# WIP: Generative vs. Traditional Computer-Aided Design—How design tools impact CAD artifact quality

Aidan Hall
*Industrial and Enterprise Systems Engineering*
*University of Illinois at Urbana-Champaign*
Urbana-Champaign, USA
aidanjh3@illinois.edu

Molly H. Goldstein, PhD
*Industrial and Enterprise Systems Engineering*
*University of Illinois at Urbana-Champaign*
Urbana-Champaign, USA
mhg3@illinois.edu

***Abstract*— This work-in-progress research paper explores the effects of generative AI design tools on engineering students' design artifacts during computer-aided design (CAD) tasks with the objective to better understand the impact of emergent design tools on design artifacts to better inform the development of curriculum surrounding generative design. Our research explored the question: to what extent does engaging in generative design produce a different quality artifact as compared to traditional design? This study utilized a mixed methods approach to compare students' (n=20) CAD artifact quality between two separate tasks; one completed with traditional parametric modeling and the other with generative design tools. Preliminary findings indicate that students produced a significantly higher quality design artifact using generative design as compared to artifacts created using traditional parametric methods.**

***Keywords*— *engineering design, generative design, computer-aided design, design artifact assessment***

## I. Introduction

The rise of generative tools powered by AI and other intelligent algorithms has caused a boom in nearly every industry that is trying to capitalize on the capabilities of these tools. Generative design "leverages the power of computationally driven artificial intelligence (AI) to automatically explore a wide design space in order to identify the best design options" [1]. Autodesk developed one of the first commercially available generative design tools within their cloud-based CAD software Fusion 360 and has been used in industry to optimize performance of design solutions [2]. Fusion 360's generative design workflow requires the designer (1) identify the critical structures of the component, (2) minimize the mass or maximize the stiffness of the structure, (3) specify a material, and (4) specify a manufacturing method [3]. Generative design workflows are being incorporated into undergraduate engineering classes [4], suggesting that these new technologies will become more commonplace in upcoming years.

Existing literature has begun to explore the concepts of generative design and how it fits into our understanding of design practices and design thinking [3-4], [6-8], but little has focused on how it is perceived among engineering students who are beginning to learn design practices as well as how those differences can be observed quantitatively from their produced artifacts. This study seeks to contribute to the growing knowledge base that will allow educators to create curriculum that ensures strong understanding of the difference between the two design modalities. The knowledge found in this study will also contribute to developing benchmarks for assessing both artifacts made during generative design and student understanding and performance.

## II. Study Design

### A. Recruitment

After approval from IRB, participants in this study were recruited via in-person pitches in courses at a large Midwestern University in the United States in which students were known to have completed introductory engineering graphics (CAD) courses. This was done to ensure that all participants had a basic knowledge of CAD software functionality, specifically Autodesk Fusion 360. Students who were selected for the study after screening were compensated with $40 gift cards after completion of the design tasks.

### B. Screening Survey

All participants completed a screening questionnaire that determined their eligibility for participation in the CAD and collected preliminary pre-task data. The data collected included basic demographic information, experience with CAD modeling, engineering self-efficacy, and perceptions of traditional vs. generative design concepts.

### C. Participants

A total of twenty (n=20) participants were selected following the screening process. A brief breakdown of their demographics can be seen in Table I. It is important to note that this population of participants does not currently include all majors that may utilize CAD software in their coursework.

TABLE I. Demographic Information

| **Participant Demographics (n = 20)** | |
|---|---|
| ***Gender*** | Male (11), Female (7), Prefer not to say (2) |
| ***Standing in School*** | Freshman (3), Sophomore (3), Junior (5), Senior (7), Graduate (2) |
| ***Major*** | Mechanical Engineering (12), Systems Engineering (4), Agricultural & Biological Engineering (2), Industrial Engineering (1), Education (1) |
| ***Self-Described CAD Skill Level*** | Beginner (1), Intermediate (6), Advanced (12), Expert (1) |

This material is based upon work supported by the National Science Foundation under Grant DUE #2207408.

## D. Design Task

During a two-hour in-person lab, participants were tasked with completing two CAD challenges, one using traditional parametric modeling and one using Fusion's generative design environment. The tasks were designed to have a similar level of difficulty and be able to be completed using both modalities. The objective of each task was to design a structural component of an existing assembly that used less material than a previous iteration of that part. The designs also needed to meet several constraints that ensure the part would fit within the existing assemblies and withstand specific loading conditions. The two design spaces can be seen in Fig. 1 and 2.

The mass of the previous part and existing assemblies were provided, but to prevent inspiring the participants' designs the image of the previous part iterations were excluded. The participants were randomly assigned to two different treatment groups, A and B (See Table II). The organization of treatment groups ensured that each participant was exposed to both tasks and both design modalities.

Students completed the design task in a lab while a researcher was present. At the end of the design session, participants uploaded their design files to a secure Box folder for analysis by the research team.

## E. Post-Task Survey

Following completion of the two CAD tasks, participants were asked to complete a post-workshop questionnaire. This questionnaire asked students to describe their design and decision-making process throughout each task as well as their preferences towards both (traditional vs. generative). The questionnaire also repeated the questions regarding perception of traditional and generative design concepts to examine whether completing the tasks had any influence on student perception of concepts.

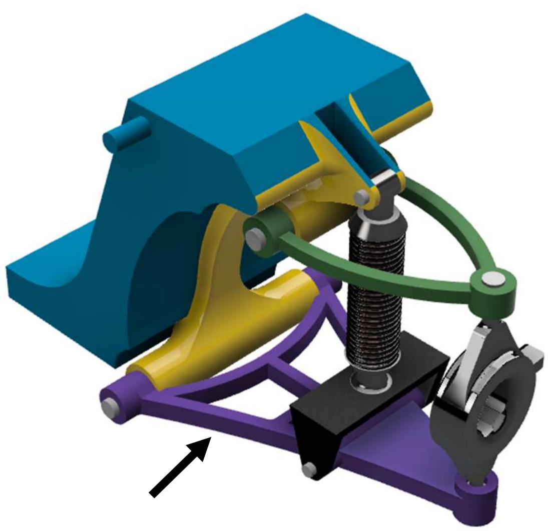

Fig. 2. Suspension Assembly design space, arrow indicates part being designed by participants.

Fig. 1. Excavator Assembly design space, arrow indicates part being designed by participants.

TABLE II. TREATMENT GROUP TASKS

| Treatment Group | *First Task* | *Second Task* |
|---|---|---|
| *A* | Traditional Suspension | Generative Excavator |
| *B* | Traditional Excavator | Generative Suspension |

# III. DATA ANALYSIS

## A. Analysis of Artifacts

The primary focus of data analysis was assessment of the submitted design artifacts to explore the research question: To what extent does engaging in generative design produce a different quality artifact as compared to traditional design? The mix of tasks and design methods created four distinct types of artifacts for comparison, which can be seen in Table III.

To effectively evaluate the broad range of design artifacts created in the workshop, a rubric was created to assign each submission a quality score. The rubric contained three primary sections: objectives and constraints, design quality, and metric rankings. The objectives and constraints assigned binary points for successfully meeting each objective and constraint set forth in the problem description. The design quality section assigned values on a qualitative scale (ranging from 0-2) for meeting various standards such as finish quality and visual interest.

The metric rankings assigned points based on the part performance relative to its peers in the following metrics: part mass, factor of safety, maximum Von Mises stress, strength to weight ratio, and material cost. The metric rankings section was adapted from previous work which developed the Trade-Off Value Protocol to evaluate design artifacts produced from an open-ended design task [5]. The submissions were analyzed in the Fusion 360 simulation environment to determine the metrics for each artifact. All artifacts were subject to the same load cases for testing. All points were then totaled to achieve a Quality Score. (Please see the appendix for rubric details)

TABLE III. ARTIFACT TYPES

| ***Group A*** | |
|---|---|
|  Traditional Suspension |  Generative Excavator |
| ***Group B*** | |
|  Traditional Excavator |  Generative Suspension |

### *B. Statistical Analyses*

The Wilcoxon Signed rank test was used to compare the Quality Scores to assess how performance varied between both tasks (suspension versus excavator) and modalities (traditional versus generative design). The Wilcoxon Signed Ranked test was selected because the quality score data was paired non-parametric data that matched the following necessary assumptions (1) the data was continuous and there were exactly two sample groups (2) the data had been obtained from representative samples of the underlying populations (3) the data was from dependent samples [9].

## IV. RESULTS AND DISCUSSION

### *A. Task Results*

Artifact quality scores in the suspension control arm design task ranged from 5.6 to 11.5 with a median of 8.9. Artifact scores in the excavator scoop arm design task ranged from 6.2 to 11.7 with a median of 9.5.

The Wilcoxon Signed Rank test was conducted to ensure that there was no statistically significant difference in the artifact quality scores between the two design tasks. The dataset generated a W-statistic $W = 97$, which is greater than the critical value 52 for $n = 20$, and a p-value of 0.78. These values confirmed that there is no significant difference in artifact quality score between the two design tasks.

This is a promising result because it suggests that student performance was not affected between the two tasks and that the difference in task prompts is likely not a confounding variable.

### *B. Modality Results*

Artifact quality scores completed in the traditional design modality ranged from 6.1 to 11.5 with a median of 8.4. Artifact scores in the generative design modality ranged from 5.6 to 11.7 with a median of 10.4.

The Wilcoxon Signed Rank test was conducted to test whether there is a significant difference in the artifact quality scores between the two design modalities. The dataset generated a W-statistic $W = 18$, which is less than the critical value 52 for $n = 20$, and a p-value of 0.031. These values confirmed that there is in fact a significant difference in artifact quality score between the two design modalities, with artifacts created in the generative design modality performing significantly better than those created with traditional design techniques.

This study suggests that computational tools show promise for increasing the quality of CAD-created design artifacts. Additionally, generative design might be a feasible tool for student designers to use to increase idea fluency which is important to address as many [10] have indicated that student designers struggle with coming up with an abundance of design ideas [11]. Indeed, Adams stated designers show “reluctance to spend the time and mental effort needed to conjure up a rich storehouse of alternatives from which to choose” [12] as cited in [10]. Generative design might complement proven divergent thinking in design techniques such as broad brainstorming techniques or specific techniques such as Design Heuristics [13][14].

### C. Limitations and Future Work

While this study offers a deeper exploration of student designers engaging in both traditional and generative design and answers the call to understand how computational tools affect individual designers [15], it has limitations that can be addressed with additional work. This study is a part of a larger project investigating how students engage in generative design. The larger study collected data including designer demographics, their conceptions of design, and engineering self-efficacy. Our future work will look to triangulate these multiple data streams to better understand profiles of designers and how they engage in the design process.

## V. Conclusion

In this study, we aimed to understand the difference in design artifact quality between traditional and generative design in engineering students. Our results showed that the design solutions created with generative design were of higher quality, as measured by meeting design objectives and constraints, finish quality, and performance metrics. The results are exciting in showing that AI-tools show promise in design, when measured in a rigorous and consistent manner. Future work will look beyond the design artifacts to understand how AI-tools like generative design impact the design process and the designer.

## Acknowledgement

We are grateful to the students who participated in this study. This material is based upon work supported by the National Science Foundation under Grant DUE #2207408. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the view of the National Science Foundation.

## REFERENCES

[1] Leake, J. M., Goldstein, M. H., and Borgerson, J. L., 2022, Engineering Design Graphics: Sketching, Modeling, and Visualization, John Wiley and Sons, Hoboken, NJ.

[2] McKnight, M., 2017, “Generative Design: What It Is? How Is It Being Used? Why It’s a Game Changer,” KnE Eng., 2(2), pp. 176–181.

[3] Barnaby-Brown, A. & Goldstein, M. H. Clay, J., Demirel, H. O., Li, Xi., Sha, Z. (2024). A study on generative design reasoning and students’ divergent and convergent thinking. Journal of Mechanical Design, 146(3): 031405. https://doi-org/10.1115/1.4064564

[4] Demirel, H. O., Goldstein, M. H., Li, X., & Sha, Z. (2024). Human-Centered Generative Design Framework: An Early Design Framework to Support Concept Creation and Evaluation. International Journal of Human–Computer Interaction, 40(4), 933–944.

https://doi-org./10.1080/10447318.2023.2171489

[5] Goldstein, M. H., Adams, R. S., & Purzer, S. (2021). Understanding Informed Design through Trade-Off Decisions With an Empirically-Based Protocol for Students and Design Educators. Journal of Pre-College Engineering Education Research (J-PEER), 11(2), Article 3. https://doi.org/10.7771/2157-9288.1279

[6] Saadi, J. I., and Yang, M. C. (February 20, 2023). "Generative Design: Reframing the Role of the Designer in Early-Stage Design Process." ASME. J. Mech. Des. April 2023; 145(4): 041411. https://doi.org/10.1115/1.4056799

[7] Elisa Koolman, John Z. Clay, Xingang Li, Rundong Jiang, Molly H. Goldstein, Charles Xie, H. Onan Demirel and Zhenghui Sha. (2024). A multi-case study of traditional, parametric, and generative design thinking of engineering students. Design Computing and Cognition Conference, July 2024, Montreal.

[8] Goldstein, M. H., Sommer, J., Trellinger Buswell, N., Li, X., Sha, Z., Demirel, H. O. (2021). Uncovering generative design rationale in the undergraduate classroom. Proceedings from the Frontiers in Education (FIE) Annual Conference, Lincoln, NE, October 2021.

[9] Matthew Damian Riina, Cassandra Stambaugh, Nathaniel Stambaugh, Kathryn E. Huber, Chapter 28 - Continuous variable analyses: t-test, Mann–Whitney, Wilcoxin rank, Editor(s): Adam E.M. Eltorai, Jeffrey A. Bakal, Daniel W. Kim, David E. Wazer, In Handbook for Designing and Conducting Clinical and Translational Research, Translational Radiation Oncology, Academic Press, 2023, Pages 153-163, ISBN 9780323884235, https://doi.org/10.1016/B978-0-323-88423-5.00070-4.

[10] Crismond, D. P., & Adams, R. S. (2012). The informed design teaching and learning matrix. Journal of engineering education, 101(4), 738.

[11] Cross, N. (2000). Engineering design methods: Strategies for product design (3rd ed.). New York: John Wiley 8c Sons.

[12] Adams, J. L. (1986). Conceptual blockbusting: A guide to better ideas/ Reading, MA: Perseus Books.

[13] Yilmaz, S., 6c Seifert, C. M. (2011). Creativity through design heuristics: A case study of expert product design. Design Studies 32(4), 384-415.

[14] Yilmaz, S., Daly, S. R., Seifert, C. M., & Gonzalez, R. (2016). Evidence-based design heuristics for idea generation. Design Studies, 46, 95-124.

[15] Gyory, J. T., Song, B., Cagan, J., and McComb, C., 2021, “Communication in AI-Assisted Teams During an Interdisciplinary Drone Design Problem,” Proceedings of the Design Society: 23rd International Conference on Engineering Design, Gothenburg, Sweden.


## APPENDIX

**Task #1 – Formula SAE Suspension Design Challenge (Traditional)**

*Problem Description:*
The Illini Formula SAE Electric team is working on their next model for the 2025 racing season and is looking at design options for the new and improved suspension system. The team is looking to improve vehicle performance by reducing the total weight of the car and every component is subject to potential weight reduction. The design team is challenging you with designing a new lower suspension arm that weighs less than last year’s model.

*Goal:* Design a new lower control arm that weighs less than the previous model to help improve performance. The previous suspension arm had a mass of **1850.729 g** and was made of **High Strength Low Alloy Steel**.

*Constraints:* The design must fit within the existing assembly. Your design will be assessed for quality by its weight, Von Mises strength, and material cost.

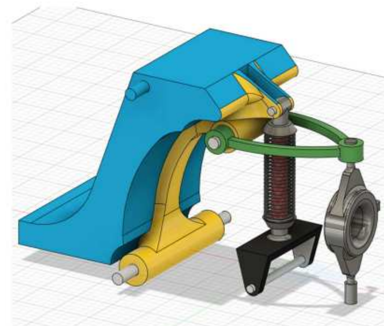

*Figure 1:* Existing Suspension Assembly

Fig. 4. Example Task Prompt

| Objective Constraints | | Yes (1) / No (0) |
|---|---|---|
| *Mass Objective* Did the part successfully reduce mass from the previous iteration? | | 1 |
| *Assembly Fit* Does the part fit within the existing assembly with no interference? | | 1 |
| *Min. Factor of Safety* Does the part have a minimum factor of safety greater than 3.0? | | 1 |
| *Material Selection* Is the part correctly assigned a material from the given list? | | 1 |
| **Design Quality** | | Exceeds (2) / Meets (1) / Does not meet (0) |
| *Finish Quality* Does the design appear complete? | | 2 |
| *Visual Interest* Does the part have a unique design that enhances the visual interest of the assembly? | | 2 |
| **Metric Rankings** | Value | Percentile Rank |
| *Part Mass* | 1158 g | 0.316 |
| *Min. Factor of Safety* | 15 | 0 |
| *Max. Von Mises Stress* | 0.72 MPa | 1 |
| *Strength/Weight Ratio* | 11.99 | 0.789 |
| *Material Cost* | $2.81 | 0.368 |
| **Total Quality Score** | | **10.473** |

Fig. 3. Example Rubric