

## خلاقیت: فرهنگ وابسته یا نه؟

### فهرست پاسخ‌های غیرخلاق تورنس در جمعیت ایرانی

سعید اکبری زردخانه<sup>۱</sup>، امیرحسین دانشمند کفترودی<sup>۲</sup>، جلیل فتح آبادی<sup>۳</sup>

#### چکیده

**هدف:** هدف پژوهش حاضر بررسی قابلیت استفاده از فهرست استاندارد پاسخ‌های غیرخلاق (۲۰۱۷) آزمون تفکر خلاق تورنس برای کارکنان ایرانی بود؛ به عبارت دیگر، تأثیر فرهنگ بر نمره اصالت و کارایی فهرست‌های استاندارد بررسی شد.

**روش:** این پژوهش بنیادی، توصیفی-پیمایشی و آمیخته بود. جامعه آماری کارکنان یک شرکت صنعتی ایرانی بودند که ۲۴۱ نفر (۱۴٪ زن) به شیوه تصادفی ساده انتخاب شدند. فهرست پاسخ‌های غیرخلاق بومی برای هر دو فرم آزمون از طریق تحلیل مضمون ایجاد و با فهرست استاندارد مقایسه شد.

**یافته‌ها:** تحلیل داده‌ها منجر به شناسایی ۱۱۱ طبقه پاسخ منحصر به فرد شد. مقایسه نمرات اصالت نشان داد که نمرات حاصل از فهرست بومی در نمره کل اصالت، به‌طور معناداری بالاتر از فهرست استاندارد بود (فرم A:  $d=0/24$ ؛ فرم B:  $d=0/12$ ). در فعالیت ۳ نیز فهرست بومی نمرات بالاتری تولید کرد (فرم A:  $d=0/55$ ؛ فرم B:  $d=0/55$ ). در مقابل، فهرست استاندارد برای فعالیت ۲ نمرات بالاتری داشت (فرم A:  $d=0/24$ ؛ فرم B:  $d=-0/82$ ). همچنین نمرات فعالیت ۱ نشان‌دهنده ناسازگاری میان دو فهرست بود.

**نتیجه‌گیری:** یافته‌ها حاکی از آن است که فهرست‌های استاندارد پاسخ‌های غیرخلاق برای ارزیابی خلاقیت در جمعیت ایرانی کارایی کافی ندارند و به کاهش مصنوعی نمرات اصالت منجر می‌شوند. بنابراین، تدوین و استفاده از فهرست‌های بومی، ضروری است.

**واژه‌های کلیدی:** آزمون تفکر خلاق تورنس، ارزیابی بین‌فرهنگی، فرهنگ وابسته، خلاقیت، ابزار بومی

<sup>۱</sup>. استادیار گروه روانشناسی کاربردی، دانشگاه شهید بهشتی، تهران، ایران. (نویسنده مسئول) (s\_akbari@mail.sbu.ac.ir)

<sup>۲</sup>. کارشناسی ارشد روان‌شناسی تربیتی، دانشکده روان‌شناسی و علوم تربیتی دانشگاه شهید بهشتی، تهران، ایران.

(amirdaneshmand08@gmail.com)

<sup>۳</sup>. دانشیار گروه روانشناسی کاربردی، دانشگاه شهید بهشتی، تهران، ایران. (j\_fathabadi@sbu.ac.ir)

## مقدمه

در دنیای امروز، با شتاب روزافزون تحولات فناوری و ظهور هوش‌های مصنوعی، بسیاری از مشاغل و فعالیت‌هایی که به مهارت‌های تکراری و روتین وابسته‌اند، به تدریج در حال ناپدید شدن یا کاهش اهمیت اقتصادی هستند (فری و آزبورن، ۲۰۱۷). این تحولات به‌ویژه در صنایعی مشهود است که خودکارسازی<sup>۱</sup> و هوش مصنوعی به‌طور گسترده‌ای در آن‌ها به کار گرفته شده‌اند (فری و آزبورن، ۲۰۱۷). بر اساس گزارش مجمع جهانی اقتصاد<sup>۲</sup> (۲۰۲۳)، تا سال ۲۰۲۷ حدود ۸۳ میلیون شغل ممکن است به دلیل خودکارسازی از بین بروند، درحالی‌که ۶۹ میلیون شغل جدید که نیازمند مهارت‌های خلاقانه هستند، ایجاد خواهند شد. این تغییرات نشان می‌دهد که مشاغل آینده به‌طور فزاینده‌ای به مهارت‌هایی مانند حل مسئله، نوآوری و خلاقیت وابسته خواهند بود (آتور میندل و رینولدز، ۲۰۲۲). در این راستا، خلاقیت به‌عنوان یکی از مهم‌ترین توانایی‌های انسانی، نقشی کلیدی در موفقیت فردی و سازمانی ایفا می‌کند (ساویر، ۲۰۱۹).

### مبانی نظری و توسعه فرضیه‌ها

با اینکه در طول چند دهه گذشته شاهد گسترش مطالعات حوزه خلاقیت برای جستجوی یک ابزار اندازه‌گیری قابل اعتماد بوده‌ایم، علی‌رغم توسعه بسیاری از ابزارها، اندازه‌گیری خلاقیت همواره با چالش‌های فراوانی همراه بوده است (بائر و مک کو، ۲۰۰۹؛ باتی، چامورو-پرموزیک و فورنهام، ۲۰۱۰؛ کافمن و همکاران، ۲۰۰۸؛ یون، ۲۰۱۷؛ شاو و همکاران، ۲۰۱۹). در این راستا، بررسی‌ها نشان می‌دهند که از ابزارهای متنوع موجود در این حوزه برخی بر سنجش شخصیت<sup>۳</sup> خلاق متمرکز هستند (مانند مقیاس فهرست صفات خلاق دومینو<sup>۴</sup>؛ دومینو، ۱۹۷۰)، برخی دیگر بر رفتارها و دستاوردهای<sup>۵</sup> خلاق (مانند فهرست رفتار خلاق هوچوار<sup>۶</sup>؛ هوچوار، ۱۹۷۹)، یا برخی بر سبک‌های<sup>۷</sup> خلاقیت (مانند فهرست انطباق-نوآوری کرتون<sup>۸</sup>؛ کرتون، ۱۹۷۶) یا محصولات<sup>۹</sup> خلاق (مانند فن ارزیابی توافقی<sup>۱۰</sup>؛ آمایل، ۱۹۸۲) تمرکز دارند. در این میان، ابزارهای متمرکز بر سنجش تفکر واگرا<sup>۱۱</sup> در پژوهش‌های آموزشی و شناسایی افراد مستعد، بیشترین کاربرد را داشته‌اند (کانلی، ۲۰۲۰).

1. automation
2. World Economic Forum
3. personality
4. Domino Creativity Scale of the Adjective Checklist
5. accomplishments
6. Hocevar Creative Behavior Inventory
7. styles
8. Kirton Adaptation-Innovation Inventory
9. products
10. Consensual Assessment Technique
11. divergen thinking

ویژگی اصلی تفکر واگرا، قالب پاسخ‌دهی باز آن‌ها است که در مقایسه با بسیاری از مفاهیم شناختی دیگر، امتیازدهی به آن را پیچیده‌تر می‌کند. گیلفورد (۱۹۶۷) تفکر واگرا را بر اساس چهار شاخص سیالی<sup>۱</sup>، اصالت<sup>۲</sup>، انعطاف‌پذیری<sup>۳</sup> و بسط<sup>۴</sup> ارزیابی کرد. سیالی به تعداد؛ اصالت به هوشمندی، ندرت و دوراز ذهن بودن؛ انعطاف‌پذیری به تعداد دسته‌ها یا خوشه‌ها؛ و بسط به جزئیات پاسخ‌ها اشاره دارد. براین اساس، گیلفورد و تیمش آزمون‌های تأثیرگذار متعددی برای ارزیابی چهار مؤلفه تفکر واگرا توسعه دادند که از آن میان آزمون تفکر خلاق تورنس (تورنس، ۱۹۷۴) یکی از محبوب‌ترین‌ها و پرکاربردترین ابزارهای سنجش خلاقیت به شمار می‌آید (اسنایدر و همکاران، ۲۰۱۹) که شواهد بسیاری در تأیید مناسبت آن وجود دارد. برای مثال، مطالعات نشان می‌دهد که این آزمون از کفایت‌های روان‌سنجی خوبی نظیر روایی پیش‌بین (تورنس ۱۹۷۴؛ هبرت و همکاران، ۲۰۰۲؛ تورنس، ۱۹۸۱) و روایی سازه (کلاپام، ۱۹۹۸؛ همبل و همکاران، ۲۰۱۸؛ کیم، ۲۰۰۶ ب؛ کیم و همکاران، ۲۰۰۶؛ کروم و همکاران، ۲۰۱۴، ۲۰۱۶؛ یون، ۲۰۱۷) برخوردار است.

دهه‌ها تحقیق در زمینه آزمون تفکر خلاق تورنس منجر به توسعه سه نسخه مختلف شده است: نسخه تصویری<sup>۵</sup>، نسخه کلامی<sup>۶</sup> و نسخه مختصر<sup>۷</sup>. از میان این سه، نسخه تصویری در تحقیقات و کاربردهای آموزشی، از جمله شناسایی جمعیت با استعداد، بیشترین استفاده را دارد (لی و همکاران، ۲۰۲۴). نسخه تصویری آزمون تفکر خلاق تورنس نسبت به نسخه کلامی<sup>۸</sup> کم‌تر تحت تأثیر توانایی زبانی مشارکت‌کننده قرار می‌گیرد (کیم، ۲۰۰۶ الف؛ یون، ۲۰۱۷). همچنین با اینکه نسخه تصویری و کلامی کاملاً مرتبط هستند؛ اما نسخه تصویری، جامع‌تر، معتبرتر و رواتر برای اندازه‌گیری خلاقیت است (کیم، ۲۰۱۷). نسخه تصویری آزمون خلاقیت تورنس دارای دو فرم A و B است و از سه فعالیت تشکیل شده است: ۱. ساخت تصویر<sup>۹</sup>، ۲. تکمیل تصویر<sup>۱۰</sup> و ۳. تصاویر (خطوط یا دایره) تکرار شده<sup>۱۱</sup> (تورنس و هنزلی، ۲۰۰۳).

اگرچه نسخه تصویری آزمون تورنس دارای چندین ویرایش است، اما اصل آزمون بدون تغییر باقی‌مانده است (کیم، ۲۰۰۶ الف). اولین نسخه آزمون تورنس (۱۹۶۶) دارای چهار مؤلفه سیالی،

1. fluency  
 2. originality  
 3. flexibility  
 4. elaboration  
 5. Figural (TTCT-F)  
 6. Verbal (TTCT-V)  
 7. Abbreviated (TTCT-A)  
 8. TTCT Verbal  
 9. picture construction  
 10. incomplete figures  
 11. repeated figures

انعطاف‌پذیری، اصالت و بسط بود (باثر، ۱۹۹۷؛ تورنس، ۱۹۶۶)؛ اما در ادامه دو مؤلفه انتزاع عنوان‌ها<sup>۱</sup> و عدم نتیجه‌گیری زود هنگام<sup>۲</sup> نیز به نسخه تصویری آن اضافه شده‌اند و مؤلفه انعطاف‌پذیری به دلیل رابطه بالا با سیالی حذف شده است (هیرت و همکاران، ۲۰۰۲).

تورنس (۱۹۶۶) برای پاسخ‌هایی که کمتر از ۲ درصد نمونه به آن‌ها اشاره کرده بودند دو امتیاز و پاسخ‌هایی که در محدوده ۲ تا ۵ درصد نمونه بودند یک امتیاز در نظر می‌گرفت؛ اما در نسخه آخر از این آزمون، امتیازدهی به اصالت بر اساس یک فهرست پاسخ‌های غیرخلاق<sup>۳</sup> که شامل پاسخ‌های پرتکرار است، صورت می‌گیرد. تورنس (۱۹۷۴) این فهرست‌ها را بر اساس نمونه‌ای ۵۰۰ نفری از مشارکت‌کنندگان که هر دو فرم A و B را تکمیل کردند، توسعه داد که به نظر فورثمن و همکاران (۲۰۲۰) برای تعیین فراوانی پاسخ‌ها در تکالیف تفکر واگرا، این حجم از نمونه، قابل‌قبول است. در این شیوه، برای فعالیت ۱ و ۲ در هر دو فرم، پاسخ‌هایی که در ۵ درصد یا بیشتر از نمونه‌ها مشاهده می‌شوند، غیراصیل تلقی می‌شوند و در فهرست پاسخ‌های غیرخلاق قرار می‌گیرند. در فعالیت ۳ فرم A، پاسخ‌هایی که در ۲۰ درصد یا بیشتر از نمونه مشاهده می‌شوند، غیراصیل هستند؛ در حالی که در فرم B این حد آستانه ۱۰ درصد است. در مجموع برای پاسخ‌هایی که در فهرست پاسخ‌های غیرخلاق قرار دارند صفر و باقی پاسخ‌ها یک امتیاز دریافت می‌کنند (آکار و همکاران، ۲۰۲۱).

با توجه به استفاده گسترده از آزمون تفکر خلاق تورنس برای شناسایی افراد با استعداد و توانمند، بررسی امتیازهای اصالت بدون استفاده از فهرست‌های پاسخ‌های غیرخلاق از پیش تعیین شده، بسیار دشوار و غیرعملی خواهد بود؛ از این رو تورنس با ارائه فهرست‌های پاسخ‌های غیرخلاق، فرایند امتیازدهی اصالت را ساده‌تر و کارآمدتر کرده است (آکار و همکاران، ۲۰۲۱). لازم به ذکر است، امتیازدهی به اصالت در تکالیف تفکر واگرا معمولاً به فهرست‌های پاسخ‌های غیرخلاق متکی است که برای هر نمونه به‌طور جداگانه تنظیم می‌شوند (فورثمن و همکاران، ۲۰۲۰)؛ لذا رویکرد تورنس در استفاده از فهرست‌های پاسخ‌های غیرخلاق استاندارد (به‌جای فهرست‌های نمونه‌محور) در میان سایر آزمون‌های خلاقیت مبتنی بر تفکر واگرا نسبتاً غیرمعمول است.

با توجه به اهمیت فهرست‌های پاسخ‌های غیرخلاق در فرایند امتیازدهی اصالت، لازم است که این فهرست‌ها به نحوی تنظیم شوند که حداقل در یک فرهنگ خاص به‌طور عمومی قابل استفاده باشند (تورنس، ۲۰۱۷). برخی پژوهشگران نگرانی‌هایی را درباره به‌روزرسانی این فهرست‌ها مطرح کرده‌اند؛ برای مثال، موشیروود و لوبارت (۲۰۰۱) اشاره کرده‌اند که نسخه فرانسوی فهرست‌های

1. abstractness of titles

2. resistance to premature closure

3. zero originality list

پاسخ‌های غیرخلاق بر اساس هنجارهای سال ۱۹۷۶ تدوین شده است. بررسی آن‌ها نشان داد که واژه "رایانه" بیش از ۵ درصد از پاسخ نمونه پژوهش آن‌ها در سال ۲۰۰۱ را شامل می‌شده که براین اساس باید به‌عنوان پاسخی غیرخلاق در نظر گرفته شود. با این حال، از آنجایی که نمونه‌ها با استفاده از هنجارهای سال ۱۹۷۶ امتیازدهی شده بودند، برخی از ایده‌هایی که در سال ۱۹۷۶ رایج نبودند (مانند رایانه) همچنان ممکن است به‌عنوان پاسخ‌های خلاق منظور شوند. در همین راستا، در یک مطالعه مروری از تغییرات نمرات تفکر خلاق در طول زمان، کیم (۲۰۱۱) هشدار داد که عدم به‌روزرسانی فهرست‌های پاسخ‌های غیرخلاق می‌تواند به‌طور مصنوعی نمرات اصالت را افزایش دهد.

در مواجهه با این چالش، فهرست پاسخ‌های غیرخلاق در سال ۲۰۱۷ با استفاده از ۱۰۲۰ شرکت‌کننده برای فرم A و ۱۲۰۰ شرکت‌کننده برای فرم B به‌روزرسانی شدند. با این حال، این فهرست‌ها هنوز از نظر سازگاری با نمونه‌های دیگر مورد آزمایش قرار نگرفته‌اند و هیچ مطالعه تجربی در مورد تأثیر این تغییرات وجود ندارد (آکار و همکاران، ۲۰۲۱). افزون بر این، فهرست‌های به‌روز شده ممکن است برای تمامی جمعیت‌ها قابل استفاده نباشند؛ چراکه اصالت می‌تواند تحت تأثیر فرهنگ قرار گیرد و در میزان محبوبیت پاسخ‌ها در میان گروه‌های قومی، نژادی و فرهنگی تفاوت ایجاد کند (کیم، ۲۰۰۶؛ میلر، ۱۹۹۵). در همین راستا استرنبرگ (۲۰۱۹) معتقد است که خلاقیت درست مانند هوش، باید به‌صورت محلی ارزیابی شود و پیشنهاد می‌کند که: "ما باید توافق کنیم که ارزیابی‌های ما از آنچه معمولاً به‌عنوان خلاقیت در نظر گرفته می‌شود نشان‌دهنده هنجارهای محلی است." همچنین سائکی و همکاران (۲۰۰۱) در مقایسه پاسخ‌های ژاپنی‌ها و آمریکایی‌ها متوجه این تفاوت در میان گروه‌های فرهنگی شدند. خارخورین و صمدپورمطلبی (۲۰۰۸) نیز در مطالعه‌ای تفاوت بین دانشجویان ایرانی، روس و آمریکایی را مشاهده کرده و این تفاوت را به تفاوت در فرهنگ نسبت دادند. تجربیات امتیازدهندگان آزمون تفکر خلاق تورنس نیز تأیید کرده است که ممکن است تفاوت‌های فرهنگی در نمره‌دهی اصالت وجود داشته باشد (کیم، ۲۰۰۶ ب)؛ لذا نتیجه می‌شود که فراوانی و ماهیت پاسخ‌های مختلف در میان افراد با فرهنگ‌های مختلف می‌تواند متفاوت باشد. در نتیجه اگر آزمون تفکر خلاق تورنس در فرهنگ‌های مختلف بدون گروه‌های هنجار مناسب برای هر جمعیت و متعاقباً فهرست پاسخ‌های غیرخلاق مختص فرهنگ استفاده نشود، نتایج گمراه‌کننده‌ای به همراه خواهد داشت (کیم، ۲۰۰۶ الف)؛ لذا مهم است که تفاوت‌های فرهنگی و جمعیتی در مورد این خصوص مدنظر قرار گیرد تا اطمینان حاصل شود که آزمون تورنس به‌طور مداوم و عادلانه در میان گروه‌های نژادی، قومی و فرهنگی امتیازدهی می‌شود (برلو و همکاران، ۲۰۱۹).

آکار و همکاران (۲۰۲۱) در پژوهشی با هدف بررسی کارایی آخرین فهرست پاسخ‌های غیرخلاقانه در گروه دانشجویان آفریقایی-آمریکایی، مشاهده کردند که نمرات اصالت حاصل از فهرستی که بر اساس نمونه پژوهش تدوین شده بود، نسبت به نمرات فهرستی ۲۰۱۷ با بهبودهایی همراه بوده است. از سوی دیگر سعیدی و شکبیا (۱۴۰۰) در پژوهشی بر روی نمونه‌ای از کودکان ایرانی دریافتند که ماهیت و فراوانی پاسخ آزمون تصویری تورنس مشارکت‌کنندگان ایرانی و آمریکایی تفاوت قابل توجهی دارند و نتیجه‌گیری کردند که نمره‌گذاری بخش اصالت پاسخ‌ها براساس فهرست آمریکایی برای مشارکت‌کنندگان ایرانی مناسبی ندارد.

باید توجه داشت که در پژوهش سعید و شکبیا (۱۴۰۰) شیوه نمره‌گذاری براساس راهنمای رسمی آزمون (چه شیوه قدیمی و چه بر مبنی فهرست پاسخ‌های غیرخلاق) صورت نگرفته است، لذا نمی‌توان بر اساس این پژوهش در رابطه با مناسب‌ترین فهرست منتشرشده (تورنس، ۲۰۱۷) تصمیم‌گیری کرد؛ با این حال، نتایج این پژوهش تأیید می‌کند که تفاوت‌های فرهنگی و زمانی بر محتوا و فراوانی پاسخ‌های جمعیت ایرانی تأثیرگذار است. از سوی دیگر، به دلیل مواجهه بخش صنعت با چالش‌های خودکارسازی و نیاز به نوآوری، به یکی از بسترهای اصلی استفاده گسترده از ابزارهای سنجش خلاقیت در فرایندهای سازمانی شده است (مام‌فورد و همکاران، ۲۰۲۴). از این رو، اطمینان از دقت و عدالت ابزارهای مورد استفاده در این بافت خاص، اهمیتی دوچندان می‌یابد.

هدف از پژوهش حاضر، تدوین فهرست بومی پاسخ‌های غیرخلاق برای نسخه تصویری آزمون تفکر خلاق تورنس و بررسی میزان سازگاری این فهرست با فهرست استاندارد ۲۰۱۷ در جمعیت کارکنان ایرانی بود. این مطالعه همچنین به دنبال ارزیابی تأثیر تفاوت‌های فرهنگی بر نمره‌گذاری و کارایی فهرست‌های استاندارد پاسخ‌های غیرخلاق ۲۰۱۷ بوده است.

### روش‌شناسی پژوهش

این پژوهش از نظر هدف بنیادی و از نظر روش گردآوری داده توصیفی پیمایشی است (سرمد، بازرگان و حجازی، ۱۳۸۰).

همچنین از حیث نوع داده‌ها، در دسته پژوهش‌های آمیخته اکتشافی<sup>۱</sup> قرار می‌گیرد (بازرگان، ۱۴۰۰). جامعه این پژوهش را کارکنان یک شرکت صنعتی که در سال ۱۴۰۳ به مرکز ارزیابی این شرکت مراجعه کرده‌اند تشکیل می‌دهد. نمونه‌گیری در دو مرحله انجام شد؛ ابتدا مشارکت‌کنندگان بر اساس فرآیندهای مرکز ارزیابی شرکت جهت اجرای آزمون انتخاب شدند. سپس از میان پاسخ‌برگ‌های تکمیل‌شده، ۲۴۱ مورد به روش نمونه‌گیری تصادفی ساده توسط

1. exploratory mixed-method

پژوهشگران انتخاب شد. ۱۲۷ نفر فرم A و ۱۱۴ نفر فرم B را تکمیل کردند. ۱۴ درصد حجم نمونه را زنان تشکیل دادند و سن مشارکت کنندگان بین ۳۰ تا ۵۵ سال متغیر بود ( $M = 43/41$ ;  $sd = 6/20$ ).

### ابزار پژوهش

**آزمون تصویری تفکر خلاق تورنس - نسخه جوانان ایرانی:** مشارکت کنندگان به عنوان 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، در حالی که فعالیت ۲ (تکمیل تصویر) شامل ده محرک و فعالیت ۳ (خطوط یا دایره‌ها در فرم A و B) شامل ۳۶ محرک تکراری هستند.

### شیوه اجرای پژوهش

این مطالعه در دو مرحله اصلی اجرا گردید: ۱. تدوین فهرست پاسخ‌های غیرخلاق؛ ۲. بررسی سازگاری<sup>۱</sup> دو فهرست (جدید و قدیمی).

### مرحله اول. تدوین فهرست پاسخ‌های غیرخلاق: این مرحله در سه گام اجرا گردید.

**۱. استخراج طبقات پاسخ:** ابتدا، تمامی پاسخ‌های ارائه‌شده شرکت کنندگان به دو فرم نسخه تصویری آزمون تفکر خلاق تورنس جمع‌آوری شدند. در این مرحله مجموعاً ۴۸۲۰ داده تصویری شناسایی گردید. سپس، به منظور استخراج طبقات پاسخ، داده‌های تصویری به روش تحلیل مضمون<sup>۲</sup> مورد تحلیل قرار گرفتند. این روش به پژوهشگر امکان می‌دهد تا به طور نظامند به بررسی و طبقه‌بندی پاسخ‌های مشارکت کنندگان پردازد (براون و کلارک، ۲۰۰۴). فرایند تحلیل مضمون شامل مراحل زیر بود:

- آشنایی با داده‌ها<sup>۳</sup>: به عنوان نخستین مرحله، ابتدا تمامی تصاویر و پاسخ‌های ثبت شده در فرم‌های آزمون به دقت مرور شدند تا پژوهشگر با محتوای کلی، محدوده و نقطه ثقل پاسخ‌های رسم شده آشنا شود. این اقدام با ایجاد دیدی کلی از محتوای پاسخ‌ها، دقت در انتخاب کدها در مرحله بعد را افزایش می‌دهد. برای مثال اصطلاحات "بته‌جقه"، "طرح اسلیمی" و "نقش قالی" سه عنوانی بودند که مشارکت کنندگان برای پاسخی واحد ارائه

1. consistency

2. thematic analysis

3. familiarization

می‌دادند؛ لذا پژوهشگر پس از آشنایی با داده‌ها تعیین می‌کند که برای تمام این پاسخ‌ها یک کد واحد ثبت کند.

• تولید کدهای اولیه<sup>۱</sup>: در این مرحله، ویژگی‌های اصلی در هر پاسخ شناسایی و به صورت کدهای اولیه ثبت شدند. این کدها نمایانگر عناصر تولیدشده از محرک‌های تصویری در هر تکلیف بودند. در این مرحله از مجموع ۴۸۲۰ داده، ۱۷۲۶ کد منحصر به فرد اولیه ثبت گردید (برای مثال مراجعه شود به جدول ۱).

• جستجوی مضمون‌ها<sup>۲</sup>: کدهای اولیه در گروه‌های مفهومی دسته‌بندی و مضمون‌های بالقوه شناسایی شدند. این مضمون‌ها نشان‌دهنده الگوهای مشترک و تفاوت‌های موجود در پاسخ‌های شرکت‌کنندگان بودند. در این مرحله ۱۷۲۶ کد اولیه در ۳۸ مضمون اصلی جای گرفتند (برای مثال مراجعه شود به جدول ۱).

• بازبینی مضمون‌ها<sup>۳</sup>: در این مرحله، مضمون‌های اصلی شناسایی شده بازبینی و پالایش شدند تا در صورت امکان به مضمون‌های فرعی تفکیک یابند تا اطمینان حاصل شود که هر مضمون به‌طور دقیق و معتبر نمایانگر داده‌ها است. در این مرحله نهایتاً ۱۱۱ مضمون فرعی تحت عنوان مضمون‌های اصلی طبقات پاسخ شناسایی شدند. لازم به ذکر است که طبقه ۱۱۱ طبقه "موارد نادر"<sup>۴</sup> نام گرفت و شامل مواردی می‌شد که در هیچ‌یک از مضمون‌ها جای نمی‌گرفتند (برای مثال مراجعه شود به جدول ۲).

• تعریف و نام‌گذاری مضمون‌ها<sup>۵</sup>: مضمون‌های نهایی، به گونه‌ای تعریف و نام‌گذاری شدند که به‌طور واضح و مختصر محتوای هر مضمون را بازتاب دهند.

• تهیه گزارش نهایی<sup>۶</sup>: در نهایت، گزارشی جامع از مضمون‌های شناسایی شده تهیه شد که شامل توضیحات مفصل درباره هر مضمون و مصداق‌هایی از پاسخ‌های شرکت‌کنندگان بود. در واقع این مضمون‌ها طبقات کل پاسخ‌ها را تشکیل می‌دادند.

## ۲. تشکیل جداول فراوانی طبقات پاسخ تکالیف: در این مرحله پس از تعیین عناوین

طبقات پاسخ‌ها، جدول فراوانی پاسخ هر تکلیف به‌طور مجزا تشکیل گردید. بدین منظور

1. generate initial codes  
2. search for themes  
3. review theme  
4. uncommon  
5. define themes  
6. reporting



فراوانی و فراوانی نسبی پاسخ‌های هر طبقه به تفکیک تکالیف محاسبه شد تا از این جداول در تدوین فهرست پاسخ‌های غیرخلاق استفاده گردد.

۳. **تدوین فهرست پاسخ‌های غیرخلاق:** در فعالیت‌های ۱ و ۲، فهرست‌ها بر اساس پاسخی که ۵ درصد یا بیشتر از فراوانی کل پاسخ‌ها را به خود اختصاص داده بودند، تشکیل شدند. تورنس در تدوین فهرست‌های فعالیت ۳ از ۲۰ درصد به‌عنوان برش در فرم  $A$  و ۱۰ درصد در فرم  $B$  استفاده کرد. باین‌حال در نمونه این پژوهش، نقطه برش‌های ۱۰ و ۲۰ درصد پاسخ‌های اصلی را از پاسخ‌های غیراصولی متمایز نمی‌کرد، زیرا هیچ پاسخی با فراوانی ۱۰ و ۲۰ درصد ارائه نشده بود. علاوه‌برآن، فرم  $A$  و فرم  $B$  فرم‌های موازی هستند و هیچ تفاوت ظاهری بین فعالیت خط و فعالیت دایره وجود ندارد. از این‌رو، در هر دو فرم  $A$  و فرم  $B$  از نقطه برش ۵ درصد استفاده شد.

**مرحله دوم. بررسی سازگاری دو فهرست (جدید و قدیمی):** پس از تدوین فهرست نهایی بر مبنای گروه نمونه، شاخص اصالت برای تمام برگه‌ها بر اساس هر دو فهرست قدیمی و جدید نمره‌گذاری شدند. به‌منظور مقایسه نمرات حاصل از دو فهرست، از نرم‌افزار آرنس‌۴.۴.۲ (گروه مرکزی آر، ۲۰۲۳) انجام گرفت. تمامی تحلیل‌ها به‌گونه‌ای انجام شدند که تغییرات صرفاً در سطح فهرست‌ها مورد ارزیابی قرار گیرند؛ بنابراین امتیازات تشویقی<sup>۲</sup> اصالت لحاظ نشدند. برای ارزیابی تفاوت میانگین نمرات اصالت به‌دست‌آمده از دو فهرست، از آزمون‌های تی وابسته استفاده شد (استودنت، ۱۹۰۸).

این آزمون‌ها برای نمرات کل و فعالیت‌های ۲ و ۳ انجام شدند که به‌صورت متغیرهای پیوسته هستند. به‌منظور بررسی تفاوت در توزیع نمرات زوجی در فعالیت ۱، از آزمون مک‌نمار (مک-نمار<sup>۳</sup>، ۱۹۴۷) استفاده گردید. برای تحلیل همبستگی میان نمرات فهرست جدید و قدیم، از ضریب همبستگی پیرسون (پیرسون، ۱۸۹۵) برای متغیرهای پیوسته شامل نمرات کل، فعالیت ۲ و فعالیت ۳ استفاده شد. برای فعالیت ۱ که نمرات زوجی دارد، از ضریب فی (یول، ۱۹۱۲) بهره گرفته شد. این ضرایب به‌منظور ارزیابی شدت ارتباط بین نمرات حاصل از دو فهرست مورد استفاده قرار گرفتند. اندازه اثر تفاوت میانگین‌ها با استفاده از ضریب دی کوهن<sup>۴</sup> (کوهن، ۱۹۸۸) برای داده‌های مکرر<sup>۵</sup> محاسبه شد. اندازه اثر به‌عنوان معیاری برای شدت تفاوت بین میانگین‌های دو فهرست مورد

1. R Core Team

2. bonus scores

3. McNemar

4. cohen's d coefficient

5. repeated

استفاده قرار گرفت. تمام تحلیل‌ها با استفاده از بسته‌های موجود در نرم‌افزار آر شامل استتس<sup>۱</sup> و بسته اف‌سایز<sup>۲</sup> (تورچیانو، ۲۰۲۰) انجام شدند. همچنین همبستگی نمرات حاصل از دو فهرست با حالت فرضی همبستگی کامل ( $r = 1$ ) مقایسه شدند. از آنجا که این مقدار فرضی همبستگی کامل از ماتریس همبستگی استخراج نشده بود، از فرمولی استفاده شد که برای مقایسه دو همبستگی در نمونه‌های مستقل طراحی شده است (اید و همکاران، ۲۰۱۱). این مقایسه با استفاده از نرم‌افزار رایگان لن‌هارد و لن‌هارد (۲۰۱۴) صورت گرفت که بر اساس فرمول ارائه‌شده توسط اید و همکاران (۲۰۱۱) طراحی شده بود.

### یافته‌های پژوهش

در فرایند تدوین فهرست پاسخ‌های غیرخلاق از ۴۸۲۰ داده تصویری اولیه ۱۷۲۶ کد منحصر به فرد استخراج گردید. این کدها نمایانگر عناصر تولیدشده از محرک‌های تصویری در هر تکلیف بودند (برای مطالعه بیشتر مراجعه شود به دانشمند، ۱۴۰۳). پس از تولید کدهای اولیه، با دسته‌بندی آن‌ها در گروه‌های مفهومی، مضمون‌های اصلی و بالقوه تشکیل شدند. در این فرایند ۳۸ مضمون اصلی ایجاد شد که نشان‌دهنده الگوهای مشترک و تفاوت‌های موجود در پاسخ‌های شرکت‌کنندگان بودند (جدول ۱).

جدول ۱. نمونه‌ای از مضمون‌های اصلی و کدهای اولیه

مضمون اصلی	مصادیق (کد اولیه)
ابزار	ارلن ظرف آزمایشگاه شیمی، اره، استوانه مدرج، بیلچه، داس، ذره‌بین، قیف
اجرام آسمانی	سیاره زحل، ماه، کسوف، منظومه راه شیری
انسان	زنی با مقنعه، دلچک، دختر بچه، دست انسان، معده انسان، نیم‌رخ چهره جادوگر
پوشاک	پوتین، عینک آفتابی، کلاه ایمنی، بند کتانی، پیراهن
تابلو و علائم	چراغ‌راهنمایی رانندگی، لوگو، نماد ارتش سرخ، پوستر (برچسب روی لباس)، تابلو جاده، فلش راهنما
حمل و نقل	بالگرد، پنجره زیر دریایی، دودکش کشتی، فضاپیما، گاری دستی، واگن قطار
ساختمان	پله، پنجره، آسانسور، پارکینگ، هایپرمارکت، سوله ذخیره گندم، سردر دانشگاه، مسجد، کلبه
طبیعت	بذر گیاه، تالاب، باران، ابر، گل، جنگل، تپه، کوهستان، موج دریا
لوازم منزل	کابینت، پرده، مبلی، تلویزیون، بشقاب، لیوان، تختخواب، کتابخانه
نماد و اصطلاحات	تسمه‌نقاله، شیر آبگیری تانکر حمل آب، قرقره بالا بر، پره ژنراتور، پروفیل سازه، غلطک نورد ورق و ...
صنعت	

در ادامه فرایند، مضمون‌های اصلی شناسایی شده بازبینی و پالایش شدند تا در صورت امکان به مضمون‌های فرعی تفکیک یابند تا اطمینان حاصل شود که هر مضمون به‌طور دقیق و معتبر

<sup>1</sup>. stats  
<sup>2</sup>. effsize

نمایانگر داده‌ها است. در این فرایند فراوانی کدها نیز در نظر گرفته شد تا از قرارگیری یک کد با فراوانی بسیار بالا در کنار کدهای دیگر در یک طبقه جلوگیری شود؛ برای مثال در مضمون اصلی "ابزار" کد داس با فراوانی ۶۳ به عنوان طبقه‌ای مستقل در نظر گرفته شد و باقی کدهای مشابه (برای مثال چنگک باغبانی با فراوانی ۴) در طبقه‌ای دیگر با نام "ابزار کشاورزی" قرار گرفتند. نهایتاً در این مرحله ۱۱۱ مضمون فرعی تحت عنوان طبقات پاسخ شناسایی شدند که نمونه‌ای از آن‌ها در جدول ۲ گزارش شده است.

جدول ۲. بخشی از طبقات پاسخ آزمون تصویری تفکر خلاق تورنس

مضمون اصلی	کد	مضمون فرعی / طبقات پاسخ	مصادیق
	۱	ابزار دوخت و دوز	قیچی، نخ و سوزن، قرقره نخ، سنجاق قفلی، کاموا، کلاف نخ، کلاف کاموا، الگوی دوخت لباس و...
ابزار	۲	ابزار عمومی علمی و آزمایشگاهی	ذره‌بین، استوانه مدرج، آهن‌ریا، عدسی محدب/ مقعر، ظروف آزمایشگاه، پریسکوپ، تلسکوپ و...
	۳	ابزار کشاورزی و باغبانی (به جز داس)	انواع چنگک باغبانی، شن کش، بیل باغبانی، بیلچه، تکیه‌گاه درخت انگور (ستون)، تیشه باغبانی، تیغه (برای شخم‌زدن) و...
انسان	۱۸	اندام‌های خارجی بدن که از بیرون قابل مشاهده هستند	دست، چشم، دهان، گوش، پا، دماغ، لب یا لبخند، انگشتان پا، چشمان کودک، دندان، سیبیل، شکم، لاله گوش و...
	۱۹	اندام‌های داخلی بدن انسان که دیده نمی‌شوند	قلب، کلیه، جنین، ریه، معده، رحم زن، تخمدان، جنین در رحم، قلو، نای گلو
جانداران (به جز انسان)	۴۳	انواع حشرات، کرم و هزارپا و اندام‌شان (به جز پروانه‌ها)	کرم، هزارپا، زنبور، کرم خاکی، کرم ابریشم، ملخ، مورچه، کفش دوزک، شفیوه، پله کرم ابریشم، چشم حشره، ملخ، سر زنبور، سنجاقک، عنکبوت، کرم ماهی، کوسه، نهنگ، دلفین، ستاره دریایی، هشت‌پا، بچه‌ماهی، جانور دریایی، دهان کوسه، دهان ماهی، گوش ماهی،
	۴۴	آبزیان و اندام‌ها و موارد مرتبط	آسانسور و پارکینگ و موارد مشابه
ساختمان	۶۶	بناهای و ساختمان‌های خاص (به جز مواردی که دقیقاً به بنا اشاره شده باشد)	آسانسور، پارکینگ از نمای بالا، پارکینگ خودرو، انباری، خط‌کشی پارکینگ اتومبیل، درب آسانسور، سوله پارکینگ عمومی، فضای پارکینگ پاساژ، راهرو مسجد، فانوس دریایی، ایستگاه اتوبوس با سایه‌بان، برج مراقبت، کلیسا، دژ یا قلعه، سر در دانشگاه، سر در مسجد، سر در ورودی شرکت، قلعه قدیمی، مناره، مونومان (بنای یادبود)، ورودی تالار، قبرستان
لوازم منزل	۹۰	فرش، قالی، پرده و موارد مرتبط	فرش، قالی، قالیچه، پرده، پرده پنجره، دار قالی، سجاده، فرش گلیم، گلیم
	۹۱	کالای خواب	تخت‌خواب، تشک، جارختی، روتشکی، لحاف، متکا (بالشت)
نماد و اصطلاحات صنعت	۱۰۶	اصطلاحات، نمادها و سازه‌های مرتبط با پایینگ	شیر کنترل، گیج فشارسنج، لوله حفاری، اتصال مخزن و لوله، برش از مقطع لوله، بلابند فلنج در صنعت لوله‌کشی، دیافراگم شیر کنترل، شیر بال ولو، شیر دروازه‌ای، شیر صنعتی، گیج فشار، قطع لوله، مسیر احداث خطوط لوله نفت، نماد شیر چک.
	۱۰۸	سازه‌های عمومی مشترک بین صنایع مختلف	تسمه‌نقاله، شیر آبیگری تانکر حمل آب، قرقره بالا بر، پره ژنراتور، پروپیل سازه، غلطک نورد ورق، غلطک چرخه تولید، محل نشیمن تیغه توربین، پره توربین و ...
موارد نادر	۱۱۱	مواردی که در هیچ یک از طبقات جای نمی‌گیرند	اله جنگ در مصر باستان، ضمیر ناخودآگاه، عصای عزرائیل، دستگاه گلاب‌گیری و...

جدول طبقات پاسخ در تدوین فهرست پاسخ‌های غیرخلاق و محاسبه نمرات اصالت مورد استفاده قرار می‌گیرد؛ لذا می‌بایست فراوانی طبقات در تمام تکالیف به صورت مجزا مشخص گردد. در این راستا مجموعاً ۲۳ جدول فراوانی طبقات توسعه یافت. یک جدول برای فعالیت ساخت تصویر در هر دو فرم آزمون (A و B)؛ ده جدول برای فعالیت تکمیل تصاویر الف در فرم A؛ ده جدول برای فعالیت ساخت تصاویر ب در فرم B؛ یک جدول برای فعالیت خطوط موازی (فرم A) و یک جدول برای فعالیت دایره‌ها (فرم B). برای مثال، بخشی از فراوانی طبقات فعالیت تکمیل تصاویر الف در جدول ۳ گزارش شده است (برای مطالعه بیشتر مراجعه شود به دانشمند و همکاران، ۱۴۰۳).

**جدول ۳. نمونه‌ای از فراوانی نسبی طبقات پاسخ در تکالیف فعالیت تکمیل تصاویر الف**

کد طبقه	تکلیف ۱	تکلیف ۲	تکلیف ۳	تکلیف ۴	تکلیف ۵	تکلیف ۶	تکلیف ۷	تکلیف ۸	تکلیف ۹	تکلیف ۱۰
۷۸	۳۴/۴۰	۱/۵۶	۰	۰/۸۱	۰	۰	۶/۶۷	۱/۵۹	۴/۳۹	۰/۸۶
۸۳	۱۶/۱۰	۳/۱۳	۰/۸۲	۰	۱/۸۵	۰/۸۲	۰	۰	۱/۷۵	۰/۸۶
۱۰۴	۶/۴۰	۰	۰	۰	۱۰/۱۹	۱/۶۴	۰/۹۵	۰	۰/۸۸	۰
۱۲	۳/۲۰	۰	۰	۰/۸۱	۴/۶۳	۰/۸۲	۰	۰	۰/۸۸	۰/۸۶
۳	۳/۲۰	۰	۰/۸۲	۰	۰	۰	۰/۹۵	۰	۰	۰
۲۰	۲/۴۰	۵/۴۷	۰/۸۲	۳/۲۵	۲۰/۳۷	۰	۱۱/۴۳	۱/۵۹	۷/۰۲	۰
۲۳	۲/۴۰	۰	۰	۰	۰	۰/۸۲	۵/۷۱	۰	۱/۷۵	۱/۷۲
۱۴	۲/۴۰	۰/۷۸	۰	۱/۶۳	۲/۷۸	۴/۱۰	۰/۹۵	۰	۱/۷۵	۰
۶۹	۲/۴۰	۰	۰	۰	۰/۹۳	۱/۶۴	۰/۹۵	۰	۰	۵/۱۷
۱	۲/۴۰	۰	۱/۶۴	۰	۰/۹۳	۲/۴۶	۰	۰	۰	۰

بر اساس جدول ۳ مشاهده می‌شود که پرتکرارترین پاسخ برای تکلیف ۱ مربوط به طبقه ۷۸ (انواع درخت یا بخشی از آن) است با فراوانی نسبی ۳۴/۴۰ که نمره اصالت صفر کسب می‌کند. از سوی دیگر طبقه ۸۳ (انواع گل) نیز فراوانی بیش‌تر از ۵ درصد را برای تکلیف ۱ ثبت کرده است که معادل با نمره صفر اصالت است. به‌عنوان مثالی دیگر، طبقه ۶۹ (انواع خانه‌های یک طبقه) با فراوانی ۵/۱۷ در تکلیف ۱۰ نمره اصالت صفر دریافت می‌کند. به‌منظور بررسی تفاوت‌های فهرست ۲۰۱۷ و فهرست تدوین شده در این پژوهش، بخش‌هایی از دو فهرست پاسخ‌های غیرخلاق در جدول ۴ گزارش شده است.

**جدول ۴. بخش‌هایی فهرست پاسخ‌های غیرخلاق جدید و قدیم**

فعالیت ۱ فرم B		محرك ۱۰ فعالیت ۲ فرم B		محرك ۶ فعالیت ۲ فرم A	
فهرست قدیم	فهرست جدید	فهرست قدیم	فهرست جدید	فهرست قدیم	فهرست جدید

کالسکه	کوبه، تپه و...	کتاب	ابر، دود و مشابه آن	حیوان
چوبه دار	کودک	پرچم	اندام‌های داخلی بدن	لویا
در، پنجره و مشابه آن	چهره و بدن انسان	طرح هندسی	انواع قایق	انواع قایق
نورهای شهری	حروف	چوبه دار	اصطلاحات پتروشیمی	تخم مرغ
پوشاک سلطنتی	اعداد	اعداد	اندام‌های خارجی بدن	کلاه
علامت سؤال	ابزار	کفش و چکمه	حمل و نقل هوایی	هات داگ
مار				

چهره یا بدن انسان

### بررسی سازگاری نمرات حاصل از دو فهرست

در جدول ۵، بررسی نمرات حاصل از فهرست جدید نشان می‌دهد که می‌توان وجود توزیعی متقارن را برای نمرات اصالت کل متصور بود (کجی<sup>۱</sup> = ۰/۳۷ و -۰/۳۲ و کشیدگی<sup>۲</sup> = ۰/۹۷ و ۰/۶۷). برای فرم A و B به ترتیب (کلاین، ۲۰۱۶). از سوی دیگر، برای نمرات فعالیت‌های ۲ (کجی = -۰/۳۵ و ۰/۶۰؛ کشیدگی = ۰/۸۵ و -۰/۵۷). برای فرم A و B به ترتیب) و ۳ (کجی = ۰/۲۳ و ۰/۱۱؛ کشیدگی = ۰/۱۱ و -۱/۱۰) نیز به طریق مشابه شاهد توزیع تقریباً متقارن هستیم. همچنین تمام مقادیر کجی و کشیدگی نمرات حاصل از فهرست قدیمی نیز در بازه قابل قبول -۲ تا +۲ قرار دارند (کلاین، ۲۰۱۶) جز شاخص‌های فعالیت ۲ در فرم B (کجی = -۲/۴۴ و کشیدگی = ۱۱/۷۸)؛ در این راستا، مقایسه شاخص کشیدگی نمره قدیم و جدید این فعالیت نشان می‌دهد که فهرست جدید با کشیدگی کم‌تر و منفی عملکرد بهتری در تشخیصی تفاوت‌های میان گروه نمونه داشته است.

جدول ۵. شاخص‌های توصیفی نمرات اصالت هر دو فهرست

	فرم B (۱۱۴)				فرم A (۱۲۷)			
	کشیدگی	کجی	انحراف استاندارد	میانگین	کشیدگی	کجی	انحراف استاندارد	میانگین
کل جدید	-۰/۶۷	-۰/۳۲	۶/۷۴	۱۷/۸۵	۰/۹۷	۰/۳۷	۵/۶۴	۱۷/۵۵
کل قدیم	-۰/۲۹	۰/۰۹	۴/۹۰	۱۷/۰۴	۱/۵۰	۰/۳۳	۴/۸۹	۱۶/۲۴
ف ۱ جدید	-۱/۸۰	-۰/۴۶	۰/۴۹	۰/۶۱	-۱/۷۷	-۰/۵۰	۰/۴۹	۰/۶۲
ف ۱ قدیم	-۰/۳۷	-۱/۲۸	۰/۴۲	۰/۷۷	-۰/۵۰	-۱/۲۳	۰/۴۳	۰/۷۶

<sup>۱</sup>. skewness

<sup>۲</sup>. kurtosis

قدیم							
ف ۲	۶/۵۸	۲/۶۰	-۰/۳۵	-۰/۸۵	۶/۹۳	۲/۵۴	-۰/۵۷
جدید							
ف ۲	۷/۱۸	۱/۸۴	-۰/۵۵	۰/۶۳	۸/۶۷	۱/۳۸	-۲/۴۴
قدیم							
ف ۳	۱۰/۳۵	۵/۱۵	۰/۲۳	-۰/۱۱	۱۰/۳۱	۵/۰۷	-۰/۱۱
جدید							
ف ۳	۸/۲۹	۴/۲۰	۰/۵۰	۰/۹۹	۷/۶۱	۴/۵۱	-۰/۹۵
قدیم							

توضیحات. جدید= نمرات حاصل از فهرست جدید؛ قدیم= نمرات حاصل از فهرست ۲۰۱۷؛ کل= نمره کل اصالت؛ ف ۱= فعالیت ۱

به منظور بررسی سازگاری نمرات حاصل از دو فهرست، ضرایب همبستگی محاسبه گردید (جدول ۶). همچنین در این جدول نتایج آزمون تی برای مقایسه میانگین تفاوت نمرات گزارش شده است.

جدول ۶. نتایج ضرایب همبستگی و آزمون تی

نمره	فرم A		فرم B	
	ضریب همبستگی	آماره t	ضریب همبستگی	آماره t
اصالت کل	۰/۹۱	۶/۳۷	۰/۸۸	۲/۵۷
اصالت ف ۳	۰/۹۱	۱۰/۷۷	۰/۸۸	۱۲/۰۳
اصالت ف ۲	۰/۸۰	-۴/۲۴	۰/۳۱	-۷/۴۶
اصالت ف ۱	-۰/۳۲	۰/۰۵	۰/۱۳	۰/۰۱

مقایسه نمرات کلی اصالت که از فهرست پاسخ‌های غیرخلاق قدیمی و فهرست پاسخ‌های غیرخلاق جدید استخراج شده‌اند، تفاوت‌های معناداری را برای هر دو فرم نشان داد. برای فرم A، فهرست پاسخ‌های غیرخلاق جدید با اندازه اثری کوچک ( $d = 0/24$ )، نمرات بالاتری تولید کرد (میانگین تفاوت =  $1/31$ ،  $t(126) = 6/37$ ،  $p < 0/001$ ). به همین ترتیب، برای فرم B، نمرات حاصل از فهرست جدید مقادیر بالاتری داشتند (میانگین تفاوت =  $2/57$ ،  $t(113) = 0/81$ ،  $p < 0/001$ )؛ اگرچه اندازه اثر ناچیز بود ( $d = 0/12$ ). از سوی دیگر تحلیل همبستگی، رابطه قوی مثبتی را بین نمرات دو فهرست برای هر دو فرم نشان داد (فرم A:  $r = 0/91$ ؛ فرم B:  $r = 0/88$ ) که نتیجه می‌شود گرچه فهرست ایرانی تمایل به دادن نمرات بالاتر دارد، با این حال در نمره اصالت کل، روندهای منعکس شده توسط فهرست قدیمی را عمدتاً بازتاب می‌دهد. همچنین مقایسه این ضرایب همبستگی با حالت فرضی همبستگی کامل نشان داد که علی‌رغم وجود همبستگی بالا اما این مقادیر به‌طور معناداری با همبستگی کامل فاصله داشتند.

برای فعالیت ۲، فهرست پاسخ‌های غیرخلاق جدید نمرات پایین‌تری نسبت به فهرست استاندارد تولید کرد. در فرم A، میانگین تفاوت برابر  $0/60$  بود ( $t(126) = -4/24$ ) و  $p < 0/001$  و اندازه اثر محاسبه شده اثری کوچک و منفی را نشان می‌دهد ( $d = -0/24$ ). در این فعالیت نیز همبستگی قوی ( $r = 0/80$ ) بین نمرات مشاهده شد که نشان می‌دهد گرچه نمرات متفاوت بودند، الگوهای نسبی نمرات بین شرکت‌کنندگان حفظ شده است. با این حال همبستگی حاصل به‌طور معناداری از همبستگی مطلوب و کامل کم‌تر است. در فرم B، فهرست جدید با اندازه اثری بزرگ و منفی ( $d = -0/82$ ) به‌طور قابل توجهی نمرات پایین‌تری تولید کرد (میانگین تفاوت =  $1/74$ ،  $-7/46 = t(113)$ )،  $p < 0/001$ . از سوی دیگر همبستگی ضعیف‌تر ( $r = 0/31$ ) در این فرم حاکی از هم‌ترازی کمتر در روند نمره‌دهی برای این فعالیت است.

در فعالیت ۳، فهرست جدید، اصالت را بالاتر از فهرست استاندارد ارزیابی کرد. برای فرم A، میانگین تفاوت  $2/06$  بود ( $t(126) = 10/77$ ) و  $p < 0/001$ ، با اندازه اثر متوسط ( $d = 0/55$ ). علاوه بر این، همبستگی بین نمرات قوی گزارش شد ( $r = 0/91$ ) که نشان‌دهنده هم‌ترازی در روند نمره‌دهی شرکت‌کنندگان است. برای فرم B، میانگین تفاوت بزرگ‌تر بود (میانگین تفاوت =  $2/70$ ،  $12/03 = t(113)$ ) و  $p < 0/001$ . با اندازه اثری متوسط ( $d = 0/55$ ). همبستگی نمرات نیز قوی باقی ماند ( $r = 0/88$ ). مقایسه هر دو ضریب همبستگی نشان داد که علی‌رغم وجود همبستگی بالا اما این مقادیر به‌طور معناداری با همبستگی کامل فاصله دارند.

نمرات اصالت فعالیت ۱ که به‌صورت زوجی ارزیابی شده است، با استفاده از آزمون مک‌نمار و ضریب همبستگی فی تحلیل شد. برای فرم A، تفاوت معناداری بین فهرست جدید و قدیم یافت شد ( $p = 0/05$ )، با همبستگی منفی ( $r = -0/32$ ) که نشان‌دهنده ناسازگاری در نحوه طبقه‌بندی اصالت است. برای فرم B نیز تفاوت معنادار بود ( $p = 0/01$ )، با این حال همبستگی ضعیف‌تر و مثبت بود ( $r = 0/13$ ) که نشان‌دهنده هم‌ترازی جزئی در طبقه‌بندی است. این یافته‌ها نشان‌دهنده اختلافات قابل توجه در نمره‌دهی اصالت برای فعالیت ۱، به‌ویژه در فرم A هستند. این ناسازگاری‌ها ممکن است ناشی از تفاوت‌های تفسیری بین فهرست جدید و قدیم، به‌ویژه در یک زمینه دوتایی باشد.

### بحث و نتیجه‌گیری

فهرست‌های پاسخ‌های غیرخلاق یکی از اجزای اصلی در امتیازدهی آزمون تفکر خلاق تورنس هستند. تورنس با ارائه فهرست‌های پاسخ‌های غیرخلاق، فرایند امتیازدهی اصالت را از طریق افزایش سرعت، ساده‌تر و کارآمدتر کرده است (آکار و همکاران، ۲۰۲۱). با این حال، با توجه به اینکه خلاقیت ریشه در هنجارها و ارزش‌های فرهنگی دارد (استرنبرگ، ۲۰۱۹)، مهم است که

پیش از استفاده از فهرست پاسخ‌های غیرخلاق، تفاوت‌های فرهنگی و جمعیتی نیز در مورد این فهرست‌ها مورد بررسی قرار گیرند تا اطمینان حاصل شود که آزمون تفکر خلاق تورنس به‌طور عادلانه در میان گروه‌های نژادی، قومی و فرهنگی امتیازدهی می‌شود (برلو و همکاران، ۲۰۱۹). در پاسخ به سوال این پژوهش؛ در مرحله نخست، به تدوین فهرست پاسخ‌های غیرخلاق ایرانی پرداخته شد؛ سپس در مرحله بعد مقایسه نمرات اصالت به‌دست‌آمده از فهرست‌های ایرانی و استاندارد با استفاده از آزمون‌های آماری از جمله تی زوجی، آزمون مک‌نمار و بررسی همبستگی‌ها، در دستور کار قرار گرفت.

برای تدوین فهرست بومی، این پژوهش از تحلیل مضمون برای ایجاد «طبقات پاسخ» بهره برد؛ رویکردی که با روش مستقیم‌تر شمارش فراوانی عناوین پاسخ‌ها در مطالعات مشابه (برای مثال، آکار و همکاران، ۲۰۲۱) متفاوت است. انتخاب تحلیل مضمون، علی‌رغم پیچیدگی اولیه، با نگاهی به آینده صورت گرفت. روش شمارش مستقیم، به‌روزرسانی‌های آتی فهرست را با چالش مواجه می‌کند، زیرا عناوین پاسخ‌ها عمیقاً به زمینه حرفه‌ای نمونه وابسته‌اند. برای مثال، پاسخ‌هایی مانند «پره توربین» یا «راکتور» در نمونه صنعتی ما رایج بودند، اما در نمونه‌های دیگر (مثلاً معلمان) بی‌معنا خواهند بود. در مقابل، طبقه‌بندی مفهومی این امکان را می‌دهد که پاسخ «مخزن صنعتی» از یک کارمند و «تخته کلاس» از یک معلم، هر دو در یک طبقه کلی تر قرار گیرند و مقایسه‌پذیر شوند. این زیرساخت مفهومی، پیش‌نیاز اصلی برای هدفی فراتر، یعنی ایجاد یک سیستم نمره‌گذاری خودکار و پویا بود. (برای مطالعه بیشتر مراجعه شود به دانشمند و همکاران، ۱۴۰۳).

رویکرد مبتنی بر تحلیل مضمون، علاوه بر ایجاد یک سیستم پویا، راهکاری مؤثر برای غلبه بر محدودیت حجم نمونه در این پژوهش نیز بود. می‌دانیم که بزرگی حجم نمونه، استحکام فهرست‌های تدوین‌شده را به همراه دارد (فورثمن و همکاران، ۲۰۲۰؛ رایتر-پالمون و همکاران، ۲۰۱۹) و نمونه این پژوهش از حد ایده‌آل ۳۰۰ تا ۴۰۰ نفر برای این منظور کمتر بود (فورثمن و همکاران، ۲۰۲۰). با این حال، تحلیل مضمون این محدودیت را از دو طریق جبران می‌کند. نخست، با گروه‌بندی پاسخ‌های مفهوماً مشابه (مانند «تخت» و «بالشت» در طبقه «کالای خواب»)، فراوانی هر طبقه را افزایش داده و اعتبار آماری آن را تقویت می‌کند. دوم، این ساختار طبقه‌بندی‌شده، یک زیربنای پایدار برای به‌روزرسانی و تقویت فهرست در طول زمان با جمع‌آوری داده‌های جدید فراهم می‌آورد. البته باید توجه داشت که اعتبار فهرست‌ها در بخش‌های مختلف آزمون یکسان نیست؛ فهرست فعالیت ۱ به دلیل تجمیع داده‌های دو فرم و محرک‌های اولیه فعالیت ۲ به دلیل ثبت پاسخ‌های بیشتر به دلیل اتمام وقت برای نیمه دوم محرک‌ها، از استحکام بالاتری برخوردارند.



تحلیل پاسخ‌ها، شواهد روشنی از تأثیر فرهنگ بر خلافت ارائه داد و در نهایت به شناسایی ۱۱۱ طبقه پاسخ منحصر به فرد انجامید. ظهور پاسخ‌های پرتکرار و بومی مانند «سردر دانشگاه تهران» و «مبدل حرارتی» از یک سو، و نادر بودن پاسخ «راهبه» (که در فهرست استاندارد تورنس (۲۰۱۷) غیرخلاق محسوب می‌شود) از سوی دیگر، ضرورت تدوین فهرست‌های محلی را به وضوح نشان داد.

این تفاوت‌های کیفی، خود را در نتایج کمی نیز نمایان ساخت: فهرست بومی، نمرات کل اصالت و نمره فعالیت ۳ را به شکل معناداری افزایش داد، در حالی که نمرات فعالیت ۲ را کاهش داد. این ناسازگاری‌ها، که احتمالاً ریشه در تفاوت‌های قومی و فرهنگی میان نمونه ایرانی و نمونه هنجاری آزمون دارد (کالتسونیس، ۱۹۷۴؛ کافمن، ۲۰۰۹؛ پرایس-ویلیامز و رامیرز، ۱۹۷۷)، مؤید آن است که فهرست استاندارد به شکل غیر عمدی موجب کم‌نمایی و کاهش مصنوعی خلافت در جمعیت ایرانی می‌شود. این یافته هم‌راستا با پژوهش خارخورین و صمدپورمطلبی (۲۰۰۸) است که تفاوت در نمرات اصالت میان دانشجویان ایرانی و غربی را گزارش کرده و بر محدودیت‌های سازگاری بین فرهنگی آزمون‌های استاندارد تأکید می‌ورزد.

نتایج نشان داد فهرست‌های پاسخ‌های غیرخلاق ۲۰۱۷ به‌طور کامل برای جمعیت بزرگسالان ایرانی قابل استفاده نیستند و نیازمند تعدیل‌هایی برای نمره‌گذاری شاخص اصالت در میان این گروه است. یکی از روش‌های پیشنهادی برای این تعدیل، بازنگری در برخی از محرک‌های آزمون است. برای مثال فعالیت ۱ در هر دو فرم و همچنین فعالیت ۲ در فرم *B* بیشترین میزان اختلاف را نشان دادند ( $p < 0/80$ ). این یافته‌ها می‌توانند برای پژوهش‌های آتی در زمینه سوگیری آزمون تورنس نیز مفید باشند. نمره‌گذاری گروه‌های قومی و نژادی ممکن است به دلیل عدم وجود فهرست‌های مناسب فرهنگی در این فعالیت‌ها دچار سوگیری بیشتری شوند. از سوی دیگر، علی‌رغم بررسی تفاوت فراوانی پاسخ‌ها به تفکیک محرک در فعالیت ۲، اما در این مطالعه نمرات تک‌تک محرک‌های فعالیت ۲ مورد بررسی قرار نگرفت، اما با ارجاع به مطالعه آکار و همکاران (۲۰۲۱) می‌توان پیش‌بینی کرد که برخی محرک‌ها در این فعالیت نسبت به برخی دیگر منبع قوی‌تری در سوگیری نمرات هستند.

این مغایرت در نمرات، ریشه در یک اصل بنیادین دارد: خلافت از فیلتر تجربیات زیسته عبور می‌کند. همان‌طور که ادبیات پژوهشی نشان می‌دهد، دانش و تجربیات فردی، انبار ایده‌های اصیل او را شکل می‌دهند (گیلهولی و همکاران، ۲۰۰۷؛ رانکو و آکار، ۲۰۱۰؛ سیلویا و همکاران، ۲۰۱۷). از آنجایی که محیط، تجربیات فرهنگی مشترکی را برای اعضای یک جامعه فراهم می‌کند، مجموعه‌ای از پاسخ‌های «رایج» و «غیرخلاقانه» نیز در آن فرهنگ شکل می‌گیرد (آپ و

همکاران، ۲۰۰۶). این وابستگی به فرهنگ، ایده یک فهرست استاندارد و جهانی را با چالشی اساسی مواجه می‌سازد. در نتیجه، ایده آل‌ترین و معتبرترین رویکرد، تدوین فهرست‌های متعدد برای گروه‌های فرهنگی یا سنی خاص است (کیم، ۲۰۰۶؛ موشیرود و لوبارت، ۲۰۰۱). اگرچه ایجاد یک فهرست جدید برای هر نمونه پژوهشی به صورت مجزا، امری دشوار و غیرعملی است، اما حرکت به سوی بومی‌سازی هنجارها، گامی ضروری برای ارزیابی عادلانه و دقیق خلاقیت در میان فرهنگ‌های مختلف محسوب می‌شود. با توجه به یافته‌های پژوهش حاضر می‌توان پیشنهاد کرد که دستورالعمل نمره‌گذاری، انعطاف‌پذیرتر و پویاتر باشد تا بتواند به ایده‌های محبوب نوظهور پاسخ دهد. به‌عنوان مثال، فهرست‌های جدید می‌توانند به صورت دیجیتالی در دسترس نمره‌دهندگان قرار گیرند و به‌طور منظم با جریان داده‌های جدید به‌روزرسانی شوند. علاوه بر این، نمره‌دهندگان که از روندها و ایده‌های محبوب آگاه هستند، می‌توانند بر اساس پاسخ‌های مشاهده شده در گسترش این فهرست‌ها مشارکت کنند. این روش مشابه استراتژی نمره‌گذاری برای شاخص انعطاف‌پذیری در نسخه کلامی آزمون تفکر خلاق تورنس است، جایی که نمره‌دهندگان دسته‌ها یا خوشه‌های اضافی که در دستورالعمل ذکر نشده‌اند را یادداشت کرده و آن‌ها را به‌عنوان دسته‌های جدید در نظر می‌گیرند (تورنس، ۲۰۰۸؛ ۲۰۱۷).

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. تکرار این مطالعه با یک نمونه‌های می‌تواند نشان دهد که چه مقدار از اختلافات مشاهده شده ناشی از روش‌های نمونه‌گیری و چه مقدار به تفاوت‌های نژادی، قومی یا حرفه‌ای مرتبط است. علاوه بر این، مطالعات آتی که از تحلیل عملکرد افتراقی آیت<sup>۱</sup> استفاده می‌کنند نیز ضروری هستند تا مشخص شود که آیا فعالیت‌ها یا زیرشاخه‌های خاصی بین این دو گروه از نظر همسانی ثابت هستند یا خیر.

از سوی دیگر، پژوهش‌های آینده باید بررسی کنند که آیا فهرست‌های پاسخ‌های غیرخلاق آزمون تورنس به‌طور یکسان برای نمونه‌های با سنین مختلف در یک فرهنگ معین قابل استفاده هستند یا خیر. اگرچه سعید و شکبیا (۱۴۰۰)، صمدی و همکاران (۱۴۰۳) و توانا و همکاران (۱۴۰۲) به بررسی طبقات پاسخ دانش‌آموزان ایرانی (ابتدایی تا دوازدهم) پرداختند؛ اما هیچ‌کدام کارایی فهرست‌های پاسخ‌های غیرخلاق را مورد بررسی قرار ندادند و همچنین در نمره‌گذاری از شیوه رسمی (تورنس، ۲۰۰۸؛ ۲۰۱۷) بهره نرفتند؛ لذا با توجه به کاربرد اصلی آزمون‌های تفکر

<sup>1</sup>. Differential Item Functioning (DIF)

خلاق تورنس در شناسایی استعداد دانش‌آموزان و کودکان، پیشنهاد می‌شود این مطالعه در سایر گروه‌های سنی تکرار گردد.

به‌علاوه در این پژوهش اعتبار و روایی نمرات بر اساس فهرست‌های تدوین‌شده مختلف ارزیابی نشد که این موضوع می‌تواند در پژوهش‌های آینده مورد بررسی قرار گیرد. همچنین در این مطالعه تنها یک نقطه برش (۵ درصد) مورد مطالعه قرار گرفت که با در نظر گرفتن نتایج حاصل نقطه برش‌های مختلف می‌توان کارایی و دقت فهرست‌های تدوین‌شده را ارتقا داد.

### سپاسگزاری

این پژوهش برگرفته از پایان‌نامه کارشناسی ارشد دانشکده روان‌شناسی و علوم تربیتی دانشگاه شهید بهشتی است. نویسندگان این پژوهش بر خود لازم می‌دانند از تمامی کارکنان محترم آن شرکت صنعتی که در این پژوهش مشارکت نمودند کمال تشکر و قدردانی را به عمل آورند.

### منابع

#### فارسی

ازرگان هرندی، عباس (۱۴۰۰). *مقدمه‌ای بر روش‌های تحقیق کیفی و آمیخته؛ رویکردهای*

*متداول در علوم رفتاری*. ویرایش پنجم. تهران: کتاب دیدآور، ۱۶۳.

توانا، زهرا و اکبری زردخانه، سعید (۱۴۰۱). *سنجش خلاقیت نوجوانان: انطباق نمره‌گذاری نسخه*

*فارسی آزمون تفکر خلاق تورنس*. *پایان‌نامه کارشناسی ارشد روان‌شناسی تربیتی*. دانشگاه

شهید بهشتی.

دانشمند کفترودی، امیرحسین؛ اکبری زردخانه، سعید و فتح‌آبادی، جلیل (۱۴۰۳). *استانداردسازی*

*اولیه آزمون تصویری تفکر خلاق تورنس برای جوانان ایرانی*. *پایان‌نامه کارشناسی ارشد*

*روان‌شناسی تربیتی*. دانشگاه شهید بهشتی.

سرمد، زهره؛ بازرگان، عباس و حجازی، الهه (۱۳۸۶). *روش‌های تحقیق در علوم رفتاری*. تهران:

نشر آگه.

سعیدی علی، شکبیا ابوالقاسم. (۱۴۰۰). *بازنگری در نمره‌گذاری آزمون تصویری فرم A خلاقیت*

*تورنس در ایران*. *رویش روان‌شناسی*. ۱۰ (۱): ۱۲۰-۱۰۹

صمدی، نرگس و اکبری زردخانه، سعید (۱۴۰۲). *سنجش خلاقیت نوجوانان: انطباق نمره‌گذاری*

*نسخه فارسی آزمون تفکر خلاق تورنس*. *پایان‌نامه کارشناسی ارشد روان‌شناسی تربیتی*.

دانشگاه شهید بهشتی.

### انگلیسی:



- Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114, 254–280.
- Gilhooly, K. J., Fioratou, E., Anthony, S. H., & Wynn, V. (2007). Divergent thinking: strategies and executive involvement in generating novel uses for familiar objects. *British journal of psychology (London, England: 1953)*, 98(Pt 4), 611–625.
- Guilford, J. P. (1967). *The nature of human intelligence*. McGraw-Hill.
- Hebert, T. P., Cramond, B., Spiers Neumeister, K. L., Millar, G., & Silvian, A. F. (2002). E. Paul Torrance: His Life, Accomplishments, and Legacy. *National Research Center on the Gifted and Talented*.
- Hocevar, D. (1979). The Development of the Creative Behavior Inventory (CBI).
- Humble, S., Dixon, P., & Mpofo, 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.
- Ip, G. W.-M., Chen, J., & Chiu, C.-Y. (2006). The relationship of promotion focus, need for cognitive closure, and categorical accessibility in American and Hong Kong Chinese university students. *The Journal of Creative Behavior*, 40(3), 201–215.
- Kaltsounis, B. (1974). Race, socioeconomic status and creativity. *Psychological Reports*, 35(1), 164-166.
- Kanlı, E. (2020). *Assessment of Creativity: Theories and methods*. In IntechOpen eBooks. <https://doi.org/10.5772/intechopen.93971>
- Kaufman, J. C., Baer, J., Cole, J. C., & Sexton, J. D. (2008). A comparison of expert and nonexpert raters using the Consensual Assessment Technique. *Creativity Research Journal*, 20, 171-178.
- Kharkhurin, A. V., & Motalleebi, S. N. S. (2008). The impact of culture on the creative potential of American, Russian, and Iranian college students. *Creativity Research Journal*, 20(4), 404–411.
- 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, 251–260.
- Kim, K. H. (2011). The Creativity Crisis: The Decrease in Creative Thinking Scores on the Torrance Tests of Creative Thinking. *Creativity Research Journal*, 23(4), 285–295.
- Kim, K.H. (2017). The Torrance Tests of Creative Thinking - Figural or Verbal: Which One Should We Use? *Creativity. Theories – Research - Applications*, 4, 302 - 321.
- Kirton, M. (1976). Adaptors and innovators: A description and measure. *Journal of Applied Psychology*, 61(5), 622–629.
- Kline, R. B. (2016). *Principles and practice of structural equation modeling (4th ed.)*. Guilford Press.

- 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., 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.
- Lenhard, W., & Lenhard, A. (2014). *Hypothesis tests for comparing correlations*. <https://www.psychometrica.de/correlation.html>.
- McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2), 153–157.
- Millar, G. W. (1995). *E. Paul Torrance: The creativity man*. Ablex.
- Mouchiroud, C., & Lubart, T. (2001). Children's Original Thinking: an empirical examination of alternative measures derived from divergent thinking tasks. *The Journal of Genetic Psychology*, 162(4), 382–401.
- Mumford, M. D., Fichtel, M., England, S., & Newbold, T. (2024). Assessing creativity in organizations: Measures and validation. In M. A. Runco & S. Acar (Eds.), *Handbook of creativity assessment* (pp. 253–269). Edward Elgar Publishing. <https://doi.org/10.4337/9781839102158>
- Pearson, K. (1895). Note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, 58(347–352), 240–242.
- Price-Williams, D. R., & Ramirez, M., III. (1977). Divergent thinking, cultural differences, and bilingualism. *Journal of Social Psychology*, 103(1), 3–11.
- R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.r-project.org/>
- Reiter-Palmon, R., Forthmann, B., & Barbot, B. (2019). Scoring divergent thinking tests: A review and systematic framework. *Psychology of Aesthetics, Creativity, and the Arts*, 13(2), 144–152. <https://doi.org/10.1037/aca0000227>
- Runco, M. A., & Acar, S. (2010). Do tests of divergent thinking have an experiential bias? *Psychology of Aesthetics, Creativity, and the Arts*, 4(3), 144–148.
- Saeedi, A., & Shakiba, A. (2021). Revisiting the Scoring of the Figural Form A of the Torrance Test of Creative Thinking in Iran. *Ravesh Psychology*, 10(1), 109–120. [In Persian]
- Saeki, N., Fan, X., & Van Dusen, L. V. (2001). A comparative study of creative thinking of American and Japanese college students. *Journal of Creative Behavior*, 35(1), 24–38

- Samadi, N., Akbari Zardkhaneh, S., & Fathabadi, J. (2023). Assessing Creativity in Adolescents: Adapting the Scoring of the Persian Version of the Torrance Tests of Creative Thinking. Master's Thesis, Educational Psychology, Shahid Beheshti University. [In Persian]
- Sarmad, Z., Bazargan, A., & Hejazi, E. (2007). Research Methods in Behavioral Sciences. Tehran: Agah Publishing. [In Persian]
- Sawyer, R. K. (2019). *The creative classroom: Innovative teaching for 21st-century learners*. Teachers College Press.
- Shao, Y., Zhang, C., Zhou, J., Gu, T., & Yuan, Y. (2019). How does culture shape creativity? A mini-review. *Frontiers in Psychology*, 10, Article 1219. <https://doi.org/10.3389/fpsyg.2019.01219>
- Silvia, P. J., Nusbaum, E. C., & Beaty, R. E. (2017). Old or new? Evaluating the old/new scoring method for divergent thinking tasks. *The Journal of Creative Behavior*, 51(3), 216–224.
- Snyder, H. T., Hammond, J. A., Grohman, M. G., & Katz-Buonincontro, J. (2019). Creativity measurement in undergraduate students from 1984–2013: A systematic review. *Psychology of Aesthetics, Creativity, and the Arts*, 13(2), 133–143.
- Sternberg, R. J. (2019). Evaluation of creativity is always local. *The Palgrave handbook of social creativity research*, 393–405.
- Student (1908). The probable error of a mean. *Biometrika*, 6(1), 1–25.
- Tavana, Z., Akbari Zardkhaneh, S., & Shahidi, S. (2022). *Assessing Creativity in Adolescents: Adapting the Scoring of the Persian Version of the Torrance Tests of Creative Thinking*. Master's Thesis, Educational Psychology, Shahid Beheshti University. [In Persian]
- Torchiano, M. (2020). *Effsize: Efficient Effect Size Computation*. R package version 0.8.1. <https://cran.r-project.org/web/packages/effsize/index.html>
- 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*. Princeton, NJ: Personnel Press.
- Torrance, E. P. (1974). *Torrance Tests of Creative Thinking Norms: Technical manual*. Personal Press/Ginn and Company.
- Torrance, E. P. (1981). Predicting the creativity of elementary school children (1958-80) —and the teacher who Made a difference. *Gifted Child Quarterly*, 25(2), 55–62.
- Torrance, E. P. (2017). *Torrance Tests of Creative Thinking: Streamlined scoring guide for Figural Forms A and B*. Scholastic Testing Services, Inc.
- 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). New York, NY: The Guildford Press.
- World Economic Forum. (2023). *The future of jobs report 2023*. <https://www.weforum.org/reports/the-future-of-jobs-report-2023>

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.

Yule, G. U. (1912). On the methods of measuring association between two attributes. *Journal of the Royal Statistical Society*, 75(6), 579–652.

## Cultural Dependence of Creativity: Myth or Reality?

Saeed Akbari-Zardkhaneh<sup>1</sup>, AmirHossein Daneshmand Kafteroudi<sup>2</sup>, Jalil Fathabadi<sup>3</sup>

### Abstract

**Objective:** This study aimed to assess the applicability of the standard Zero Originality List (ZOL) of the Torrance Tests of Creative Thinking-Figural (TTCT-Figural) to an Iranian employee sample. Specifically, it investigated the impact of culture on originality scores.

**Method:** The research was fundamental in purpose, descriptive-survey in data collection, and exploratory mixed-method in design. The sample consisted of 241 industrial employees (14% women) selected using simple random sampling. The responses from TTCT-Figural were coded and categorized through thematic analysis.

**Results:** A total of 111 unique response categories were identified. Comparisons revealed that originality scores derived from the culturally adapted ZOL were significantly higher for total originality (Form A:  $d = 0.24$ ; Form B:  $d = 0.12$ ) and Activity 3 (Form A:  $d = 0.55$ ; Form B:  $d = 0.55$ ). Conversely, the standardized ZOL produced higher scores for Activity 2 (Form A:  $d = -0.24$ ; Form B:  $d = -0.82$ ). Activity 1 exhibited inconsistencies between the two lists, with correlations showing low alignment.

**Conclusion:** The findings suggest that standardized ZOLs may not be fully applicable for evaluating creativity in Iranian samples and could lead to artificially lower originality scores.

**Keywords:** Creativity, Cross-Cultural Assessment, Cultural Influence, Torrance Tests of Creative Thinking (TTCT)

<sup>1</sup>. Assistant Professor, Department of Applied Psychology, psychology and Educational Science Faculty, Shahid Beheshti University, Tehran, Iran. (s\_akbari@mail.sbu.ac.ir)

<sup>2</sup>. MA of Educational Psychology, psychology and Educational Science Faculty, Shahid Beheshti University, Tehran, Iran. (amirdaneshmand08@gmail.com)

<sup>3</sup>. Associate Professor, Department of Applied Psychology, psychology and Educational Science Faculty, Shahid Beheshti University, Tehran, Iran. (j\_fathabadi@sbu.ac.ir)