Engineering Educators Bringing the World Together
2025 ASEE Annual Conference & Exposition
Palais des congrès de Montréal, Montréal, QC • June 22–25, 2025
Paper ID #47272

Exploring Creative Productivity: Development of an Engineering Creativity Assessment Tool (ECAT)

Dr. Zeynep Gonca Akdemir-Beveridge, University of Connecticut

Zeynep G. Akdemir-Beveridge is a Postdoctoral Research Associate at the University of Connecticut. Her current research focuses on exploring the creative productivity of engineering students and promoting strength-based approaches in engineering education to better support neurodiverse learning profiles in STEM/engineering learning environments. Her doctoral research centered on K–12 curriculum development specifically tailored to quantum information science and engineering, as well as student motivation and engagement in science education. A third area of her work involves the application of artificial intelligence (AI) in science and engineering education, particularly in the context of personalized learning. She is the recipient of the Best of Computers in Education award for her co-authored work presented at the 2022 ASEE Annual Conference & Exposition.

Prof. Arash Esmaili Zaghi P.E., University of Connecticut

Arash E. Zaghi is a Professor in the Department of Civil and Environmental Engineering at the University of Connecticut. He received his PhD in 2009 from the University of Nevada, Reno, and continued there as a Research Scientist. His latest

Ms. Connie Syharat, University of Connecticut

Connie Syharat is a Ph.D. student in Engineering Education and a Research Assistant at the University of Connecticut as a part of two neurodiversity-centered NSF-funded projects. As the Program Manager of a Revolutionizing Engineering Departments (NSF:RED) project titled, ”Beyond Accommodation: Leveraging Neurodiversity for Engineering Innovation”, she has co-facilitated a range of Neuroinclusive Teaching Institutes and workshops for STEM instructors and Teaching Assistants. As a graduate researcher, she conducts qualitative research related to the experiences of neurodiverse graduate students in STEM fields through an Innovations in Graduate Education (NSF:IGE) project titled, ”Encouraging the Participation of Neurodiverse Students in STEM Graduate Programs to Radically Enhance the Creativity of the Professional Workforce”. Previously, she spent eight years as a K-12 teacher in Connecticut, where she maintained a focus on providing a varied learning environment and differentiated instruction for all types of learners.

Exploring Creative Productivity: Development of an Engineering Creativity Assessment Tool (ECAT)

Abstract

This work-in-progress study presents the Engineering Creativity Assessment Tool (ECAT), designed to systematically evaluate the creative productivity of engineering students by considering both the novelty of their ideas and the practicality of their implementations. Grounded in the 4P Model of Creativity, ECAT emphasizes the synergy between “seed” creativity and the structured, context-driven processes required to transform imaginative potential into innovative solutions. While existing engineering education assessments often focus on technical precision, the ECAT framework highlights the multifaceted nature of creativity, recognizing how fluency, originality, flexibility, elaboration, resistance to premature closure, and creative strengths collectively shape creative productivity.

ECAT is adapted from the Consensual Assessment Technique (CAT), a widely respected but underutilized method in engineering contexts. CAT leverages expert judgment to evaluate creativity and ensures reliability through inter-rater assessments, making it particularly well-suited for the nuanced, domain-specific evaluations of engineering projects. By incorporating multiple quasi-experts (advanced graduate engineering students), ECAT integrates professional-level criteria with authentic student experiences, reflecting real-world engineering challenges. During calibration sessions, these evaluators honed their scoring consistency and refined the assessment dimensions, balancing subjective insights with product-based measures.

Students’ creative outputs (n = 199) in this study were physical models constructed from basic materials within a constrained timeframe, coupled with short video explanations. This approach captures both the iterative process of engineering design and tangible outcomes. Although the data collection and evaluation are ongoing, the ECAT framework shows promise for producing a reliable measure of creative productivity. Once finalized, inter-rater reliability and various validity checks (content, convergent, and discriminant) will confirm the robustness of the instrument.

By offering a structured yet flexible assessment that values both imaginative exploration and practical function, ECAT hopes to serve as a diagnostic tool that could shift how engineering curricula address creativity. The long-term goal is to empower educators to identify gaps, refine pedagogical strategies, and ultimately cultivate engineers who can generate, refine, and implement innovative solutions with genuine impact.

Introduction

Creativity is widely recognized as an individual's capacity to generate original ideas or solutions by using imagination and ingenuity. Scholars argue that creativity encompasses certain cognitive and neural mechanisms not only involving divergent thinking but also exploring multiple perspectives and solutions to a problem [1]. For example, in the context of engineering, creativity is not just about artistic innovation but also the ability to navigate constraints, think outside of the box, elaborate on ideas, and optimize solutions. However, creativity alone, without purposeful direction, remains an untapped resource [2]. For instance, consider how Thomas Edison applied his creativity purposefully to invent the light bulb. His imaginative vision was supported by a structured, iterative process of trial and error, testing thousands of materials to find the ideal filament. Without this purposeful direction, the creative idea might have remained a theoretical concept, never realizing its full functionality in life. Therefore, it is crucial to realize that creativity serves as the seed of innovation, but it requires cultivation. This is where the concept of *creative productivity* becomes critical.

Creative productivity requires the systematic training of productive engineers who should implement creative ideas in existing circumstances so that innovations are not just conceptualized but also realized and function effectively. The synergy between creativity and productivity is especially important in Science, Technology, Engineering, and Mathematics (STEM) disciplines, where success hinges on both the generation of innovative ideas and their practical implementation. For example, during the COVID-19 pandemic we were able to see how engineers used their creative productivity for the rapid development and deployment of ventilators and other medical equipment under resource-constrained conditions. Engineering teams around the world harnessed their creativity to design and produce necessary medical equipment (i.e., critical care ventilators developed by General Motors and Ventec Life Systems, 3D printing of personal equipment and face shields) quickly. Despite the known importance of engineering creativity for daily life, 75% of engineering graduates report that their engineering education prioritizes technical proficiency over their engineering creativity [3] which we see as a reflection of the limitations inherent in the current structure of engineering education. Therefore, in this work-in-progress study, we aim to develop an Engineering Creativity Assessment Tool (ECAT) that aims to systematically quantify creative productivity of engineering students and to holistically understand their sources of creative productivity.

Background and Motivation

This study is grounded in the 4P Model of Creativity [4] [5] which identifies four key components essential for understanding and fostering creativity in any context. These components are Person, Process, Product, and Press. In the context of engineering education, we position that the Product, Process, and Press dimensions are particularly vital in evaluating creative productivity as they highlight the transition from ideation to implementation.

We can think of the Person in this model as the seed. A seed contains the innate potential for growth and holds the essential genetic blueprint for the tree to grow. Similarly, the Person dimension represents raw creativity including traits like imagination, curiosity, and divergent thinking that should serve as the starting point for innovation. Without the seed, the tree (or the entire creation) cannot exist. While the remaining three components serve a different purpose which is more contributing to the productivity aspects.

The Process is like the act of nurturing the seed such as watering it, ensuring it gets sunlight, and pruning it as it grows. It represents the structured actions, problem-solving, and iterative efforts required to transform raw creativity into something tangible. The Press is the environment where the quality of the soil, the weather, and the surrounding ecosystem determines whether the seed can thrive. It reflects the external factors, such as support, collaboration, and challenges, that influence the success of creative efforts. Lastly, the Product is the fully grown tree, representing the outcome of creativity or the entire creation. It is the tangible result of the combined efforts of the person, process, and press or an innovative solution or design that has grown from the seed of creativity into something functional and impactful.

The beginning (Person) and the end (Product) of the creative process are pivotal, but they do not stand alone. The Product dimension is critical for understanding the effectiveness of the creative process as a whole. It provides tangible evidence of how the initial creative potential (Person) was nurtured (Process) and supported (Press) to achieve an impactful outcome. By using the Product as the reference point, we can comprehensively evaluate and improve creative practices in fields like engineering. However, engineering education scholars overemphasize on technical mastery [6], [7] often neglecting the balanced integration of all components of the 4P model, thereby hindering the comprehensive development of creative productivity in engineering students.

Rationale

Engineering education scholars and educators have long recognized the importance of creativity in terms of innovative problem-solving and design thinking. However, measuring creative productivity in this domain remains complex due to its multifaceted nature, blending cognitive, technical, and contextual elements. Traditionally, engineering creativity has been assessed through structured methods such as standardized tests, rubrics evaluating design processes, and task-based assessments emphasizing technical accuracy over imaginative exploration [8]. These approaches often focus on assessing technical proficiency and correctness rather than the originality or novelty of ideas, thus offering a limited view of creativity. To address these limitations, contemporary assessment methods have evolved to incorporate subjective evaluations, such as reflective self-assessments and expert reviews, alongside product-based metrics like prototype evaluations and functional testing. However, what remains missing in these contemporary methods is an integration of expert-lead subjectivity with product-based evaluation. While subjective assessments effectively capture the internal cognitive processes and emotional engagement driving creativity, they often lack the tangible validation offered by product-based evaluations. Conversely, product-based approaches focus heavily on the outcomes but may overlook the innovative thinking and iterative process that led to these results. Without a combined framework that aligns the imaginative aspects of creativity with its practical application, assessments risk providing an incomplete picture, failing to fully appreciate the interplay between ideation and execution.

To address this gap, we developed the Engineering Creativity Assessment Tool (ECAT) aiming to explore creative productivity in engineering products by adapting the well-established Consensual Assessment Technique (CAT). By tailoring CAT to engineering contexts, ECAT will integrate subjective expert evaluation with product-based metrics, providing a framework for assessing not just the final product but also the creative processes underlying its development.

Consensual Assessment Technique (CAT)

Unlike rigid rubrics or predefined scoring systems, the CAT developed by [9] relies on the judgment of domain-specific experts who assess artifacts such as products, ideas, or solutions. In engineering, creativity often manifests through divergent thinking, innovative combinations of ideas, and the contextual appropriateness of solutions. These qualities are challenging to quantify but are crucial to solving complex engineering problems. CAT can address this challenge by leveraging expert evaluations to ensure that assessments remain relevant to the specific domain. Unlike objective tools, which are adept at measuring technical correctness or efficiency, CAT excels in capturing the nuanced and multi-faceted aspects of the assessed construct.

While CAT is a subjective assessment method, it incorporates mechanisms that ensure reliability and minimize bias, making it both valid and robust. The use of multiple independent evaluators is a key feature of CAT; experts evaluate artifacts without external influence or collaboration. Their judgments are aggregated to produce a consensus-based evaluation, reducing the impact of individual biases. Additionally, studies using CAT frequently measure inter-rater reliability, which validates the consistency of expert judgments and reinforces the credibility of the assessments [10].

In engineering education, CAT has been applied to assess the novelty and functionality of prototypes in settings such as capstone projects and design challenges [11] [12]. However, the concept of creative productivity in college-level engineering education was not explored through an adapted CAT before. Therefore, this study develops the Engineering Creativity Assessment Tool (ECAT) to address the unique challenges of creativity evaluation in this field. Unlike objective tools, which may oversimplify or overlook the complex interplay between imagination and functionality, ECAT hopes for a comprehensive evaluation that encompasses both. By combining expert judgment with engineering-relevant criteria, ECAT aims to build on the strengths of CAT to ensure that assessments reflect both the creative thought processes and the practical implementation of innovative ideas.

Proposed Framework for ECAT

Given the importance of creativity in engineering, this study leveraged the flexibility of the original CAT and tailored it for use in an engineering context. Historically, the focus on creativity in engineering has been linked to the invention of the concept of divergent thinking [13]. However, the notion of attributing innovative problem-solving in engineering design scenarios solely to a single cognitive ability, divergent thinking, has recently been conceptually [14] [15] [16] [17] and empirically [18] [19] challenged. Rather than presuming a universal creative ability with divergent thinking as its core component, domain-specific perspectives on creativity propose [14] [15] that a diverse range of skills relevant to a particular domain or task are essential for achieving creative outcomes within that context. This perspective acknowledges the potential role of divergent thinking in creative achievement but emphasizes the importance of understanding its contribution within the specific context of a domain. Instead of treating divergent thinking scores as definitive indicators of an individual's creativity, domain-specific perspectives suggest that the extent to which divergent thinking contributes to creative behavior in a given domain should be determined through empirical investigation [12]. This rationale is our motivation for the development of ECAT that will evaluate the domain-specific nature of engineering products.

ECAT Subscales

The initial design of ECAT has six distinct components: *fluency*, *originality*, *flexibility*, *elaboration*, *resistance to premature closure (RPC)*, and *creative strengths (CS)*. Although creativity studies using CAT have mainly followed three clusters of dimension types (creativity, technical strength, and aesthetic appeal) [9] [12] we expanded this structure by adding standardized measures of divergent thinking (i.e., flexibility, creative strengths) and adapted them to better fit the context of engineering.

Fluency

Fluency assesses the quantity and diversity of ideas, including the number of distinct components and functional features, as well as the variation in design approaches. It is broken down into three key subcategories, each with a five-point scale based on specific criteria:

- Number of Distinct Items (F1): Evaluators assess how many different materials or components are integrated.
  Example:
  - Score 1: The product is made using only one material.
  - Score 5: The product includes five or more materials, showing high material variety and resourcefulness.
- Number of Functional Features (F2): Evaluators look at the variety of distinct functional elements in the product.
  Example:
  - Score 1: The product performs a single function.
  - Score 5: The product includes five or more functional elements, showing a high level of creative capability.
- Variation in Product (F3): This subcategory captures how much the product explores diverse design approaches.
  Example:
  - Score 1: The product remains a single concept with little variation.
  - Score 5: The product demonstrates paradigm-shifting ideas, incorporating multiple applications.

Originality

Originality evaluates how novel and innovative a product is, extending beyond basic functionality to assess creative and unconventional thinking. This dimension includes three subcategories:

- Uniqueness (O1): Evaluates how much the product's core idea diverges from standard engineering designs.
  Example:
    - Score 1: The product is highly conventional.
    - Score 5: The product represents a significant innovation.
- Innovation (O2): Examines whether the product introduces new or improved functions.
  Example:
    - Score 1: The product is a basic design with no additional functional enhancements.
    - Score 5: The product integrates advanced functionalities.
- Surprise (O3): Assesses whether the product defies expectations in its material usage or design.
  Example:
    - Score 1: The product follows predictable material use.
    - Score 5: The product completely defies expectations.

Flexibility

Flexibility measures the range of design strategies and how well the product adapts to constraints. This dimension has two subcategories, both scored on a five-point scale:

- Range of Design Approaches (FL1): Assesses whether the product incorporates diverse design methods.
  Example:
    - Score 1: The product follows a single, rigid design.
    - Score 5: The product explores multiple materials and techniques.
- Adaptation to Constraints (FL2): Evaluates whether the product adjusts to limitations and unexpected issues.
  Example:
    - Score 1: The student fails to adapt when materials run out, leading to an unstable final product.
    - Score 5: The student redesigns the product due to constraints.

Elaboration

Elaboration assesses the detail, completeness, and depth of thought behind the product. It includes three subcategories:

- Functional Completeness (E1): Examines whether the product fully meets its intended purpose.
  Example:
    - Score 1: The product is incomplete or barely functional.

        - Score 5: The product is highly refined, with exceptional attention to detail and no flaws.
- Depth of Explanation (E2): Evaluates how well the student explains their design process.
  Example:
    - Score 1: The explanation lacks depth or reasoning behind design choices.
    - Score 5: The student provides an in-depth narration, explaining why certain materials were chosen and how challenges were overcome.
- Uniqueness of the Problem (E3): Assesses whether the product addresses a novel or challenging issue.
  Example:
    - Score 1: The problem is trivial or generic.
    - Score 5: The product addresses a complex, innovative challenge that pushes engineering boundaries.

Resistance to Premature Closure (RPC)

RPC evaluates whether students continue exploring and refining ideas rather than settling on the first workable solution. Unlike the other dimensions, RPC is scored on a descriptive scale (Complete, Satisfactory, Unsatisfactory). Evaluators examine:

- Number of Iterations (RPC1):
  Example: Students who revise their designs multiple times receive a higher score.
- Continuous Development of Ideas (RPC2):
  Example: Products that evolve significantly during the design process score higher than those that remain unchanged.
- Willingness to Address Challenges (RPC3):
  Example: Products where students actively problem-solve and improve their design based on limitations receive higher ratings.

Creative Strengths (CS)

CS adapts elements from the Torrance Tests of Creative Thinking (TTCT) and focuses on the practicality, aesthetic appeal, and effective use of color in the product. This dimension is also assessed on a descriptive scale (Complete, Satisfactory, Unsatisfactory).

- Usefulness (CS2): Evaluates whether the product serves a practical function.
  Example: A product that performs multiple useful functions scores higher than one with limited application.
- Aesthetics & Colorfulness (CS3&5): Judges whether the product is visually appealing and well-designed.
  Example: A product that is symmetrically balanced, colorful, and thoughtfully constructed scores higher than one that appears rough or unstructured.

Student Participants

Before the ECAT evaluation, a questionnaire was administered requesting demographic information from student participants (n = 199). This included their age, sex, and college major. Data about race/ethnicity was not collected. The sample was comprised of students from Biomedical Engineering (17.59%), Civil Engineering (12.56%), Chemical Engineering (7.54%), Computer Science & Engineering (12.56%), Electrical Engineering (5.03%), Environmental Engineering (4.02%), Manufacturing & Engineering Management (7.54%), Materials Science & Engineering (5.53%), Mechanical Engineering (24.12%), Engineering Physics (.50%), Dual Major (1.01%), Undeclared (1.51%), and Unknown (0.50%). Demographics data indicated that 199 undergraduate engineering students consisting of 56.3% males and 43.7% females ranging in age from 18-33 at a northeastern public U.S. university participated in a creative engineering activity. They were prompted to create their own product in 20 minutes by using different types and colors of simple materials such as paper clips, buttons, stir sticks, and string. At the end of the production process, participants were asked to explain what they made and show what it does. Students' explanations were video recorded (with resulting videos ranging in length from about 30 seconds to 2 minutes) and then transcribed. Photographs of the product were also taken from multiple angles. Participants received a $35 gift card for their participation in the study. Participants were coded and given a number for identification purposes according to the Institutional Review Board (IRB) approval procedures.

Selection of Evaluators

To ensure a reliable assessment of domain-specific creativity, we assembled an appropriate group of evaluators by recruiting five quasi-experts from the engineering field to evaluate the creative potential of student engineering projects. This group included four advanced graduate-level engineering students. Following Amabile's recommendation [20], we confirmed that the evaluators possessed a significantly higher level of expertise than the students who created the projects. Additionally, based on prior research [21] we determined that quasi-experts were well-suited for this study, as their expertise meaningfully overlapped with the context of the products being evaluated.

Calibration Meetings

We conducted a structured calibration process involving training and iterative refinement with quasi-expert evaluators. This process spanned three calibration meetings designed to align evaluators' scoring practices and refine the assessment.

Familiarization and Hands-On Experience

The first calibration meeting introduced evaluators to the concept of creativity in engineering education and the specific dimensions of ECAT. Each dimension was discussed in detail, supported by examples of engineering products to illustrate their relevance. Evaluators also participated in a short hands-on workshop, where they created simple engineering products using materials similar to those available to students. This activity encouraged empathy for student challenges and provided practical experience to internalize the ECAT dimensions.

Feedback and Refinement

The second meeting focused on refining the tool based on evaluators' feedback. Evaluators were provided with three sample student products representing low, medium, and high creativity levels, which they evaluated using a Qualtrics form. A week later, score distributions were analyzed to identify significant discrepancies. A group discussion followed, where evaluators explained their reasoning, resolving misunderstandings and ambiguities in scoring. Insights and data from this session informed further refinement of ECAT, ensuring evaluators shared a consistent understanding of the rubric.

Consistency Testing

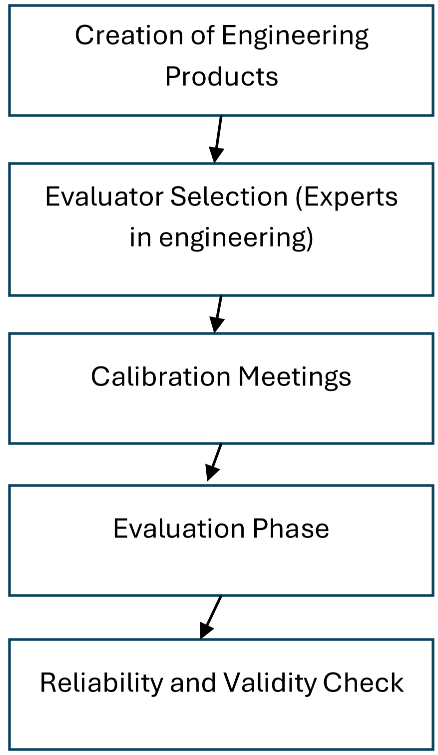
The final meeting assessed evaluators' consistency in applying ECAT. Evaluators individually scored five additional samples with varying creativity levels, focusing on achieving alignment in their assessments. The session confirmed that evaluators were largely consistent, with only minor disagreements, marking the conclusion of the calibration process and readiness for ECAT evaluation.

The results from the calibration session confirmed that evaluators reached a sufficient level of alignment, allowing us to proceed confidently with the actual evaluation process. A high proportion of partial agreement (up to 71.4%) suggested that evaluators had a shared understanding of the scoring criteria, even if they did not always assign identical scores. The level of disagreement was within an acceptable range (within 28.6% – 35.3%), confirming that ECAT dimensions were interpretable but still flexible enough to accommodate diverse expert perspectives.

Evaluation Phase

The ECAT was employed to assess the creative productivity of student generated products developed between Fall 2018 and Spring 2019. Each evaluator was assigned an individual rater ID to ensure anonymity and consistency throughout the evaluation process. Evaluators accessed the ECAT via a Qualtrics link. The evaluation period commenced in November 2025 and spanned three months, providing raters with autonomy in pacing while adhering to the deadline set by the research team.

Figure 1 shows the structured flowchart of ECAT testing methodology.



ECAT's Reliability and Validity

To establish the reliability and validity of ECAT, we examined inter-rater reliability, internal consistency, content validity, convergent validity, discriminant validity, and factor structure analyses.

Reliability

To ensure consistent evaluator interpretation, we conducted calibration meetings where experts aligned their scoring strategies for creative engineering products. Given the subjective nature of creativity assessment, achieving absolute agreement was not expected; instead, the goal was to ensure consistency in scoring interpretations.

During calibration sessions, inter-rater agreement reached 71.4%, allowing us to proceed with the full evaluation phase. To quantify inter-rater reliability, we computed Kendall's Tau-b correlations between sum scores of each subscale across 827 evaluations. The results indicated moderate to strong correlations ($p < 0.01$), supporting consistent evaluator judgment across ECAT dimensions. This aligns with prior research, which suggests that agreement among independent raters reinforces construct validity in creativity assessment [12].

After the evaluation phase, we assessed internal consistency by computing Cronbach's Alpha (α) across ECAT subscales. The reliability coefficient demonstrated acceptable internal consistency ($\alpha = .89$) indicating that ECAT subscales consistently measure their intended constructs.

Validity

Content Validity

To ensure content validity, we conducted structured calibration sessions with expert evaluators. The process began with a systematic item review, in which experts assessed whether ECAT's six subscales comprehensively captured key aspects of engineering creativity. Rather than using individual content validity index (CVI) scores, we employed group discussions to elicit qualitative feedback, which was documented and systematically analyzed. Based on expert recommendations, certain items were refined or removed to enhance ECAT's clarity and construct representation.

Convergent Validity

To evaluate convergent validity, we compared ECAT scores with TTCT scores from the same sample of students. We assessed whether ECAT dimensions correlate with TTCT's established measures of divergent thinking, including verbal fluency, verbal originality, and figural elaboration.

The results provided partial support for convergent validity, demonstrating that ECAT's originality ($r = 0.26–0.32$), elaboration ($r = 0.27$), and flexibility ($r = 0.29$) moderately correlated with TTCT's verbal creativity measures. These findings suggest that ECAT captures similar aspects of creative productivity as TTCT verbal subscales.

As expected, correlations between ECAT and TTCT figural measures were lower ($r < 0.20$), reinforcing that engineering creativity may involve domain-specific manifestations distinct from traditional TTCT assessments. This aligns with prior research indicating that engineering creativity assessments emphasize problem-solving and functional originality, whereas TTCT figural tasks assess visual fluency and elaboration [22].

Discriminant Validity

To test discriminant validity, we examined Pearson correlations among ECAT subscales to determine whether each dimension captures a distinct aspect of engineering creativity.

The results confirmed that fluency, originality, and flexibility are related but remain statistically distinct constructs, as indicated by moderate correlations between fluency and originality (F2–O1: $r = 0.52$, F3–O2: $r = 0.72$) and weaker correlations between fluency and flexibility (F1–FL2: $r = 0.22$). This distinction supports the theoretical framework that fluency represents idea quantity, originality reflects uniqueness, and flexibility captures adaptability in engineering design [22].

Furthermore, elaboration was found to be independent from fluency and originality, reinforcing that depth and refinement of ideas do not necessarily align with idea generation or novelty. This is evidenced by lower correlations between elaboration and fluency (F1–E1: $r = 0.26$, F2–E1: $r = 0.47$) and moderate correlations with originality (O1–E1: $r = 0.42$, O2–E2: $r = 0.45$).

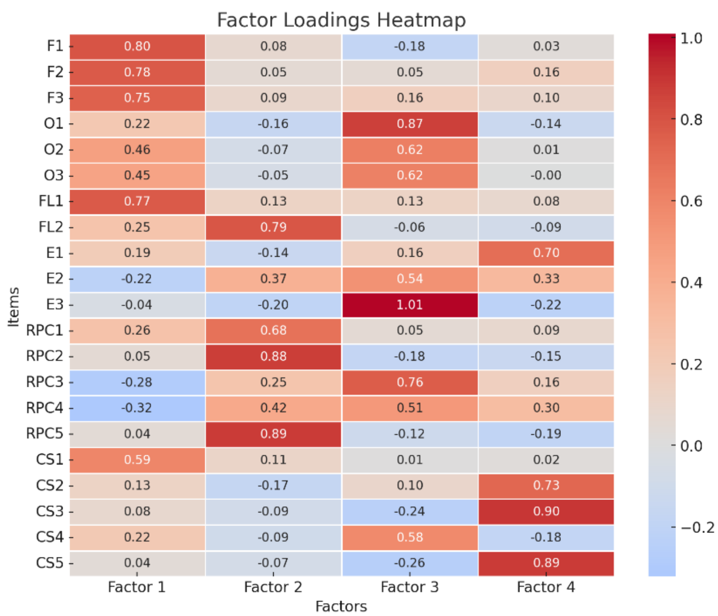
Additionally, resistance to premature closure (RPC) emerged as a separate construct, exhibiting weaker associations with fluency (F1–RPC1: $r = 0.30$, F2–RPC1: $r = 0.39$) and flexibility (FL1–RPC1: $r = 0.45$). Lastly, creative strengths (CS) remained statistically distinct from fluency and flexibility, with aesthetic appeal (CS4) showing stronger associations with originality (O1–CS4: r = 0.44, O2–CS4: r = 0.45).

These findings confirm that ECAT's dimensions measure theoretically distinct constructs.

Factor Structure Analyses

Exploratory Factor Analysis (EFA)

To determine the underlying structure of the dataset and identify the factor loadings of each subscale in the ECAT, we conducted an EFA. Factor loadings were extracted using principal axis factoring with an oblique rotation to account for potential correlations among factors. All statistical analyses were performed using Python. Figure 2 presents a heatmap of the factor loadings.



The heatmap in Figure 2 illustrates the four-factor solution identified through EFA. The magnitude and direction of each cell's color and numerical value indicate the strength and sign, respectively, of the item's loading on each factor. Consistent with established guidelines in behavioral and social sciences [23], we considered loadings above ±0.40 to be substantive. Items with the highest absolute-value loadings per row were used to define the factor structure. The results suggested a meaningful thematic grouping:

- Factor 1 primarily consists of F and FL1 items
- Factor 2 centers around RPC and FL2 items.
- Factor 3 includes O items along with E3 (and partially RPC3).
- Factor 4 groups E1 with several CS items.

The observed overlap between RPC and FL2 can suggest a conceptual relationship between RPD and the adaptation to constraints (FL2). Consequently, Factor 2 was renamed "Cognitive Flexibility", integrating flexibility and RPC dimensions. The differentiation between Factors 1 and 3 further confirms that ECAT's fluency and originality are distinct constructs. To validate this factor structure, we proceeded with Confirmatory Factor Analysis (CFA).

Confirmatory Factor Analysis (CFA)

The initial CFA model was assessed using key fit indices to evaluate how well the hypothesized four-factor model fit the data. The initial model demonstrated an acceptable but suboptimal fit, as indicated by the following values:

- Comparative Fit Index (CFI) = 0.88
- Normed Fit Index (NFI) = 0.86
- Adjusted Goodness-of-Fit Index (AGFI) = 0.80
- Tucker-Lewis Index (TLI) = 0.84
- Root Mean Square Error of Approximation (RMSEA) = 0.127
- $\chi^2 = 199.47$, $p < 0.01$

To improve model fit, we refined Factor 2 by integrating flexibility and RPC into a single construct and removed E3 from Factor 3 due to its inflating effect. The revised CFA model demonstrated significantly improved fit indices, as follows:

- CFI = 0.95
- NFI = 0.93
- AGFI = 0.90
- TLI = 0.93
- RMSEA = 0.089
- $\chi^2 = 97.42$, $p < 0.01$

The improved CFA results confirm the validity of the refined four-factor model, reinforcing the theoretical basis of ECAT's structure. These findings indicated that the revised model more accurately captures the latent constructs, supporting its use in engineering creativity assessment. Future work should include validation with external datasets to further ensure model generalizability.

Limitations

We have adapted CAT, which is inherently subjective. Although we do not see this aspect as a limitation, we acknowledge that ECAT does not aim for absolute objectivity but rather structured subjectivity, where evaluators apply clear criteria to reach a shared understanding of creative value of engineering student products. While employing multiple evaluators and measuring inter-rater agreement enhances reliability, individual biases and varying interpretations of creative productivity may still influence the results.

Additionally, we recognize that the products being evaluated are created within a limited time frame (20 minutes) using simple materials and prompt-based instructions. These constraints may not reflect the broader complexities and real-world pressures involved in engineering design processes. Lastly, this study focuses on a specific sample of engineering students. The extent to which our findings can be generalized to other contexts (such as different academic levels, other STEM fields, or professional engineering settings) remains to be determined.

Significance and Expected Contributions

Our ECAT framework aims to enhance the assessment of engineering students' creative productivity by extending the traditional CAT. It achieves this by blending expert-led subjective judgments with product-based criteria, offering a dual emphasis that provides a more holistic view of creative productivity. This approach captures both the novelty and practicality of engineering solutions. Additionally, the framework emphasizes domain-specific engineering creativity, addressing a key weakness of existing creativity assessments in the field [23]. Unlike many current tools that focus solely on divergent thinking or final prototypes, ECAT's multiple dimensions highlight the entire journey from ideation to functional implementation. This comprehensive perspective is particularly relevant for preparing students to engage in innovative problem-solving within professional engineering environments.

The ECAT framework also holds significant potential for educational improvement. Once fully validated, it can function as a diagnostic tool for engineering educators and curriculum developers, helping them identify gaps in student creativity and productivity. This feedback can inform pedagogical interventions, project-based learning strategies, and assessment methods designed to foster well-rounded engineering skills.

Acknowledgments

This study is based upon work supported by the National Science Foundation (NSF) under Grant CAREER - 1653854 and Grant IUSE/PFE: RED - 1920761. Any opinions, findings, conclusions, expressed in this study are those of the authors and do not necessarily reflect the views of the NSF. Additionally, we acknowledge Dr. Connor Beveridge, who is a postdoctoral researcher at Purdue University for their support for the data analysis and Python assistance during the development of this study.

References

[1] C. Chen, H. Moradi, L. Vahid and R. Khalil, "Editorial: Exploring goal-directed behavior through creativity: perspectives from psychology, neuroscience, and psychiatry," *Frontiers Psychology,* vol. 15, 2024.

[2] L. H. Malinin, "How radical is embodied creativity? Implications of 4E approaches for creativity research and teaching.," *Frontiers in Psychology,* p. https://doi.org/10.3389/fpsyg.2019.02372, 2019.

[3] R. M. Brooks and S. Shinde, "Unlocking the Creativity Potential," in *2020 ASEE Virtual Annual Conference Content Access*, 2020, June.

[4] M. Rhodes, An analysis of creativity, Phi Delta Kappan, 1961.

[5] J. C. Kaufman, Creativity 101, Newyork: Springer, 2016.

[6] D. &. C. A. J. Cropley, "Recognizing and fostering creativity in technological design education.," *International Journal of Technology and Design Education,* vol. 20, no. 3, pp. 345-359, 2010.

[7] A. J. Cropley, Fostering creativity: A diagnostic approach for higher education and organizations., Springer, 2009.

[8] C. Charyton and J. A. Merrill, "Assessing General Creativity and Creative Engineering Design in First Year Engineering Students," *Journal of Engineering Education,* vol. 98, no. 2, pp. 145-156, 2013.

[9] T. M. Amabile, The social psychology of creativity, Springer-Verlag, 1983.

[10] P. Barth and G. Stadtmann, "Creativity Assessment over Time: Examining the Reliability of CAT Ratings," *Journal of Creative Behavior,* vol. 55, no. 2, pp. 396-409, 2020.

[11] J. K. Buelin-Biescker and C. D. Denson, "Consensual Assessment: A Means of Creativity Evaluation for Engineering Graphics Education," in *ASEE Engineering Design Graphics Division 68th Mid-Year Conference*, Worcester, MA, October 20-22, 2013.

[12] C. D. Denson, J. K. L. M. D. Buelin and D. D'Amico, "Developing Instrumentation for Assessing Creativity in Engineering Design," *Journal of Technology Education,* vol. 27, no. 1, pp. 23-40, 2015.

[13] J. P. Guilford, The Nature of Human Intelligence, New York: McGraw-Hill, 1967.

[14] J. Baer, "Domain Specificity and the Limits of Creativity Theory," *Journal of Creative Behavior,* vol. 46, no. 1, pp. 16-29, 2012.

[15] J. Baer, "Creativity Doesn't Develop in a Vacuum," *New Directions for Child and Adolescent Development,* vol. 151, pp. 9-20, 2016.

[16] J. Baer, "Content matters: Why nurturing creativity is so different in different domains," *reative contradictions in education: Cross disciplinary paradoxes and perspectives,* pp. 129-140, 2017.

[17] M. A. Runco and S. Acar, "Divergent thinking," in *The Cambridge handbook of creativity*, Cambridge University Press, 2019, pp. 224-254.

[18] W. A. Carpenter, "Engineering creativity: Toward an understanding of the relationship between perceptions and performance in engineering design (Doctoral dissertation).," University of Akron, 2016.

[19] K. H. Kim, "The creativity crisis: The decrease in creative thinking scores on the Torrance Tests of Creative Thinking.," *Creativity Research Journal,* vol. 23, no. 4, pp. 285-295, 2011.

[20] T. M. Amabile, Creativity in context, Boulder, CO: Westview Press, 1996.

[21] J. C. Kaufman, J. Baer, D. H. Cropley, R. Reiter-Palmon and S. Sinnett, "Furious activity vs. understanding: How much expertise is needed to evaluate creative work?," *Psychology of Aesthetics, Creativity, and the Arts,* vol. 7, no. 4, 2013.

[22] R. Reiter-Palmon, B. Forthmann and B. Barbot, "Scoring divergent thinking tests: A review and systematic framework.," *Psychology of Aesthetics, Creativity, and the Arts,* vol. 13, no. 2, pp. 144-152, 2019.

[23] D. H. Cropley, "Promoting Creativity and Innovation in Engineering Education," *Psychology of Aesthetics, Creativity and the Arts,* vol. 9, no. 2, pp. 161-171, 2015.