Title: Development and Preliminary Validation of a Divergent Thinking-Based Game for Creativity Assessment: A Behavioral-Cognitive Approach

Authors: Shaghayegh Chitsaz[1,*], Hamed Asadollahi[1], Meysam Sadeghi[2], Hadi Moradi Sabzevar[3], Reza Kazemi[4]

1. *Department of Cognitive Design, Institute for Cognitive Science Studies, Tehran, Iran.*
2. *Department of Cognitive Psychology, Institute for Cognitive Science Studies, Tehran, Iran.*
3. *University of Tehran, Tehran, Iran.*
4. *Department of Entrepreneurship Development Faculty of Entrepreneurship, University of Tehran Farshi Moghadam, Tehran, Iran.*

***Corresponding Author**: Shaghayegh Chitsaz, Department of Cognitive Design, Institute for Cognitive Science Studies, Tehran, Iran. Email: chitsaz.sh@gmail.com

To appear in: **Basic and Clinical Neuroscience**

Received date: 2025/07/25

Revised date: 2026/06/1

Accepted date: 2026/07/22

Please cite this article as:
Chitsaz, S., Asadollahi, H., Sadeghi, M., Moradi Sabzevar, H., Kazemi, R. (In Press). Development and Preliminary Validation of a Divergent Thinking-Based Game for Creativity Assessment: A Behavioral-Cognitive Approach. Basic and Clinical Neuroscience. Just Accepted publication Jul. 10, 2026. Doi: http://dx.doi.org/10.32598/bcn.2026.8191.1
DOI: http://dx.doi.org/10.32598/bcn.2026.8191.1

Abstract

Introduction: Creativity is a fundamental cognitive ability that plays a crucial role in problem-solving, innovation, and adaptive behavior. Traditional creativity assessments often lack ecological validity and fail to capture the dynamic and context-driven nature of creative thought. This study introduces *Creatopia*, a gamified simulation designed to assess divergent thinking.

Objective: The aim of this study was to evaluate the psychometric performance of *Creatopia* and examine its ability to measure key dimensions of divergent thinking (fluency, originality, flexibility, and elaboration) within an interactive gaming environment.

Methods: In a pilot study, 90 participants (4۰ males and 4۰ females) were randomly selected from the general population of Tehran. Participants first completed the Torrance Tests of Creative Thinking (TTCT) and then engaged with *Creatopia*. The assessments involved both qualitative and quantitative evaluations of the interactive tasks within the game, including real-time behavioral tracking. Data collected included the number of ideas generated, novelty and variety of responses, and the level of detail in the drawings. Statistical analyses, including factor analysis and correlation with TTCT scores, were used to evaluate concurrent validity.

Results: The results showed that *Creatopia* reliably measures fluency, originality, and flexibility, with significant correlations to TTCT scores. However, the elaboration component showed mixed results, with lower content validity ratings compared to the other dimensions. Factor analysis supported a unidimensional structure, explaining 75.89% of the total variance.

Conclusion: *Creatopia* demonstrates promising psychometric performance as a gamified tool for assessing divergent thinking and can serve as a viable alternative to traditional creativity assessments. However, further refinement is needed, particularly in the measurement of elaboration. This study highlights the potential of digital games to enhance the validity and applicability of creativity assessments in educational and research contexts.

Keywords: Divergent Thinking, Gamified Simulation, Creativity Assessment, Game-Based Learning, Digital Assessment Tools

Introduction

In today's increasingly complex and rapidly evolving world, creativity is recognized as a vital cognitive capacity that enables individuals to adapt and respond effectively to novel and unpredictable challenges. Creativity defined as the ability to generate original and valuable ideas (Runco, 2014; Sawyer & Henriksen, 2024). The broad applicability has established creativity as a fundamental skill for the 21st century and has sparked growing interest in how it can be reliably assessed and cultivated.

California State University provides a practical and accessible perspective, defining creativity as "the tendency to generate or identify ideas, alternatives, or possibilities that may be useful in solving problems, communicating with others, and entertaining oneself and others" (Dietrich, 2015). More broadly, creativity is often understood to encompass three essential components: the production of novel or original ideas, the relevance or usefulness of those ideas in addressing a need or problem, and the recognition or acceptance of those ideas within a specific social or cultural context (Puryear & Lamb, 2020). This multifaceted definition underscores the importance of creativity not only as an internal cognitive process but also as a socially situated and goal-oriented activity, making its valid and reliable assessment increasingly critical. Creativity manifests across a wide range of domains, including artistic expression, scientific innovation, everyday problem_solving, and social interaction (Cropley, 2019; Cropley, 2015; Croply, 1999). Foundational theoretical frameworks have shaped our current understanding of creative cognition by highlighting both structured processes and dynamic interactions. Wallas's four-stage model (1926) —comprising preparation, incubation, illumination, and verification—proposes a sequential pathway for creative thinking (Wallas, 1926). Guilford's Structure of Intellect model emphasizes divergent thinking as a critical component of intelligence, laying the groundwork for psychometric approaches to creativity (Guilford, 1950; J. P. Guilford, 1967). Rhodes's 4P model (1961)—Person, Process, Product, and Press—provides a comprehensive lens for analyzing the multifaceted nature of creativity (Rhodes, 1961). More recently, Kaufman and Beghetto's 4C model (2009) introduced a developmental perspective, distinguishing between Mini-C (personal insights), Little-C (everyday creativity), Pro-C (professional-level expertise), and Big-C (eminent, culture-shaping creativity) (Kaufman & Beghetto, 2009). Collectively, these models underscore the complex, iterative, and contextual nature of creativity, offering a theoretical foundation for assessing it in both structured and ecologically valid formats.

According to Runco and Acar (2024), "divergent thinking assessments constitute the largest and most widely used category of creativity assessments" (M. Runco & S. Acar, 2012; Runco & Acar, 2024). Among these, the Torrance Tests of Creative Thinking (TTCT) (Torrance, 1966, 1974), the Alternative Uses Task (Guilford et al., 1978), and the Remote Associates Test (RAT) (Mednick & Halpern, 1968) have emerged as foundational tools in creativity research. The TTCT, in particular, has demonstrated exceptional reliability and validity over the past six decades, having been employed in more than 2,000 studies across 35 languages (Millar, 2002). Its adaptability and cross-cultural utility have been confirmed through validated translations in Arabic (Said-Metwaly et al., 2021), Persian (Karimi et al., 2010), Kiswahili (Humble et al., 2018), Turkish (Bart et al., 2017), Spanish (Krumm et al., 2016), Korean (Yoon, 2017), and Japanese (Higuchi et al., 2012). These applications underscore the TTCT's robustness in assessing creative potential across diverse cultural and linguistic contexts while minimizing biases related to culture, language, and intelligence (Torrance, 1966, 1974). Additionally, the

TTCT has been shown to predict both current and future creative performance, reinforcing its relevance as a long-term indicator of individual creative potential (Erwin et al., 2022; Runco et al., 2010).

Despite their widespread use and established validity, traditional creativity assessments have faced several criticisms. A primary concern is that these assessments are typically administered in controlled, often artificial environments—such as laboratories—designed to maximize internal validity. However, such settings may inadvertently introduce confounding factors, including test anxiety (Schram & Ule, 2019) and the Hawthorne effect, where participants alter their behavior simply because they are aware of being observed (McCambridge et al., 2014). These psychological influences can distort authentic creative performance. Furthermore, many participants perceive these assessments as abstract or lacking personal relevance, which may diminish intrinsic motivation and engagement (Valladares-Rodríguez et al., 2016). Another limitation is their reliance on simple, often decontextualized tasks that focus solely on evaluating the final output. This approach overlooks the underlying cognitive processes and strategies that contribute to creative thinking (Schulte-Mecklenbeck et al., 2019). Additionally, these assessments are resource-intensive, frequently requiring trained human raters to score responses, which raises concerns about scalability, efficiency, and inter-rater reliability (Cseh & Jeffries, 2019). These challenges underscore the need for more engaging, process-oriented, and ecologically valid approaches to assessing creativity. If the number of reviewers is limited, their judgments may be influenced by individual tastes, preferences, and characteristics. While employing multiple raters and measuring inter-rater reliability can mitigate this issue to some extent, it requires additional resources. Furthermore, raters who participate in the scoring process for extended periods may experience fatigue and priming effects, which can undermine the accuracy and reliability of their assessments (Cseh & Jeffries, 2019). Another significant consideration is that each creativity test appears to measure different aspects of the creative process, and the results obtained from various tests typically do not correlate highly with one another (i.e., they exhibit low convergence). Runco et al. (2016) concluded that relying solely on a divergent thinking test cannot yield generalizable results (Runco et al., 2016). In contrast to traditional creativity assessment tools, well-designed digital games offer a unique combination of scalability and the capacity to collect precise data.

In the field of creativity assessment, relatively few digital games have been integrated into academic or serious game applications. However, game-based approaches present significant potential advantages over traditional methods. First, digital games can elicit more natural and spontaneous behaviors than those typically observed in standard experimental settings, thereby reducing the influence of test anxiety and researcher bias, which enhances the ecological validity of the assessment (Lumsden et al., 2016). Furthermore, games are inherently more motivating and engaging for participants, often surpassing the appeal of conventional paper-and-pencil tasks (Gee, 2005). This increased engagement can facilitate entry into a state of "flow"—a cognitive-emotional condition characterized by focused attention, intrinsic motivation, and a sense of control when individuals are immersed in an activity with clear goals and immediate feedback (Csikszentmihalyi, 1997). Flow states are particularly conducive to creativity, providing an optimal context for observing innovative thinking and problem-solving behaviors.

Additionally, digital games facilitate richer and more dynamic forms of interaction compared to static tests. They can simulate complex environments and multifaceted challenges

that reflect real-world contexts, providing opportunities to assess not only creative outputs but also the underlying cognitive strategies and behavioral patterns.

Furthermore the other advantages of digital games is their ability to unobtrusively collect detailed behavioral data, including real-time records of users' actions during problem-solving or idea generation tasks (Hart et al., 2017). This capability allows researchers to attain a more profound, process-oriented understanding of the cognitive mechanisms that underpin creativity, including the emergence and evolution of novel ideas over time.

These characteristics facilitate the collection of data from large and diverse populations, both online and offline, thereby enhancing the external validity and generalizability of creativity research outcomes. Importantly, open-access game platforms can lower participation barriers and contribute to more equitable opportunities for assessing creativity across socioeconomic and cultural contexts. Collectively, these features highlight the significant potential of digital games as innovative, scalable, and inclusive tools for creativity assessment (Rafner et al., 2022). Recent research underscores the necessity for creativity assessment tools that encompass a diverse range of domains and task types to improve the generalizability of findings concerning individual creative potential.

This paper aims to design and initially validate the ***Creatopia*** Creativity Assessment Game, developed to address the limitations of traditional creativity assessment methods. The primary focus is on utilizing interactive, design-based tasks to assess various dimensions of divergent and convergent thinking in real-time. This approach seeks to enhance both the validity of the assessment and the overall user experience.

Method

A. Game Design and Development

The development of the ***Creatopia*** creativity assessment game was guided by the principles of Design Thinking, a human-centered, iterative framework focused on creative problem-solving. Design Thinking has gained prominence as a powerful methodology for developing products and services that align with user needs and cognitive behavior patterns. In the context of this study, this approach was applied to ensure both the usability and ecological validity of the assessment tool.

This study was conducted in accordance with established ethical principles Committee of Institute for Cognitive Sciences studies (IR.UT.IRICSS.REC.1403.002), including the acquisition of written informed consent from all participants, assurance of confidentiality, provision of comprehensive information regarding the research procedures, and the participants' right to withdraw from the study at any time without penalty.

The design process consisted of five iterative phases: empathize, define, ideate, prototype, and test. The initial stages focused on engaging with potential users—addressing both cognitive and experiential aspects—to identify key challenges in traditional creativity assessments. This engagement informed the definition of essential user and system requirements for the game. Subsequently, a variety of ideas for interactive tasks and problem scenarios were generated to stimulate both divergent and convergent thinking. These concepts were then transformed into low- and high-fidelity prototypes, which were evaluated through iterative user testing. The final version of ***Creatopia*** incorporates real-time behavioral tracking, intuitive interaction mechanics,

and adaptive feedback loops, facilitating both standardized data collection and an engaging user experience Figure 1.

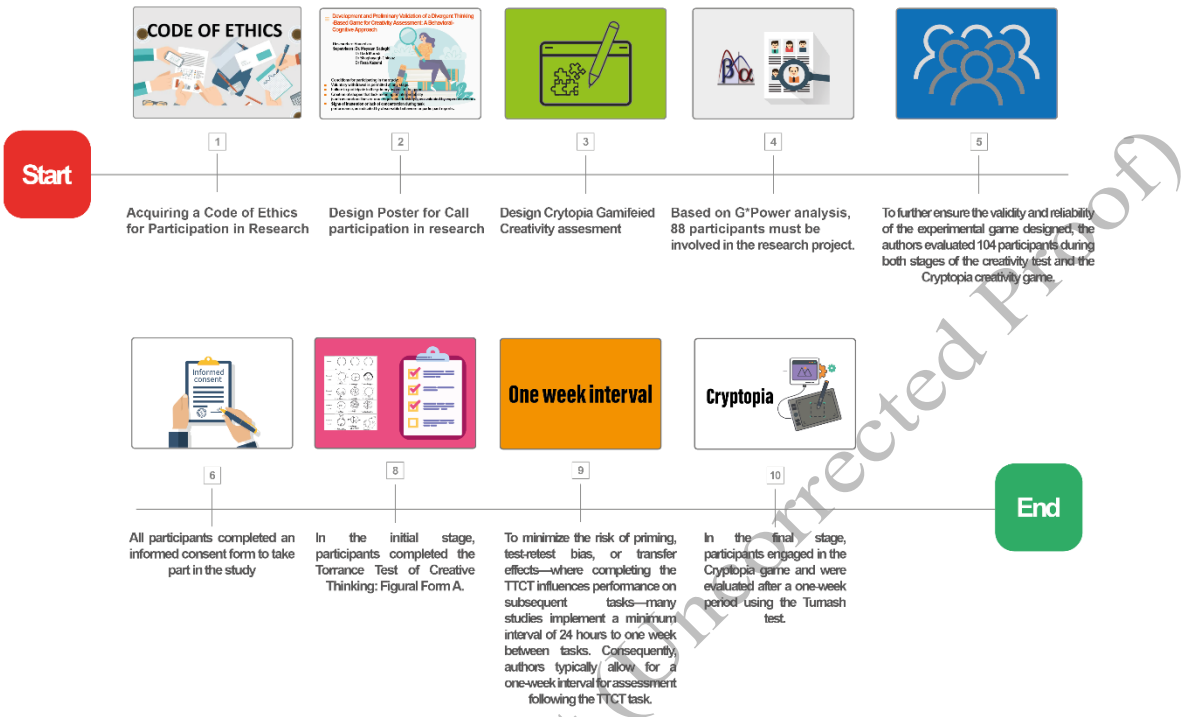

Figure 1. *Procedural Timeline of the Creatopia Creativity Assessment Study.* This visual timeline outlines the sequential steps in the development, ethical preparation, participant recruitment, and evaluation process for the *Creatopia* gamified creativity assessment. The study began with ethics approval and participant outreach, followed by the design of the *Creatopia* simulation and sample size determination using G*Power. Participants first completed the Torrance Test of Creative Thinking (TTCT), followed by a one-week interval to mitigate priming effects. Subsequently, they engaged in the *Creatopia* game, allowing for the comparative evaluation of divergent thinking components across both tools.

This systematic design approach ensured that the tool aligned with principles of cognitive science while maintaining user engagement and minimizing extraneous variables that could confound the measurement of creativity.

The development process of the ***Creatopia*** game is outlined in four main sections:

1. A general overview of the game's core concept and framework.
2. The theoretical foundations and cognitive principles that informed the design.
3. The narrative structure and storyline that enhance player engagement and immersion.
4. The interactive mechanisms and game elements employed to evaluate creativity.

Each section highlights how behavioral and cognitive design principles were integrated into the game to enhance its validity as a creativity assessment tool. Following the description of the game development, we provide detailed information about the study participants, instruments, and research methodology employed to evaluate the tool's performance and psychometric properties.

The developed game, ***Creatopia***, has been implemented as a standalone digital application for the Windows operating system. The game consists of three sequential stages,

each featuring interactive tasks aimed at eliciting and assessing creative thinking through visual expression. In each stage, players encounter narrative-driven challenges and are prompted to draw original objects that are crucial for advancing the storyline. These drawing tasks are specifically designed to engage both divergent and convergent thinking processes, which are central to the cognitive mechanisms underlying creativity.

The complete gameplay experience lasts approximately 30 minutes, offering ample engagement time without causing cognitive fatigue. Upon finishing all stages, the player's visual outputs undergo both qualitative and quantitative analysis, with scoring criteria based on established creativity metrics such as originality, elaboration, and fluency. The results of this analysis produce a composite creativity score, providing a standardized yet nuanced assessment of individual creative performance.

١. Game Design

The design of ***Creatopia*** is based on Guilford's divergent thinking model (J. P. Guilford, 1967), a foundational framework in creativity research. This model defines creativity as consisting of four core cognitive components: fluency (the ability to generate a large number of ideas), originality (the uniqueness of responses), flexibility (the capacity to produce ideas across diverse categories), and elaboration (the degree of detail and refinement in an idea). These dimensions have been directly incorporated into the game's tasks and scoring system to ensure alignment with well-established cognitive constructs of creative thinking.

To further enhance both the theoretical and practical validity of the creativity assessment, the game's design was informed by the structure and evaluation criteria of the Torrance Tests of Creative Thinking—Figural Form (TTCT) (Torrance, 1974). The TTCT is widely recognized as one of the most extensively validated instruments in creativity research and has been utilized across diverse cultural and linguistic contexts. Drawing from the TTCT's figural creativity format, ***Creatopia*** incorporates similar open-ended, figural prompts that encourage freeform drawing responses. The game's scoring rubric adapts TTCT-inspired indicators to evaluate participants' figural outputs, thereby ensuring that creativity is measured according to empirically supported and internationally accepted standards.

٢. Game Narrative

The narrative framework of ***Creatopia*** places the player in the role of an astronaut who crash-lands on a mysterious, uncharted planet inhabited by an alien civilization. As the story unfolds, the protagonist faces a series of culturally and cognitively complex challenges that necessitate creative problem-solving skills. Central to the gameplay is the requirement for players to draw imaginative, functional objects—such as communication devices, symbolic tools, or mission-critical artifacts—that enable interaction with the alien environment.

The narrative is crafted to immerse players in a dynamic and responsive storyline, where creativity is not only encouraged but essential for advancement. The game's progression is directly influenced by the player's creative contributions, making their drawings pivotal to the development of the story. This integration of narrative and task enhances engagement while embedding assessment within a meaningful context.

As the plot reaches its climax, the astronaut is captured by the planet's ruler. Recognizing the player's creative ingenuity, the ruler assigns the astronaut a covert mission: to return to Earth as a secret envoy. The game concludes with the protagonist reentering Earth's atmosphere, now

transformed into a creative agent or “spy,” symbolizing the internalization and application of creative thinking skills in real-world contexts.

۳. Game Stages

Stage 1: Creative Request Prior to Execution

In the initial stage of ***Creatopia***, the narrative places the astronaut protagonist in a perilous situation: having been captured by an alien civilization, the player learns that the character has been sentenced to execution. However, by the customs of the alien culture, the astronaut is granted a final request. At this juncture, the player is invited to draw any object of their choosing—an imagined item they wish to receive—using only their creativity and intuition. This stage is intentionally open-ended and free from contextual constraints, allowing participants to explore a wide range of potential solutions. The activity is time-limited to eight minutes, which encourages spontaneous ideation while minimizing the tendency to overthink. The task is designed to measure key aspects of divergent thinking, including originality, fluency, and elaboration, by evaluating the novelty, variety, and detail of the participant's drawing. This unstructured prompt serves as an initial diagnostic tool for assessing the participant's creative potential in a high-stakes yet imaginatively rich environment Figure 2.

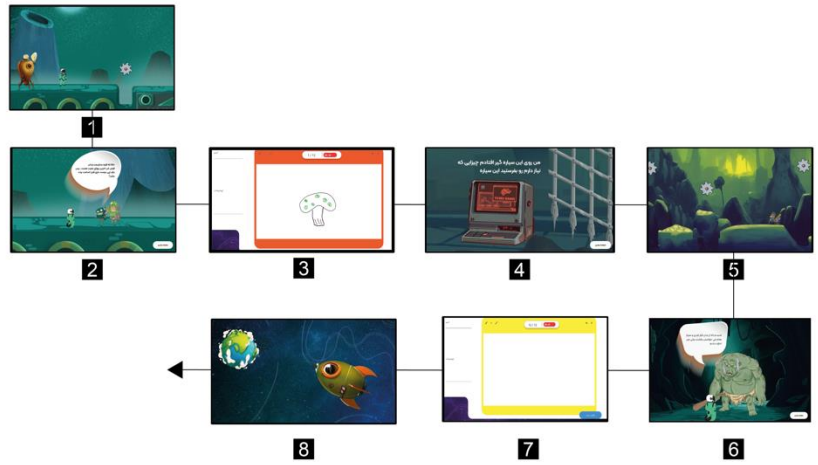

Figure 2. *Narrative Progression and Gameplay Sequence in the Creatopia Simulation.* This visual illustrates key stages in the *Creatopia* creativity assessment game. The storyline begins with the player’s arrival on an alien planet (1), followed by interaction with alien characters (2) and a drawing task requesting the creation of an object to influence their fate (3). A communication device is then discovered (4), prompting players to proceed through a problem-solving journey (5) culminating in negotiation with the alien ruler (6). Additional drawing tasks (7) assess idea generation under time constraints. The narrative concludes with the protagonist’s transformation into a covert agent and return to Earth (8). This sequence integrates storytelling, visual prompts, and interactive tasks to elicit divergent thinking behaviors within an ecologically valid digital environment.

Stage 2: Sending a Message to Earth

In the continuation of the story, the astronaut discovers a rudimentary device that allows him to send a message to Earth. He must describe the tools he needs to save himself. During this stage,

the player encounters a collection of incomplete and abstract shapes, which he must creatively transform into ten distinct images of the necessary tools. This stage lasts for 10 minutes and emphasizes flexibility and the expansion of creativity, as the player is required to generate diverse ideas and provide richer details.

Stage 3: Preparing for Espionage

After failing to escape, the astronaut is presented with an offer from the king: in exchange for saving his life, he must return to Earth as a secret agent. In this phase, the player is required to sketch the tools necessary for this covert mission and the journey back to Earth. The time allocated for this phase is six minutes. The primary objective is to assess fluency and originality, as the player must generate a variety of innovative ideas.

B. Measuring the Components of Divergent Thinking in the Game

As previously outlined, the theoretical framework supporting ***Creatopia*** is grounded in Guilford's model of divergent thinking (J. P. Guilford, 1967), which identifies four essential components: fluency, originality, flexibility, and elaboration. These constructs were operationalized through the game's drawing-based tasks and evaluated through a structured analysis of participants' figural responses.

Fluency was assessed by the number of unique ideas or visual elements generated by the participant throughout the game stages.

Originality was assessed by comparing the uniqueness of each drawing to those of other participants, identifying statistically rare or novel features.

Flexibility was evaluated based on the variety of conceptual categories represented in the drawings, which reflects the participant's ability to transition between different types of ideas.

Elaboration was assessed based on the level of detail, refinement, and complexity evident in each drawing.

Each of these components was evaluated using a standardized scoring rubric adapted from established creativity assessment protocols (e.g., Torrance, 1974), ensuring consistency and comparability across participants. This multimodal approach facilitated a comprehensive and ecologically valid assessment of divergent thinking within an interactive, game-based environment.

Table 1. Operational Definitions and Measurement Methods of Divergent Thinking Components in the Game Environment

DIVERGENT THINKING COMPONENT	OPERATIONAL DEFINITION	MEASUREMENT METHOD IN THE GAME
FLUENCY	The ability to produce a large number of relevant ideas or solutions	Number of distinct designs generated during the time window of stages 2 and 3
ORIGINALITY	The uniqueness or novelty of ideas or responses	Evaluation of each design based on its novelty and inventiveness compared to benchmark samples
FLEXIBILITY	The ability to shift between categories or types of ideas	Number of distinct conceptual categories represented across a participant's set of drawings
ELABORATION	The level of detail, decoration, or development of an idea	Degree and richness of visual detail or functional explanations added to each design

C. Participants

Given the exploratory and preliminary nature of this study, the research was conducted as a pilot investigation to assess the feasibility and initial validity of the game-based creativity assessment tool. The required sample size was determined using G*Power statistical software. Based on a linear correlation model with a Type I error rate of 5% ($\alpha = 0.05$), a statistical power of 80% ($1 - \beta = 0.80$), and a medium effect size ($r = 0.25$), the minimum recommended number of participants was calculated to be 88 individuals.

A total of 90 participants were randomly selected from the general population of Tehran. The sample consisted of 4۰ males (5۰%) and 4۰ females (۰۰%), with a mean age of 34.11 years (*SD* = 7.33). Regarding the highest educational degree, the majority held a master's degree (56.7%, n = 51), followed by a bachelor's degree (30%, n = 27), and a doctoral degree (13.3%, n = 12). In terms of academic discipline, over half of the participants (58%) were from humanities fields, 35% from engineering, 6% from arts, and 1% from medicine. Participants were recruited through a convenience sampling method using direct verbal requests to accessible individuals within the community. Eligibility criteria required individuals to be between 20 and 50 years of age and to have completed or be currently pursuing higher education (i.e., a bachelor's degree or above). Out of an initial pool of 130 individuals who were approached, 90 met the eligibility criteria and consented to participate. Those who expressed initial interest were provided with a digital informed consent form, which outlined the study's purpose, procedures, and their rights as participants. Upon providing consent, they were formally enrolled in the study on a voluntary basis. As an incentive, each participant received a complimentary personalized report detailing their creativity test results upon completion of the study. data collection was conducted over a six-month period, from October 2024 to March 2025. To enhance the generalizability of the findings within the context of a pilot study, the sample was intentionally diverse in terms of educational background, occupational status, and professional experience. This heterogeneity facilitated a more representative evaluation of the game's applicability across various demographic groups.

Inclusion Criteria:

- Age range: 20 to 50 years.

- Fluent in Persian.
- No severe mental disorders were present at the time of the study.
- No previous experience with formal creativity assessment tests.

These conditions were established to ensure cognitive maturity, a shared linguistic and cultural understanding, and to prevent performance bias resulting from prior familiarity with creativity assessment tools.

Exclusion Criteria:

- Voluntary withdrawal is permitted at any stage.
- Failure to participate in the primary aspect of the game.
- Creation of shapes that lack meaning or interpretability (such as random lines or unrecognizable drawings), as evaluated by expert assessors.
- Signs of inattention or lack of concentration during task performance, as indicated by observable behaviors or participant reports.

Results

To evaluate the psychometric properties of the Creatopia assessment game and investigate gender differences in creativity performance, several statistical tests were conducted. The results indicated that three components (fluency, originality, and flexibility) demonstrated acceptable content validity, unidimensional factor structure, and significant correlations with the benchmark Torrance Test of Creative Thinking (TTCT). In contrast, the elaboration component received lower expert ratings for content validity, yet interestingly showed the strongest correlation with the TTCT.

An independent samples t-test revealed that male participants scored significantly higher than females in both the elaboration component and the overall Creatopia score (Gtotal). Furthermore, male scores showed greater variance, particularly in elaboration, with more males occupying the upper quartile, consistent with the Greater Male Variability Hypothesis (GMVH) in figural creativity. This hypothesis suggests that while average scores may not differ significantly between genders, males are more likely to be represented at both the high and low ends of the distribution.

Content Validity

To evaluate the content validity of the preliminary version of Creatopia, ratings were obtained from 10 experts (6 people in creativity and 4 people in game design). Using a 5-point Likert scale assessing clarity and relevance, the Content Validity Index (CVI) was calculated for each component. A minimum CVI of 0.62 (Lawshe, 1975) was used as the cutoff for acceptable validity.

Table 2. Content Validity Indices (CVI)

Component	*Clarity*	*Relevance*
Fluency	0.70	0.80
Originality	0.70	0.70
Flexibility	0.70	0.80
Elaboration	0.50	0.50

Only elaboration fell below the minimum threshold in both criteria.

Variance Analysis

Levene's tests were conducted to examine homogeneity of variances.

Table 3. Levene's Tests and Variance Ratios

COMPONENT	LEVENE F	P-VALUE	MALE/FEMALE VARIANCE RATIO
*GFLUENCY	2.645	0.107	0.643
GORIGINALITY	2.567	0.113	0.748
GFLEXIBILITY	0.351	0.555	0.722
GELABORATION	1.674	0.199	1.582
GTOTAL	0.035	0.852	1.072

* The components that start with the letter G are related to the **Creatopia**.

Elaboration and Gtotal showed greater male variance, aligning with GMVH.

Gender Differences and Effect Sizes

Table 4. Independent t-Tests and Effect Sizes

VARIABLE	T	P	EFFECT SIZE (D)	NUMBER OF MEN	NUMBER OF WOMEN
***TFLUENCY**	0.699	0.487	0.147	45	45
TORIGINALITY	0.229	0.819	0.048	45	45
TFLEXIBILITY	1.371	0.174	0.289	45	45
TELABORATION	1.525	0.131	0.322	45	45
TTOTAL	1.411	0.162	0.297	45	45
GFLUENCY	0.116	0.908	0.024	45	45
GORIGINALITY	1.159	0.249	0.244	45	45
GFLEXIBILITY	0.943	0.348	0.199	45	45
GELABORATION	2.248	0.027	0.474	45	45
GTOTAL	2.05	0.043	0.432	45	45

* The components that start with the letter T are related to the **TTCT**.

Significant gender differences were observed in elaboration and total scores, favoring males.

Factor Analysis

Exploratory factor analysis was conducted on the Creatopia subscales.

Table 5. Factor Loadings and Variance Explained

INDEX	CREATOPIA (G)
KMO	0.79
BARTLETT'S X^2 ($P < .001$)	298.76
TOTAL VARIANCE EXPLAINED	75.89%
FACTOR 1 – ELABORATION	30.45%
FACTOR 2 – FLUENCY	20.78%
FACTOR 3 – ORIGINALITY	14.92%
FACTOR 4 – FLEXIBILITY	9.74%

According to the results of Tables 5, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy (values > 0.7 are acceptable) was 0.79, and the Bartlett's test of sphericity ($p < 0.05$), which assesses the significance of correlations among variables for factor analysis, also indicated acceptable values. In this analysis, four factors were identified that collectively explained 75.89% of the total variance, accounting for a substantial portion of the changes; therefore, the model provides a good fit for the data.

Reliability

Table 6. Cronbach's Alpha (Internal Consistency)

COMPONENT	A (CREATOPIA SCORES)
FLUENCY	0.85
ORIGINALITY	0.82
FLEXIBILITY	0.81
ELABORATION	0.87
TOTAL SCORE	0.89

To assess the reliability of the Creatopia game, Cronbach's alpha values above 0.8 were recorded for all components, indicating desirable reliability (Table 6).

Concurrent Validity

Table 7. Correlations Between Creatopia and TTCT Scores

PAIR	R	P-VALUE
TFLUENCY – GFLUENCY	0.541	< 0.001
TORIGINALITY – GORIGINALITY	0.255	0.015
TFLEXIBILITY – GFLEXIBILITY	0.500	< 0.001
TELABORATION – GELABORATION	0.637	< 0.001
TTOTAL – GTOTAL	0.569	< 0.001

Additionally, the results of the correlation analysis between the Creatopia game and the Torrance Test show that the highest correlation is related to the **Elaboration** component, which, despite its weakness in content validity, demonstrated a strong correlation with the Torrance Test. This outcome may be due to an overlap in style or the sensitivity of this component to individual differences.

Comparative Features of Creativity Assessment Games

Table 8 summarizes key characteristics of 13 digital games and platforms designed or adapted for creativity assessment. The comparison includes gamification type, user involvement, assessed creativity dimensions, embedded game features, and evaluation methodologies.

Table 8. This table summarizes key characteristics of 13 digital games and platforms designed or adapted for creativity assessment.

GAME TITLE	GAMIFICATION TYPE	USER INVOLVEMENT	CREATIVITY DIMENSIONS	GAMIFICATION FEATURES	EVALUATION METHODS
KODETU (VARIANT 5)	Puzzle-based Programming	Active Coding Tasks	Fluency, Originality	Level-based challenges, timers	Expert coding analysis
ARCTIC QUEST 2 (VARIANT)	Puzzle/Game	Visual puzzle solving	Fluency, Flexibility, Elaboration	Shape-matching, creative reconstruction	Fuzzy logic scoring system
MINECRAFT (VARIANT)	Sandbox/Creative	Free-form building	Originality, Flexibility	Unlimited resources, open environment	Behavioral observation, self-report
CREA.BLENDER	Co-Creative AI Game	Image generation with AI	Fluency, Originality, Flexibility	Collaborative AI, visual output	Expert ratings, comparison to norm
BUZZWORDS	Word Association Game	Prompt-response input	Fluency, Originality, Humor	Timed prompts, scoring rare responses	Algorithmic scoring
PHYSICS PLAYGROUND	Physics Puzzle Game	Drawing-based interaction	Originality, Flexibility	Stealth assessment, in-game feedback	In-game data mining, expert scoring
DROODLE	Visual Humor Generation	Caption generation	Originality, Humor/Surprise	Funny image + caption combo	Human judgment (crowdsourced)
CREATIVE FORAGING GAME	Exploration Game	Search in shape space	Fluency, Originality, Insight	Freedom of navigation	Behavioral trajectory analysis
OPEN CREATIVITY APP	App-based Tasks	Prompt-based input	Fluency, Flexibility	Creative prompts, minimal UI	Quantitative coding of responses
CREATIVITY GAME-BASED EVALUATION SYSTEM	Gamified Evaluation Tool	Task-based response	Originality, Flexibility	Story-based missions	Expert rating, scoring rubrics
IMMUNE DEFENSE	Tower Defense Strategy	Tool placement strategy	Fluency, Flexibility, Elaboration	In-game decisions	Behavioral tracking and rubric
GEDTA	Gamified application for creativity training	Co-designed with users via Participatory Design (PD)	Focus on divergent thinking in open-ended challenges	XP points, challenges, guilds, rewards, badges	Iterative user testing, preference surveys, expert reviews
CREATOPIA	Serious game for creativity assessment	Content and validation created by experts + participants	Fluency, flexibility, originality, (elaboration weak)	Drawing tasks, challenges, scoring, feedback	Expert CVI, factor analysis, correlation with TTCT

Discussion

This study aimed to explore the potential of *Creatopia*, a gamified simulation tool, as a reliable and engaging assessment for divergent thinking. The results demonstrate that *Creatopia* effectively measures core components of creativity, such as fluency, originality, and flexibility, through interactive and narrative-driven tasks. These findings align with the results of traditional creativity assessments, such as the Torrance Tests of Creative Thinking (TTCT), while also overcoming many of the ecological validity and engagement limitations inherent in conventional tests. One of the key findings of this study was the strong psychometric performance of *Creatopia* for fluency, originality, and flexibility. These dimensions have long been recognized as central to the process of divergent thinking (Joy P Guilford, 1967), and the fact that they are reliably measured through a digital game-based platform is a promising step toward enhancing the validity and applicability of creativity assessment tools. Additionally, the significant correlation between *Creatopia* and the TTCT for these components strengthens the argument that digital games can provide a valid alternative to traditional creativity assessments. However, the elaboration component showed mixed results. While it demonstrated strong correlations with the TTCT, it was also rated lower in terms of content validity, indicating a potential area for further refinement. Elaboration is often considered a more complex aspect of creativity, requiring greater cognitive processing and attention to detail. The discrepancy in its measurement between *Creatopia* and traditional tests may suggest that additional design elements, such as incorporating more specific tasks that require detailed development and refinement of ideas, could enhance this component's validity in the game context.
A notable aspect of this study was the observed gender differences, particularly in the elaboration component. Male participants scored higher in elaboration, as well as in the overall *Creatopia* score. This finding is consistent with the Greater Male Variability Hypothesis (GMVH), which suggests that males tend to exhibit more variability in creative performance (Taylor et al., 2024). These gender differences, however, were not observed in the TTCT scores, indicating that *Creatopia* may be more sensitive to individual differences in creativity expression, particularly in relation to cognitive flexibility and elaboration. Future studies could further explore the role of gender in creative performance and whether game-based assessments provide a more nuanced understanding of this variability. Furthermore, the use of a gamified simulation to assess creativity offers several advantages over traditional assessment methods. The immersive and interactive nature of *Creatopia* allowed participants to engage in creative problem-solving tasks in a context that closely resembles real-world scenarios. This approach not only increased participant engagement but also captured behavioral data that would be difficult to obtain through conventional paper-and-pencil tasks. As digital games continue to evolve, they offer an opportunity to collect more detailed, real-time data about the cognitive processes underlying creativity, providing a deeper understanding of how individuals approach open-ended problems and generate novel ideas. In terms of practical implications, *Creatopia* offers a scalable and accessible platform for assessing creativity in educational, research, and professional settings. Its ability to capture real-time behavioral data allows for the monitoring of individual progress, which could be beneficial in educational environments where creativity is a key focus. Moreover, *Creatopia* offers potential for cross-cultural application, as its digital format and narrative-driven design can be easily adapted to different languages and cultural contexts, broadening the scope for creativity assessment worldwide.

Conclusion

The current study offers empirical validation for ***Creatopia,*** a gamified, simulation-based platform designed to evaluate divergent thinking through interactive, narrative-driven visual tasks. The results not only confirm but also expand upon findings from previous research in several critical areas of creativity assessment and cognitive science.

First, the strong performance in the components of fluency, originality, and flexibility supports previous studies indicating that these are core dimensions of divergent thinking that can be reliably measured across both traditional and digital formats) M. A. Runco & S. Acar, 2012; Silvia et al., 2008 (. The positive correlations between ***Creatopia*** scores and the Torrance Tests of Creative Thinking (TTCT) further align with earlier findings regarding the convergent validity of gamified creativity tools (Shute & Ventura, 2013), confirming that creativity can be accurately assessed in ecologically valid digital environments.

Second, this study contributes to the expanding body of research that underscores the significance of elaboration in creative cognition. Despite receiving lower content validity scores, elaboration exhibited the strongest correlation with TTCT scores. This finding aligns with the conclusions drawn by Baer & Kaufman (2005) who posited that elaboration may indicate deeper cognitive processing and a focus on detail that traditional scoring methods frequently overlook (Baer & Kaufman, 2005). Our results suggest that improving the design and measurement of elaboration in digital formats could facilitate more nuanced assessments of creative style.

Third, the observed gender-based variability—particularly the wider performance range in males and the higher male scores in elaboration—aligns with the Greater Male Variability Hypothesis (GMVH), as previously documented in creativity research (Feingold, 1994; Sitorus, 2016). However, the lack of gender differences in overall TTCT scores reinforces the broader consensus that creativity potential is not inherently biased by gender but may vary in expression depending on task format and cognitive dimension.

Finally, the methodological integration of behavioral simulation, narrative context, and expert-validated scoring positions ***Creatopia*** within an emerging movement toward context-rich, gamified assessments in cognitive neuroscience (Hamari et al., 2014; Lumsden et al., 2016);This approach enhances ecological validity, boosts participant engagement, and facilitates real-time data capture—features that are increasingly essential for studying complex cognitive functions such as creativity.

(Hamari et al., 2014***)*** proves to be a valid interactive, simulation-based tool for assessing divergent thinking. Its findings not only support previous research on creativity and gender variability but also emphasize the importance of integrating behavioral science, game design, and psychometric rigor. Future developments should focus on exploring adaptive scoring systems, cross-cultural replication, and longitudinal monitoring to further enhance the tool's application in both research and practice.

In conclusion, This study highlights the effectiveness of *Creatopia* as a gamified tool for assessing divergent thinking. The results indicate that *Creatopia* successfully measures key dimensions of creativity such as fluency, originality, and flexibility, offering promising psychometric performance. While the elaboration component requires further refinement, the overall findings suggest that *Creatopia* provides a reliable, engaging alternative to traditional creativity assessments with potential applications in both educational and research settings.

References

Baer, J., & Kaufman, J. C. (2005). Bridging generality and specificity: The amusement park theoretical (APT) model of creativity. *Roeper Review*, *27*(3), 158-163.

Bart, W. M., Hokanson, B., & Can, I. (2017). An investigation of the factor structure of the torrance tests of creative thinking. *Educational Sciences: Theory and Practice*, *17*(2), 515-528.

Cropley, A. (2019). Definitions of Creativity. In. https://doi.org/10.1016/B978-0-12-809324-5.23524-4

Cropley, D. H. (2015). Promoting creativity and innovation in engineering education. *Psychology of Aesthetics, Creativity, and the Arts*, *9*(2), 161-171. https://doi.org/10.1037/aca0000008

Croply, A. (1999). Definition of creativity. *Encyclopedia of Creativity*, *1*, 511-524.

Cseh, G. M., & Jeffries, K. K. (2019). A scattered CAT: A critical evaluation of the consensual assessment technique for creativity research. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(2), 159.

Csikszentmihalyi, M. (1997). Flow and the psychology of discovery and invention. *HarperPerennial, New York*, *39*, 1-16.

Dietrich, A. (2015). How creativity happens in the brain.

Erwin, A. K., Tran, K., & Koutstaal, W. (2022). Evaluating the predictive validity of four divergent thinking tasks for the originality of design product ideation. *PLoS One*, *17*(3), e0265116.

Feingold, A. (1994). Gender differences in personality: a meta-analysis. *Psychological bulletin*, *116*(3), 429.

Gee, J. P. (2005). Learning by design: Good video games as learning machines. *E-learning and Digital Media*, *2*(1), 5-16.

Guilford, J. P. (1950). Creativity. *American Psychologist*, *5*(9), 444-454. https://doi.org/10.1037/h0063487

Guilford, J. P. (1967). Creativity: Yesterday, today and tomorrow. *The Journal of Creative Behavior*, *1*(1), 3-14.

Guilford, J. P. (1967). Creativity: Yesterday, today, and tomorrow. *The Journal of Creative Behavior*, *1*(1), 3-14. https://doi.org/10.1002/j.2162-6057.1967.tb00002.x

Guilford, J. P., Christensen, P. R., Merrifield, P. R., & Wilson, R. C. (1978). Alternate uses.

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work?--a literature review of empirical studies on gamification. 2014 47th Hawaii international conference on system sciences,

Hart, Y., Mayo, A. E., Mayo, R., Rozenkrantz, L., Tendler, A., Alon, U., & Noy, L. (2017). Creative foraging: An experimental paradigm for studying exploration and discovery. *PLoS One*, *12*(8), e0182133.

Higuchi, T., Miyata, K., & Yuizono, T. (2012). Creativity improvement by idea-marathon training, measured by torrance tests of creative thinking (TTCT) and its applications to laboratories. 2012 Seventh International Conference on Knowledge, Information and Creativity Support Systems,

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.

Karimi, L., Ramezani, V., Ahmadi, M., Heshmati, R., & Jafar, E. (2010). Psychometric properties of Torrance test (Persian version) of creative thinking (A form). *Procedia-Social and Behavioral Sciences*, *5*, 1429-1433.

Kaufman, J., & Beghetto, R. (2009). Beyond Big and Little: The Four C Model of Creativity. *Review of General Psychology - REV GEN PSYCHOL*, *13*. https://doi.org/10.1037/a0013688

Krumm, G., Aranguren, M., Arán Filippetti, V., & Lemos, V. (2016). Factor structure of the torrance tests of creative thinking verbal form B in a Spanish-speaking population. *The Journal of Creative Behavior*, *50*(2), 150-164.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel psychology*, *28*(4).

Lumsden, J., Edwards, E. A., Lawrence, N. S., Coyle, D., & Munafò, M. R. (2016). Gamification of cognitive assessment and cognitive training: a systematic review of applications and efficacy. *JMIR serious games*, *4*(2), e5888.

McCambridge, J., Witton, J., & Elbourne, D. R. (2014). Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. *J Clin Epidemiol*, *67*(3), 267-277. https://doi.org/10.1016/j.jclinepi.2013.08.015

Mednick, M. T., & Halpern, S. (1968). Remote associates test. *Psychological Review*.

Millar, G. W. (2002). *The torrance kids at mid-life: Selected case studies of creative behavior*. Ablex Publishing.

Puryear, J. S., & Lamb, K. N. (2020). Defining creativity: How far have we come since Plucker, Beghetto, and Dow? *Creativity Research Journal*, *32*(3), 206-214.

Rafner, J., Biskjær, M. M., Zana, B., Langsford, S., Bergenholtz, C., Rahimi, S., Carugati, A., Noy, L., & Sherson, J. (2022). Digital games for creativity assessment: Strengths, weaknesses and opportunities. *Creativity Research Journal*, *34*(1), 28-54.

Rhodes, M. (1961). An Analysis of Creativity. *The Phi Delta Kappan*, *42*(7), 305-310. http://www.jstor.org/stable/20342603

Runco, M. (2014). *Creativity: Theories and Themes: Research, Development, and Practice*. https://doi.org/10.1016/C2012-0-06920-7

Runco, M., & Acar, S. (2012). Divergent Thinking as an Indicator of Creative Potential. *Creativity Research Journal - CREATIVITY RES J*, *24*, 66-75. https://doi.org/10.1080/10400419.2012.652929

Runco, M. A., Abdulla, A. M., Paek, S. H., Al-Jasim, F. A., & Alsuwaidi, H. N. (2016). Which test of divergent thinking is best. *Creativity. Theories–Research-Applications*, *3*(1), 4-18.

Runco, M. A., & Acar, S. (2012). Divergent thinking as an indicator of creative potential. *Creativity research journal*, *24*(1), 66-75.

Runco, M. A., & Acar, S. (2024). *Handbook of creativity assessment*. Edward Elgar Publishing.

Runco, M. A., Millar, G., Acar, S., & Cramond, B. (2010). Torrance tests of creative thinking as predictors of personal and public achievement: A fifty-year follow-up. *Creativity Research Journal*, *22*(4), 361-368.

Said-Metwaly, S., Van den Noortgate, W., & Barbot, B. (2021). Torrance test of creative thinking-verbal, Arabic version: Measurement invariance and latent mean differences across gender, year of study, and academic major. *Thinking Skills and Creativity*, *39*, 100768.

Sawyer, R. K., & Henriksen, D. (2024). *Explaining creativity: The science of human innovation*. Oxford university press.

Schram, A., & Ule, A. (2019). Introduction to the Handbook of Research Methods and Applications in Experimental Economics. In *Handbook of Research Methods and Applications in Experimental Economics* (pp. 1-7). Edward Elgar Publishing.

Schulte-Mecklenbeck, M., Kühberger, A., & Johnson, J. G. (2019). *A handbook of process tracing methods*. Routledge New York, NY.

Shute, V., & Ventura, M. (2013). *Stealth assessment: Measuring and supporting learning in video games*. The mit press.

Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., Martinez, J. L., & Richard, C. A. (2008). Assessing creativity with divergent thinking tasks: exploring the reliability and validity of new subjective scoring methods. *Psychology of Aesthetics, Creativity, and the Arts*, *2*(2), 68.

Sitorus, J. (2016). Students' creative thinking process stages: Implementation of realistic mathematics education. *Thinking Skills and Creativity*, *22*, 111-120.

Taylor, C. L., Said-Metwaly, S., Camarda, A., & Barbot, B. (2024). Gender differences and variability in creative ability: A systematic review and meta-analysis of the greater male variability hypothesis in creativity. *Journal of Personality and Social Psychology*, *126*(6), 1161.

Torrance, E. P. (1966). Nurture of creative talents. *Theory Into Practice*, *5*(4), 167-173.

Torrance, E. P. (1974). Retooling education for creative talent: How goes it? *Gifted Child Quarterly*, *18*(4), 233-239.

Valladares-Rodríguez, S., Pérez-Rodríguez, R., Anido-Rifón, L., & Fernández-Iglesias, M. (2016). Trends on the application of serious games to neuropsychological evaluation: A scoping review. *J Biomed Inform*, *64*, 296-319. https://doi.org/10.1016/j.jbi.2016.10.019

Wallas, G. (1926). *The art of thought*. London.

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.