Exploring Challenges and Opportunities to Support Designers in Learning to Co-create with AI-based Manufacturing Design Tools

Frederic Gmeiner
HCI Institute
Carnegie Mellon University
Pittsburgh, PA, USA
gmeiner@cmu.edu

Humphrey Yang
HCI Institute
Carnegie Mellon University
Pittsburgh, PA, USA
hanliny@cs.cmu.edu

Lining Yao
HCI Institute
Carnegie Mellon University
Pittsburgh, PA, USA
liningy@andrew.cmu.edu

Kenneth Holstein
HCI Institute
Carnegie Mellon University
Pittsburgh, PA, USA
kjholste@cs.cmu.edu

Nikolas Martelaro
HCI Institute
Carnegie Mellon University
Pittsburgh, PA, USA
nikmart@cmu.edu

ABSTRACT

AI-based design tools are proliferating in professional software to assist engineering and industrial designers in complex manufacturing and design tasks. These tools take on more agentic roles than traditional computer-aided design tools and are often portrayed as "co-creators." Yet, working effectively with such systems requires different skills than working with complex CAD tools alone. To date, we know little about how engineering designers learn to work with AI-based design tools. In this study, we observed trained designers as they learned to work with two AI-based tools on a realistic design task. We find that designers face many challenges in learning to effectively co-create with current systems, including challenges in understanding and adjusting AI outputs and in communicating their design goals. Based on our findings, we highlight several design opportunities to better support designer-AI co-creation.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**;
• **Applied computing** → **Computer-aided design**.

KEYWORDS

computational co-creation, generative AI, human-AI collaboration, think-aloud study, group cognition, team learning

1 INTRODUCTION

Modern manufacturing processes allow designers to produce complex parts optimizing strength-to-weight or leveraging new materials such as shape-changing plastics, yet their creation often surpasses the designers' cognitive capabilities. Recently, computer-aided design (CAD) tools have begun incorporating AI-based features to generate part designs based on a designer's myriad optimization goals [54]. For example, Japanese electric vehicle manufacturer WHILL used Autodesk Fusion 360 Generative Design to optimize material economy, strength, and sustainability for an electric wheelchair component [27]. AI tools are also helping designers generate parts using emerging manufacturing processes such as shape-changing smart material structures [80]. In both examples, working with AI allows designers to create designs that would be extremely tedious or infeasible without AI support.

These generative AI tools take more agency and autonomy in parts of the human-AI design process and are often referred to as "co-creators" [18]. However, effective and practical co-creation presents a significant learning curve for designers, as they are required to work and think collaboratively with AI agents that operate differently than human collaborators or complex CAD tools.

Studies have shown that effectively working with professional feature-rich *non-AI* design software already requires substantial and continual learning as such software becomes more capable [38, 49]. Consequently, instead of simplifying the software's interfaces, an active field within HCI studies how the learning of such complex software systems can be better supported through interactive interfaces [24, 44, 51]. However, working with AI "co-creators" is different from working with conventional CAD tools. Designers do not directly manipulate 3D geometry but rather formulate design goals for the AI system to build from. Yet, little is known about how to support designers in learning to work with AI tools that take on this more active and collaborative role.

In this study, we ask how designers can be better supported in learning to co-create with AI design tools. We guide our investigations by a recent wave of HCI literature that looks to inform the design of human-AI collaboration based on the mechanisms that make *human-human* collaborations effective—such as *grounding in communication* or *shared mental models* [5, 6, 10, 31, 40, 77, 85].

Furthermore, we seek inspiration from *team learning* [79] which models what actions help people *learn* to collaborate effectively with each other.

To gain insights on improving human-AI design collaboration, we study how engineering and industrial designers without prior AI co-creation experience learn to work with AI CAD tools in the context of advanced manufacturing design tasks. We chose this domain because of its increasing design task complexity, for which designers often *require* AI assistance. Through our studies, we aim to generate insights to inform future support interfaces of AI design tools. In particular, we investigate the following research questions:

RQ1a *What challenges do designers face when learning to co-create with computational AI tools?*

RQ1b *How do designers overcome these challenges?*

To answer these questions, we conducted a series of think-aloud studies observing how trained engineering and architectural designers (tried to) learn to co-create with two different computational AI tools on complex manufacturing design tasks (Study 1). Based on analyses of their interactions with the systems and retrospective interviews, we found that they generally valued the AI's assistance but faced challenges in learning to effectively co-create with the tools and interpret the design outputs. Those who were able to produce feasible and satisfying designs learned to co-create with the tool by systematically testing the boundaries of its capabilities early on, by self-explaining AI behaviors they observed, and by sketching and reflecting on design issues.

After learning about these challenges, we then explored how designers could be supported to better co-create by asking:

RQ2 *What are effective strategies to support designers in learning to co-create with computational AI tools?*

To answer this question, we took inspiration from prior work on human-human collaboration [9, 53, 68]. We conducted a human-human collaboration study to see how human guides would assist new users of AI tools in learning to co-create and how the new users learned with human assistance (Study 2). The observed effective support strategies included providing step-by-step instructions, prompting design reflection, and suggesting alternative strategies and goals for the design task. We also observed that the human guides relied heavily on multi-modal communication (e.g., screen annotations and mouse gesturing) to communicate more effectively with designers.

Lastly, to inform design opportunities for new support tools we asked:

RQ3 *What are designers' needs and expectations for human-AI co-creation?*

Synthesizing the results from both studies, we learned that many participants felt unable to communicate their design goals with the AI and wished for more conversational interactions and contextual awareness from the tool. We discuss potential support implications and future work from these needs and expectations.

In short, this study makes three main contributions:

(1) providing a set of observed challenges that engineering and architectural designers face when learning to collaborate with AI on complex co-creation tasks in the context of advanced manufacturing design;

(2) advancing our understanding of designers' needs and expectations for human-AI co-creative tools;

(3) highlighting design opportunities to better support designers in learning to co-create.

2 RELATED WORK

2.1 AI-based design tools for manufacturing

AI-based design support tools use various computational methods for generating 2D and 3D design options based on constraints and objectives set by designers [1, 72]. In 3D architectural, industrial, and mechanical design, new generative design tools have helped designers create consumer goods [43], building layouts [60], and lightweight automotive and airplane components [59, 61]. In the context of emerging advanced materials, AI design tools assist designers in creating structures out of shape-changing or elasticity-changing materials [29, 80]. Such AI tools use multiple techniques to generate designs from a set of goals and requirements, including constraint-based solvers [73], style transfer [2, 28], simulation and optimization [80], and genetic algorithms [59]. Such techniques are becoming commercially available in 3D CAD design tools such as Siemens NX, Solidworks, and Autodesk Fusion360 [3, 17, 69]. Many of these tools operate as black boxes where designers first set objectives and then review generated designs. However, this interaction can make it hard for designers to quickly develop a mental model of how the tool works, limiting their creative use.

Recent research has developed generative design interfaces for interactively exploring multiple design options [39, 52, 83] or more iterative engagement between the designer and the tool through real-time design generation and assessment [14, 19, 37]. However, few empirical studies exist that evaluate how engineers and designers *learn* to work with AI design tools on realistic tasks. Some existing work has measured the performance impact of AI agents on human engineering teams [84] while other work has investigated what role professional makers expect for involving AI in digital fabrication workflows [81]. This study provides empirical observations on how engineering, industrial and architectural designers learn to work and co-create with computational AI-based design tools.

2.2 Learning complex software

Prior HCI research has looked to evaluate and improve the learnability of complex software systems. Past studies explore how people of different ages learn a feature-rich notetaking tool [49], how professional engineers learn 3D design software [38], or how casual designers learn professional motion graphics software [33]. Often, people learn by searching web forums or asking knowledgeable colleagues for help [38]. Research on interfaces to support people in learning complex software has proposed dynamic feedforward tool tips [44], guided tutorial systems [24], and widgets that support self-directed trial and error learning [51].

While prior work has explored lenses such as self-directed learning for working with complex software [13], it remains an open question of how best to support self-directed learning for co-creation with AI systems that take an active role in the design process. For example, prior studies in Human-AI collaboration show that the black box nature of AI systems introduces new challenges where users grapple with non-transparent and non-intuitive

system behavior, hindering coordination and communication when completing "collaborative" tasks [10]. To address these issues, various strategies like explainable AI or intelligibility are aimed at helping users refine their mental models of AI systems [55, 70].

However, even as AI models become more intuitive for users, we expect that there will always remain a need for learning to work effectively with AI systems to, for example, develop *shared mental models* [35, 40]. Consequently, as we discuss in the next section, supporting humans in learning to effectively co-create with AI requires bringing in additional theoretical lenses.

2.3 Human-human collaboration as a lens for studying co-creative systems

To design effective human–AI collaboration, researchers have suggested drawing lessons from studying what makes human–human collaboration effective [5, 6, 10, 31, 40, 77, 78, 85]. While it remains an open question to what extent scaffolds for human-AI collaboration should mirror the designs of supports for human–human collaboration [77, 85], human-AI interaction researchers suggest that theories and findings from psychology, education, and the learning sciences are currently underutilized. For instance, Koch and Oulasvirta [40] note that *group cognition*—the study of how agents relate to other agents' decisions, abilities, beliefs, common goals, and understandings—provides powerful concepts for work on human–AI collaboration, yet is rarely referenced within this literature. Group cognition comprises phenomena such as *grounding in communication* [16] (creating mutual sense through verbal and non-verbal communication) and *theory of mind* [22] (the ability of agents to be aware of their own and the other's beliefs, intentions, knowledge, or perspectives). Similarly, Kaur et al. [35] argue that like human-human collaboration, effective collaborations between humans and AI may require *shared mental models* between people and the AI to enable mechanisms such as adaptive coordination of actions among team members [16, 56]. These may include shared representations of the task to be accomplished, of each other's abilities and limitations, or of each other's goals and strategies [20, 25, 66, 74]. A line of work addressing these opportunities has begun to explore how humans might be supported in developing and maintaining more accurate mental models of an AI collaborator's capabilities and limitations [5, 6, 41]. However, compared to concepts of human-human collaboration, honing only users' mental models is not sufficient enough for effective collaboration, which requires *shared mental models* between the user and system [35].

To date, little work has explored how best to support humans in learning to collaborate with AI on authentic tasks, such as design tasks, despite growing recognition of the need for such supports [10, 45, 50, 85]. Design tasks represent compelling challenges for human–AI collaboration, given that design problems are often ill-defined and require teams to navigate and negotiate both the problem and solution space [21] through an iterative process of generating ideas, building prototypes, and testing [32].

In this study, we investigate human-AI collaboration for emerging manufacturing design tasks—an area where successful task performance sometimes requires human–AI collaboration, yet where effective collaboration may be challenging to achieve without strong supports [45, 85].

2.4 Team learning

While phenomena such as *grounding in communication, theory of mind* and *shared mental models* provide useful concepts to explain which cognitive and social phenomena enable collaboration among a group of agents [35, 40], these theories do not explain how groups of individual agents *learn* to effectively collaborate. To address this gap, *team learning* emerged to study what actions and conditions contribute to how human groups learn to effectively collaborate together [79]. For example, team learning studies suggest that the development of effective shared mental models is supported through an *active process of negotiation* between team members, involving "*constructive*" forms of *conflict, argumentation*, and *resolution* [30, 34, 74]. However, to date, team learning has been under-utilized as a lens for studying human-AI co-creation. In this work, we draw upon concepts from team learning, such as *active processes of communication, joint information processing*, and *coordination of actions* [53], to study what actions and support strategies can help designers learn to co-create with AI-based design tools.

3 STUDY 1: THINK-ALOUD DESIGN SESSIONS

We conducted a series of think-aloud studies [75] with trained designers new to working with AI, where they worked with an AI design tool to complete a realistic advanced manufacturing design challenge. Think-aloud studies have people verbalize their thoughts while performing a task so that researchers can understand their cognitive processes (e.g., forming mental models [15], learning [82]). In our study, we use the think-aloud method to see how AI novices, who encounter a real learning challenge and are less biased than experts, learn to co-create with the AI tools. Participants first completed a 30-minute moderated think-aloud session where a member of the research team observed them working and listened to what they said they were thinking and doing while working. Half of the participants had a mechanical engineering background and designed a light and strong mounting bracket for a ship engine while considering the optimal manufacturing method and material combination using Autodesk's Fusion360 Generative Design (based on topology optimization that generates multiple options) [52]. The other half of the participants with a background in architecture or industrial design designed a bike bottle holder made with shape-changing materials—a complex design task that is challenging to complete without computational support. Designers working on the bottle holder task worked with SimuLearn [80], a machine learning-based research tool built on top of Rhino3d that helps designers create structures from shape-changing materials.

After completing the design task, participants submitted their designs and joined a semi-structured interview to reflect on their experience of working with the design tools. Across the study, we collect the following data:

- Video and audio recordings and machine-generated transcripts of the open-ended think-aloud design sessions
- Audio recordings and machine-generated transcripts of the post-task interviews
- 3D designs created during the think-aloud sessions

Table 1: Overview of study 1 participants. P-F11* and P-S11* are the same participant who had experience in using both Fusion360 and SimuLearn. Explanation for gaps in participant IDs: Some participants had dropped out after the first design session, or participants were assigned to Study 2 (see Section 6).

ID	Group	Age Years	Gender	Domain	Occupation	CAD Exp. Years	Indus. Exp. Years
P-F01	Fusion360	27	M	Civil & Environ. Engin.	Student / MA	>5	2 − 5
P-F02	Fusion360	27	M	Mechanical Engineering	Student / PhD	>5	2 − 5
P-F03	Fusion360	25	M	Mechanical Engineering	Student / PhD	>5	2 − 5
P-F04	Fusion360	26	M	Mechanical Engineering	Student / MA	>5	1 − 2
P-F05	Fusion360	19	M	Architecture, Mathematics	Student / BA	2 − 4	0
P-F11*	Fusion360	64	M	Mechanical Engineering	Contractor	>10	>30
P-F12	Fusion360	59	M	Mechanical Engineering	Designer	>10	>30
P-S01	SimuLearn	21	F	Architecture	Student / BA	2 − 4	0
P-S03	SimuLearn	23	F	Computational Design	Student / MA	2 − 4	0
P-S04	SimuLearn	21	M	Architecture	Student / BA	>5	1 − 2
P-S06	SimuLearn	23	F	Architecture	Student / MA	>5	<1
P-S07	SimuLearn	23	M	Architecture	Student / MA	>5	0
P-S10	SimuLearn	33	F	Industrial Design	Researcher	>5	6 − 10
P-S11*	SimuLearn	64	M	Mechanical Engineering	Contractor	>10	>30

3.1 Participants

We recruited 14 designers (4 female / 10 male, aged 19 to 64 *(M = 32.5, SD = 16.6))* with backgrounds in Architecture, Industrial Design, or Mechanical Engineering (Table 1). Most participants were recruited from our institution's student body, but we also recruited three professional designers via an online forum for designers who work with Fusion360 [64]. Participants had a minimum of two years of experience using CAD (Fusion360 or Rhino 3D) but no experience working with the studied AI design tools, determined via a screening questionnaire. Participants included mostly undergraduate and Ph.D. students and three engineers with > 30 years of industry experience (Table 1). We recruited participants familiar with either Fusion360 or Rhino3d so that they could focus on learning to work with the AI co-creation features rather than on learning the CAD tool's user interface. Before the study, all participants signed a consent form approved by our institution's IRB *(STUDY2021_00000202)*. Participants were paid 20 USD per hour.

3.2 Study context: AI-based design tools and tasks

To gather generalizable insights into designers' challenges, needs, and expectations around designer-AI co-creation, we observed how designers tried to learn to co-create with *two different computational AI tools* for digital manufacturing tasks. Participants were given a non-trivial, realistic design task to work on during the study. We selected a mechanical engineering design task concerning multi-dimensional optimization and an industrial design task exploring the use case of shape-changing materials. Both tasks are too complex to accomplish without AI and also have functional AI tools already developed. We collaborated with engineers and advanced manufacturing experts to identify and pilot the tasks to ensure they were adequately complex but not overwhelming for our target population. The tasks required participants to generate design solutions within a few hours over multiple sessions.

3.2.1 ***Mechanical design support tool****.* Mechanical designers worked with the "Generative Design" feature of Autodesk Fusion360[3], which helps designers to create lightweight and strong parts through topology optimization and genetic algorithms. In this task (Figure 1A-D), the designer is asked to design a material-efficient and structurally-sound engine mounting bracket by considering the optimal manufacturing and material combination from a large pool of possibilities. While designing mounting brackets is common for mechanical engineers, optimizing designs for different manufacturing methods and materials is difficult without simulation and AI support. Traditionally, engineers would first build a part and then gradually remove or add material based on structural analysis to derive a weight-optimized part. Exploring different manufacturing options would be necessary for every material and manufacturing constellation—which is time-consuming and tedious. In contrast, Generative Design can automatically generate many different design options based on specified high-level requirements, which the designer can explore and choose from.

Participants were provided a starter file containing the geometric constraints and needed to specify the mechanical design criteria (e.g., loads, bolt connection clearance, boundary condition). Participants then ran the solver and evaluated the AI-generated solutions to identify three viable designs for submission (Figure 1E). If none of the outcomes were deemed satisfactory, the user might choose to iterate the design by adjusting the input criteria.

3.2.2 ***Industrial design support tool****.* Industrial and architectural designers worked with SimuLearn [80], a research system built on Rhino3D that uses ML-driven simulation and optimization to enable designers to rapidly create objects out of shape-changing materials. This manufacturing process creates 3D-printed flat grids out of PLA plastic that can transform into a volumetric shape when heated. The transformation and the resulting shape can be controlled by tuning the grid geometry and the portion of the active transformation element (i.e., actuator ratio) within the beams. This
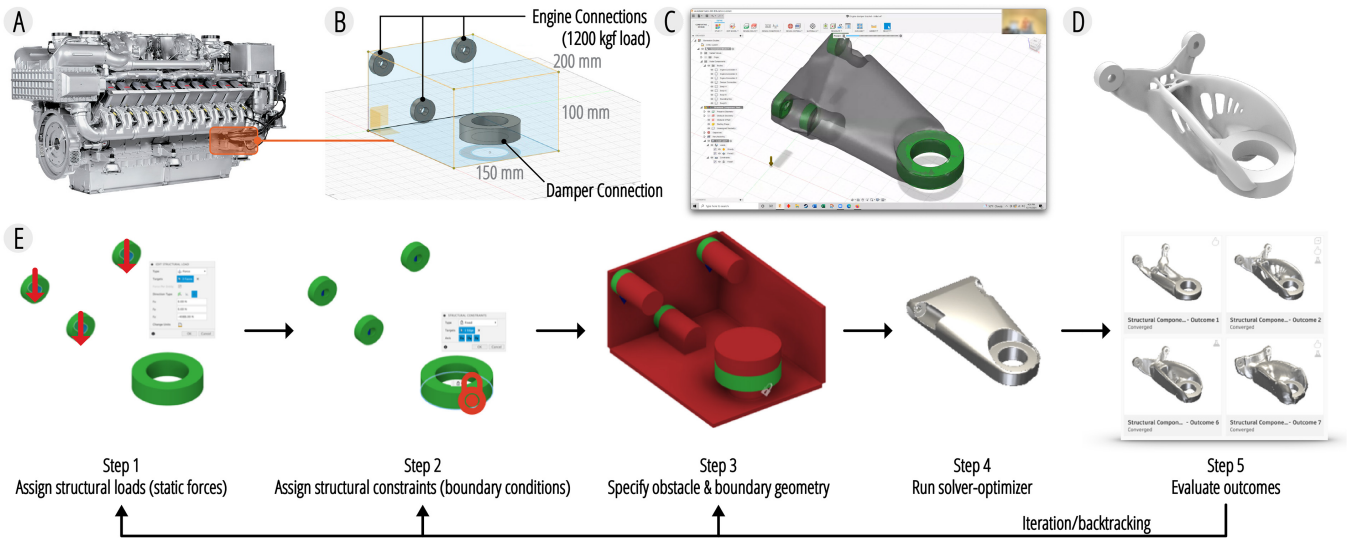
Figure 1: The Fusion 360 design task and workflow. The task involves (A) designing an engine bracket that connects the engine to a damper. (B) A starter file containing connection holes and bounding dimensions is provided to the users to initiate the design in (C) Fusions 360. The user is prompted to create (D) a viable design while minimizing weight. (E) The workflow involves five steps, and based on the AI system's solutions, the user may iterate the design by adjusting the design constraints and criteria to produce new solutions. (Image A: © Rolls-Royce Solutions America Inc.)

technique is envisioned to reduce production waste and shipping costs (e.g., flat packaging, reduced support material).

While new materials are being developed to manufacture shape-changing structures, designing shape-changing components poses unique challenges that designers are ill-equipped to handle. Unlike 2D and 3D design, designing with shape-changing materials involves a non-intuitive mismatch between the final target (3D shapes) and the design input (often 2D). Effectively designing for shape-changing materials requires an understanding of (often complex) spatiotemporal, self-assembling material behaviors that may push against the limits of what humans can mentally simulate. At the low level, designing such materials requires modifying volumes voxel-by-voxel, which is infeasible for complex structures if done manually. AI-driven tools allow designers to create complex artifacts that would otherwise be impossible to create by hand. SimuLearn aids the design process by providing real-time simulation and optimization to iterate designs toward the desired morphing behavior.

In this task (Figure 2A-E), participants design a bike bottle holder using morphing grids. A starter file containing the bottle and bike frame geometry was provided to contextualize the design. This task was more open-ended than the mechanical engineering task since each designer may assemble the morphed grids in different ways to create the holder. To use the tool (Figure 2F), the designer models the grid geometry and assigns bending actuator ratios. Next, participants simulate the design, observe the predicted transformation, and iterate the design by changing the grid model and actuator assignment. Alternatively, participants may opt to use functions to optimize the grids toward a targeted transformed shape. The optimization process can be either autonomous or interactive (i.e., the tool suggests edits for the user to choose from). To effectively

work with the tool, users need to learn to work with the different levels of AI assistance to produce a satisfactory design iteratively.

3.3 Choosing design tools

We specifically study these systems for two reasons. First, while both tools support advanced manufacturing tasks, they represent computational systems with distinct purposes and interaction paradigms. Fusion360's Generative Design module assists engineering designers with the generally familiar task of creating light and structural solid parts. The AI system helps designers to navigate a large design space and explore opportunities while adhering to specified requirements and constraints. SimuLearn, on the other hand, supports designers in working with an emerging material and manufacturing process unfamiliar to most designers. SimuLearn's AI tool provides rapid simulations of the shape-changing material and offers different levels of design assistance—from manual, over interactive, to autonomous optimization/iteration.

Second, each tool represents a different interaction style and synchronicity. In Fusion360, users follow a structured sequence of steps to set up the parameters and acquire solutions. It may take a few hours to generate new solutions, and the designer may export the generated models at any time or iterate the design by adjusting the parameters and rerunning the solver. By contrast, SimuLearn's solver runs two to three magnitudes faster (5-180 seconds), and users interact with the system without a predefined workflow. Participants may also freely switch between the three levels of AI support at any point.
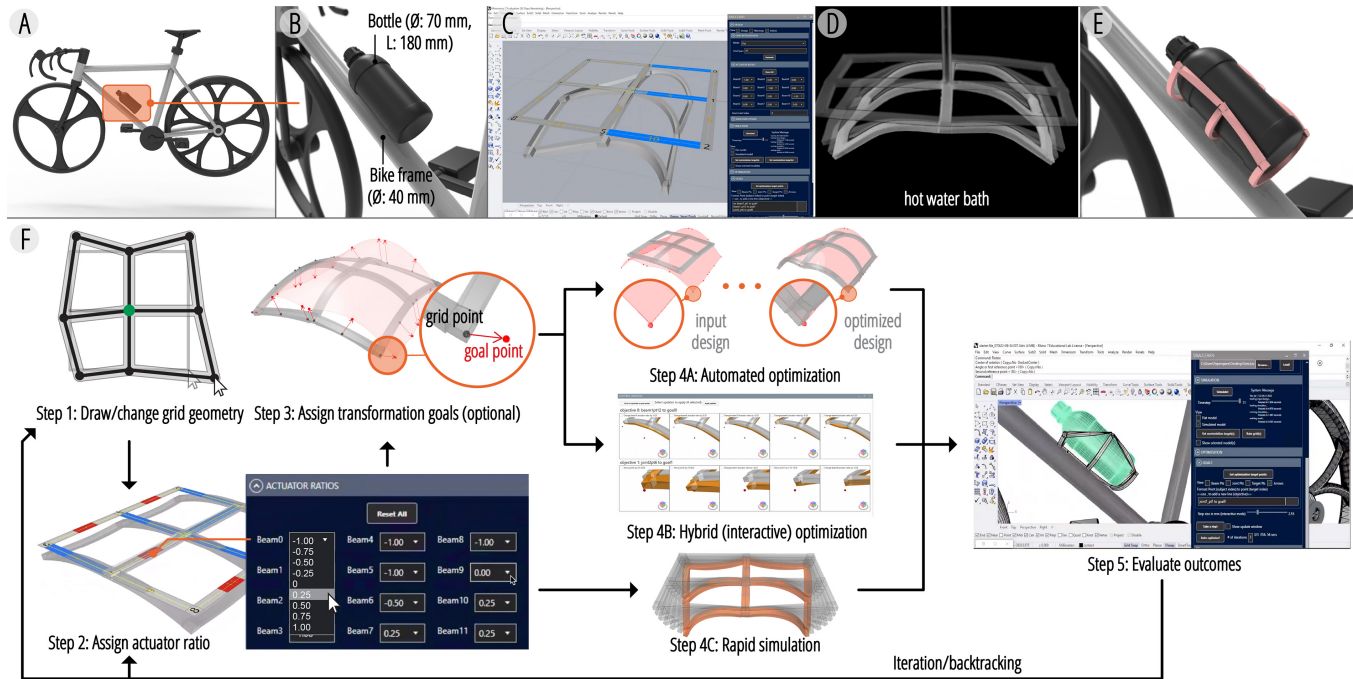
Figure 2: The SimuLearn design task. (A) The user is prompted to design a bike bottle holder using the provided (B) starter file and (C) the SimuLearn tool to create (D) morphing grid structures. (E) The transformed grid should assemble into the holder and fit around the bottle and bike frame. (F) The SimuLearn tool provides three types of design workflows. After initializing a morphing grid design, the user can choose between AI-assisted optimization functions or rapid simulations to iterate the design.

4 STUDY DESIGN

4.1 Study procedure

The study was structured into four phases (see Figure 3):

1) On-Boarding: Before the first session, participants received instructions by email on how to access the design tools running on a dedicated remote machine. They also watched a 25-minute

Figure 3: Overview of Study 1 think-aloud design sessions procedure. Participants were first introduced to the design tool and task, then worked while thinking aloud across multiple sessions. They completed an interview after turning in their design.

video tutorial demonstrating the tools' core functionalities with a step-by-step example.

2) Intro Design Session: At the beginning of the first session, a research team member introduced the participants to the study, design brief, task, and starter file. Sessions were conducted over video conference (Zoom) with audio and video recording. Participants then worked while sharing their screens and thinking aloud. Participants were allowed to use all available support resources, such as internal help files, external video tutorials, or user forums. The researcher quietly observed the participants setting the parameters of the computational design system and took notes. Due to the complex and open-ended nature of the tasks, the research team interrupted the task after 30 minutes and asked participants to continue working on their own as a compensated "homework assignment" without any time limit.

3) Homework Sessions: Following the initial session, participants continued working independently for as long as needed to iterate and produce the final design submission. Participants used the same remote machine and joined a personalized video conference meeting with automatic recording to document their work while thinking aloud. We also deployed a simple web application to prompt the users to verbalize their thoughts. The application analyzed the speech input from the microphone and reminded the user to "keep talking" after twenty seconds of silence. Once the designer felt satisfied with the result, they submitted the 3D files to our team.

4) Post-Task Interview: Within two days of submitting their result, participants completed a one-hour semi-structured interview with a research team member. Participants were asked to reflect
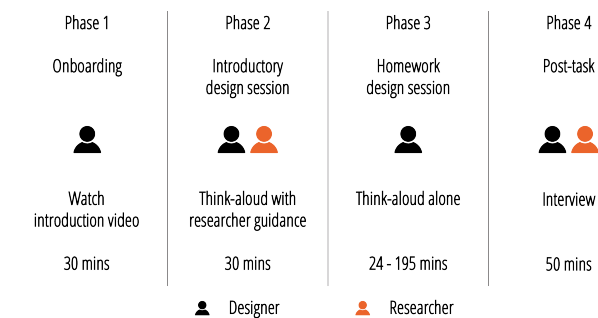
on their experience working with the AI tool. The interviews were conducted remotely over video conferencing. The interviewer took notes, and the interview audio and video were recorded. The interview protocol contained 36 questions clustered into three topics: *collaboration with the tool*, *design process*, and *learning process* (see the Appendix for interview protocol). The topics and questions clustered under *collaboration with the tool* were inspired by measures from team learning literature on assessing collaboration quality and effectiveness of human teams [9, 34, 53, 74]. These include perceived team roles and coordination, communication between user and tool, conflict resolution, timing, and (shared) mental models.

4.2 Analysis

To gain insight into research questions **RQ1a** *What challenges do designers face when learning to co-create with computational AI tools?* and **RQ1b** *How do designers overcome these challenges?* we (1) evaluated the design outcomes and analyzed more than 40 hours of think-aloud videos and 17 hours of interview recordings using (2) video interaction analysis of think-aloud videos, and (3) reflexive thematic analysis of think-aloud sessions (videos, transcripts) and interview transcripts.

4.2.1 **Evaluation of design outcomes**. We evaluated the effectiveness of Designer-AI collaboration by measuring the time required to complete the task and designer satisfaction with their results as rated on a three-point Likert scale (satisfied, neutral, unsatisfied) during the post-task interview. We also measured product feasibility for the mechanical engineering task by checking the designed engine brackets against the requirements in the design brief. The structural soundness was validated using finite element analysis (FEA), and the used material was checked by measuring part volume. We also checked the models for shape requirements (i.e., clear bolt holes, body within the bounding box). Since the bottle holder was a more free-form and aesthetic design task, we only checked if the user submitted their design and primarily relied on their self-reported satisfaction with the outcome.

4.2.2 **Video interaction analysis**. We used *video interaction analysis* [7] of the think-aloud recordings to understand participants' learning process while working with the AI features. To understand how well participants learned over time to use the AI features effectively, we tracked their interactions with the AI features relevant to the design task and documented whether the actions would produce satisfactory outcomes. For Fusion360, we tracked how participants specified structural loads, mechanical constraints, and the obstacle geometry feature to control the bracket's bolt clearance and overall size. For SimuLearn, we tracked how participants used different AI-assisted features (hybrid and automated optimization) throughout the think-aloud sessions.

4.2.3 **Reflexive thematic analysis**. To understand participant's challenges, needs, and expectations when learning to co-create with the AI system, we performed a *reflexive thematic analysis* [8] of the interview data (transcripts) and the think-aloud sessions (video, transcripts). We followed an iterative inductive coding process and generated themes through affinity diagramming. We used ATLAS.ti to analyze transcripts, audio, and video.

In the initial coding, the think-aloud and interview transcript data were equally distributed among two researchers who generated preliminary codes utilizing both a *semantic* (what people said) and *latent* (our interpretations of the data) coding strategy. Next, the research team collectively identified initial codes and themes. We generated themes in a bottom-up manner. However, we looked at the data with a mindset of collaboration between the designer and the tool—inspired by previous studies on human-human collaboration, co-creation, and team learning [9, 53]. We also tried our best to identify and separate usability issues from the codes and themes to avoid confoundment.

The two researchers then coded the think-aloud recordings to document where designers specified system parameters or evaluated design outcomes. These moments allowed us to find many of the problems that designers faced. We also coded non-verbal expressions like mouse gesturing or screen annotations that showed how designers attempted to communicate.

Finally, we created *summary videos* highlighting specific situations related to co-creation with the tool (e.g., designers confused by AI-generated outcomes). The video clips were annotated with a time code, participant ID, and a contextual description of the situation to share and discuss with the entire research team (for an example, please see the video figure in the supplementary material). The research team collectively analyzed the think-aloud summary videos in a half-day session and discussed the themes. We completed the qualitative analysis by iteratively reviewing and revising codes and themes until we identified a stable network of coherent and rich themes.

Table 2: Evaluation of design outcomes for the engine bracket design task (left) and bottle holder design task (right). The designer's satisfaction with the outcome is rated with green=satisfied, yellow=neutral, red=unsatisfied. For the engine bracket task, meeting structural and shape requirement checks are rated as X=fail, check mark=pass.

ID	Durat. Min.	Num. Iter.	Des. Satisf.	Struct. Req.	Shape Req.	ID	Durat. Min.	Des. Satisf.
P-F01	61	3	☹	X	✔	P-S01	97	☹
P-F02	85	2	☺	X	X	P-S03	190	😐
P-F03	54	2	☹	X	X	P-S04	225	☹
P-F04	105	9	☹	X	X	P-S06	166	😐
P-F05	140	6	☺	✔	✔	P-S07	63	☹
P-F11	122	6	😐	X	X	P-S10	189	☺
P-F12	160	5	😐	X	X	P-S11	151	☹
	Engine Bracket Task						**Bottle Holder Task**	

5 RESULTS

Overall, participants expressed seeing potential value in the AI systems to support their design process—especially that the tools would enable them (at least in theory) to create and explore more complex designs in a shorter time than without AI. However, most faced unresolvable challenges in learning to effectively co-create with the tools. In the following subsections, we provide a brief overview of participants' performance on the design tasks and then explore

Figure 4: Overview of engine bracket designs created by participants. P-F05 met both the shape and structural requirements. P-F01 only met the shape requirements (the bracket has holes for bolts with sufficient clearance and respects the specified dimensions but is structurally too weak). All other brackets missed both requirements and were either too heavy, weak, larger than the specified dimensions, or had not enough bolt clearance. For simplicity, we only show one design option per participant. Please see the Appendix for all submitted design options.
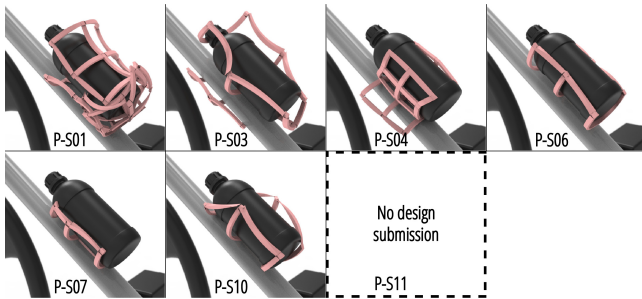

Figure 5: Overview of bottle holder designs created by participants. Participants came up with different designs using the shape-changing grid material in various ways. Participant P-S01 was not able to control the shape as intended and improvised by stacking grids together. While some designs (P-S06, P-S07) would not be able to hold a bottle, some designers (P-S04, P-S10) utilized the shape-changing grid structure to create promising bottle holder designs. Participant P-S11 stopped working on the task out of frustration and did not submit a design.

what challenges they faced in co-creating with them (RQ1a). We then highlight what some successful learning strategies looked like (RQ1b).

5.1 Design Tasks Outcome Summary
In the engine bracket task, all participants were familiar with designing similar mechanical components by considering forces and

constraints. Generally, such a task is a standard exercise in engineering education, and our task was comparable to the example provided by Autodesk in the introduction video that the participants watched. Participants required between 54 and 160 minutes *(M=104, SD=39.7)* to complete the task (see Table 2 left). No participant mentioned in the post-interview that the task itself was too difficult for them. Yet, only participant P-F05 was able to produce a self-satisfactory design that met both shape and structural requirements (see Figure 4 and Appendix for additional designs). P-F02 was also satisfied with their design but opted to manually refine the generated geometry that did not meet the requirements (i.e., using excessive materials and blocked bolt holes). We were surprised to find that few engineering participants produced satisfying results, even though they were familiar with the type of design task. Designers struggled to perform this otherwise familiar design task when they attempted to do so with AI assistance.

For our industrial or architectural designers, designing a bottle holder in 3D was not perceived as difficult. However, working with shape-changing material structures was new and everyone expressed in the post-interview that working with the shape-changing material was "unintuitive" and "challenging." Participants worked on the task between 63 and 255 minutes in total *(M=154, SD=56.6)*. All but one participant submitted a bottle holder design (Figure 5). This participant stopped working on the project after 151 minutes because he felt he could not control the material well, even with the AI. In the end, almost all designers were either dissatisfied with their final design or had a neutral opinion (see Table 2).

Table 3 shows that only a few designers understood how to set relevant parameters for the design task to produce satisfying results on the first attempt in both tasks. In Fusion360, most designers learned over time to successfully specify the structural constraints and obstacle geometry for bolt clearances. However, many designers failed to correctly specify structural loads and the boundary box. As a result, some submitted designs were too heavy, weak, or larger than the specified dimensions. In particular, one designer (P-F02) decided to manually refine the design after the first iteration because they felt more in control this way. For SimuLearn, all participants quickly learned how to control the grid shape using manual adjustments and rapid simulations (Table 3 right). However, many avoided using the other two AI-assisted features after an unsuccessful first attempt and continued to work manually (see Section 5.2.2 for further explanation).

5.2 Challenges designers faced when learning to co-create with AI (RQ1a)
We clustered challenges designers faced when learning to co-create into three themes: *Understanding and fixing AI outputs*, *working "collaboratively" with the tools*, and *communicating their design goals to the AI*.

5.2.1 Challenges in understanding and adjusting AI outputs.
Designers were often confused about the generated results and had difficulty understanding the cause and remedy of "weird" outcomes. This often occurred when generated solutions contained minor aesthetic flaws such as surface bumps, holes, or slightly twisted geometry. Designers wondered about the AI's underlying design rationale and were unable to determine whether such design features were intended or caused by algorithmic glitches.

Table 3: Schematic overview of each participant's learning process of design task-relevant features of the AI system. Left: Engine Bracket Task. Right: Bottle Holder Task. Correct use of input parameters or AI tools is shown with check marks. Increasing numbers of check marks from the first iteration to the last iteration suggest participants learning to work with the AI system.

ID	Structural Loads	Structural Constraints	Obstacle Geometry **Bolt Clearance**	Obstacle Geometry **Boundary Box**	Structural Loads	Structural Constraints	Obstacle Geometry **Bolt Clearance**	Obstacle Geometry **Boundary Box**	ID	Manual Actuator Setup	Auto Optimizer	Hybrid Optimizer	Manual Actuator Setup	Auto Optimizer	Hybrid Optimizer
	First Iteration				*Last Iteration*					*First Feature Usage*			*Last Feature Usage*		
P-F01	✗	✗	✗	✗	✓	✗	✓	✓	P-S01	✗	✓	✗	✓	•	•
P-F02	✗	✓	✗	•	Task completed without AI assistance				P-S03	✓	•	•	✓	•	•
P-F03	✗	✗	•	✗	✗	✗	✗	✗	P-S04	✓	✓	✓	✓	✓	✓
P-F04	✗	✓	•	•	✗	✓	✓	✗	P-S06	✗	✗	✗	✓	✓	•
P-F05	✗	✗	✗	✗	✓	✓	✓	✓	P-S07	✗	✓	✗	✓	•	•
P-F11	✗	✓	•	•	✗	✓	✓	•	P-S10	✓	✗	✓	✓	✓	✓
P-F12	✗	✗	•	•	✗	✓	•	•	P-S11	✓	✓	•	✓	•	•
	Engine Bracket Task									**Bottle Holder Task**					

• Feature not used
✗ Feature incorrectly used
✓ Feature correctly used

Furthermore, designers were unsure how to correct the generated designs (e.g., identifying the parameters that led to the problem). Some designers were hesitant to manually refine the generated geometry because they felt uncomfortable changing the optimized structure:

> *"...I also realized that if I was making any change in the mesh maybe I was changing the stress that that area will have. So I didn't feel very comfortable changing stuff because I knew that was optimized for my loads and my material."* P-F04

Designers were also unable to make sense of apparent structural issues in the generated designs, like when the generated parts were unreasonably thin or thick, provided little clearance, or produced confusing grid transformations. As verbalized by this designer while evaluating a generated engine bracket:

> *"it just doesn't make sense that, like this region here, which is obviously pretty thick... well I guess it's not. I just don't think that it can hold up the weight. It says the factor of safety is five. That just seems ridiculous to me."* P-F01 Think-aloud

At other moments, designers were uncertain whether the AI or the user was responsible for fixing the problem. In some cases, designers accepted imperfect results and attributed the flaws to the system. On one occasion, a designer verbalized their concerns when evaluating a generated engine bracket with insufficient bolt hole clearance:

> *"It's a goofy-looking bracket. I would never design it like that, but this thing thinks it can do it like that […] It just seems [that there needs] to be bolt heads and stuff, so I just don't see that bolt fitting in this area. But this is what it gave me, so I don't have a problem with that."* P-F13 Think-aloud

5.2.2 ***Challenges in working "collaboratively".*** Designers frequently expressed challenges in sharing control with AI-based design tools, noting that it often felt like these tools were dominating the design process. In those situations, designers either gave up and accepted unsatisfying results, improvised 'hacky' strategies to work around the AI or abandoned the AI assistance altogether and proceeded to work manually. For instance, one designer using SimuLearn expressed frustration about having insufficient control over the design process and decided to accept imperfect results:

> *"I feel like the collaborative process […] it seemed a little difficult to control. I felt that SimuLearn had more control over it than I did."* P-S01 Interview

Similarly, another participant pointed out the lack of participation opportunities and conversation throughout the design process:

> *"I would say no, that it's not co-creative. The whole program creates the thing but based on my limits. […] I don't feel like I interact in that creation of the shape. I just worked at the beginning and then I let the program do the rest. I missed that once there is one option, I cannot change anything, I cannot interact with the solutions that the program is giving me."* P-F04 Interview

In other cases, designers developed their own 'hacky' workarounds to attain feasible results. In Fusion360, this happened when designers tried to use simple boxes as obstacle geometries to limit the material growth within the specified perimeter. However, the algorithm often found ways to "squeeze" in additional, unwanted material through the gaps between obstacle geometries. A Fusion360 user expressed they felt like tricking the system when they were unsuccessful at blocking material generation:

> *"I think the software did kind of dominate the design process a little bit in that I was making things to satisfy the software instead of it kind of adapting to my needs*

[...] I think for something that's like supposed to be so smart and easy to use I wish it gave me more options instead of [me] trying to find these like little tricks. So I don't want to use little tricks. I don't want to use a hidden kind of 'I'll trick the software into thinking that it's correct.' P-F05 Interview

Finally, designers often avoided AI assistance when they could not learn to co-create with it effectively. This issue was particularly frequent when designers used SimuLearn, given that this tool explicitly allows designers to switch between different levels of AI assistance. When comparing which SimuLearn features designers used to complete the tasks, all designers learned over time to manually adjust, simulate, and iterate the design (see Table 3 right). However, only two designers (S04, S10) were able to work with the AI-assistance features (i.e., auto and hybrid optimizer) on the first try. All others expressed that they could not make sense of the AI output, even when they used the feature correctly. These participants subsequently avoided using the AI features and continued to work manually with rapid simulation.

5.2.3 Challenges in communicating design goals to the AI.
Designers often had difficulty communicating design goals to the AI system. For example, designers were unsure about the use and implication of certain parameters. Furthermore, many designers recognized their knowledge gaps of parameters that defined manufacturing processes or materials. E.g., a designer was unfamiliar with the specificity of a manufacturing method and kept using default parameter values:

"I'll have to see if they have five-axis milling... minimum tool diameter 10 millimeters... uh I don't know, I'm not really a tooling guy... tool shoulder length... [laughs] minimum tool diameter... all right I guess I don't know... I just hit okay on that." P-F12 Think-aloud

Designers often relied on the AI system's default settings or made assumptions about their effect when they were unsure about the parameters' meanings. Instead of seeking clarification from helpful resources, they often tried to determine a parameter's effect on the final result independently. However, it was oftentimes hard or impossible to notice and trace back parameter influences from the final results.

In the think-aloud sessions, we also observed several occasions where designers had different interpretations of the parameters. For example, in the Fusion360 task, half of the users made mistakes when converting the loads into the correct unit expected by the system. Similarly, when applying loads to multiple targets (i.e., bolt holes), the system applied the same load to each of the targets instead of equally distributing the load across the targets, which the designers had anticipated [1]. This mismatch led to higher load assignments and unnecessarily strong and bulky bracket designs. Interestingly, most designers verbalized their uncertainty about the load distribution when specifying parameters, as exemplified by this think-aloud comment:

"All the loads... I remember being a little wonky... so... I said three... let's see...what would be the case here? I don't know if all of these three forces are the same... that's the issue. I don't know if this is applying to each

one in particular... like if it's 12,000 here, here, and here... or if it's split evenly? I hope it's being split evenly... that's what I'm assuming." P-F03 Think-aloud

Although participants were aware that the system might interpret the load assignments differently, only a few were actually able to figure out and correct the mistake.

5.3 Learning strategies among successful designers (RQ1b)

Here we present findings related to *how designers overcame the previously reported learning challenges* **(RQ1b)**. We observed that all designers (after watching the introduction tutorial video) tried to learn to work with the tools through an iterative trial-and-error process. We also observed that participants sporadically consulted different support resources, including software tooltips and help files, and external resources like video tutorials, online user forums, and in some cases asking colleagues for help. Designers sought help from these support resources primarily after encountering interface or usability issues, which they often resolved. However, despite the available support resources, most designers struggled to learn to co-create with the tools effectively. Nonetheless, some designers employed successful strategies that helped them in learning to work better with the AI systems:

5.3.1 Systematically exploring AI's limitations and capabilities.
We observed that, early on in their interactions, two designers (P-F11, P-S10) deliberately and systematically experimented with the AI tools to develop a better intuition of the AI's behavior, capabilities, and limitations. These designers conducted tests to understand what effects different parameter values would have on the final result and documented the value-result correspondence to create a mental model. For example, PS-10 realized that their initial design sketches were not feasible with SimuLearn and the shape-changing materials, thus decided to systematically test different extreme grid shapes to hone their mental model of the AI's behavior:

"Even though I tried sketching some stuff, I think it just didn't work. So I thought it's better if I just go into the tool and see if I will be able to do this. I tried stuff like folding one corner upwards and one corner downwards or stuff like that. I took lots of screenshots and those really helped me to understand like 'if I do this, then it's gonna behave like that' so I think initially it was a lot of trying to form a mental model and like what's the capability of this tool." P-S10 Interview

5.3.2 Sketching, explaining and reflecting on design issues.
Another strategy that helped designers overcome flawed outcomes was to actively abstract and explain the problem. In the think-aloud sessions, we observed when facing similar fundamental challenges like over-constraining the engine bracket (such that the loads had no effect), some designers were able to understand and overcome the issue by sketching out the bracket's free-body diagram and explaining the acting forces and constraints to themselves. We observed similar strategies in SimuLearn, where participants understood the significant influence of gravity during the transformation process by explaining the process to themselves:

---

[1]This is a known issue in Fusion360 that many users have discussed in the user forum. See [4].

Table 4: Overview of Study 2: Learning with a Peer Guide participants. Six additional participants (3xFusion360, 3xSimuLearn) were paired with peer guides who had worked successfully with the AI design tools in Study 1.

ID	Group	Age Years	Gender	Domain	Occupation	CAD Exp. Years	Ind. Exp. Years	Guided by Participant
P-F07	Fusion360	27	M	Mechanical Engineering	Student / PhD	>5	1 − 2	P-F05
P-F08	Fusion360	27	M	Mechanical Engineering	Student / PhD	>5	1 − 2	P-F05
P-F13	Fusion360	19	M	Mechanical Engineering	Student / BA	2 − 4	0	P-F11
P-S05	SimuLearn	26	F	Architecture	Student / MA	>5	2 − 5	P-S04
P-S08	SimuLearn	20	F	Architecture	Student / BA	2 − 4	0	P-S04
P-S09	SimuLearn	23	F	Architecture	Student / BA	2 − 4	2 − 5	P-S04

"Whoa, I was not expecting that at all... uh... is that just because of gravity? And there's something crazy going on here... there's a lot of gravity... oh is it because I made the thing so big? Yikes, that is not at all what I expected."
P-S11 Think-aloud

6 STUDY 2: LEARNING WITH A PEER GUIDE (RQ2)

To gain insights into *how designers can be better supported in learning to co-create with computational AI tools* **(RQ2)**, we conducted additional think-aloud sessions where designers were paired with experienced peers to support them during the task (i.e., guided sessions). The guide had prior experience using the AI tools and provided support as needed to help participants to more effectively co-create with AI-driven features. Motivated by human-human collaboration, we aimed to derive insights into when and how to effectively support users in learning to co-create with AI by observing the support strategies, pedagogical moves, and communication patterns of the human guide.

6.1 Method

We recruited six additional participants (Table 4) following the same criteria as described in Section 3.1. The guides were recruited from the pool of participants who had completed Study 1 and demonstrated a thorough understanding of the domain, tool, and task (P-F05, P-F11, P-S04). We asked these guides to support the designers in learning to co-create with the tool. No script was provided to the guides because we intended to find possible support strategies from their natural interactions.

All guided sessions followed the same procedure as that of the unguided sessions, except that the homework sessions were limited to 50 minutes and the guides were present to help the designers. The designers and the guides communicated with each other through audio, screen sharing, and screen annotations. A researcher quietly observed and took notes. After the design session, we conducted separate 15-minute semi-structured interviews with the guide and designer. All sessions and interviews were recorded (video and audio) and automatically transcribed.

We conducted a reflexive thematic analysis to identify situations and themes on how peer guides supported designers to overcome challenges in learning to work with AI tools. We specifically focused on aspects of collaboration and knowledge transfer such as communication, joint information processing, and coordination

of actions [53]. We coded the video, think-aloud transcripts, and interview transcripts then generated themes with a focus on communication and learning by analyzing the guide's actions and support strategies that helped designers overcome challenges previously observed in Study 1.

6.2 Results – Guided Sessions

The guides supported the designer in using the tools, understanding the AI's behaviors, capabilities, and limitations, and sometimes suggesting and discussing alternative design goals or strategies. Guides primarily reacted to designers' verbalization and actions when they asked for help, expressed uncertainty, or when guides observed common mistakes. To get an impression of conversational dynamics, please see the video figure in the supplementary material. We portray five of the most common peer support strategies that helped designers to learn better to co-create with AI systems:

6.2.1 **Guide providing step-by-step walk-through instructions***. Guides often provided designers with step-by-step instructions for setting specific parameters. Such instructions were provided in response to designers' actions, such as showing confusion or struggle, but sometimes designers also specifically requested such assistance.

6.2.2 **Guide reacting to designers' expressions of uncertainty** *. We observed that guides were especially sensitive to moments when designers signaled uncertainty or when designers verbalized knowledge gaps with hedging expressions such as *"I don't know"*, *"maybe"*, or *"I assume."* In these situations, the guides often intervened and offered support or suggested alternative design strategies. Here, the guide suggested creating an obstacle geometry in Fusion360 to prevent material build-up at the bottom of the part in response to the designer who was wondering how to keep the part within the specified dimensions.

> DESIGNER: *"...but because the bolt is not here, I don't know where it would be..."*
> GUIDE: *"I mean maybe on the bottom face, right? That's what's resting on the body of the ship effectively."*
> DESIGNER: *Yeah, let's go ahead and choose that bottom face."*
> (Designer P-F13, Guide P-F11)

6.2.3 **Guide prompting designer reflection on generated designs***. We often observed situations where the guides prompted

feedback from designers each time SimuLearn's simulation or Fusion360's solver had finished. In those moments, peer guides often asked the designer *"Is that what you envisioned?"* or *"Is that what you wanted?"* or even *"Yeah, there you go! Is that how you want it to bend?"*. These prompts triggered designers to reflect on the generated designs, which helped the guides better understand how to provide support.

6.2.4 ***Guide suggesting alternative means and goals***. Beyond supporting tool operation and technical troubleshooting, the guides frequently suggested alternative means and goals to the designer, as in this dialogue:

> DESIGNER: *"It's going for the green part, but it's not able to figure out like a perfect way to get there without like touching this obstacle geometry."*
> GUIDE: *"Right, well I mean just make your obstacle geometry really long, say, all the way back past the connection to the ship, and then you're saying 'no material is allowed to go here' and that would make sure that you can always get a bolt in. See what I'm saying?"*
> DESIGNER: *"Yeah, I do"*
> (Designer P-F13, Guide P-F11)

Here the guide suggested an alternative way to achieve the designer's goal by enlarging the existing obstacle geometry. This strategy helped designers to better communicate design goals to the AI system and develop an intuition for harnessing the AI's capabilities.

6.2.5 ***Guide and designer making use of screen annotations and mouse gesturing to discuss design strategy***. An essential part of building understanding between the designer and the peer guide was through nonverbal communication, such as screen annotations, sketches, or mouse gestures. Guides frequently used the screen annotation feature built into Zoom to highlight elements they spoke about by circling or drawing arrows. We also observed that all designers naturally used the mouse cursor to emphasize design features through circling or pointing gestures when explaining something to the guide. Both behaviors are illustrated in this situation where the guide and the designer discussed a strategy to achieve a specific bottle holder shape:

> GUIDE: *"So in this case, maybe I suggest that you move these two points in particular more towards the center."*
> `[Guide draws arrows from points towards the center of the grid]`
> DESIGNER: *"And this one seems to have dropped downwards. Even this point here... The beam seems to be going downward."*
> `[Designer points with mouse at different beams]`
> GUIDE: *"oh this beam right here?"* `[Guide draws an arrow pointing at beam]`
> DESIGNER: *"Yeah."*
> (Designer: P-S08, Guide: P-S04)

Many designers and guides also annotated generated designs and sketched to clarify or illustrate their ideas. In summary, different forms of nonverbal communication helped designers and peer guides develop better shared mental models of the task and collectively overcome design issues.

7 DESIGNERS' NEEDS AND EXPECTATIONS FOR CO-CREATING WITH AI-BASED DESIGN TOOLS (RQ3)

Based on our observations and interviews from both the unguided and guided sessions, we highlight four themes that capture the *needs and expectations that designers expressed around co-creating with AI-based design tools* **(RQ3)**.

7.1 Designers expect the AI system to have more contextual awareness about the design problem at hand

In both systems, participants missed the kind of contextual awareness that a human collaborator might have about a design task, such as a part's function or how the part interfaces with other elements in the environment. Such a lack of contextual awareness was one of the main reasons people thought working with the tool was not collaborative. This lack of context also meant that the tool could not support the designers more proactively like a human partner, as described by this participant:

> *"Certainly it would have saved me some time if at the beginning the software would have said 'oh I see that these are your connection points. Can you actually get a bolt in there?' [...] Things like that would have felt really much more collaborative and helpful."* P-F11 Interview

Others expected the tool to offer more intelligent manufacturing and material suggestions or help them anticipate real-world design issues, as this participant expressed:

> *"I would like to see a design tool that would show me simulations of the water bottle in action, like 'oh is there enough friction' or 'will it actually stay in place while a cyclist is on the bike' and then provide suggestions of how to alleviate those problems."* P-S04 Interview

7.2 Designers desire a more conversational form of interaction with the tool

Most participants complained about the lack of reciprocal interaction between them and the tool. Participants compared designing with the AI systems to *"programming"* or *"working with a skilled teammate who is not listening to you."* Participants wished for a more conversational interaction with the tools, as desired by this designer:

> *"More like a tool that I can have a conversation with while I'm always sure that everything that I'm making is fulfilling the expectation of the piece and the loads and materials and everything."* P-F04 Interview

7.3 Designers require support in thinking through design problems

Across all guided sessions, designers appreciated that the peer guide helped them learn to operate the tool but also to *think through design problems*—a feature they would eventually also expect from a co-creative AI tool. As summarized by this designer who reflects on working with their peer guide:

> *"I think working together with the peer guide was actually really helpful because he had a lot of insight.*

I felt like a couple of things that I didn't really think about just from a fundamental engineering standpoint of how the thing can actually be made. I think that was really beneficial." P-F13 Interview

7.4 Designers expect an AI tool to provide project-relevant work examples

Designers frequently suggested that a co-creative tool should suggest similar task- and project-related examples created by other users. Such examples could help designers learn the system's capabilities and provide creative inspiration, as suggested by this SimuLearn user:

"Let's say I'm modeling some form and in real-time, the tool is searching some sort of database to show me some possibilities that other people have previously done, just by the similarity of the shape, and then I'm like 'yeah this shape can go there and go there.' So that's actually a creative input that can help—a bit more like you're designing with someone. It doesn't necessarily have to generate, it can pull up from other people and tell you 'here is how some other designers work with this.'" P-S10 Interview

8 DISCUSSION

Co-creative AI design tools have the potential to amplify the abilities of engineering and industrial designers. However, we found that designers face major challenges in (learning to) effectively co-create through understanding and adjusting system outputs, communicating design goals, and working "collaboratively" with the AI. Designers who overcame challenges did so by systematically exploring AI's limitations and capabilities, and by explaining and reflecting on their design issues. The observed support strategies of peer guides comprised step-by-step walk-through instructions, reacting to designers' expressions of uncertainty, prompting designer reflection on generated designs, suggesting alternative means and goals, and making use of screen annotations and mouse gestures to discuss design strategies. Overall, designers expected the AI system to have more contextual awareness about the design problem at hand, desired a more conversational form of interaction with the tool, asked for more support in thinking through design problems, and expected project-relevant work examples from the tool.

We discuss our findings in the context of prior research studying how users learn to work with complex software tools and relate our findings to concepts and empirical findings from research on human collaboration. We highlight design opportunities (Table 5) to better support designer-AI co-creation by scaffolding designers in actively exploring the boundaries of AI capabilities and limitations, prompting designers to actively reflect on design problems and observed AI behaviors, enhancing AI systems' contextual awareness of designers' tasks and objectives, and supporting more conversational forms of multi-modal communication between designers and AI systems.

8.1 New learning challenges for human–AI co-creation

Based on our findings, we believe that the challenges designers faced when learning to co-create with the tools go beyond learning the tools' interfaces. Firstly, all participants were experienced designers familiar with the CAD software's interface and had watched a step-by-step instructional video demonstrating the operation of the AI features. Furthermore, designers used different support resources, including tooltips, help files, video tutorials, online user forums, and asking colleagues for help to overcome interface issues they faced. However, even with the help of these resources, most designers could not produce feasible and satisfying outcomes with AI assistance. While this result may have been due, in part, to the limitations of the AI tools themselves, it is clear that this was not the full story. Indeed, a few of our participants were able to overcome challenges and learned to co-create effectively with the AI features. Hence we believe that the challenges participants faced when working with the AI systems (such as communicating design goals or understanding AI outputs) were partly *new learning challenges* due to the tool's more active role in the design process. These learning challenges go beyond learning the tool's interface. For example, a major challenge we identified relates to how designers specify all required parameters upfront instead of modeling and testing a part step-by-step. This workflow requires designers to think through the design problem in advance, which is challenging and different from the usual iterative design process. Our findings suggest that these new learning challenges require new support strategies for example, by prompting designers' reflection in response to expressions of uncertainty or suggesting alternative design goals that align with the AI's capabilities.

8.2 Toward models of human collaboration as lenses for studying and designing co-creative systems

Participants had trouble learning to predict how the AI might behave in response to the specified parameters. They struggled to make sense of the AI system's reasoning and struggled to correct unwanted design issues. Prior literature on group cognition suggests that to achieve effective collaboration group members should be able to interpret each other's reasoning and predict roughly how their partner might behave in response to their own actions [40]. Similarly, from a team learning perspective, our findings suggest that designers who systematically explored the AI's limitations and capabilities early on were better at predicting the tool's actions in response to their own and produced more satisfactory results. This result is in line with studies on human-AI collaboration in decision-making, suggesting that users learn to better predict the machine's behavior through inductive mechanisms (i.e., via concrete examples and hands-on testing) than via general, declarative information about internal processes [11]. While explainable AI research focuses primarily on directly communicating information about the AI system to the user, recent research has suggested that more engaging and longer forms of learning and deliberate practice might improve human-AI collaboration [36]. However, in addition to supporting honing the user's mental model of the AI's capabilities and limitations, it is equally important for the AI system to have an understanding of the user's capabilities, limitations, and task context to enable more effective human-AI collaboration. Hence, this would require the AI system to have better contextual awareness of the user and the current task at hand. We further discuss the resulting design opportunities in section 8.3.3.

Table 5: Overview of design opportunities and example applications in relation to group cognition and team learning.

Support opportunity	Example Application	Relation to Models of Human-Human Collaboration
Scaffolding inductive learning of AI's capabilities and limitations		
Offer guided interactive mini experiments for exploring AI behavior *Seen in 5.3.1 and 6.2.2*	*System offers designers explorative mini experiments to develop intuition of the AI's responses to specific (extreme) parameter constellations.*	Refining mental model of AI's behavior and capabilities through experiential, hands-on learning; Learning to predict the tool's actions in response to own actions.
Present similar work produced by AI design tool *Seen in 7.4*	*System proactively suggests task-related project examples of designers creations with AI system as learning resource and source of inspiration.*	Refining mental model of AI's capabilities and limitations by learning from examples; Aligning designer's expectations with tool's capabilities; Getting inspired by co-creative AI's suggestions.
Prompting designers' planning and reflection		
Support designers in thinking through design problems *Seen in 6.2.4 and 7.3*	*System asks designers questions to help reframe or further explore problem and solution space, like "Have you considered other alternatives?", "Could the bottle holder also be placed at a different location on the bike?" System could also suggest alternative strategies and goals to designer.*	Helping designers to stay cognitively engaged in co-creative design process with AI.
Prompt designers' feedback and reflection *Seen in 5.3.2 and 6.2.3*	*Prompting designers' feedback when evaluating generated designs, like "Is this design what you envisioned?" Based on the feedback, system could offer support guidance on how to correct generated designs.*	Helping designers to identify their own mental state and communicate goal to AI system; Information pooling.
Improving co-creative AI tools' contextual awareness of designers' tasks and objectives		
Equip AI with increased contextual awareness of designer, task, and project *Seen in 7.1*	*System probes designers to share information about their background, skills and project, like high level design goals and manufacturing considerations.*	Fostering shared mental models of task and objectives by asking for explicit feedback or implicit social inference.
Respond to designers' uncertainties and knowledge gaps *Seen in 6.2.2*	*Enable designers to communicate uncertainties about parameters and own knowledge gaps to system by "flagging" parameters. System could track designers' "level of certainty."*	Foster shared mental models of each other's knowledge and goals. Increasing the feeling of designers' ownership of the mutual activity with AI system.
Multi-modal interactive communication for human–AI co-creation		
Allow a more interactive and conversational interaction between designer and AI system *Seen in 7.2*	*Utilizing conversational agentive interfaces for designer-AI co-creation*	Facilitating grounding in communication through turn-taking, acknowledgments, back-channeling, error correction; Mutual confirmation of the reception and interpretation of objectives.
Enable non-verbal communication between designer and AI system *Seen in 6.2.5*	*Designers provide feedback to AI system by sketching design goals or annotating generated designs for refinement (like in a design critique)*	Facilitating grounding in communication by selecting a domain and task-appropriate communication medium.

Most designers felt the tools were uncollaborative and had more control over the design process than they would have preferred. As a consequence, they accepted imperfect results, developed improvised workarounds, or avoided AI assistance altogether. Previous studies on group cognition and team learning suggest that group members feel more ownership of the mutual activity when the group learns to coordinate cognitive capabilities among participants—united by their interpretations of each other's mental states [57]. Groups can reach this level of collaboration by an active process of communication, joint information processing, and coordination of actions [53].

Our findings from the guided sessions show that the guides' active communicative support strategies, such as reacting to designers' uncertainties and providing step-by-step instructions, helped designers learn to work confidently with the AI. Furthermore, when peer

guides prompted designers' feedback and reflection on generated designs, the designers were required to articulate their intentions explicitly. As a result, both the guides and designers were able to discuss and better coordinate further actions to improve the outcome.

8.3 Design Opportunities and Future Work

The following section highlights several design opportunities we identified in our findings to support designers in learning to co-create with computational AI tools.

*8.3.1 **Scaffolding inductive learning of AI's capabilities and limitations** .* As discussed, some designers learned the AI system's capabilities and limitations by testing and documenting the effectual correspondence between various parameters and the generated result (see 5.3.1). In the guided sessions, peer guides also frequently provided step-by-step instructions for setting up parameters or walked designers through the sequence of steps (see 6.2.2). These inductive learning strategies helped designers to better predict the AI's behavior and understand its capabilities. Previous work on novice-AI music co-creation has also found that users systematically tested AI limitations to hone their mental model of the system's behavior [48]. Going further, to better support users in this learning activity, future co-creative systems may offer designers a set of hands-on mini-guided 'experiments' to better understand the system's responses to specific (extreme) parameter inputs [42]. Systems may also offer designers opportunities to view sets of examples of input-output pairs to help designers develop useful mental models of an AI tool's generative capabilities and limitations (cf. [58]). A co-creative tool may also proactively recommend similar tasks and project-related examples created by other human-AI teams (see 7.4) to help designers learn the system's capabilities and provide creative inspiration.

*8.3.2 **Prompting designers' planning and reflection** .* Participants who were more successful at co-creating with the AI tools did so by abstracting and explaining their problems—either to themselves during think-aloud sessions or to the peer guides (see 5.3.2 and 6.2.3). Literature from the learning sciences shows that self-explanation positively affects understanding and problem-solving [76]. In addition, participants from the guided sessions appreciated the guides' prompt for reflection on AI behaviors and their suggestions for alternative design goals or strategies (see 6.2.4 and 7.3). Such actions helped designers think through the tasks and plan actions with AI tools. Conversely, designers who did not reflect on the design problems were unable to learn to understand the system's behavior, capabilities, and limitations well enough and failed to produce satisfactory outcomes. This observation was more prevalent in Fusion360 than in SimuLearn since Fusion360's long simulation time doesn't support the kind of rapid interactive adjustments as SimuLearn did, thus requiring the user to strategize their actions in advance. Hence, supporting users in thinking through the design problem for specifying parameters ahead would be especially beneficial for AI tools with longer processing time.

One additional explanation for why participants failed to produce satisfactory outcomes might be that AI systems can lead designers to over-rely on their support, creating an *"illusion of success"* that reduces their effort in solving the design problem [84], something we saw when designers accepted results even when they appeared

unfeasible (see Section 5.2.1). To compensate for this tendency, a co-creative system may help designers reframe the problem or further explore the solution space by suggesting alternative goals or asking *generative design questions* [23] like *"What could other alternatives look like?".* Moreover, actively reflecting on the design process is an essential part of professional design practices [67]. Much like the guides, a co-creative system could prompt feedback and active reflection on observed AI behaviors or generated designs by asking deep reasoning questions about the results, such as *"Is this generated design what you envisioned?"* (cf. [12]). Based on the feedback, the system may then offer support and help to coordinate further actions. This strategy would be complementary to the inductive learning support described in the previous section (8.3.1).

*8.3.3 **Improving co-creative AI tools' contextual awareness of designers' tasks and objectives**.* Our results show that designers felt a lack of the tool's awareness about the design context and therefore missed the kind of proactive support a human partner might provide (see 7.1). While building contextual awareness into AI systems has long been a tradition in HCI research, it also presents many technical challenges. However, in the context of co-creative design tools, promising directions are being explored. For example, the system could derive its user model through explicit and implicit mechanisms to develop a shared mental model of the context by asking the user for information about the specific design task (i.e., parsing a written design brief for context and goals [71]) or infer design goals from user behaviors [46].

From a group cognition and team learning perspective, contextual awareness also includes an understanding of the other team member's existing or missing knowledge about the design task. Based on our findings (see 6.2.2), a co-creative tool might learn about the designers' knowledge by responding to verbalized knowledge gaps such as *"I don't know"* or responding to hedging expressions such as *"maybe"* or *"I assume."* This observation is in line with research showing student learning is positively affected by human tutors' responses to their expressions of uncertainty [26]. Based on this phenomenon, literature has explored how intelligent tutoring systems can detect and respond to student hedging [26, 62]. A co-creative system may also allow designers to communicate uncertainties about parameters and their own knowledge gaps. For example, designers may flag an *"???"* checkbox next to a parameter's input field to signal uncertainty. The AI system could then track designers' "level of certainty" for each parameter and provide reactive help or re-surface those parameters later in the design process to identify possible reasons for unexpected outcomes.

*8.3.4 **Multi-modal interactive communication for human–AI co-creation**.* Our findings show that designers felt communicating with the AI was like giving instructions without receiving much feedback (see 7.2). Results from the guided sessions show that peer guides used conversational strategies like confirming the reception of each other's goals or asking for clarifications when they were unsure about the other's intentions. As conversation is widely seen as a vital mode of designing [67] and empirical work suggests that much of design work lies within conversations between collaborators and stakeholders [47], co-creative systems should consider using more back-and-forth conversation as an interaction interface. Furthermore, studies on team learning show that the forming of effective,

shared mental models is strengthened through an active process of iterative negotiation between team members, involving "constructive" forms of conflict, argumentation, and resolution [30, 34, 74]. Such a strategy may also prove useful for negotiating design goals.

Our findings from the guided sessions also suggest that non-verbal communication may support design partners in developing better shared mental models of design goals (see 6.2.5). Designers and guides discussed goals and strategies by pointing at features with the mouse cursor or sketching with the screen annotation feature, similar to a *'spatial-action language'* described by Schön [67] which explains gesturing and drawing along with verbal expressions as typical forms of communication in traditional design critique sessions. Such non-verbal interaction is still an underutilized medium in human-AI co-creation. Allowing designers to use sketching, annotation, or gesturing atop generated results may help them communicate design goals to an AI system.

8.4 Human–AI co-creation beyond manufacturing

Although we identify opportunities to support learning to co-create with AI systems in the context of manufacturing, many of our findings could also apply to other human-AI co-creation domains such as image, music, or text generation. Given recent advancements in generative image AI models (such as DALL-E [63] or Stable Diffusion [65]) with fast release cycles of new tools and capabilities, supporting creative professionals in learning to effectively co-create with such tools might become increasingly important. Furthermore, many prompt-based AI models like DALL-E expect users to express their goals through text prompts, which is an unfamiliar modality for most of today's visual designers. Consequently, our findings suggest that AI tool users across many domains could be supported in learning to better co-create with AI systems by scaffolding inductive learning of the AI's capabilities, prompting users' planning and reflection, improving the tools' contextual awareness of tasks and objectives, and facilitating multi-modal interactive communication between tool and user. Future work might further explore interfaces for supporting learning to co-create with computational AI tools in other domains beyond manufacturing.

8.5 Limitations of the study

We highlight three limitations of this study: first, our participants only represent a subset of engineering, architectural, and industrial designers. Although all participants had relevant training in their design fields and worked with 3D CAD software, most had minimal industry experience. We also included three professionals with substantial industry experience to compensate for this imbalance, however, even these participants struggled. Further, the self-selected participants in our study were presumably interested and open to the idea of co-creating with an AI system. Thus, some of our findings may be reflective of this openness to co-creative work. Second, both AI tools are still early in development, and we noticed user experience issues that could benefit from improved UIs. To our best extent, we isolated UI issues from more fundamental challenges in learning to co-create with AI systems. Although the tools are relatively new, we believe our findings provide value in exploring new opportunities for co-creative design systems. Finally, despite

our effort to make sure the design tasks were realistic, designers knew that they were in a research study and that the designs would not be manufactured. In a professional context, participants might have spent longer learning the tool to produce feasible designs.

9 CONCLUSION

In this paper, we presented an empirical study to understand how engineering and architectural designers learn to work with AI-based manufacturing design tools, documenting their challenges in working with an AI and probing their needs and expectations for co-creating with AI systems. We identified several support opportunities with an eye toward learning from effective human-human teams to improve future designer-AI co-creation. Overall, we aim to inspire others to explore untapped support opportunities and to work toward future co-creative tools that combine the strength of both human and AI systems to achieve complex designs that neither could achieve alone.

REFERENCES

[1] Grégoire Allaire, François Jouve, and Anca-Maria Toader. 2002. A level-set method for shape optimization. *Comptes Rendus Mathematique* 334, 12 (Jan. 2002), 1125–1130. https://doi.org/10.1016/S1631-073X(02)02412-3

[2] Gantugs Atarsaikhan, Brian Kenji Iwana, and Seiichi Uchida. 2020. Guided neural style transfer for shape stylization. *PLOS ONE* 15, 6 (June 2020), e0233489. https://doi.org/10.1371/journal.pone.0233489

[3] Autodesk. 2020. Fusion 360 Generative Design Extension. https://www.autodesk.com/products/fusion-360/generative-design-extension

[4] Autodesk. 2021. Generative Design - Structural Load - Force per Entity? https://forums.autodesk.com/t5/fusion-360-design-validate/generative-design-structural-load-force-per-entity/td-p/10380547 Section: Fusion 360 Design, Validate & Document.

[5] Gagan Bansal, Besmira Nushi, Ece Kamar, Walter S. Lasecki, Daniel S. Weld, and Eric Horvitz. 2019. Beyond Accuracy: The Role of Mental Models in Human-AI Team Performance. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing* 7, 1 (Oct. 2019), 2–11. https://ojs.aaai.org/index.php/HCOMP/article/view/5285

[6] Gagan Bansal, Besmira Nushi, Ece Kamar, Daniel S. Weld, Walter S. Lasecki, and Eric Horvitz. 2019. Updates in Human-AI Teams: Understanding and Addressing the Performance/Compatibility Tradeoff. *Proceedings of the AAAI Conference on Artificial Intelligence* 33 (July 2019), 2429–2437. https://doi.org/10.1609/aaai.v33i01.33012429

[7] Eric P.S. Baumer and Bill Tomlinson. 2011. Comparing Activity Theory with Distributed Cognition for Video Analysis: Beyond "Kicking the Tires". In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 133–142. https://doi.org/10.1145/1978942.1978962 event-place: Vancouver, BC, Canada.

[8] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 11, 4 (Aug. 2019), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

[9] Jean-Marie Burkhardt, Françoise Détienne, Anne-Marie Hébert, Laurence Perron, Stéphane safin, and Pierre Leclercq. 2009. An Approach to Assess the Quality of Collaboration in Technology-Mediated Design Situations. In *European Conference on Cognitive Ergonomics: Designing beyond the Product — Understanding Activity and User Experience in Ubiquitous Environments (ECCE '09)*. VTT Technical Research Centre of Finland, FI-02044 VTT, FIN. event-place: Helsinki, Finland.

[10] Carrie J. Cai, Samantha Winter, David Steiner, Lauren Wilcox, and Michael Terry. 2019. "Hello AI": Uncovering the Onboarding Needs of Medical Practitioners for Human-AI Collaborative Decision-Making. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 1–24. https://doi.org/10.1145/3359206

[11] Arjun Chandrasekaran, Deshraj Yadav, Prithvijit Chattopadhyay, Viraj Prabhu, and Devi Parikh. 2017. It Takes Two to Tango: Towards Theory of AI's Mind. http://arxiv.org/abs/1704.00717 arXiv:1704.00717 [cs].

[12] Catherine C. Chase, Jenna Marks, Laura J. Malkiewich, and Helena Connolly. 2019. How teacher talk guidance during Invention activities shapes students' cognitive engagement and transfer. *International Journal of STEM Education* 6, 1 (Dec. 2019), 14. https://doi.org/10.1186/s40594-019-0170-7

[13] Rimika Chaudhury, Philip J. Guo, and Parmit K. Chilana. 2022. "There's no way to keep up!": Diverse Motivations and Challenges Faced by Informal Learners of ML. In *2022 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. 1–11. https://doi.org/10.1109/VL/HCC53370.2022.9833100 ISSN: 1943-6106.

[14] Xiang 'Anthony' Chen, Ye Tao, Guanyun Wang, Runchang Kang, Tovi Grossman, Stelian Coros, and Scott E. Hudson. 2018. Forte: User-Driven Generative Design. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–12. https://doi.org/10.1145/3173574.3174070

[15] Michelene T. H. Chi. 2006. Laboratory Methods for Assessing Experts' and Novices' Knowledge. In *The Cambridge Handbook of Expertise and Expert Performance*, K. Anders Ericsson, Neil Charness, Paul J. Feltovich, and Robert R. Hoffman (Eds.). Cambridge University Press, Cambridge, 167–184. https://doi.org/10.1017/CBO9780511816796.010

[16] Herbert H. Clark and Susan E. Brennan. 1991. Grounding in communication. In *Perspectives on socially shared cognition.*, Lauren B. Resnick, John M. Levine, and Stephanie D. Teasley (Eds.). American Psychological Association, Washington, 127–149. https://doi.org/10.1037/10096-006

[17] Dassault Systèmes. 2021. Solidworks | Enhance Your Design Workflow with Generative Design. https://www.solidworks.com/media/enhance-your-design-workflow-generative-design

[18] Nicholas Davis, Chih-Pin Hsiao, Yanna Popova, and Brian Magerko. 2015. An Enactive Model of Creativity for Computational Collaboration and Co-creation. In *Creativity in the Digital Age*, Nelson Zagalo and Pedro Branco (Eds.). Springer London, London, 109–133. https://doi.org/10.1007/978-1-4471-6681-8_7 Series Title: Springer Series on Cultural Computing.

[19] Nicholas Davis, Chih-PIn Hsiao, Kunwar Yashraj Singh, Lisa Li, Sanat Moningi, and Brian Magerko. 2015. Drawing Apprentice: An Enactive Co-Creative Agent for Artistic Collaboration. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition*. ACM, Glasgow United Kingdom, 185–186. https://doi.org/10.1145/2757226.2764555

[20] Leslie A. DeChurch and Jessica R. Mesmer-Magnus. 2010. Measuring shared team mental models: A meta-analysis. *Group Dynamics: Theory, Research, and Practice* 14, 1 (2010), 1–14. https://doi.org/10.1037/a0017455

[21] Kees Dorst and Nigel Cross. 2001. Creativity in the design process: co-evolution of problem–solution. *Design Studies* 22, 5 (Sept. 2001), 425–437. https://doi.org/10.1016/S0142-694X(01)00009-6

[22] David Engel, Anita Woolley, Lisa Jing, Christopher Chabris, and Thomas Malone. 2014. Reading the Mind in the Eyes or Reading between the Lines? Theory of Mind Predicts Collective Intelligence Equally Well Online and Face-To-Face. *PloS one* 9 (Dec. 2014), e115212. https://doi.org/10.1371/journal.pone.0115212

[23] Ozgur Eris. 2003. Asking generative design questions: a fundamental cognitive mechanism in design thinking. In *Proceedings of the international conference on engineering design*.

[24] Jennifer Fernquist, Tovi Grossman, and George Fitzmaurice. 2011. Sketch-Sketch Revolution: An Engaging Tutorial System for Guided Sketching and Application Learning. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11)*. Association for Computing Machinery, New York, NY, USA, 373–382. https://doi.org/10.1145/2047196.2047245 event-place: Santa Barbara, California, USA.

[25] Stephen M. Fiore and Travis J. Wiltshire. 2016. Technology as Teammate: Examining the Role of External Cognition in Support of Team Cognitive Processes. *Frontiers in Psychology* 7 (Oct. 2016). https://doi.org/10.3389/fpsyg.2016.01531

[26] Kate Forbes-Riley and Diane J. Litman. 2009. Adapting to Student Uncertainty Improves Tutoring Dialogues. (July 2009), 33–40. MAG ID: 1568422853.

[27] formlabs. 2020. Generative Design 101. https://formlabs.com/blog/generative-design/

[28] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. 2016. Image Style Transfer Using Convolutional Neural Networks. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Las Vegas, NV, USA, 2414–2423. https://doi.org/10.1109/CVPR.2016.265

[29] Aldair E. Gongora, Siddharth Mysore, Beichen Li, Wan Shou, Wojciech Matusik, Elise F. Morgan, Keith A. Brown, and Emily Whiting. 2021. Designing Composites with Target Effective Young's Modulus using Reinforcement Learning. In *Symposium on Computational Fabrication*. ACM, Virtual Event USA, 1–11. https://doi.org/10.1145/3485114.3485123

[30] George Head. 2003. Effective collaboration: Deep collaboration as an essential element of the learning process. *The Journal of Educational Enquiry* 4, 2 (2003).

[31] Kenneth Holstein, Vincent Aleven, and Nikol Rummel. 2020. A Conceptual Framework for Human–AI Hybrid Adaptivity in Education. In *Artificial Intelligence in Education*, Ig Ibert Bittencourt, Mutlu Cukurova, Kasia Muldner, Rose Luckin, and Eva Millán (Eds.). Vol. 12163. Springer International Publishing, Cham, 240–254. https://doi.org/10.1007/978-3-030-52237-7_20 Series Title: Lecture Notes in Computer Science.

[32] Ivan Hybs and John S. Gero. 1992. An evolutionary process model of design. *Design Studies* 13, 3 (July 1992), 273–290. https://doi.org/10.1016/0142-694X(92)90216-W

[33] Amir Jahanlou, William Odom, and Parmit Chilana. 2021. Challenges in Getting Started in Motion Graphic Design: Perspectives from Casual and Professional Motion Designers. In *Proceedings of Graphics Interface 2021 (GI 2021)*. Canadian Information Processing Society, 35 – 45. https://doi.org/10.20380/GI2021.06 ISSN: 0713-5424 event-place: Virtual Event.

[34] Heisawn Jeong and Michelene T. H. Chi. 2007. Knowledge convergence and collaborative learning. *Instructional Science* 35, 4 (May 2007), 287–315. https://doi.org/10.1007/s11251-006-9008-z

[35] Harmanpreet Kaur, Alex C Williams, and Walter S Lasecki. 2019. Building Shared Mental Models between Humans and AI for Effective Collaboration. (2019), 7.

[36] Anna Kawakami, Nikos Arechiga, Luke Guerdan, Matthew Lee, Yang Cheng, Scott Carter, Anita Sun, Haiyi Zhu, and Alison Hu. 2022. Towards a Learner-Centered Explainable AI. *Workshop on Human-Centered Explainable AI (HCXAI) at the ACM Conference on Human Factors in Computing Systems, CHI 2022* (2022), 7.

[37] Rubaiat Habib Kazi, Tovi Grossman, Hyunmin Cheong, Ali Hashemi, and George Fitzmaurice. 2017. DreamSketch: Early Stage 3D Design Explorations with Sketching and Generative Design. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. ACM, Québec City QC Canada, 401–414. https://doi.org/10.1145/3126594.3126662

[38] Kimia Kiani, Parmit K. Chilana, Andrea Bunt, Tovi Grossman, and George Fitzmaurice. 2020. "I Would Just Ask Someone": Learning Feature-Rich Design Software in the Modern Workplace. In *2020 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. 1–10. https://doi.org/10.1109/VL/HCC50065.2020.9127288 ISSN: 1943-6106.

[39] Hannah Kim, Jaegul Choo, Haesun Park, and Alex Endert. 2016. InterAxis: Steering Scatterplot Axes via Observation-Level Interaction. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (Jan. 2016), 131–140. https://doi.org/10.1109/TVCG.2015.2467615

[40] Janin Koch and Antti Oulasvirta. 2018. Group Cognition and Collaborative AI. In *Human and Machine Learning*, Jianlong Zhou and Fang Chen (Eds.). Springer International Publishing, Cham, 293–312. https://doi.org/10.1007/978-3-319-90403-0_15 Series Title: Human–Computer Interaction Series.

[41] Rafal Kocielnik, Saleema Amershi, and Paul N. Bennett. 2019. Will You Accept an Imperfect AI?: Exploring Designs for Adjusting End-user Expectations of AI Systems. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–14. https://doi.org/10.1145/3290605.3300641

[42] Kenneth R. Koedinger, Albert T. Corbett, and Charles Perfetti. 2012. The Knowledge-Learning-Instruction Framework: Bridging the Science-Practice Chasm to Enhance Robust Student Learning. *Cognitive Science* 36, 5 (2012), 757–798. https://doi.org/10.1111/j.1551-6709.2012.01245.x _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1551-6709.2012.01245.x.

[43] Sivam Krish. 2011. A practical generative design method. *Computer-Aided Design* 43, 1 (Jan. 2011), 88–100. https://doi.org/10.1016/j.cad.2010.09.009

[44] Benjamin Lafreniere, Parmit K. Chilana, Adam Fourney, and Michael A. Terry. 2015. These Aren't the Commands You're Looking For: Addressing False Feedforward in Feature-Rich Software. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. ACM, Charlotte NC USA, 619–628. https://doi.org/10.1145/2807442.2807482

[45] Matthew V. Law, Nikhil Dhawan, Hyunseung Bang, So-Yeon Yoon, Daniel Selva, and Guy Hoffman. 2019. Side-by-Side Human–Computer Design Using a Tangible User Interface. In *Design Computing and Cognition '18*, John S. Gero (Ed.). Springer International Publishing, Cham, 155–173. https://doi.org/10.1007/978-3-030-05363-5_9

[46] Matthew V. Law, Amritansh Kwatra, Nikhil Dhawan, Matthew Einhorn, Amit Rajesh, and Guy Hoffman. 2020. Design Intention Inference for Virtual Co-Design Agents. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*. ACM, Virtual Event Scotland UK, 1–8. https://doi.org/10.1145/3383652.3423861

[47] Bryan Lawson and Shee Ming Loke. 1997. Computers, words and pictures. *Design Studies* 18, 2 (April 1997), 171–183. https://doi.org/10.1016/S0142-694X(97)85459-2

[48] Ryan Louie, Andy Coenen, Cheng Zhi Huang, Michael Terry, and Carrie J. Cai. 2020. Novice-AI Music Co-Creation via AI-Steering Tools for Deep Generative Models. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–13. https://doi.org/10.1145/3313831.3376739

[49] Shareen Mahmud, Jessalyn Alvina, Parmit K. Chilana, Andrea Bunt, and Joanna McGrenere. 2020. Learning Through Exploration: How Children, Adults, and Older Adults Interact with a New Feature-Rich Application. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu

HI USA, 1–14. https://doi.org/10.1145/3313831.3376414

[50] Mary Lou Maher, Brian Magerko, Dan Venura, Douglas Fisher, Rogelio E. Cardona-Rivera, Nancy Fulda, John Gero, Minwoo Lee, David Wilson, James C. Kaufman, Maithilee Kunda, Michael Muller, Rachel K.E. Bellamy, Maya Ackerman, and Evangelia Chrysikou. 2022. A Research Plan for Integrating Generative and Cognitive AI for Human Centered, Explainable Co-Creative AI. In *CHI 2022 Workshop on Generative AI and HCI*. https://generativeaiandhci.github.io/

[51] Damien Masson, Jo Vermeulen, George Fitzmaurice, and Justin Matejka. 2022. Supercharging Trial-and-Error for Learning Complex Software Applications. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3491102.3501895 event-place: New Orleans, LA, USA.

[52] Justin Matejka, Michael Glueck, Erin Bradner, Ali Hashemi, Tovi Grossman, and George Fitzmaurice. 2018. Dream Lens: Exploration and Visualization of Large-Scale Generative Design Datasets. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–12. https://doi.org/10.1145/3173574.3173943

[53] Anne Meier, Hans Spada, and Nikol Rummel. 2007. A rating scheme for assessing the quality of computer-supported collaboration processes. *International Journal of Computer-Supported Collaborative Learning* 2, 1 (April 2007), 63–86. https://doi.org/10.1007/s11412-006-9005-x

[54] Achim Menges and Sean Ahlquist (Eds.). 2011. *Computational design thinking* (1. publ ed.). Wiley, Chichester.

[55] Tim Miller. 2019. Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence* 267 (Feb. 2019), 1–38. https://doi.org/10.1016/j.artint.2018.07.007

[56] Susan Mohammed and Brad C. Dumville. 2001. Team mental models in a team knowledge framework: expanding theory and measurement across disciplinary boundaries. *Journal of Organizational Behavior* 22, 2 (March 2001), 89–106. https://doi.org/10.1002/job.86

[57] Susan Mohammed and Erika Ringseis. 2001. Cognitive Diversity and Consensus in Group Decision Making: The Role of Inputs, Processes, and Outcomes. *Organizational Behavior and Human Decision Processes* 85, 2 (July 2001), 310–335. https://doi.org/10.1006/obhd.2000.2943

[58] Hussein Mozannar, Arvind Satyanarayan, and David Sontag. 2021. Teaching Humans When To Defer to a Classifier via Exemplars. *arXiv:2111.11297 [cs]* (Dec. 2021). http://arxiv.org/abs/2111.11297 arXiv: 2111.11297.

[59] Danil Nagy, Dale Zhao, and David Benjamin. 2018. Nature-Based Hybrid Computational Geometry System for Optimizing Component Structure. In *Humanizing Digital Reality*, Klaas De Rycke, Christoph Gengnagel, Olivier Baverel, Jane Burry, Caitlin Mueller, Minh Man Nguyen, Philippe Rahm, and Mette Ramsgaard Thomsen (Eds.). Springer Singapore, Singapore, 167–176. https://doi.org/10.1007/978-981-10-6611-5_15

[60] Nelson Nauata, Kai-Hung Chang, Chin-Yi Cheng, Greg Mori, and Yasutaka Furukawa. 2020. House-GAN: Relational Generative Adversarial Networks for Graph-Constrained House Layout Generation. In *Computer Vision – ECCV 2020*, Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm (Eds.). Vol. 12346. Springer International Publishing, Cham, 162–177. https://doi.org/10.1007/978-3-030-58452-8_10 Series Title: Lecture Notes in Computer Science.

[61] Ahmed K. Noor. 2017. AI and the Future of the Machine Design. *Mechanical Engineering* 139, 10 (Oct. 2017), 38–43. https://doi.org/10.1115/1.2017-Oct-2

[62] Heather Pon-barry, Karl Schultz, Elizabeth Owen Bratt, Brady Clark, and Stanley Peters. 2006. Responding to Student Uncertainty in Spoken Tutorial Dialogue Systems. *International Journal of Artificial Intelligence in Education* 16 (2006), 171–194.

[63] Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever. 2021. Zero-Shot Text-to-Image Generation. http://arxiv.org/abs/2102.12092 arXiv:2102.12092 [cs].

[64] reddit.com. 2022. r/Fusion360. https://www.reddit.com/r/Fusion360/

[65] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. 2022. High-Resolution Image Synthesis with Latent Diffusion Models. https://doi.org/10.48550/arXiv.2112.10752 arXiv:2112.10752 [cs].

[66] Matthias Scheutz, Scott A. DeLoach, and Julie A. Adams. 2017. A Framework for Developing and Using Shared Mental Models in Human-Agent Teams. *Journal of Cognitive Engineering and Decision Making* 11, 3 (Sept. 2017), 203–224. https://doi.org/10.1177/1555343416682891 Publisher: SAGE Publications.

[67] Donald A. Schön. 1983. *The reflective practitioner: how professionals think in action*. Basic Books, New York.

[68] Donghoon Shin, Sachin Grover, Kenneth Holstein, and Adam Perer. 2021. Characterizing Human Explanation Strategies to Inform the Design of Explainable AI for Building Damage Assessment. (2021). https://doi.org/10.48550/ARXIV.2111.02626 Publisher: arXiv Version Number: 1.

[69] Siemens Software. 2021. Generative Engineering: Instantly create multivariations of your designs that meet or exceed your requirements. https://www.plm.automation.siemens.com/global/en/products/mechanical-design/generative-design.html

[70] Jiao Sun, Q. Vera Liao, Michael Muller, Mayank Agarwal, Stephanie Houde, Kartik Talamadupula, and Justin D. Weisz. 2022. Investigating Explainability of Generative AI for Code through Scenario-based Design. In *27th International Conference on Intelligent User Interfaces*. ACM, Helsinki Finland, 212–228. https://doi.org/10.1145/3490099.3511119

[71] Fuwen Tan, Song Feng, and Vicente Ordonez. 2019. Text2scene: Generating compositional scenes from textual descriptions. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 6710–6719.

[72] Nurcan Gecer Ulu and Levent Burak Kara. 2015. DMS2015-33: Generative interface structure design for supporting existing objects. *Journal of Visual Languages & Computing* 31 (Dec. 2015), 171–183. https://doi.org/10.1016/j.jvlc.2015.10.016

[73] Nobuyuki Umetani, Takeo Igarashi, and Niloy J. Mitra. 2012. Guided exploration of physically valid shapes for furniture design. *ACM Transactions on Graphics* 31, 4 (Aug. 2012), 1–11. https://doi.org/10.1145/2185520.2185582

[74] Piet Van den Bossche, Wim Gijselaers, Mien Segers, Geert Woltjer, and Paul Kirschner. 2011. Team learning: building shared mental models. *Instructional Science* 39, 3 (May 2011), 283–301. https://doi.org/10.1007/s11251-010-9128-3

[75] Maarten van Someren. 1995. The think aloud method: A practical guide to modelling cognitive processes. *Information Processing & Management* 31, 6 (Nov. 1995), 906–907. https://doi.org/10.1016/0306-4573(95)90031-4

[76] Kurt VanLehn, Randolph M. Jones, and Michelene T.H. Chi. 1992. A Model of the Self-Explanation Effect. *Journal of the Learning Sciences* 2, 1 (Jan. 1992), 1–59. https://doi.org/10.1207/s15327809jls0201_1

[77] Dakuo Wang, Elizabeth Churchill, Pattie Maes, Xiangmin Fan, Ben Shneiderman, Yuanchun Shi, and Qianying Wang. 2020. From Human-Human Collaboration to Human-AI Collaboration: Designing AI Systems That Can Work Together with People. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–6. https://doi.org/10.1145/3334480.3381069

[78] Qiaosi Wang, Koustuv Saha, Eric Gregori, David Joyner, and Ashok Goel. 2021. Towards Mutual Theory of Mind in Human-AI Interaction: How Language Reflects What Students Perceive About a Virtual Teaching Assistant. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–14. https://doi.org/10.1145/3411764.3445645

[79] Christopher W. Wiese and C. Shawn Burke. 2019. Understanding Team Learning Dynamics Over Time. *Frontiers in Psychology* 10 (June 2019), 1417. https://doi.org/10.3389/fpsyg.2019.01417

[80] Humphrey Yang, Kuanren Qian, Haolin Liu, Yuxuan Yu, Jianzhe Gu, Matthew McGehee, Yongjie Jessica Zhang, and Lining Yao. 2020. SimuLearn: Fast and Accurate Simulator to Support Morphing Materials Design and Workflows. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*. ACM, Virtual Event USA, 71–84. https://doi.org/10.1145/3379337.3415867

[81] Nur Yildirim, James McCann, and John Zimmerman. 2020. Digital Fabrication Tools at Work: Probing Professionals' Current Needs and Desired Futures. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–13. https://doi.org/10.1145/3313831.3376621

[82] Kirsty Young. 2009. Direct from the source: the value of 'think-aloud' data in understanding learning. *The Journal of Educational Enquiry* 6 (2009).

[83] Loutfouz Zaman, Wolfgang Stuerzlinger, Christian Neugebauer, Rob Woodbury, Maher Elkhaldi, Naghmi Shireen, and Michael Terry. 2015. GEM-NI: A System for Creating and Managing Alternatives In Generative Design. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, Seoul Republic of Korea, 1201–1210. https://doi.org/10.1145/2702123.2702398

[84] Guanglu Zhang, Ayush Raina, Jonathan Cagan, and Christopher McComb. 2021. A cautionary tale about the impact of AI on human design teams. *Design Studies* 72 (Jan. 2021), 100990. https://doi.org/10.1016/j.destud.2021.100990

[85] Rui Zhang, Nathan J. McNeese, Guo Freeman, and Geoff Musick. 2021. "An Ideal Human": Expectations of AI Teammates in Human-AI Teaming. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (Jan. 2021), 1–25. https://doi.org/10.1145/3432945
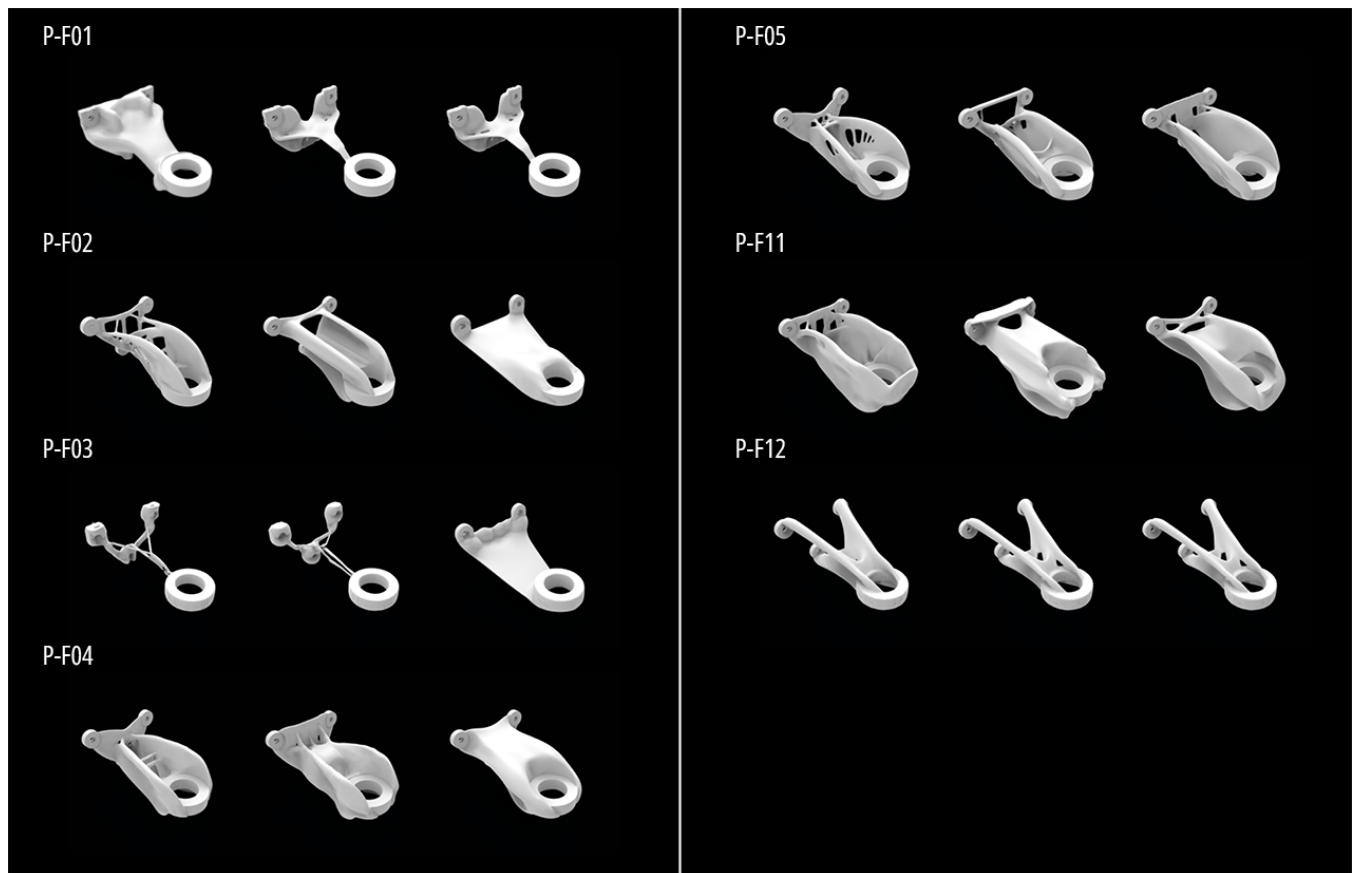
A ADDITIONAL MATERIALS

Figure 6: Overview of all engine mounting brackets created in the Fusion360 task. Every participant submitted three design options. Since the AI system generated the three options based on the same parameter values, each trio either met or missed the same criteria.

Table 6: Interview protocol with questions of the semi-structured post-task interview.

Theme	Topic		Question
Collaboration with the tool	*General Experience*	Q3	Could you tell me what it was like to work with the design tool in general?
	General Expectations	Q4	Before starting this project, what role did you expect the design tool to take? And what role did it actually take during the design process? Did the design tool match your expectations?
	Roles – Coordination	Q5	Do you feel that there was a collaborative interaction between you and the design tool? Can you give us some specific examples?
		Q6	Using your own words, what metaphors or analogies would you use to characterize the tool and the collaboration with it?
		Q7	Would you say that the process of working with the design tool is co-creative? Why?
		Q8	How was the workflow with the design tool different from your workflow of working with a human teammate?
	Roles – Expectations	Q9	What role do you want the AI to take during your process?
		Q10	How would you describe an ideal collaborative design tool for that task?
		Q11	What would a design tool need to do so that you would perceive it as "collaborative" and "co-creative"?
	Mental Model	Q12	Do you remember situations in which you found the behavior of the design tool confusing or irritating?
	Shared Mental Models	Q13	Do you think the AI feature successfully understands your design intent and design practice?
		Q14	Do you think the design tool collected enough information to support you most effectively? Why?
	Shared Mental Model – Expectations	Q15	Are there any aspects from your design process where you hope a design tool like this could learn more from you to better support you?
	Communication	Q16	How would you characterize the dialogue between you and the design tool?
		Q17	Can you describe situations in which the communication worked well and not so well?
		Q18	Do you find the interface helpful in guiding you through the whole process while getting started?
	Conflict Resolution	Q19	Do you remember situations of conflict with the design tool? How did you solve the conflict?
	Timing / Flow	Q20	How did the simulation time influence your design process?
Design Process	*Satisfaction with design outcome*	Q21	Are you satisfied with the final design in general? How closely does it match the design brief?
	Surprise	Q22	Do you remember situations in which suggestions from the design tool surprised you? If so, could you give me an example of that? Do you find the surprises productive or not?
	Design Negotiation	Q23	How did you find that the design tool changed the way you design? Did you ever adjust your design goal/strategy according to the tool's capabilities?
	Exploration	Q24	Did you explore more design options because of the design tool?
	Productivity	Q25	Do you think you can achieve the same output level (quality and quantity) without the help of an AI feature given the time constraints? Why?
	Trust / reassurance	Q26	How much do you trust the results from the design tool?
		Q27	Is there any other sort of analysis you were hoping the software would do that it didn't? Why?
Learning Process	*Learning Process*	Q28	How would you describe the learning process for using the tool and collaborating with it?
		Q29	Which situations of the design process did require the most guidance?
		Q30	How did you learn about the capabilities of the design tool?
		Q31	Which aspects of the design tool were challenging to understand or learn about?
	Learning Resources	Q32	Which learning resources did you find most useful? Why?
	Learning Expectations	Q33	How could the tool better guide an experienced designer who is starting to use the tool?
		Q34	What would the ideal collaborative learning process look like, in your opinion?
General Expectations		Q35	What would you want the design tool to do more of?
		Q36	What would you want the design tool to do differently?