی، روایی سازه نسخه تصویری ابزار تفکر خلاق تورنس مورد ارزیابی قرار گرفت. شاخص‌های برازش عبارت بودند از: آماره خی دو[3]، شاخص برازش مقایسه‌ای[4] (بنتلر[5]، ۱۹۹۰)، ریشه میانگین مربعات خطای تقریب[6] (استیگر[7]، ۱۹۹۰)، شاخص تاکر – لوئیس[8] (تاکر و لوئیس، ۱۹۷۳) و ریشه میانگین مجذور باقی‌مانده استاندارد[9] (هو و بنتلر، ۱۹۹۸). جدول ۲ نشان می‌دهد مدل سوم، در تمام شاخص‌ها، برازش بهتری نسبت به مدل‌های رقیب کسب کرده است؛ این نتیجه در تحلیل عاملی تأییدی مستقل بر روی هر فرم نیز تکرار شده است. ازسوی‌دیگر در نتایج فرم B، به نظر می‌رسد که شاهد نوعی بیش‌برازش[10] در شاخص‌های مدل سوم و چهارم هستیم، به طور ویژه در شاخص‌های برازش مقایسه‌ای (۱/۰۰)، تاکر – لوئیس (۱/۰۰) و ریشه میانگین مربعات خطای تقریب (۰/۰۰). لازم به ذکر است که از بین چهار مدل رقیب، ماتریس کوواریانس متغیرهای نهفته مدل چهارم، معین مثبت[11] نبوده؛ از این رو، به دلیل مدل چهارم از گزارش تحلیل مربوط به مجموع دو فرم حذف شد.

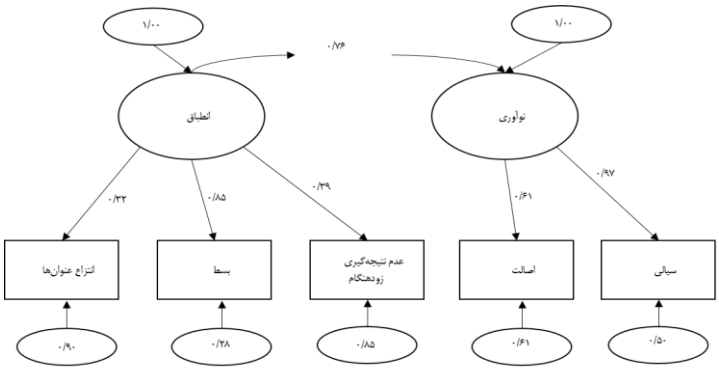
جدول ۲

شاخص‌های برازش مدل‌های رقیب

فرم	مدل	تعداد	X^2	df	X^2/df	CFI	TLI	RMSEA	SRMR
A و B	۱	۱	۱۹/۰۴	۵	۳/۸۱	۰/۹۵	۰/۹۰	۰/۱۱	۰/۰۴
	۲	۲	۱۷/۵۰	۴	۴/۳۷	۰/۹۵	۰/۸۸	۰/۱۲	۰/۰۴
	۳	۲	۹/۶۱	۴	۲/۴۰	۰/۹۸	۰/۹۵	۰/۰۸	۰/۰۴
	۴	۲	–	–	–	–	–	–	–
A	۱	۱	۱۴/۸۳	۵	۲/۹۷	۰/۸۹	۰/۷۷	۰/۱۲	۰/۰۶
	۲	۲	۱۴/۰۰	۴	۳/۵	۰/۸۸	۰/۷۱	۰/۱۴	۰/۰۶
	۳	۲	۸/۹۹	۴	۲/۲۵	۰/۹۴	۰/۸۶	۰/۱۰	۰/۰۶
	۴	۲	–	–	–	–	–	–	–
B	۱	۱	۵/۸۰	۵	۱/۱۶	۰/۹۹۶	۰/۹۹۲	۰/۰۳۸	۰/۰۳۳
	۲	۲	۵/۱۲	۴	۱/۲۸	۰/۹۹۴	۰/۹۸۶	۰/۰۵۰	۰/۰۳۳
	۳	۲	۲/۰۸	۴	۰/۵۲	۱/۰۰۰	۱/۰۰۰	۰/۰۰۰	۰/۰۲۱
	۴	۲	۱/۰۹	۳	۰/۳۶	۱/۰۰۰	۱/۰۰۰	۰/۰۰۰	۰/۰۱۸

[1] linear regression analysis
[2] Latent structure
[3] Chi_Square statistic
[4] Comparative Fit Index (CFI)
[5] Bentler
[6] Root mean square error of approximation (RMSEA)
[7] Steiger
[8] Tucker-Lewis Index (TLI)
[9] Standardized Root Mean Square Residual (SRMR)
[10] Overfitting
[11] Positive definite

نمودار مسیر (شکل ۳) روابط ساختاری بین عوامل و مؤلفه‌های آزمون را در مدل سوم نشان می‌دهد. آنچه از نتایج بر می‌آید، نقش تعیین‌کننده سیالی (۰/۹۷)، اصالت (۰/۶۱) و بسط (۰/۸۵) را در پیش‌بینی عوامل نوآوری و انطباق نشان می‌دهد. همچنین، انتزاع عنوان‌ها و عدم نتیجه‌گیری زودهنگام به عنوان پیش‌بینی‌کننده‌های دیگر (۰/۳۲ و ۰/۳۹ به ترتیب) برای عامل انطباق عمل می‌کنند. همچنین شاهد کوواریانس قوی بین نوآوری و انطباق (۰/۷۶) هستیم.



تصویر۲

نمودار مسیر مدل سوم

برای بررسی هم‌ارزی عملکرد مدل سوم در دو فرم موازی (A و B)، تغییرناپذیری اندازه‌گیری[1] آن‌ها به روش‌های پیکربندی[2]، متریک[3] و اسکلار[4] مدنظر قرار گرفت. جدول ۳ نشان می‌دهد که تغییرناپذیری پیکربندی، با تأیید سازگاری ساختار عاملی دو فرم، برازش بهتری نشان می‌دهد. این مدل به‌عنوان مبنایی در مقایسه سایر هم‌ارزی‌ها در نظر گرفته شد. تغییرناپذیری متریک نیز با محدودیت برابری بارهای عاملی در دو فرم و درنظرگرفتن معیارهای (چن[5]، ۲۰۰۷) نشان‌دهنده برازش مناسبی بود (ΔCFI ≤−۰/۰۱؛ ΔRMSEA≤۰/۰۱۵؛ ΔSRMR≤ ۰/۰۳)؛ اما تغییرناپذیری اسکلار با اعمال محدودیت عرض از مبدأ نتایج برازش خوبی نشان نداد (ΔCFI≥−۰/۰۱).

جدول ۳

نتایج بررسی هم‌ارزی دو فرم

مدل	χ^2 (df)	CFI	TLI	RMSEA	SRMR	ΔCFI	ΔRMSEA	ΔSRMR
پیکربندی	۱۱/۰۶ (۸)	۰/۹۸۹	۰/۹۷۳	۰/۰۵۶	۰/۰۳۶	–	–	–
متریک	۱۶/۲۳ (۱۱)	۰/۹۸۲	۰/۹۶۶	۰/۰۶۳	۰/۰۴۶	۰/۰۰۷	۰/۰۱	۰/۰۱
اسکلار	۱۸/۵۷ (۱۴)	۰/۹۸۴	۰/۹۷۷	۰/۰۵۲	۰/۰۵۰	۰/۰۰۲	۰/۰۱۱	۰/۰۰۴

ضرایب همگونی درونی دو فرم A و B و همچنین مجموع دو فرم (جدول ۴) نشان داد برای عامل نوآوری مجموع دو فرم، ضریب آلفای کرونباخ و امگا برابر با ۰/۷۵ و در دامنه قابل‌قبول است (نونالی و برنشتاین[6]، ۱۹۹۴). در فرم A نیز ضرایب آلفا و امگا هر دو ۰/۶۱ و در فرم B به ۰/۸۵ افزایش یافتند. این نشان‌دهنده بهبود قابل‌توجه همگونی درونی در فرم B است. برای عامل انطباق، در مجموع دو فرم، ضریب آلفا ۰/۴۷ و ضریب امگا ۰/۵۸ به‌دست آمد که نشان‌دهنده همگونی درونی پایین این عامل است. در فرم A نیز هر دو ضریب پایین‌تر از سطح پذیرفته‌شده (۰/۷) بودند. در فرم B، ضرایب به ۰/۵۸ و ۰/۶۴ افزایش یافتند که نشان‌دهنده بهبود نسبی است؛ اما همچنان زیر سطح مطلوب قرار دارند. در مجموع دو فرم، ضریب آلفا برای خلاقیت کل ۰/۷۱ و ضریب امگا ۰/۸۲ بود که نشان‌دهنده همگونی درونی قابل‌قبول و بالاتر برای این شاخص است. در فرم A، ضرایب آلفا و امگا به ترتیب ۰/۶۱ و ۰/۸۰ بودند، اما در فرم B این ضرایب به ۰/۷۹ و ۰/۸۴ افزایش یافتند که بالاترین سطح همگونی درونی در میان تمامی محاسبات را نشان می‌دهد.

جدول ۴

ضرایب همگونی درونی مدل نهایی

عامل	مجموع دو فرم	A	B
	آلفا (امگا)	آلفا (امگا)	آلفا (امگا)
نوآوری	۰/۷۵ (۰/۷۵)	۰/۶۱ (۰/۶۱)	۰/۸۵ (۰/۸۵)
انطباق	۰/۴۷ (۰/۵۸)	۰/۳۵ (۰/۵۲)	۰/۵۸ (۰/۶۴)
شاخص کلی	۰/۷۱ (۰/۸۲)	۰/۶۱ (۰/۸۰)	۰/۷۹ (۰/۸۴)

[1] Measurement invariance
[2] Configural invariance
[3] Metric invariance
[4] Scalar invariance
[5] Chen
[6] Nunnally & Bernstein

بحث و نتیجه‌گیری

در امتداد پژوهش اکبری زردخانه و همکاران (۱۴۰۳) که به تدوین فهرست پاسخ‌های غیرخلاق در جمعیت ایرانی پرداختند، پژوهش حاضر با دو هدف اصلی انجام شد: نخست، بررسی روایی سازه و اعتبار جدیدترین شیوه نمره‌گذاری نسخه تصویری آزمون تفکر خلاق تورنس (فرم‌های A وB) برای جمعیت بزرگسال ایرانی؛ و دوم، پاسخ به این پرسش که آیا ساختار دوبعدی خلاقیت که در پژوهش‌های بین‌المللی گزارش شده است، در جمعیت ایرانی نیز معتبر است یا خیر.

در گام نخست، بررسی همبستگی میان دو شاخص سیالی و انعطاف گواه از رابطه بسیار قوی میان این دو بود که می‌تواند حاکی از شباهت زیاد این دو شاخص به یکدیگر باشد (شوبر[1] و همکاران، ۲۰۱۸). از سوی دیگر نتیجه تحلیل رگرسیون خطی میان این دو شاخص، نشان داد که مدل حاصل قادر است تا ۸۳ درصد از واریانس را تبیین کند؛ این هم‌پوشانی را می‌توان به ساختار نمره‌گذاری این دو شاخص نسبت داد، سیالی تعداد ایده‌های تولید شده را اندازه‌گیری می‌کند و انعطاف‌پذیری میزان تنوع آن ایده‌ها را، درنتیجه تولید ایده‌های بیشتر با احتمال گستردگی طبقات پاسخ و نهایتا افزایش نمره انعطاف‌پذیری همراه است؛ لذا می‌توان همسو با پیشنهاد تورنس و بال (۱۹۸۴) شاخص انعطاف را از محاسبات نسخه ایرانی آزمون کنار گذاشت؛ چرا که گزارش آن در جمعیت ایرانی فاقد ارزش‌افزوده است.

مقایسه مدل‌های رقیب چهارگانه نشان داد که مدل دوبعدی نسبت به مدل تک‌بعدی برتری دارد. در میان مدل‌های دوبعدی، نیز مدل سوم به دلیل برتری تمام شاخص‌ها نسبت به دیگر مدل‌ها، به‌عنوان مدل نهایی انتخاب گردید. برخلاف یافته‌های مطالعاتی همچون (هاهم[2] و همکاران، ۲۰۱۹؛ هامبل و همکاران، ۲۰۱۸؛ کروم و همکاران، ۲۰۱۴؛ کروم و همکاران، ۲۰۱۶؛ یون، ۲۰۱۷) که نشان می‌دهند مؤلفه عدم نتیجه‌گیری زودهنگام هم در عامل ابتکار و هم در عامل انطباق وجود دارد، این مطالعه نشان داد که در جمعیت مورد پژوهش حاضر، مؤلفه عدم نتیجه‌گیری زودهنگام تنها بر عامل انطباق بارگذاری دارد. همچنین، مدل‌هایی که از بارگذاری این مؤلفه بر عامل نوآوری حمایت می‌کردند (مدل دوم و یکم) در برخی از شاخص‌ها فاقد برازش حداقلی بودند. این نتیجه همسو است با یافته‌های سید متوالی و همکاران (۲۰۱۸) و آکار و همکاران (۲۰۲۳) که مدل دوبعدی، عدم نتیجه‌گیری زودهنگام همراه با بسط و انتزاع عنوان‌ها بر عامل انطباق‌پذیری و سیالی و اصالت بر عامل نوآورانه را به‌عنوان مدل نهایی تأیید کردند. در ارتباط با عامل نوآورانه، جای تعجب نیست که سیالی و اصالت در یک عامل ادغام شوند، زیرا این دو بعد رابطه قوی با یکدیگر دارند (هوسلر و تامپسون[3]، ۱۹۸۸؛ کیم، ۲۰۰۶). علاوه بر این، افرادی که تعداد زیادی ایده تولید می‌کنند، معمولاً از اصالت بیشتری برخوردار هستند (تورنس و سافتر[4]، ۱۹۹۹)؛ لذا این عامل بیشتر با سبکی از پاسخ‌های سریع و اصیل مرتبط است که در آن ترجیح بر نوآوری است، نه کار بر روی موقعیت‌های فعلی (کیرتون[5]، ۱۹۷۶).

ارتباط عدم نتیجه‌گیری زودهنگام با عامل انطباق ممکن است به عوامل انگیزشی[6] درگیر در فرایند خلاقیت، مانند پشتکار و تعهد به انجام تکالیف مرتبط باشد (لوبارت و زنازنی[7]، ۲۰۱۰). تمایل به اصلاح یک ایده با جزئیات و عناوین غیربدیهی، ویژگی مشترکی است که در انتزاع عنوان‌ها و بسط مورد نیاز است که این امر مستلزم صبر و پشتکار در انجام‌وظیفه است. این ویژگی بیشتر در افرادی مشاهده می‌شود که می‌توانند در برابر میل به تکمیل سریع اما ناقص کار مقاومت کنند. با این حال، این ارتباطات نیاز به بررسی‌های بیشتری در مطالعات آینده دارند که در آن همبسته‌های شناختی، شخصیتی و هیجانی هر دو عامل بررسی شوند. چنین اطلاعاتی می‌تواند به استفاده مؤثرتر از داده‌های ارائه شده توسط ابزار کمک کند و امکان ارائه کمک‌های فردی مبتنی بر تحلیل نیمرخ با استفاده از نمرات دو مقیاس را فراهم کند. نظر به آنکه در این پژوهش، مشارکت‌کنندگان فرم‌های آزمون را به عنوان بخشی از فرایند ارزیابی شغلی تکمیل کرده‌اند، می‌توان استنباط کرد که نسبت به مشارکت‌کنندگان در محیط‌های پژوهشی، انگیزه و تعهدی بیشتری در تکمیل فرم‌ها داشتند؛ لذا می‌توان بخشی از علت تفاوت نتایج این پژوهش و کیم (۲۰۰۶) را در این امر جستجو کرد.

مقایسه شاخص‌های برازش مدل سوم در دو فرم نشان می‌دهد که فرم A در تمام شاخص‌های برازش جز دو شاخص تاکر – لوئیس و ریشه میانگین مربعات خطای تقریب از برازش قابل‌قبولی برخوردار است. در مقابل، فرم B با عملکردی بهتر مقادیر قابل‌قبولی در تمامی شاخص‌های برازش حاصل کرده است. در همین راستا نتایج بررسی هم‌ارزی پیکربندی، سازگاری ساختار عاملی در هردو فرم را تأیید کرد؛ لذا می‌توان نتیجه‌گرفت که هر دو فرم آزمون ساختار مشابهی از سازه خلاقیت را اندازه‌گیری می‌کنند (ساختار دوبعدی). بنابراین، می‌توان از هر دو فرم برای بررسی و تحلیل ساختار عاملی استفاده کرد و اطمینان داشت که این ساختار در هر دو فرم ثابت است. همچنین بررسی هم‌ارزی متریک نشان داد که قدرت رابطه بین آیتم‌ها و عوامل در هر دو فرم معادل است. این موضوع فرصت استفاده از هر دو فرم به‌منظور بررسی تغییرات در افراد یا گروه‌های مختلف را فراهم می‌کند، زیرا مدل اندازه‌گیری یکسان است. با این حال، هم‌ارزی اسکلار دو فرم تأیید نشد؛ لذا نمی‌توان به مقایسه مستقیم نمرات خام یا میانگین‌های بین دو فرم پرداخت؛ به‌بیان‌دیگر در جمعیت بزرگسالان ایرانی نمی‌توان از دو فرم A و B به عنوان فرم‌های موازی پیش‌آزمون و پس‌آزمون در ارزیابی آموزش‌های ارتقا خلاقیت بهره گرفت (پوتنیک و بورنشتاین، ۲۰۱۶).

[1] Schober
[2] Hahm
[3] Heausler & Thompson
[4] Safter
[5] Kirton
[6] Conative factors
[7] Lubart & Zenasni

در بررسی شاخص‌های برازش فرم B به نظر می‌رسد که شاهد نوعی بیش‌برازشی نگران‌کننده در شاخص‌های مدل سوم و چهارم هستیم، به طور ویژه در شاخص‌های برازش مقایسه‌ای، تاکر – لوئیس و ریشه میانگین مربعات خطای تقریب حاصل شد. فرمول محاسبه هر سه این شاخص‌ها مشتمل است از پارامتر نامرکزی[1] (تفاضل خی‌دو و درجه آزادی)؛ لذا اگر مقادیر خی‌دو یک مدل به درجه آزادی آن نزدیک شود، همانند آنچه در برآوردهای فرم B حاصل شد، پارامتر نامرکزی صفر می‌شود؛ لذا شاخص‌های برازش مقایسه‌ای و تاکر – لوئیس نیز برابر با ۱ و شاخص ریشه میانگین مربعات خطای تقریب برابر صفر می‌شود (مولایک[2]، ۲۰۰۹). در نتیجه حصول این مقادیر حاکی از برازش بسیار خوب مدل سوم برای فرم B است.

جدول نتایج خلاصه‌ای که توسط شرکت ارزیابی‌های آموزشگاهی (ناشر آزمون تفکر خلاق تورنس) و نسخه‌های موجود فارسی ارائه می‌شود، تمامی شاخص‌های امتیازدهی را در قالب یک شاخص خلاقیت ترکیب می‌کند. این شاخص فرض می‌کند که آزمون دارای یک ساختار تک‌بعدی است، درحالی که نتایج این پژوهش و ادبیات موجود نشان می‌دهد این فرض به‌ندرت درست بوده است، حتی زمانی که روش‌های استخراج عاملی نامطلوب (مانند تحلیل مؤلفه‌های اصلی) استفاده شده است (برای مثال کلافام، ۱۹۹۸)؛ لذا بر اساس یافته‌های این پژوهش پیشنهاد می‌شود که در تدوین گزارش نتایج این آزمون، چارچوب امتیازدهی اتخاذ شود که ساختار دوبعدی را منعکس نماید. چرا که این رویکرد امکان تحلیل نیمرخ فردی را به نحوی فراهم می‌کند که از جنبه‌های خاص توسعه خلاقیت حمایت شود. به این دلیل که نباید ابزارهای ارزیابی صرفاً با هدف دستیابی به یک امتیاز توسعه داده شوند، بلکه باید به گونه‌ای باشند که با آموزش، مداخلات و توسعه نیز هماهنگ شوند (کالاهان[3] و همکاران، ۲۰۲۲؛ گوبینز[4] و همکاران، ۲۰۲۱؛ لی[5] و همکاران، ۲۰۲۰). با اتخاذ چارچوب امتیازدهی دوبعدی، مدیریان، مربیان، معلمان و به طور کلی سیاست‌گذاران می‌توانند از نمرات این آزمون برای پرورش و حمایت بهتر از رشد خلاقیت استفاده کنند (ترفینگر[6]، ۲۰۰۴). از این رو پیشنهاد می‌شود که پژوهشگران به جای یا علاوه بر استفاده از یک امتیاز کلی، از دو نمره مقیاس نیز (عامل نوآوری و انطباق) استفاده کنند. به این ترتیب، فرضیه‌ها و روابط خاص‌تری می‌توانند مورد آزمایش و بررسی قرار گیرند.

نتایج نشان داد که شاخص کلی خلاقیت از اعتبار بالاتری برخوردار است (امگا: ۰/۸۰ تا ۰/۸۴)؛ لذا نمره ترکیبی خلاقیت که عملکرد کلی در آزمون را نشان می‌دهد، می‌تواند برای بیشتر کاربردها (از اهداف پژوهشی گرفته تا ارزیابی‌های حساس آموزشی و استخدامی) به عنوان خوب در نظر گرفته شود (کوهن و سوردلیک[7]، ۲۰۱۸). همچنین بالابودن قابل توجه اعتبار عامل نوآوری (امگا ۰/۶۱ تا ۰/۸۵) نسبت به انطباق (امگا ۰/۵۲ تا ۰/۶۴) یافته ارزشمندی بود. بر اساس معیارهای کوهن و سوردلیک (۲۰۱۸)، برآورد اعتبار عامل انطباق می‌تواند به عنوان به سختی قابل‌قبول[8] تا ردشدنی[9] در نظر گرفته شود. این یافته‌ها با مطالعه اخیر آکار (۲۰۲۳) همخوانی دارد که در سه مورد از چهار مدل دو بعدی با بهترین تناسب، اعتبار قوی‌تری برای عامل نوآوری گزارش کرده است. همچنین اعتبار نوآوری را می‌توان در سطح متوسط برای مجموع دو فرم، نامناسب در فرم A و تا خوب در فرم B در نظر گرفت؛ بنابراین، هنگام استفاده از نمرات انطباق باید احتیاط کرد، زیرا این نمرات شامل درجه بالاتری از خطای اندازه‌گیری هستند. بارگذاری ضعیف‌تر مولفه‌ها بر انطباق، اعتبار پایین‌تر این عامل را توضیح می‌دهد و تغییر در روش‌های امتیازدهی شاخص‌ها می‌تواند در مطالعات آینده مورد بررسی قرار گیرد. تحلیل‌های راشِ[10] یوروک[11] (۲۰۱۶) به چنین مسائلی در بسط اشاره کرده و یک روش امتیازدهی جایگزین (مانند مقیاس ۱ تا ۶ امتیازی برای هر محرک) را آزمود. ازآنجا که یک پاسخ می‌تواند برای سنجش توانایی‌های گوناگون به روش‌های مختلفی نمره‌گذاری شود، مطالعات آینده می‌توانند شاخص‌های امتیازدهی بالقوه دیگری را بررسی کنند که ممکن است اعتبار عامل انطباق را بهبود بخشد. در مجموع نتایج بررسی اعتبار نشان می‌دهد که نمرات ترکیبی نسخه تصویری آزمون تورنس و همچنین نمرات عامل نوآوری برای استفاده مناسب‌تر از نمرات عامل انطباق هستند.

از سوی دیگر اعتبار بالاتر فرم B به معنای دقت بیشتر و اختلاف کمتر بین نمرات مشاهده‌شده و حقیقی[12] است. در نتیجه، اعتبار قوی‌تر می‌تواند با حساسیت بالاتری افراد توانمند را شناسایی کند (مک‌بی[13] و همکاران، ۲۰۱۶؛ پیترز[14] و همکاران، ۲۰۲۳). علاوه بر این، به دلیل اعتبار بالاتر، استفاده از فرم B برای پژوهش و اهداف حساس مانند ارزیابی‌های استخدامی توصیه می‌شود، زیرا اعتبار پایین ممکن است میزان روابط بین سازه‌های مورد بررسی را پنهان کند. بهتربودن اعتبار فرم B در این پژوهش در تضاد با نتایج پژوهش آکار و همکاران (۲۰۲۴) است که فرم A را به عنوان فرم بهتر معرفی کردند. کوهن و سوردلیک (۲۰۱۸) تفاوت در اعتبار دو فرم موازی را به نمونه‌گیری آیتم‌ها (محرک‌ها در این مثال) نسبت دادند. همچنین به اعتقاد آکار و همکاران (۲۰۲۴) عامل دیگری که می‌تواند بر این تفاوت اثر گذاشته باشد مبنای نمره‌گذاری اصالت است که از فهرست پاسخ‌های غیرخلاق گرفته شده است، آن‌ها دریافتند که تغییرات ایجادشده در فهرست‌های جدید پاسخ‌های غیرخلاق برای فرم A موفق‌تر از فرم B بودند. به این معنا که فهرست‌ها در فرم A دقیق‌تر بوده و لذا نمره‌گذاری مسنجم‌تر به ضرایب اعتبار بالاتر می‌انجامد. به طریق مشابه و با استناد به پژوهش اکبری‌زردخانه و همکاران (۱۴۰۳)، نظر به عملکرد بهتر فهرست پاسخ‌های غیرخلاق ایرانی در فرم B، می‌توان کسب ضرایب اعتبار بالاتر این فرم را تبیین کرد.

[1] Noncentrality parameter
[2] Mulaik
[3] Callahan
[4] Gubbins
[5] Lee
[6] Treffinger
[7] Cohen & Swerdlik
[8] Hardly acceptable
[9] Rejectable
[10] Rasch Measurement Theory (RMT)
[11] Yoruk
[12] Observed and true score
[13] McBee
[14] Peters

منبع دیگر تفاوت اعتبار دو فرم را می‌توان به عدم وجود روایی تفکیکی[1] بین دو شاخص عامل نوآوری (سیالی و اصالت) نسبت داد (ضرایب همبستگی بالا). تأثیر آلوده‌کننده سیالی بر اصالت و انعطاف‌پذیری پدیده‌ای شناخته‌شده است (آکار، ۲۰۲۳؛ آکار و همکاران، ۲۰۲۲؛ کلارک و میرلس[2]، ۱۹۷۰؛ فورثمن و همکاران، ۲۰۲۰؛ هوچوار، ۱۹۷۹؛ سیلویا، ۲۰۰۸) و اغلب زمانی مشاهده می‌شود که از تجمیع جمع‌بندی‌شده[3] برای به دست آوردن نمرات در سطح پاسخ‌دهنده استفاده می‌شود. آکار (۲۰۲۳) این مسئله را در نسخه تصویری آزمون تورنس شناسایی کرد که مشاهدات او نشان داد با حذف فعالیت ۳ از نمرات کل می‌توان از این مسئله اجتناب کرد. فورثمن (۲۰۲۲) هشدار داد که این مشکل تحقیقات تحلیل عاملی را زیر سؤال می‌برد، زیرا این عوامل ممکن است صرفاً مستخرج از همبستگی‌های ساختگی[4] بین سیالی و اصالت یا انعطاف‌پذیری باشند. در نگاه برخی از پژوهشگران وجود چنین همبستگی‌هایی میان سیالی و دیگر شاخص‌ها ممکن است بر روایی تفکیکی شاخص‌های آزمون‌های تفکر واگرا تأثیر سوء بگذارد (هوچوار و مایکل[5]، ۱۹۷۹) تا حدی که سودمندی تمام شاخص‌ها جز سیالی را با تردید همراه کند (آکار و همکاران، ۲۰۲۲)؛ با این حال آکار و همکاران (۲۰۲۳) در فراتحلیل خود بر روی ماتریس‌های ضرایب همبستگی شاخص‌های نسخه تصویری آزمون تفکر خلاق تورنس، نشان دادند که وجود یا عدم وجود چنین همبستگی‌های بر برازش سازه ابزار تأثیری ندارد. با این حال اگر دیدگاه فورثمن (۲۰۲۲) را تا حدودی اثرگذار بدانیم، می‌توان انتظار داشت که اعتبار زمانی بالاتر باشد که سیالی و اصالت همبستگی بالایی داشته باشند. همان‌طور که در این مطالعه همبستگی اصالت و سیالی در فرم B (۰/۷۳) بیشتر از فرم A (۰/۴۴) بود که می‌توان بخشی از تفاوت در اعتبار دو فرم را به این مسئله نسبت داد.

یکی از محدودیت‌های احتمالی این مطالعه این است که نمونه مورد بررسی از کارکنان یک شرکت صنعتی تشکیل شده و فاقد گروه مقایسه‌ای، مانند نمونه‌ای از دانشجویان، مشاغل یا قوم و نژاد دیگر است. همچنین با توجه به فراوانی بسیار کم زنان در میان کارکنان این شرکت صنعتی، پیشنهاد می‌شود که در مطالعات بعدی نمونه‌ای با توازن بهتر میان جنسیت انتخاب گردد. از دیگر محدودیت‌های این پژوهش می‌توان به عدم استفاده از روش‌های دیگر سنجش اعتبار مانند اعتبار آزمون–بازآزمون اشاره کرد که به ارزیابی ثبات نمرات ابزار در طول زمان می‌پردازند (دولیس، ۲۰۱۷). اگرچه این آزمون در دو فرم توسعه‌یافته است؛ اما نتایج این پژوهش نشان داد که نمی‌توان به مقایسه میانگین نمرات بین دو فرم پرداخت؛ لذا بررسی اعتبار آزمون–بازآزمون می‌تواند اطمینان از استفاده واحد از هر فرم را در مطالعاتی نظیر ارزیابی اثربخشی برنامه‌های آموزشی افزایش دهد؛ زیرا از این طریق اطمینان حاصل می‌شود که تغییرات نمرات مشارکت‌کنندگان ناشی از اجرای مداخله است و نه تأثیر عواملی مانند خطای اندازه‌گیری.

سپاس‌گزاری

نویسندگان این مقاله بر خود لازم می‌دانند از تمام کارکنان محترم مرکز ارزیابی شرکت صنعتی که در این پژوهش مشارکت نمودند و همچنین از سرکار خانم شبنم نژدامیدواری‌پور و آقایان واعظی و ارسلان احمدپور که در نمره‌گذاری برگه‌ها همکاری داشتند کمال تشکر و قدردانی را به عمل آورند.

ملاحظات اخلاقی

تمامی افرادی که در این تحقیق حضور داشتند با پر کردن فرم رضایت‌نامه، مشارکت خود را اعلام کردند و پژوهشگران نیز تضمین کردند که اطلاعات حاصل از این مطالعه به صورت محرمانه نگهداری خواهد شد.

حامی مالی

تمام منابع مالی و هزینه پژوهش و انتشار مقاله تماماً بر عهده نویسندگان بوده و هیچ‌گونه حمایت مالی دریافت نشده است.

مشارکت نویسندگان

این مقاله از پایان‌نامه کارشناسی ارشد استخراج شده است. نویسنده مسئول به عنوان استاد راهنما فعالیت داشته است. نویسنده دوم دانشجوی کارشناسی ارشد و مجری اصلی پایان‌نامه بوده است. نویسنده سوم به عنوان مشاور پایان‌نامه و نویسنده چهارم به عنوان مدیر سازمان، فرآیند جمع‌آوری داده‌ها را هماهنگ کرده است.

تعارض منافع

نویسندگان اعلام می‌دارند که هیچ گونه تعارض منافعی ندارد.

منابع

اکبری‌زردخانه، سعید.، دانشمند کفترودی، ا. ح. و فتح‌آبادی، ج (در دست چاپ). خلاقیت: فرهنگ‌وابسته یا نه؟ فهرست پاسخ‌های غیرخلاق تورنس در جمعیت ایرانی.

[1] Discriminant validity
[2] Clark & Mirels
[3] Summative aggregation
[4] Artifactual correlation
[5] Michael

اوبالاسی، ا. و حسینی نسب، س. د. (۱۳۹۴). بررسی تأثیر آموزش روش مونته‌سوری بر خلاقیت کودکان پیش‌دبستانی ۴ و ۵ ساله شهر تبریز. *آموزش و ارزشیابی (علوم تربیتی)*، ۷(۲۸)، ۸۱-۹۸.

بابایی اوصالو، ح.، سهرابی، ز.، و خواجه آزاد، م. (۱۳۹۴). بررسی خلاقیت اساتید علوم پایه پزشکی براساس نظریه تورنس؛ دانشگاه علوم پزشکی ایران. *دوماهنامه علمی- پژوهشی راهبردهای آموزش در علوم پزشکی*. ۸ (۳)، ۱۴۳-۱۳۷.

پیرخائفی، ع. ر.، برجعلی، ا.، دلاور، ع و اسکندری، ح. (۱۳۸۸). تاثیر آموزش خلاقیت بر مؤلفه‌های فراشناختی تفکر خلاق دانشجویان. *رهبری و مدیریت آموزشی*، ۳(۲)، ۵۱-۶۱.

توانا، ز.، اکبری زردخانه، س. و شهیدی، ش. (۱۴۰۱). *سنجش خلاقیت نوجوانان: انطباق نمره‌گذاری نسخه فارسی آزمون تفکر خلاق تورنس*. پایان‌نامه کارشناسی ارشد روان‌شناسی گرایش تربیتی، دانشگاه شهید بهشتی.

حمیدی، ف.، شهمیر، س.، محمدی، ن و دهنوی، ا. (۱۳۹۱). مقایسه رویکرد خلاقیت بین دانش‌آموزان مدارس عادی و هنرستان‌ها بر اساس آزمون خلاقیت تورنس. *ابتکار و خلاقیت در علوم انسانی*، ۲(۳)، ۲۰۵-۲۱۸.

دانشمند کفترودی، ا. ح.، اکبری زردخانه، س و فتح آبادی، ج. (۱۴۰۳). *استانداردسازی اولیه آزمون تصویری تفکر خلاق تورنس برای جوانان ایرانی*. پایان‌نامه کارشناسی ارشد روان‌شناسی تربیتی. دانشگاه شهید بهشتی.

دائمی، حمیدرضا و مقیمی بارفروش، فاطمه (۱۳۸۳). هنجاریابی آزمون خلاقیت. *تازه‌های علوم‌شناختی*. ۶ (۳ و ۴)، ۱-۸.

ذاکری، س. و برقی، ع. (۱۳۹۴). تاثیر آزمون تفکر خلاق تورنس بر سیگنال نرخ ضربان قلب. *دوماهنامه علمی- پژوهشی راهبردهای آموزش در علوم پزشکی*. ۸ (۶)، ۱-۶.

رضایی، س. و منوچهری، م. (۱۳۸۷). بررسی اعتبار، روایی و هنجاریابی آزمون خلاقیت تورنس در بین دبیران دبیرستان‌های تهران. *روانشناسی و علوم تربیتی*، ۳۸(۳)، ۴۷-۶۸.

رضوانی، ع. ر و برادران توکلی، د. (۱۴۰۳). بررسی تاثیر مهارت موسیقی بر خلاقیت طراحی دانشجویان معماری (بر اساس آزمون تورنس و ارزیابی طراحی). *مطالعات هنر اسلامی*، ۲۱(۵۳)، ۳۱۴-۳۳۰.

سعیدی، علی، و شکیبا، ا. (۱۴۰۰). بازنگری در نمره‌گذاری آزمون تصویری فرم A خلاقیت تورنس در ایران. *رویش روان‌شناسی*. ۱۰ (۱)، ۱۲۰-۱۰۹.

صمدی، ن.، اکبری زردخانه، س. و فتح آبادی، ج. (۱۴۰۲). *سنجش خلاقیت کودکان: انطباق نمره‌گذاری نسخه فارسی آزمون تفکر خلاق تورنس*. پایان‌نامه کارشناسی ارشد روان‌شناسی گرایش تربیتی، دانشگاه شهید بهشتی.

اقدسی، م. ت.، فتحی رضائی، ز و عباس‌پور، ک.، (۱۴۰۰). اثربخشی زمینه‌ی محیطی از دیدگاه بوم‌شناختی بر رشد حرکتی و خلاقیت کودکان. *اندیشه‌های نوین تربیتی، ۱۷* (۲)، ۲۷۳-۲۴۹.

References

Abedi, J. (2002). A latent-variable modeling approach to assessing reliability and validity of a creativity instrument. *Creativity Research Journal*, 14(2), 267–276.

Acar, S., Lee, L. E., & Scherer, R. (2024). A reliability generalization of the Torrance Tests of Creative Thinking-Figural. *European Journal of Psychological Assessment, 40*(5), 396–411

Adair, W. L., & Xiong, T. X. (2018). How Chinese and Caucasian Canadians conceptualize creativity: The mediating role of uncertainty avoidance. *Journal of Cross-Cultural Psychology,* 49(3), 223–238.

Aghdasi, M. T., Fathirezaie, Z. and Abbaspour, K. (2021). The Effect of Environmental Contexts from Ecological Perspective on Motor Development and Creativity of Children. *The Journal of New Thoughts on Education*, *17*(2), 249-273. [Perstion]

Aranguren, M. (2014). Validez de constructo del Test de Pensamiento Creativo de Torrance en una muestra de jóvenes argentinos. *Anuario de Psicología/The UB Journal of Psychology,* 44(1), 55–70. Retrieved from.

Babaei Osalou, H., Sohrabi, Z., & Khajehazad, M. (2015). Creativity analysis in faculties of biomedical courses according to Torrance creativity theory: Iran University of Medical Sciences. *Education Strategy in Medical Sciences*, *8*(3), 137–143. [Perstion]

Baer, J., & McKool, S. S. (2009). Assessing creativity using the consensual assessment technique. In C. S. Schreiner (Ed.), *Handbook of research on assessment technologies, methods, and applications in higher education* (pp. 65–77).

Bandalos, D. L. (2018). *Measurement theory and applications for the social sciences*. Guilford Press.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin,* 107(2), 238–246.

Callahan, C. M., Azano, A., Park, S., Brodersen, A. V., Caughey, M., & Dmitrieva, S. (2022). Consequences of implementing curricular-aligned strategies for identifying rural gifted students. *Gifted Child Quarterly,* 66(4), 243–265.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. Structural Equation Modeling*: A Multidisciplinary Journal*, 14(3), 464–504.

Clapham, M. M. (1998). Structure of Figural Forms A and B of the Torrance Tests of Creative Thinking. *Educational and Psychological Measurement*, 58(2), 275–283.

Clark, P. M., & Mirels, H. L. (1970). Fluency as a pervasive element in the measurement of creativity. *Journal of Educational Measurement*, 7(2), 83–86.

Cohen, R. J., & Swerdlik, M. E. (2018). *Psychological testing and assessment: An introduction to tests and measurement* (9th ed.). McGraw-Hill Education.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika,* 16(3), 297–334.

Crutzen, R., & Peters, G. J. Y. (2015). Scale quality: Alpha is an inadequate estimate and factor-analytic evidence is needed first of all. *Health Psychology Review*, 11(3), 242–247.

Daemi, H., & Moghimi Barforoush, F. (2004). Normalization of the Creativity Test. *Advances in Cognitive Science*, 6(3–4), 1–8. [Perstion]

Daneshmand Kaftaroudi, A. H., Akbari Zardkhaneh, S., & Fathabadi, J. (2024). Preliminary standardization of the Torrance Figural Creativity Test for Iranian youth (Master's thesis, Shahid Beheshti University). [Persian]

DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Sage.

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology,* 105(3), 399–412.

Forthmann, B., Szardenings, C., & Holling, H. (2020). Understanding the confounding effect of fluency in divergent thinking scores: Revisiting average scores to quantify artifactual correlation. *Psychology of Aesthetics, Creativity, and the Arts*, 14(1), 94–112.

Frasier, M. M. (1990). Torrance verbal and figural tests: Measuring general creative thinking processes. *Paper presented at the Annual Meeting of the American Educational Research Association*, San Francisco, CA.

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, 19(1), 72–91.

Gravetter, F. J., & Wallnau, L. B. (2014). *Essentials of statistics for the behavioral sciences* (8th ed.). Wadsworth.

Gubbins, E. J., Siegle, D., Ottone-Cross, K., McCoach, D. B., Langley, S. D., Callahan, C. M., Brodersen, A. V., & Caughey, M. (2021). Identifying and serving gifted and talented students: Are identification and services connected? *Gifted Child Quarterly*, 65(2), 115–131.

Hahm, J., Kim, K. K., & Park, S. H. (2019). Cortical correlates of creative thinking assessed by the figural Torrance Test of Creative Thinking. *Neuroreport*, 30(18), 1289–1293.

Hamidi, F., Shahmir, S., Mohammadi, N., & Dehnavi, E. (2016). Comparing creativity approach between normal schools and vocational schools based on the Torrance Creativity Test. Innovation Creativity in Human Science, 2 (3), 205–2018. [Perstion]

Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. But…. *Communication Methods and Measures*, 14(1), 1–24.

Heausler, N. L., & Thompson, B. (1988). Structure of the Torrance Tests of Creative Thinking. *Educational and Psychological Measurement*, 48(2), 463–468.

Hébert, T. P., Cramond, B., Neumeister, K. L. S., Millar, G., & Silvian, A. F. (2002). Paul Torrance: His life, accomplishments, and legacy. *The National Research Center on the Gifted and Talented* (NRC/GT).

Hocevar, D., & Michael, W. B. (1979). The effects of scoring formulas on the discriminant validity of tests of divergent thinking. *Educational and Psychological Measurement,* 39(4), 917–921.

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3(4), 424–453.

Humble, S., Dixon, P., & Mpofu, E. (2018). Factor structure of the Torrance Tests of Creative Thinking Figural Form A in Kiswahili-speaking children: Multidimensionality and influences on creative behavior. *Thinking Skills and Creativity*, 27, 33–44.

Kaufman, J. C., Lee, J., Baer, J., & Lee, S. (2007). Captions, consistency, creativity, and the consensual assessment technique: New evidence of reliability. *Thinking Skills and Creativity*, 2(2), 96–106.

Kim, K. H. (2006a). Can we trust creativity tests? A review of the Torrance Tests of Creative Thinking (TTCT). *Creativity Research Journal*, 18(1), 3–14.

Kim, K. H. (2006b). Is creativity unidimensional or multidimensional? Analyses of the Torrance Tests of Creative Thinking. *Creativity Research Journal*, 18(3), 251–260.

Kim, K. H. (2017). The Torrance Tests of Creative Thinking - Figural or Verbal: Which one should we use? Creativity. *Theories – Research - Applications*, 4(2), 302–321.

Kim, K. H., Cramond, B., & Bandalos, D. (2006). The latent structure and measurement invariance of scores on the Torrance Tests of Creative Thinking-Figural. *Educational and Psychological Measurement*, 66(3), 459–474.

Kirton, M. (1976). Adaptors and innovators: A description and measure. Journal of Applied Psychology, 61(5), 622–629.

Krumm, G., Arán Filippetti, V., Lemos, V., Koval, J., & Balabanian, C. (2016). Construct validity and factorial invariance across sex of the Torrance Test of Creative Thinking—Figural Form A in Spanish-speaking children. *Thinking Skills and Creativity*, 22, 180–189.

Krumm, G., Lemos, V., & Arán Filippetti, V. (2014). Factor structure of the Torrance Tests of Creative Thinking Figural Form B in Spanish-speaking children: Measurement invariance across gender. *Creativity Research Journal*, 26(1), 72–81.

Lee, L. E., Ottwein, J. K., & Peters, S. J. (2021). Eight universal truths of identifying students for advanced academic interventions. In J. L. VanTassel-Baska & A. Stambaugh (Eds.), *Methods and materials for teaching the gifted* (pp. 61–79). Routledge.

Lee, L. E., Rinn, A. N., & Rambo-Hernandez, K. E. (2024). What happens after nomination? Evaluating the probability of gifted identification with the Torrance Test of Creative Thinking. *Gifted Child Quarterly*, 68(2), 119–136.

Lubart, T., & Zenasni, F. (2010). A new look at creative giftedness. *Gifted and Talented International*, 25(1), 53–57.

McBee, M. T., Peters, S. J., & Miller, E. M. (2016). The impact of the nomination stage on gifted program identification: A comprehensive psychometric analysis. *Gifted Child Quarterly*, 60(4), 258–278.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum Associates.

McNeish, D. (2017). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, 23(3), 412–433.

Mulaik, S. A. (2009). *Linear causal modeling with structural equations*. Chapman and Hall/CRC.

Muthén, L. K., & Muthén, B. O. (1998–2021). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Nijstad, B. A., De Dreu, C. K., Rietzschel, E. F., & Baas, M. (2010). The dual pathway to creativity model: Creative ideation as a function of flexibility and persistence. *European Review of Social Psychology*, 21(1), 34–77.

Obalasi, A., & Hoseini-Nasab, S. D. (2015). *Investigating the effect of Montessori teaching method on the creativity of 4- and 5-year-old preschool children in Tabriz. Instrucation and Evaluation (Educational Sciences), 7*(28), 81–98. [Persion]

Peters, S. J., Stambaugh, T., Makel, M. C., Lee, L. E., McBee, M. T., McCoach, D. B., & Johnson, K. R. (2023). The CASA criteria for evaluating gifted and talented identification systems: Cost, alignment, sensitivity, and access. *Gifted Child Quarterly*, 67(2), 137–150.

Pirkhaefi, A. R., Borjali, A., Delavar, A., & Eskandari, H. (2009). Investigating the effect of creativity instruction on metacognition components of creative thinking among university students. *Journal of Educational Leadership & Administration*, *3*(2), 51–61. [Perstion]

Plucker, J. A., Qian, M., & Wang, S. (2011). Is originality in the eye of the beholder? Comparison of scoring techniques in the assessment of divergent thinking. *The Journal of Creative Behavior*, 45(1), 1–22.

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.

Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement*, 21(2), 173–184.

Revelle, W. (2023). *psych: Procedures for psychological, psychometric, and personality research*. Northwestern University.

Rezaei, S., & Manouchehri, M. (2008). The study of reliability, validity and normalization of Torrance Creativity Test among Tehran high school teachers. *Journal of Psychology and Education*, 38(3), 47–68. [Perstion]

Rezvani, A., & Baradaran Tavakoli, D. (2024). Effects of musical skill on creative design among architecture students: Based on the Torrance Test and design evaluation. *Islamic Art Studies*, 21(53), 314–330. [Perstion]

Saeidi, A., & Shakiba, A. (2021). Review of Torrance Test of Creative Thinking Figural Form A scoring in Iran. Rooyesh, 10(1), 109–120. [Perstion]

Said-Metwaly, S., Fernández-Castilla, B., Kyndt, E., & Van den Noortgate, W. (2018). The factor structure of the Figural Torrance Tests of Creative Thinking: A meta-confirmatory factor analysis. *Creativity Research Journal*, 30(4), 352–360.

Samadi, N., Akbari Zardkhaneh, S., & Fathabadi, J. (2023). Assessing children's creativity: Adaptation of scoring procedures for the Persian version of the Torrance Test of Creative Thinking (Master's thesis, Shahid Beheshti University, Educational Psychology). [Persian]

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variable analysis: Applications to developmental research* (pp. 399–419). Sage.

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia & Analgesia*, 126(5), 1763–1768.

Shao, Y., Zhang, C., Zhou, J., Gu, T., & Yuan, Y. (2019). How does culture shape creativity? A mini-review. *Frontiers in Psychology*, 10, 1219.

Silvia, P. J. (2008). Creativity and intelligence revisited: A latent variable analysis of Wallach and Kogan (1965). *Creativity Research Journal*, 20(1), 34–39.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, 25(2), 173–180.

Tavana, Z., Akbari Zardkhaneh, S., & Shahidi, Sh. (2022). *Assessing adolescent creativity: Adaptation of the scoring system for the Persian version of the Torrance Test of Creative Thinking* [Master's thesis, Shahid Beheshti University, Educational Psychology program]. [Perstion]

Torrance, E. P. (1966). *The Torrance Tests of Creative Thinking-Norms-Technical Manual Research Edition-Verbal Tests, Forms A and B-Figural Tests, Forms A and B.* (Research ed.). Personnel Press.

Torrance, E. P. (1974). *The Torrance Tests of Creative Thinking-Norms-Technical Manual Research Edition-Verbal Tests, Forms A and B-Figural Tests, Forms A and B*. Personnel Press/Ginn and Copany.

Torrance, E. P. (1981). Empirical validation of criterion-referenced indicators of creative ability through a longitudinal study. *Creative Child and Adult Quarterly*, 6(3), 136–140.

Torrance, E. P. (1998). *The Torrance Tests of Creative Thinking norms—Technical manual figural (streamlined) forms A & B*. Scholastic Testing Service, Inc.

Torrance, E. P., & Ball, O. E. (1984). *Torrance Tests of Creative Thinking Streamlined* (Revised) Manual, Figural A and B. Scholastic Testing Service.

Torrance, E. P., & Haensly, P. A. (2003). Assessment of creativity in children and adolescents. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence, aptitude, and achievement* (2nd ed., pp. 584–607). Guilford Press.

Torrance, E. P., & Safter, H. T. (1999). *Making the creative leap beyond: Revision of the search for satori and creativity*. Creative Education Foundation Press.

Treffinger, D. J. (2004). Research on creativity. In D. J. Treffinger (Ed.), *Creativity and giftedness* (pp. 87–96). Corwin Press.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38(1), 1–10.

Xie, G., & Paik, Y. (2019). Cultural differences in creativity and innovation: Are Asian employees truly less creative than Western employees? *Asia Pacific Business Review*, 25(1), 123–147.

Yoon, C. H. (2017). A validation study of the Torrance Tests of Creative Thinking with a sample of Korean elementary school students. *Thinking Skills and Creativity*, 26, 38–50.

Yoruk, S. (2016). Application of Rasch measurement theory to the Torrance Tests of Creative Thinking, Figural Form A (*Doctoral dissertation, University of Georgia*).

Zakeri, S., & Barqi, I. (2016). Effect of the Torrance Creative Thinking Test on heart rate signal features. *Educational Strategy in Medical Sciences*, 8(6), 1–6. [Perstion]