A Hybrid Prototype Method Combining Physical Models and Generative Artificial Intelligence to Support Creativity in Conceptual Design

HONGBO ZHANG, Zhejiang University, China
PEI CHEN*, Zhejiang University, China
XUELONG XIE, Zhejiang University, China
ZHAOQU JIANG, Zhejiang University, China
ZIHONG ZHOU, Zhejiang University, China
LINGYUN SUN, Zhejiang University, China

Conceptual design is an essential stage in the design process, and its ultimate success largely depends on designers' creativity. Both physical and digital prototypes are commonly adopted by designers to support ideation and creativity, providing intuitive perception and rapid iteration, respectively. In recent advancements, large-scale generation models are able to offer data-enabled creativity support by generating high-quality solutions comparable to human designers. This opens up an imaginary space for designers and brings new possibilities for design tools. In this study, we proposed a hybrid prototype method that synergistically combines physical models and generative artificial intelligence (AI) in the conceptual design stage. Correspondingly, we developed a hybrid prototype system to implement the proposed method. We conducted a comparative user study with 45 designers who completed a design task using the physical prototype method, standalone generative AI, and the hybrid prototype method, respectively. Our results verified the effectiveness of the hybrid prototype method and investigated its mechanism in supporting creativity. Finally, we discussed the application value and optimisation space of the hybrid prototype method.

CCS Concepts: • **Human-centered computing** → **Interactive systems and tools**.

Additional Key Words and Phrases: Hybrid Prototype, Generative AI, Co-creative Systems, Physical Prototype, Creativity Support Tool (CST), Applications of Large-scale Generation Models

ACM Reference Format:
Hongbo Zhang, Pei Chen, Xuelong Xie, Zhaoqu Jiang, Zihong Zhou, and Lingyun Sun. 2024. A Hybrid Prototype Method Combining Physical Models and Generative Artificial Intelligence to Support Creativity in Conceptual Design. *ACM Trans. Comput.-Hum. Interact.* 1, 1, Article 1 (January 2024), 34 pages. https://doi.org/10.1145/3689433

1 INTRODUCTION

Conceptual design is a fundamental and essential stage in the design process, where designers seek to start with vague ideas and define the product outline and layout [31, 94]. Its ultimate

*The corresponding author.

Authors' addresses: Hongbo Zhang, hongbozhang@zju.edu.cn, Zhejiang University, Hangzhou, China; Pei Chen, Zhejiang University, Hangzhou, China, chenpei@zju.edu.cn; Xuelong Xie, Zhejiang University, Hangzhou, China, xuelongxie@zju.edu.cn; Zhaoqu Jiang, Zhejiang University, Hangzhou, China, zhaoqujiang@zju.edu.cn; Zihong Zhou, Zhejiang University, Hangzhou, China, zihongzhou@zju.edu.cn; Lingyun Sun, Zhejiang University, Hangzhou, China, sunly@zju.edu.cn.

ACM 1073-0516/2024/1-ART1
https://doi.org/10.1145/3689433

success is greatly contingent upon the designers' creativity. Creativity is the ability to generate and refine ideas, which involves coming up with new approaches to problems, original resolutions to conflicts, or fresh insights into existing solutions [38]. In the conceptual design process, prototypes are commonly used to understand the design space, learn about design problems, supplement designers' mental models, and discover unexpected phenomena [52]. They are conceived as the material embodiment or manifestation of designers' ideas [25, 49, 73].

The physical prototype serves as an instrumental approach in conceptual design, enabling designers to effectively externalise and simulate their concepts by tangible materials. It affords several distinct advantages for creativity support. First, physical models augment the accessible information and sensorial stimuli, fostering design reasoning [8, 29, 31, 67]. Second, embodied interaction with tangible models enables the body to participate actively in the creative process, mitigating cognitive overload and boosting creativity [44, 90]. Third, tangible models help designers identify defects in design solutions. The physical prototype method facilitates problem visualisation, draws attention to flawed design assumptions, and highlights significant discrepancies between the actual solutions and their conceptual predictions [37, 54]. Owing to its manifold merits for bolstering creativity, the physical prototype is a popular design method in conceptual design for idea exploration.

Utilising digital prototypes is another approach to support creativity in conceptual design. They are digital mock-ups, models or simulations, with the ability to create and test virtual design schemes [41]. Digital prototype tools often have graphical user interfaces for interactions [70]. They can support complex computation and provide more flexibility and fidelity during prototyping. The incorporation of artificial intelligence (AI) into digital prototypes for design ideation has become pervasive [45]. It employs the data-fitting capabilities of machine learning models to help designers derive insights from extensive datasets. For example, FashionQ [38] delivers an array of analytic outcomes on fashion attributes derived from a dataset encompassing 302,772 images, fostering both divergent and convergent cognitive processes during ideation. Rico [18] supports the design of UI layouts for mobile applications by analysing design elements across 72,000 popular designs.

Recent advancements in generative AI have further enhanced the data-enabled design process. Large-scale generation models have demonstrated capabilities to produce high-quality outputs, indistinguishable from human-created artefacts [85]. This enables AI to become a partner in the ideation process, transforming the way designers work [17]. Generative AI can proactively offer design schemes, making the conceptual design process demonstrate an inclination to be probabilistic due to generative variability [25]. This feature can open an imaginary space and support extensive design exploration. In addition, generative AI provides data-enabled understandings, which might offer insights that escape designers' relevance sense and awareness [25]. These distinct strengths of generative AI might bring the possibility of developing more powerful digital prototype tools and efficient creativity support tools (CSTs).

However, there are still inherent limitations in using physical models and standalone generative AI tools independently to support conceptual design. On the one hand, low-fidelity physical models are inadequate for presenting a complete design scheme, while detailed physical models demand long-time costs [66]. On the other hand, although generative AI can create inexhaustible design schemes, it usually lacks meaningful and interpretable controls for designers, making it challenging to express design intentions to AI intuitively [50]. Besides, as the majority of AI-based tools are typically proposed by developers, leaving designers one step behind them, designers lack appropriate prototype tools that meet actual design needs and support efficient work [20].

Recently, researchers have attempted to merge physical and digital materials to combine their collective strengths and ameliorate their respective drawbacks [80]. Inspired by this, we proposed a hybrid prototype method combining the physical model and generative AI to support creativity in conceptual design. With the support of the hybrid prototype method, designers are empowered to engage in the ideation process utilising physical materials. Simultaneously, AI comprehends the underlying design intentions through physical models and further refines them based on its generative capabilities. To assess the effectiveness of the proposed method, we conducted an empirical comparison study. 45 participants majoring in design were randomly assigned to the hybrid prototype group, the traditional physical model group, and the standalone generative AI group to complete the same conceptual task. We investigated the following research questions:

– ***RQ1***: *Does the hybrid prototype method enhance creativity support in the conceptual design stage?*
– ***RQ2***: *How do designers generate new ideas with the support of the hybrid prototype method?*
– ***RQ3***: *What are the distinct strengths of the hybrid prototype method in aiding conceptual design?*

To answer the ***RQ1***, expert evaluations were adopted to objectively and comprehensively evaluate participants' outcomes. A subjective questionnaire was also used to obtain the creativity differences when designers used different CSTs. To answer the ***RQ2***, the Think-aloud method was employed, and participants' design behaviour was coded. To answer the ***RQ3***, a semi-structured interview was conducted after the entire design task to obtain participants' feedback.

Our research makes three contributions. First, we proposed a hybrid prototype method combining physical models and generative AI to support creativity in conceptual design. Second, to verify the effectiveness of the proposed method, we accordingly developed a hybrid prototype system. Third, we conducted an empirical user study and provided a fundamental understanding of how generative AI can be integrated as a design material in conceptual design. We provided insights into designing interactions within the co-creation system that involves both human designers and generative AI.

2 RELATED WORK

2.1 Physical Prototype in Conceptual Design

The physical prototype has been employed for various purposes like exploring basic shapes, evaluating physical properties, and confirming analysis based on virtual models [41]. It has unique advantages in supporting creativity. First, physical models can increase the number of sensory stimuli and promote design reasoning. They present more visual cues, allowing designers to work with more subtle design schemes and enhancing the available information for design reasoning [31]. Tangible models also offer more multisensory stimuli than texts and images [29], making the design environment more engaging and intuitive [31]. Second, the tangibility of physical prototypes can enhance embodied sensemaking [70]. As research in cognitive theory pointed out, thinking is a distributed and interactive process, and body movement can literally be part of thinking [70]. Interacting with tangible models during ideation involves physical engagement in cognition, which eases the cognitive load and supports creativity [31]. Third, there is more room for interpretation with physical models than with digital design tools [70], allowing designers to assign meaning to situations from their own perspective [35] and develop into a plausible narrative [64]. Fourth, the use of physical models helps to reveal flaws in design. The tangible nature of physical models allows designers to more effectively evaluate the shape and function of their design schemes, as they are better at showing what the design looks like and uncovering incorrect assumptions [37, 46, 54]. Beyond these primary benefits, the physical prototype method can expand designers' viewpoints [5] and boost confidence in their decisions [7], further enhancing creative potential.

Nevertheless, the physical prototype method has some inherent limitations. Specifically, its fabrication time is a disadvantage, which often increases with the fidelity and materials required [41]. Designers often face a trade-off between quality and time when selecting prototype methods [66]. Rapid prototyping and iteration are essential in the conceptual design stage, physical modelling may not be the best choice for a design task where precision or fidelity is necessary. Additionally, it is difficult to keep track of changes with physical models to manage design versions [41], limiting the shareability and searchability of the design outcomes [70]. Consequently, it is crucial to thoroughly assess limitations of the physical prototype method and explore strategies to leverage its strengths, thereby enhancing creativity at the conceptual design stage.

2.2 AI-driven Creativity Support Tools in Conceptual Design

The advent of deep generative models has facilitated significant progress in the integration of AI within the design process. The essence of generation models lies in learning the distribution patterns of extensive datasets, thus mastering abstract knowledge automatically and ultimately gaining the capability to create novel visual and textual content. Especially, large-scale generation models such as Midjourney [16], Stable Diffusion [63], and DALL-E [60, 61] have been renowned for synthesising high-quality images from textual descriptions. These models exhibit a remarkable proficiency in generating realistic artefacts that often possess a high quality comparable to those created by human beings [85].

AI has transformed from a mere auxiliary to a synergistic collaborator for humans, and multiple AI-driven art generative systems [16, 53, 85, 87] have been developed to assist the art and design domain. This transition has a subtle impact on creativity. First, AI-driven design enriches the source of visual reasoning. By leveraging the generative variability, designers are presented with the opportunity to continuously obtain a diverse range of outputs by adjusting input parameters simply [86]. It provides designers with an infinite number of design schemes for design reasoning. Second, AI-driven design opens up an imaginative space for designers. In the interaction of co-creation systems, the value of design exploration is found in its potential to harness the probabilistic results generated by AI, in other words, to improvise rather than prescribe [25]. Third, the employment of AI-driven design expands the knowledge space and augments the scope of problem insights for designers. Generative AI can unpredictably offer insights that escape human relevance sense and awareness [26], thereby helping designers to problematise the design space rather than locking designers into ideas of predictability and normativity [25].

Many studies have discussed challenges when integrating AI as design materials [20, 88, 89]. On the one hand, the generative variability might cause unstable outputs, which lack common sense and often reduce designers' expectations [20, 86]. On the other hand, many generative design tools operate as black boxes where designers set prompts and then obtain generated results. It is challenging for designers to share design intentions with AI and iterate design ideas in a step-by-step procedure [14, 28]. Although the underlying co-creation mechanisms between human and generative AI remain unclear, existing evidence highlights the pivotal role of interaction innovation in co-creation systems [12]. Therefore, it is valuable to further explore the interaction modes and applied scenarios in co-creation systems. Moreover, it is also important to investigate how to effectively apply the AI's variability and uncertainty to support creativity.

2.3 Hybrid Prototype in Design Process

In numerous design scenarios, the physical creation unfolds independently of any digital counterparts, with the digital phase exclusively conducted using computer-based tools. Previous research has highlighted the inadequacy of both purely physical and purely digital methods [70]. Establishing a closer integration between the digital and physical domains within the design process might be an important requirement for CSTs [30]. The hybrid prototype bridges the gap between digital information and physical objects [80], striving to harness a potential that transcends the mere summation of its digital and physical components [19]. The salient characteristics exhibited by digital tools within hybrid systems encompass heightened prototype precision, facilitated capabilities for scheme preservation and dissemination, as well as robust support for complex computing tasks. Consequently, the HCI community has paid attention to the hybrid framework and attempted to meld the strengths of both the physical and digital spheres. Barbieri et al. [4] introduced a hybrid prototype with the intuition of physical models and high precision of digital models through mixed reality technology. This hybrid method engaged the visual, auditory, and tactile senses to analyse the inter-relationships between the physical form and behaviour of industrial products [4]. Tao et al. [76] developed a mixed reality system named 4Doodle to assist novices in mastering the manual skills of physical 4D printing through interaction on a tablet computer. Their findings suggested that the hybrid prototype reduced the prototype barrier and had great potential for creative production and spatial ability enhancement. Peng et al. [58] developed Robotic Modelling Assistant (RMA), in which the designer used an augmented reality controller to sketch 3D designs. The RMA then quickly realised the designer's scheme through a 3D printing pen, which strengthened the designer's intuitive perception of the scheme and sped up the prototyping. Kim et al. [43] proposed a hybrid prototype system that integrated hand motion and pen stroke in a complementary manner. It reduced the difficulty of spatial drawing through physical activities and improved the accuracy and expression of sketches through digital refinement. Against that background, we attempt to synergistically combine the traditional physical prototype method and the generative AI in a hybrid prototype environment, combining their collective strengths to support conceptual design.

3 HYBRID PROTOTYPE METHOD

3.1 Introduction of Hybrid Prototype Method

In this study, a hybrid prototype method combining the traditional physical prototype and generative AI was proposed to support creativity in the conceptual design stage. As shown in Figure 1, with the support of the hybrid prototype method, the designer is able to conceive and prototype ideas through physical models. The hybrid prototype method facilitates embodied idea exploration, rapid concept externalisation, and structure iteration. While the designer interacts with physical models, AI comprehends design intentions through not only textual requirements and restrictions but also the physical prototype. By assimilating and processing multi-modal inputs, AI can generate digital candidate schemes that adhere to the specified constraints while offering new perspectives. Enabled by the hybrid prototype method, designers can draw inspiration from multiple sources. On the one hand, they derive creative stimuli from embodied interaction with tangible models. On the other hand, they collaborate with AI and are inspired by diverse generated outputs.

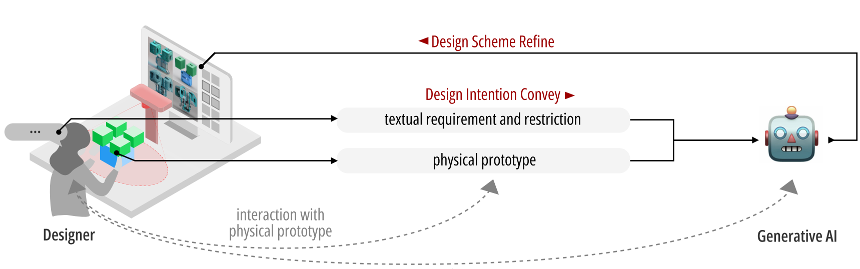

Fig. 1. Introduction to the hybrid prototype method.

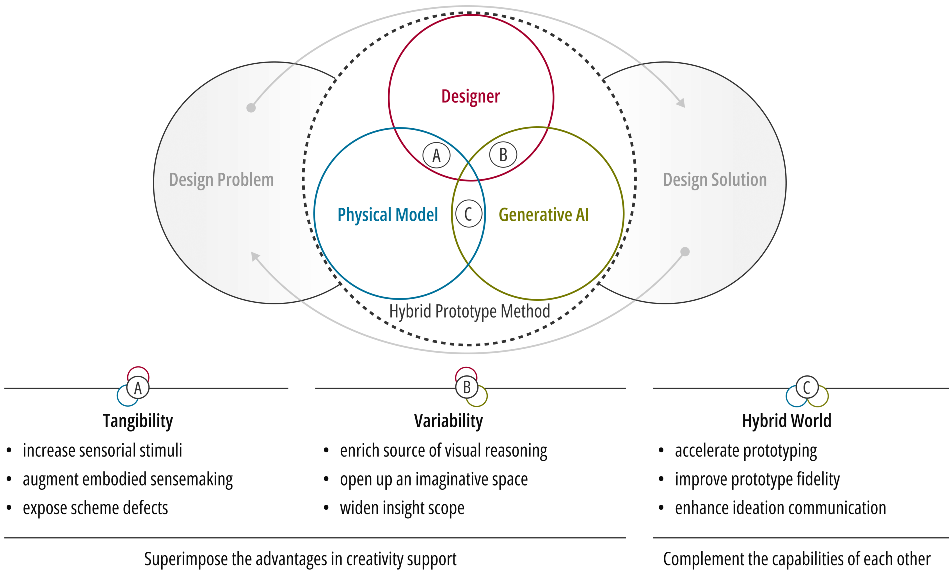

Fig. 2. The framework of the proposed hybrid prototype method.

3.2 Combined Capabilities and Expected Strengths of Hybrid Prototype Method

This section clarifies the combined capabilities and expected strengths of the hybrid prototype method. As it is an integrated method of the traditional physical prototype method and standalone AI tool, we divide the combined capabilities into superposition ability and complementary ability, as presented in Figure 2.

In the hybrid prototype method, we attempt to integrate tangibility in physical models and variability in generative AI to maximise their impact in facilitating creativity (Ⓐ and Ⓑ in Figure 2). Specifically, the tangibility has been proven to increase sensorial stimuli [29, 31] and augment embodied sensemaking [44], while the generative variability might open up an imaginative space [25, 86] and widen insight scope [26] during ideation. The tangibility and variability can be activated in the hybrid prototype method to provide creativity support.

The hybrid prototype can potentially complement the physical prototype and generative AI capabilities (Ⓒ in Figure 2). Specifically, the long fabrication time is an inherent disadvantage of the physical prototype method [41] while generative AI is good at producing diverse schemes efficiently and accelerating iteration [68, 86]. Therefore, with the support of the hybrid prototype method, AI can exploit its generative potential to refine coarse physical models, thereby enhancing the scheme fidelity and reducing the manufacturing duration. This method strikes an optimal balance between enhanced fidelity and expedited manufacturing processes. Besides, designers consider that AI is a challenging material due to its interaction complexity [48, 89] and natural language prompts are not expressive enough to convey design intentions to AI [14]. We are trying to combine AI interaction modes with common design materials familiar to designers, thereby enhancing detailed and intuitive designer-AI communication.

3.3 Implementation of Hybrid Prototype Method

We implemented our hybrid prototype method in a developed system using a computer and camera. To demonstrate the characteristics of the hybrid system more clearly, we also compared the

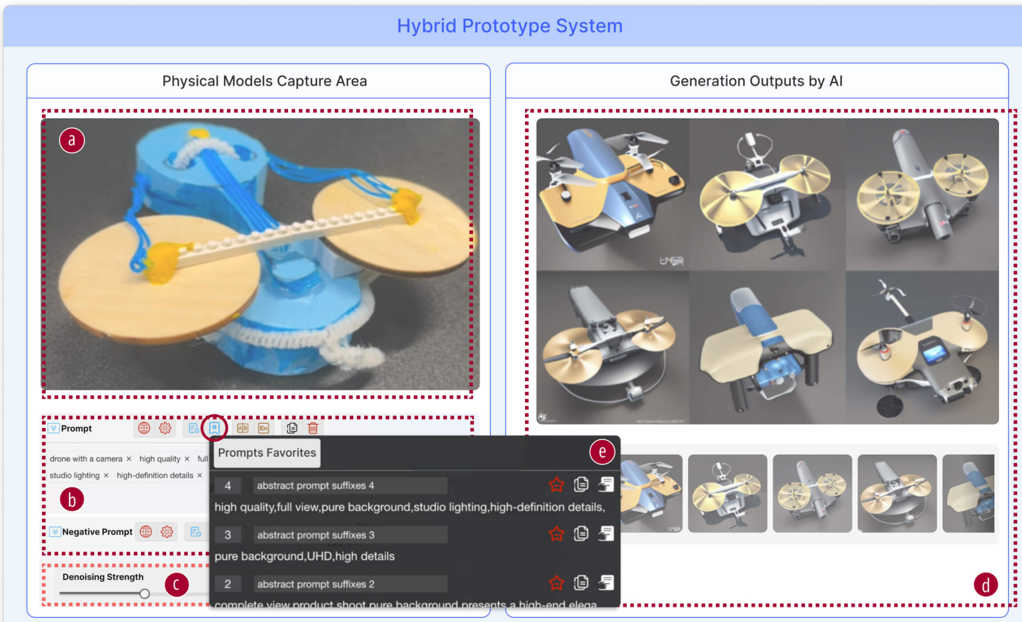

Fig. 3. The user interface of the hybrid prototype system.

traditional physical prototype system and the standalone generative AI system. As we conducted an experimental task focused on "drone design" in a user study, we employs related design examples.

- The environment of *traditional physical prototype* encompassed a purely physical space supporting prototype through tangible materials.
- The environment of *standalone generative AI tool* contained a computer-supported system running generation models in a purely digital space.
- The environment of *hybrid prototype method* contained the proposed hybrid prototype system, which had capabilities to capture photographs of physical models and produce AI-refined schemes.

The Stable Diffusion v2.1 was adopted in our hybrid generative system as the generative model [63, 72] . The user interface was modified based on the Stable Diffusion WebUI [3] (Figure 3). Our system requires no customised hardware to support it, which can be set up in any typical design workspace. Specifically, ⓐ shows the image of the physical model captured by the camera, and the designer can save and print it for iteration management. ⓑ presents a prompt area that allows users to describe requirements and constraints through text. ⓒ regulates the weight control of the prompt on the generated outputs. ⓓ presents the generated outputs to support designers in browsing, supporting further saving and printing. We have made the following modifications based on the original Stable Diffusion WebUI [3] to the implemented system.

First, we simplified the interactive components in Stable Diffusion WebUI. Through the investigation before organising the experiment, we knew that designers thought the original interface layout was not friendly enough and too complex. Therefore, we simplified the interactive components and kept the three most important control conditions, the *prompt*, *negative prompt*, and *denoising strength*, which help to determine how closely the generated image will be influenced by the captured image of physical models. Other parameters that control the generation, such as *Sampling Steps*, *Batch Count*, and *Sampling Method* were built into the program and not visible to designers on the interface. Second, we developed favourites of prompt words of recommended prompt words for designers' reference (ⓔ in Figure 3). These prompt words in favourites were intended to serve as abstract suggestions for quality control, such as "*UHD*", "*high details*", and "*best quality*", and would not contain concrete words that influence the design intentions or schemes. Specifically, after designers input the prompt words related to the design theme (e.g., "*drone with a camera*"), they can click the favourites of prompt words to add a set of abstract prompt suffixes randomly, making the prompt words more complete and complicated to improve the generation quality (e.g., "*drone with a camera, high quality, full view, pure background, studio lighting, high-definition details*"). This was achieved through the development of web plugins. Third, since most of the participants we recruited were not native English speakers, to avoid the obstacle of language to the design process, we provided a translator for prompt words to support designers in communicating with generative AI in their most familiar language.

We have carefully considered the influences of the above modifications. For example, the reduction of interactive components might reduce the generated output quality. Additionally, the limitation of the translator's accuracy might also influence the generated outputs. However, our focus is exploring creativity support in conceptual design, thus high-quality and high-precision outputs are not the most necessary considerations. Conversely, if the designers' focus is directed toward tool learning as opposed to creative activities, our results may be more influenced. In addition, applied plug-ins, such as the prompt favourite and translator, were called from the Kitchen Theme [23] and integrated into the Stable Diffusion WebUI [3]. They were incorporated solely to enhance participants' comprehension and interaction with the informational content. They did not produce substantive semantic information nor impact the generated outputs, thereby avoiding any influence on the research outcomes.

3.4 Design Workflow Showcase in Hybrid Prototype Method

We present a step-by-step design example to show the design workflow under our hybrid prototype method (Figure 4). Designers' thinking and understanding of the design process are also articulated in the form of text to show the idea-development process.

4 USER STUDY

This section introduces an empirical study to explore how the hybrid prototype affects creativity during conceptual design. A contrast experiment was conducted, whereby participants were randomly assigned to design groups with the support of the hybrid prototype method, the traditional physical prototype, and the standalone generative AI tool to complete the same conceptual task. The solution outputs of conceptual design were presented completely on a design solution answer sheet, and the whole creative process of conceptual design was required to use the Think-Aloud protocol. Through analysing the results and processes of the conceptual design, we explored three research questions (***RQ1*** & ***RQ2*** & ***RQ3*** described in *Section* 1).

4.1 Participants

We recruited 45 designers (24 female and 21 male, aged 19 to 27) with backgrounds in Industrial Design. Most participants were recruited from our university. All participants had a minimum of two years' experience in design. This also includes students who had actively participated in practical design projects and had completed over two years of education encompassing practical design courses, in addition to possessing their own design internship experience. They were required to be proficient in designing through physical prototype materials and using generative AI tools,

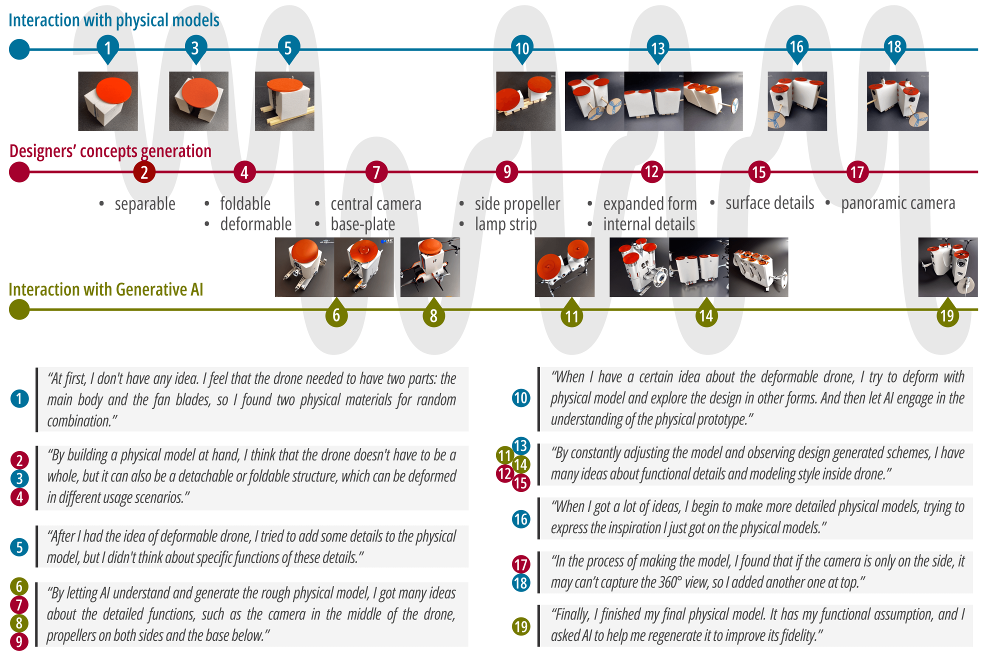

Fig. 4. A step-by-step design showcase with the support of the hybrid prototype method. ⓐ shows the development of the design process. ⓑ shows the designer's thinking extracted from the Think-aloud method.

such as Midjourney [16], Stable Diffusion [63], and DALL-E [60, 61], which was determined via a registration questionnaire. Its aim was to ensure participants focused on the conceptual design process instead of learning the use of various tools. Three experiment groups, each composed of 15 participants, were formed randomly and assigned to use traditional physical prototype (*Group PP*), standalone Generative AI tools (*Group AI*), and hybrid prototype method (*Group HP*) respectively to complete the same conceptual design task. In the process of grouping, the gender, age, design experience, and education level of the designers were balanced to eliminate individual differences. The related information of recruited participants is presented in Table 1. All participants signed a consent form approved by our institution.

Table 1. Overview information of participants.

Group	Participant ID	Age	Gender	Occupation	Design Exp
Group PP	PP01 - PP15	M = 22.2 SD = 1.5	8 Female 7 Male	14 Students (6 BA, 6 MA, 2 Ph.D) 1 Designer	M = 3.5 SD = 1.6
Group AI	AI01 - AI15	M = 22.3 SD = 2.2	8 Female 7 Male	14 Students (5 BA, 7 MA, 2 Ph.D) 1 Designer	M = 3.3 SD = 1.4
Group HP	HP01 - HP15	M = 23.1 SD = 2.0	8 Female 7 Male	14 Students (7 BA, 5 MA, 2 Ph.D) 1 Designer	M = 3.4 SD = 1.1
Total	/	M = 22.5 SD = 2.0	24 Female 21 Male	42 Students (18 BA, 18 MA, 6 Ph.D.) 3 Designer	M = 3.4 SD = 1.4

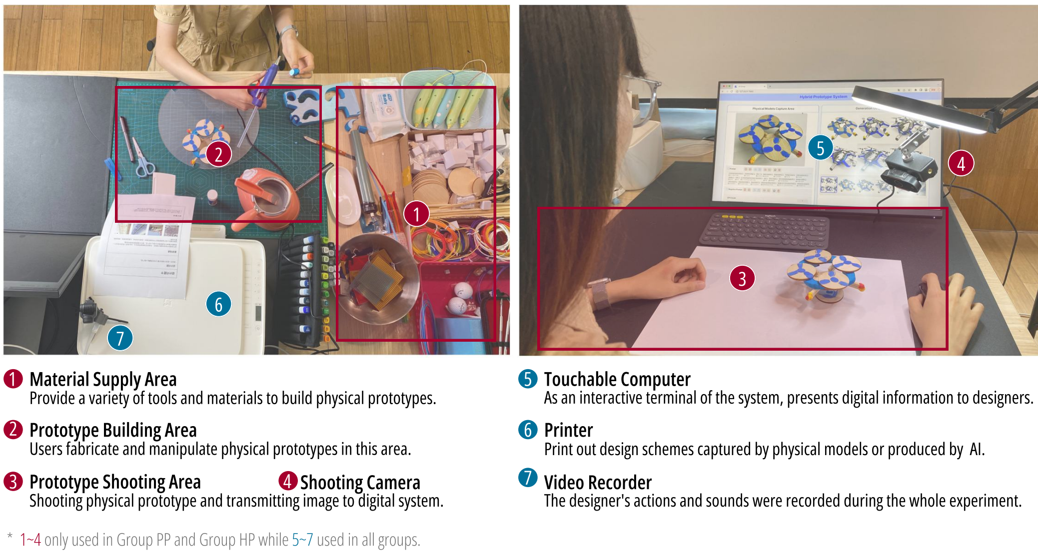

Fig. 5. The experimental environment layout.

4.2 Development of Experimental Environment and System

To fulfill the prerequisites of the experimental environment defined above, an experimental workspace was set up (Figure 5). For participants in *Group PP* and *GroupHP*, there was the *Material Supply Area* (①) providing various physical materials. Designers fabricated in the *Prototype Building Area* ②) and captured the image of physical models in the *Prototype Shooting Area* (③), which set up a shooting camera (④) and helped designers record the prototype process and version management. A touchable computer (⑤) was provided to all groups as the carrier of information presentation and interaction. A printer (⑥) was provided to all groups to help them express design solutions on proposal paper. A video recorder (⑦) was used to record designers' behaviours.

We also developed another two interactive systems for control groups in order to support the implementation of the contrast experiment.

- For *Group PP*, although participants completed the design task in the pure physical space, in order to record the design ideas and support the final scheme expression, a system that can shoot and record the tangible models was provided to help designers save process information.
- For *Group AI*, a generative AI system slightly modified by Stable Diffusion WebUI [3] was provided to *Group AI*, which was the same as the system in *Group HP*, but it could not capture photographs of physical models.

4.3 Experiment Task

Participants from all groups were requested to complete the same conceptual design task: *the conceptual innovation design of a traffic drone*. A potential design theme was provided in the experimental task, so that designers could have a similar ability to investigate the design background and complete the design task as they usually do in actual design activities [47]. The design task was carefully selected by researchers. First, the traffic scenario is universally known, which could facilitate the utilisation of prior knowledge and experience in design [42]. Second, all participants have been exposed to intelligent product design involving drones as part of their practical curriculum, possessing a foundational understanding of pertinent design principles and technology. Therefore,

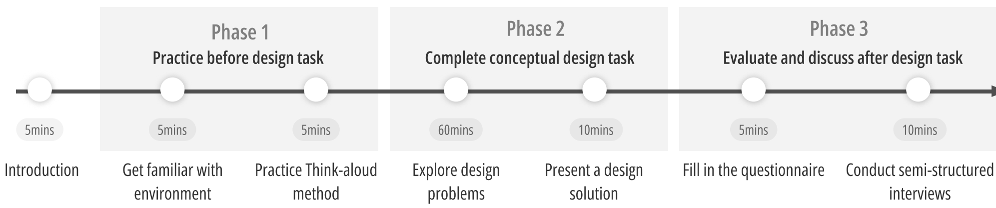

Fig. 6. The procedure of the experiment.

they could independently execute design tasks without the need for online searches or external resources. A design problem card was provided to participants, which defined and elaborated the design problem, background, and requirements (present in Appendix A).

Participants were requested to propose design solutions considering originality, practicality, and feasibility. Throughout the design exploration phase, emphasis was placed on encouraging participants to approach design problems from an innovative standpoint and to propose novel design solutions. Following the conclusion of the design exploration stage, participants were required to organise their design ideas and submit a complete design solution using a design solution answer sheet. The use instructions for the solution answer sheet and the structure of the scheme description were unified through the introduction before the experiment (detailed in Appendix D).

4.4 Procedure

Participants from all groups followed the experiment procedure shown in Figure 6. At the beginning of the experiment, researchers introduced the study aim, design brief, and experiment task to participants. After the informed consent confirmation, participants were instructed to get familiar with the experimental environment and prototype tools, and then practiced designing with the Think-aloud method. The conceptual design task lasted 70 minutes, consisting of 60 minutes for exploring and 10 minutes for completing the solution answer sheet. The time duration allowed for design exploration was suggested as sufficient for idea generation [78]. After the conceptual design, there was a phase to answer the questionnaire and engage in the semi-structured interview. The entire experiment took approximately 100 minutes.

4.5 Evaluation and Analysis

As summarised in Table 2, we collected data from *the results of conceptual design task*, *the process of conceptual design task*, and *the feedback after conceptual design task* to conduct qualitative and quantitative analysis among three experimental groups. To answer ***RQ1***, a multifaceted approach was employed. The expert evaluation was conducted to provide an objective assessment of the designers' outcomes. The Creativity Support Index (CSI) [13] was also utilised to investigate differences in creativity resulting from the use of diverse CSTs. To address ***RQ2***, the Think-aloud method [1] was employed during the design process. Participants' design behaviours were analysed to investigate creative thinking processes and idea generation patterns. Finally, ***RQ3*** was answered through the semi-structured interviews, which were utilised to obtain participant perspectives on the user experience. All analysis metrics are detailed in the following section.

*4.5.1 Analyse the result of conceptual design task (**RQ1**).* We initially evaluated the differences in design results among three experimental groups. Specifically, we invited six design experts possessing a minimum of ten years of professional design experience to participate in an expert evaluation together offline. They were asked to evaluate the *utility*, *novelty*, *feasibility*, and *structural*

Table 2. The summary of analysis perspectives and evaluation metrics.

Standpoint	Focused Problem	Research Method	Metrics	RQ
The result of conceptual design	Differences in conceptual design results	Expert objective evaluation (quantitative analysis)	Utility, Novelty, Feasibility, Structural counts	***RQ1***
The process of conceptual design	Patterns of creativity generation	Behaviour analysis by Thinking-Aloud protocol (qualitative analysis)	/	***RQ2***
The feedback after conceptual design	Subjective feelings in design process	Evaluation by CSI questionnaire (quantitative analysis)	CSI total score, score of Collaboration, Enjoyment, Exploration, Expressiveness, Immersion, and Results Worth Effort	***RQ1***
	Advantages and disadvantages in three comparative tools; The cooperation in co-creation systems	Semi-structured interview (qualitative analysis)	/	***RQ3***

counts of design solutions based on the answer sheets completed by the participants. Before the evaluation began, experts were informed of the evaluation procedure, metrics, methods, and standards in detail, while the different settings among groups were not informed in advance to avoid experts' subjective influence. Two test-solutions were also provided to experts for practice before the formal evaluation to ensure a consensus on the evaluation process and standards.

The procedure of the expert evaluation is presented in Figure 7, and it contains four steps. Initially, the design solution was interpreted and decomposed by experts to identify and extract all functions together. To avoid order influence, the design solutions were displayed in a random and counter-balanced way. Next, experts independently evaluated whether each function was novel or not. We provided all experts with the criteria for judging novel functions defined by Sarkar and Chakrabarti [65] as a reference. Eventually, experts scored several indicators related to technical feasibility relying on their professional knowledge and experience. They were also asked to identify and extract structural functions in design solutions.

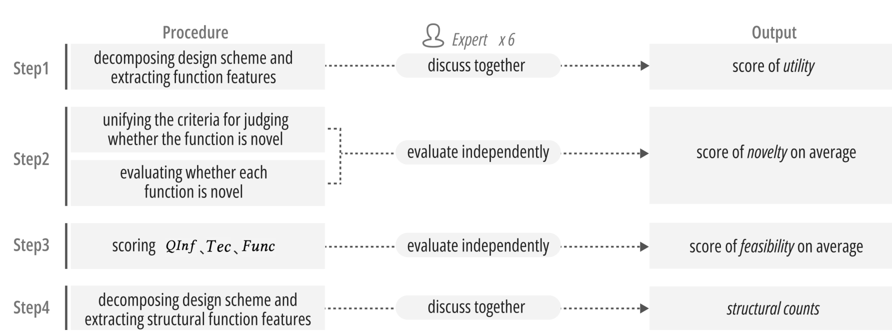

Fig. 7. The procedure of expert evaluation.

The measurement method of the *utility*, *novelty*, and *feasibility* was adopted with reference to metrics proposed by previous studies [22, 31, 65, 71]. Specifically, the *utility* is related to the total amount of functions included in the solution. It refers to the ability of the idea to add value and to satisfy the design requirements [31]. It is calculated as the sum of the different functions that can be performed by the solution. The function is defined as a specific attribute or module in a design scheme that performs a distinct functionality or serves a particular purpose in the expert evaluation. However, as the function identification is relatively subjective and ambiguous, we allow experts to discuss functions in a scheme together, resolve differences, and reach a consensus on function counts, serving to mitigate the impact of varying expert interpretations. For example, after the expert discussion, they usually regard the combined function, such as the "solar panel with lights", as one function instead of two independent functions.

The *novelty* metric assesses the extent to which the functions of a given solution are novel relative to current or similar technologies utilised to address the problem at hand. Specifically, let $F = \{f_1, f_2, \ldots, f_N\}$ be the set of functions in a design solution, with $F_n = \{f \in F | Nv(f) = 1\}$ representing the set of novel functions. Here, $Nv(f) = 1$ indicates a function with a novelty aspect, while $Nv(f) = 0$ indicates no novelty. The novelty (Nov) score is defined as the proportion of innovative functions to total functions:

$$Nov = \frac{|F_n|}{|F|} \tag{1}$$

The feasibility metric is defined as the extent to which a proposed solution can be effectively implemented in the subsequent design phase, contingent upon its practical realisability. In order to calculate the feasibility (Fea) score, experts were asked to evaluate three perspectives: the information quantity ($QInf$), technological maturity (Tec), and relevant functions ($Func$) [32]. Specifically, the $QInf$ is the amount of available information in the design solution, assessing the extent to which the proposed design principles can be fully understood based on descriptions provided by participants. The assigned scores correspond to 0 (for incomplete information), 1 (for partially complete information), or 2 (for complete information) depending on the comprehensiveness of the solution description presented. The Tec represents the technological maturity of design solution, as it pertains to the potential for implementation success of the technology proposed. The available score ranges from 0 (for solutions that are impossible to evaluate) to 1 (for solutions involving obsolete technology), 2 (for new technologies with high potential for failure), 3 (for new technologies with good potential for success), and 4 (for mature technologies that are commercially available). The $Func$ evaluates whether the described functions are adequate for the overall operation of the function. The possible assigned scores range from 0 (for solutions that fail to describe any relevant function) to 1 (for solutions that describe some relevant functions), and 2 (for solutions that comprehensively describe all necessary relevant functions). After evaluating these three scores, the feasibility (Fea) is calculated by Equation 2 based on the standard of Grubisic's [32].

$$Fea = (0.41 \times QInf) + (0.33 \times Tec) + (0.26 \times Func) \tag{2}$$

*4.5.2 Analyse the process of conceptual design task (**RQ2**).* We adopted the concurrent Think-aloud protocol [1, 39] during the experimental process to discover idea generation patterns, including how designers spark new ideas, what they think and perform, what patterns they follow during the design process. Think-aloud studies require participants to verbalise their thoughts while performing a task thereby allowing researchers to investigate and comprehend underlying mechanisms. Two researchers reviewed the entire design process, observed what they said when thinking and what they did while working, and then coded their behaviour. The coding scheme was modified based on the creative process model summarised by Howard et al. [34] and Liao et al. [47].

- Building action to prototype (Build): building physical materials, such as the action of fabrication, adding connection, and manipulation;
- Generating through generative AI (Generate): interacting with a generative system to obtain generated outputs, such as typing prompts;
- Analysing facts in prototyping (Analyse): exploring design solutions and analysing problems, such as analysing design requirements and exploring the design problem space;
- Evaluating and comparing in prototyping (Evaluate): looking back, evaluating, and relating the idea to others, such as assessing and implementing structural rationality;
- Sparking new idea (Spark): proposing factors or descriptions of the design idea.

Before the coding process, three different ten-minute-long video segments were used to ensure the inter-rater agreement for the coding scheme between the two researchers. The agreement among researchers was calculated on a minute-by-minute basis and was ensured to be greater than 0.60 indicating good agreement [33].

*4.5.3 Analyse the feedback after conceptual design task (**RQ1** & **RQ3**).* To minimise the potential influence of the bias of researchers and design experts on study results, all participants' subjective feedback was documented for subsequent data analysis, which consists of subjective questionnaire (***RQ1***) and semi-structured interview (***RQ3***).

In this study, the ability of a CST to assist a user engaged in creative work was measured by the CSI questionnaire [13]. It measures six dimensions of creativity support: *Collaboration*, *Enjoyment*, *Exploration*, *Expressiveness*, *Immersion*, and *Results Worth Effort*. The CSI evaluation was conducted in two stages: the agreement statement session and the paired-factor comparison section. In the agreement statement session, 12 statements representing six factors were randomly presented on a questionnaire (Appendix B). Each agreement statement was answered on a scale of "Highly Disagree" (1) to "Highly Agree" (10). Within the paired-factor comparison section, each factor was paired with every other factor comparison. In this process, participants selected a factor description in response to the prompt: "*When doing this task, it's most important that I'm able to...*". The *CSI Score* and each *Weighted Factor Score* of six dimensions were calculated to evaluate the subjective creativity support among three experimental groups.

In addition, all participants were invited to participate in semi-structured interviews. We paid attention to the subjective evaluation of design process under specific support system (for all groups) and the user experience of cooperation with generative AI (only for *Group AI* and *Group HP*). Specifically, we conducted interviews around four key issues: *advantages of the support tool*, *disadvantages of the support tool*, *cooperation with AI*, and *AI influence on design ideas*. The interview structure and outlines are presented in Appendix C. We used thematic analysis [75] to analyse raw interview data. Two researchers independently assigned different statements to these themes. The coding process strictly pertained to the semantic content of the interview data instead of the assumptions of coders. For example, the statement "*I think this system can help me generate richer design expressions and help me express my design ideas more clearly and concretely*" (*HP12*) was coded as "*improve the expressiveness*". Next, the two researchers shared their codes, discussing inconsistent codes to resolve disagreements and merging similar codes until they reached a consensus.

4.6 Statistical Analysis

For the evaluation index of quantitative analysis, the statistical analysis was conducted among three experimental groups in SPSS (V22.0, International Business Machines Corporation, Armonk, NY, USA). A Shapiro-Wilk normality test was run at the significance level of 0.05 before statistical analysis. The homogeneity of variances assumption was assessed using Levene's test. If the assumption of homogeneity of variances was met, one-way ANOVA was performed. However, when the assumption of homogeneity of variances was violated, the Welch test was used instead. Post hoc multiple analyses were also performed. Bonferroni and Tamhane corrections were used to adjust for the p-value to identify significant differences for homogeneous results and heterogeneous results, respectively. The level of statistical significance for all of these analyses was set at 0.05.

5 RESULTS AND FINDINGS

5.1 Display of Design Output During Design Task

Examples of design solution answer sheets are shown in Figure 8. They were randomly chosen from three experimental groups. Participants were instructed to use the left area of the answer sheet to provide a visual rendering of their design scheme using pictures and notes. Additionally, they were required to describe their design solutions in words on the right side of the answer sheet.

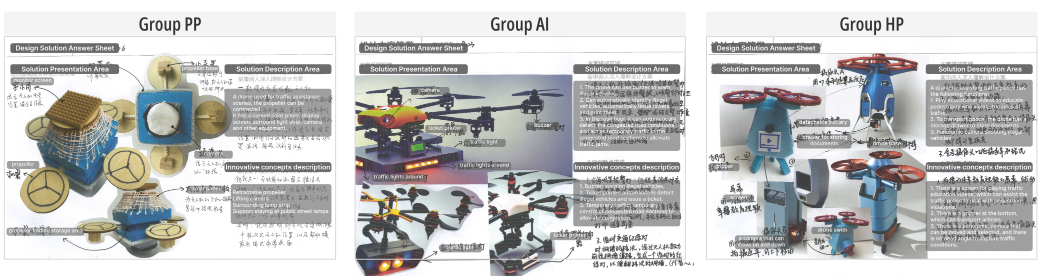

Fig. 8. Examples of design solution answer sheets from three experimental groups. Best viewed magnified on screen.

5.2 Differences in Objective Evaluation by Experts (*RQ1*)

The *utility*, *novelty*, and *feasibility* of 45 design solutions were assessed by six expert evaluators. The statistical analysis results are reported in Table 3 and Figure 9. Specifically, a significant difference was seen in the *utility* ($p = 0.002$). The post hoc test showed that the *utility* of *Group HP* were significantly higher than those of *Group AI* ($p = 0.005$), which indicated that designers produced more functions with the support of hybrid prototype method than the standalone AI tools. In addition, there was a significant difference in *novelty* among the three experimental groups ($p < 0.001$), and the post hoc test indicated that the *novelty* in *Group HP* was greater than *Group PP* ($p = 0.03$) and *Group AI* ($p < 0.001$) while the *novelty* in *Group PP* was also greater than it in *Group AI* ($p < 0.001$). The results showed potential of the hybrid prototype method to increase idea novelty. For the difference in the *feasibility* score, there was a significant difference ($p < 0.001$), and the post hoc test showed that the *feasibility* in *Group PP* and *Group HP* was significantly higher than it in *Group AI* ($p < 0.001$). It indicated that the engagement of physical models enhanced the

Table 3. The statistical result of the objective evaluation by experts among three experimental groups.

	$M(SD)$			*F value*	*P value*	Post hoc
	PP	AI	HP			
Utility	5.98 (1.44)	5.00 (1.25)	7.13 (1.98)	6.79	**0.002**	HP > AI
Novelty	0.69 (0.09)	0.49 (0.10)	0.80 (0.06)	50.55	**< 0.001**	HP > PP > AI
Feasibility	2.26 (0.26)	1.73 (0.21)	2.20 (0.16)	27.06	**< 0.001**	HP, PP > AI
Structural Counts	3.20 (1.32)	1.13 (0.74)	3.07 (0.96)	27.06	**< 0.001**	HP, PP > AI

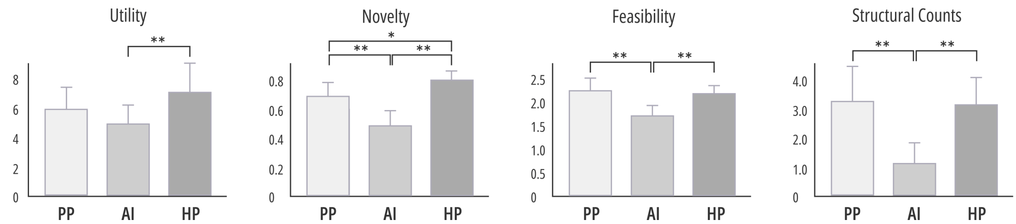

Fig. 9. The difference in the objective evaluation among the three experimental groups. Statistical significance was represented by asterisks (* & **) with p-values less than 0.05 and 0.01, respectively.

technical feasibility of design schemes. Experts also extracted the structures in the scheme, such as the rotating camera, folding solar panel, and telescopic bracket. The count of structures in *Group PP* and *Group HP* was significantly higher than those in *Group AI* ($p < 0.001$).

Besides, in *Group HP*, experts paid special attention to the influence of AI refinement on the structure changes in physical models. In addition to the total number of structures in the scheme, experts respectively focused on the number of structures in the physical model, the number of structures in the generated scheme, and the number of structures that AI successfully understood and refined based on the physical model. Specifically, according to the structural counts in *Group HP*, the structural counts presented by the physical model and the generated image were 2.27 ($SD = 1.1$) and 0.8 ($SD = 0.86$) respectively. In other words, the amount that AI successfully understood the structures in the physical model and refined it into a digital structure accordingly only accounted for 22.11% of total structural counts in *Group HP*. The findings revealed that the incorporation of physical models within the hybrid prototype approach facilitated designers' exploration of structural configurations. However, there were challenges for the hybrid prototype method in comprehending and effectively translating structural demands into digital resolutions.

5.3 Findings of Design Behaviour Differences and Idea Generation Patterns (*RQ2*)

After two researchers reviewed the design process of all participants, the design process was coded. For example, Figure 10 shows the design process of six participants selected randomly from *Group PP*, *Group AI*, and *Group HP*.

Initially, we quantified the total number of design idea-sparking times during the design process. On the whole, a total of 289 idea sparking moments were recorded, including 91 in *Group PP*, 76 in *Group AI*, and 122 in *Group HP*. This finding was consistent with the expert evaluations' *utility*. Moreover, we summarised the idea generation patterns in three experimental groups and reported their proportion (Figure 11). n represents the total number of sparks while p represents its proportion.

For *Group PP*, there were three patterns in the idea generation:

(1). spark an (spontaneous) idea <Spark> ; ($n = 7, p = 7.69\%$)
(2). infer by analysis after physical modelling action <Build-Analyse-Spark>; ($n = 69, p = 75.82\%$)
(3). infer by evaluation and connection with other ideas after modelling action <Build-Evaluate-Spark>. ($n = 15, p = 16.48\%$)

For *Group AI*, there were three patterns in the idea generation:

(1). spark an (spontaneous) idea <Spark>; ($n = 6, p = 7.89\%$)
(2). infer by analysis after observing generated outputs <Generate-Analyse-Spark>; ($n = 22, p = 28.95\%$)

(3). infer by evaluation and connection with other ideas after observing generated outputs <Generate-Evaluate-Spark>. ($n = 48, p = 63.16\%$)

For *Group HP*, as *Group HP* combined the interactive mode and creativity support ability of the other two groups, the idea generation patterns under *Group PP* and *Group AI* could also be found in *Group HP*. Besides, two additional idea generation patterns were also found. Specifically:

(1). spark an (spontaneous) idea <Spark>; ($n = 8, p = 6.56\%$)
(2). infer by analysis after physical modelling action <Build-Analyse-Spark>; ($n = 9, p = 7.38\%$)
(3). infer by evaluation and connection with other ideas after modelling action <Build-Evaluate-Spark>; ($n = 8, p = 6.56\%$)
(4). infer by analysis after observing generated outputs <Generate-Analyse-Spark>; ($n = 36, p = 29.51\%$)
(5). infer by evaluation and connection with ideas after observing generated outputs <Generate-Evaluate-Spark>; ($n = 16, p = 13.11\%$)
(6). infer by analysis after observing generated outputs based on physical model <Build-Generate-Analyse-Spark>; ($n = 27, p = 22.13\%$)
(7). infer by evaluation and connection with other ideas after observing generated outputs based on physical model <Build-Generate-Evaluate-Spark>. ($n = 18, p = 14.75\%$)

To further gain insight into the interplay between idea generation and development, we superimposed the moments when new ideas were sparked and visualised a diagram overlaying moments of idea emergence across all three experimental groups (Figure 12). According to the frequency of idea generation, we indicated the design stage with intensive idea generation on the time axis, which was defined as the *creative surge stage* on the diagram. We unexpectedly found the difference in *creative surge stage* among three experimental groups. The *creative surge stage* occurred in the middle and late stages of the whole design process with the support of the physical prototype method. While the *creative surge stage* occurred in the early stage of design with the support of standalone AI tools, with idea generation occurring as early as the third minute of the design task. There was a relatively longer *creative surge stage* with the support of the hybrid prototype method in *Group HP*.

We also focused on the iteration and development of ideas. Two researchers coded all participants' idea iteration, dividing ideas into new ideas and iterated ideas based on the method proposed by Kim et al. [42] (Figure 13 presents idea iteration modes of six participants for examples). The proportion of iterated ideas in the total ideas was 51.48% ($SD = 12.63\%$), 21.97% ($SD = 20.84\%$),

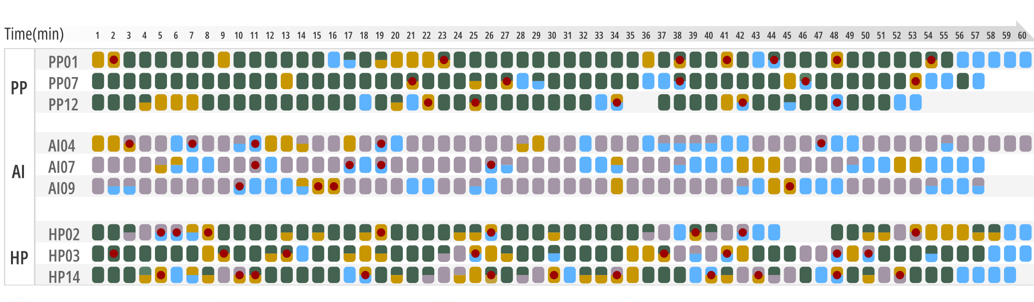

Fig. 10. The design process of six participants representative annotated by the Think-aloud protocol and Howard's model. The blank denotes a participant who remains silent and does not perform any action during an interval.

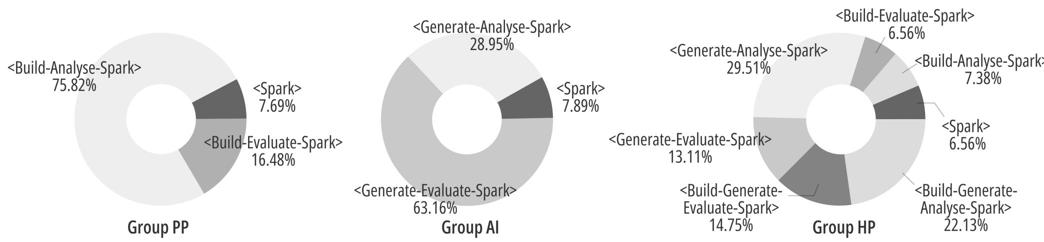

Fig. 11. The proportion of idea generation patterns among three experimental groups.

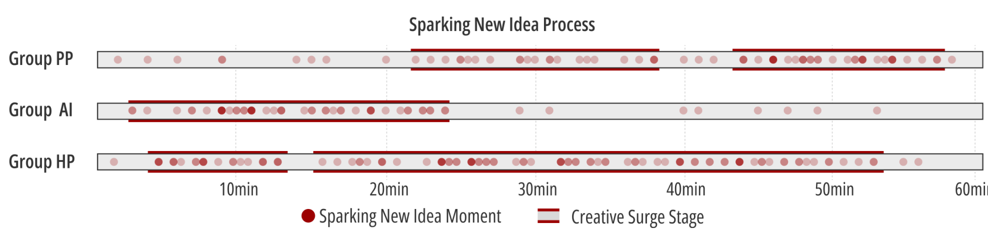

Fig. 12. The idea emergence visualisation during 60-min design task. The red dots represent the time of idea emergence.

and 56.88% (SD = 9.73%) in *Group PP*, *Group AI*, and *Group HP* respectively. We found that the proportion of new ideas was high in *Group AI*, which indicated that the generative AI provided a lot of new inspiration stimulation during the conceptual design process. While the proportion of iterated ideas was high in *Group PP* and *Group HP*, highlighting the participation of physical models in the hybrid prototype method that promoted the iteration and development of original ideas. The influence of the hybrid prototype method on idea generation and iteration will be discussed in detail in *Section* 6.3.2.

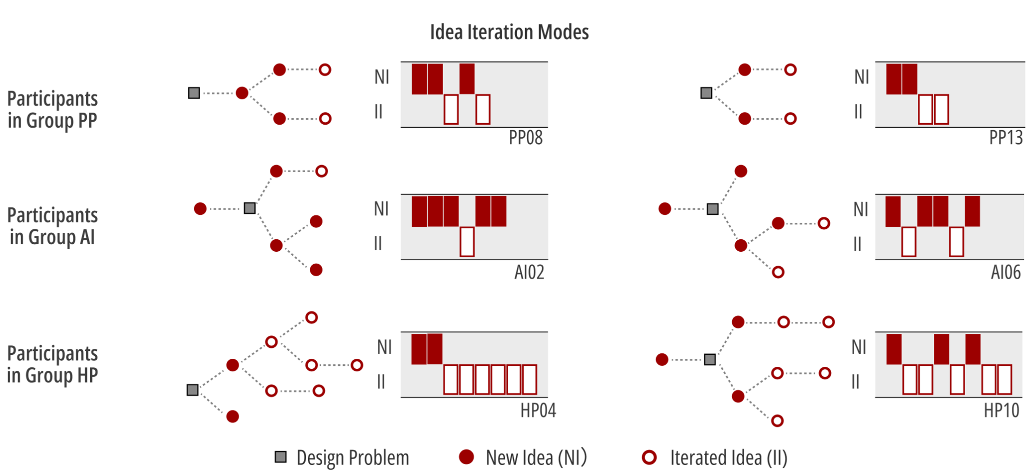

Fig. 13. The idea iteration modes of six participants under different creativity support tools.

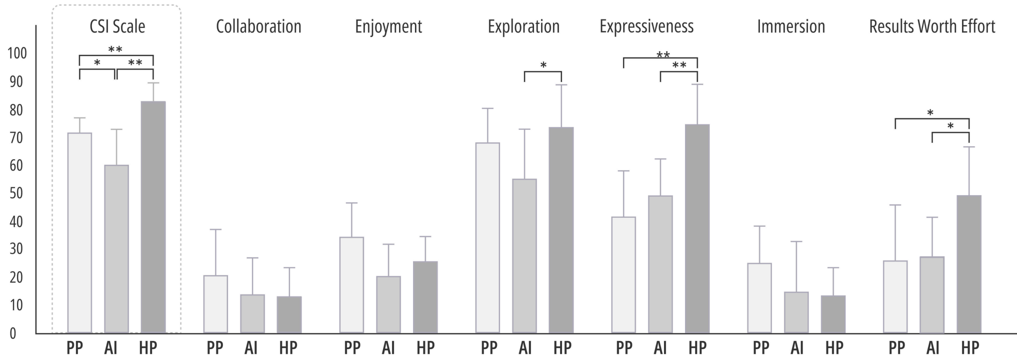

Fig. 14. The difference in the CSI among the three experimental groups. Statistical significance is represented by asterisks (* & **) with p-values less than 0.05 and 0.01, respectively.

5.4 Differences in Subjective Evaluation for Creativity Support (*RQ1*)

The total *CSI scor*e and six factors of CSI were analysed (Table 4 and Figure 14). The *CSI score* of three experimental groups was 72.03 (*Group PP*), 60.43 (*Group AI*), and 83.13 (*Group HP*) respectively. The statistical analysis showed that there was a significant difference in total *CSI score* ($p < 0.001$). The post hoc test indicated that the *CSI score* of *Group HP* was significantly higher than *Group PP* ($p = 0.002$) and *Group AI* ($p < 0.001$). The results proved the effectiveness of the hybrid prototype method in creativity support from a subjective point of view.

We further broke down the analysis and investigated six individual factors. Specifically, there was a significant difference in the score of *exploration* ($p = 0.032$), *expressiveness* ($p < 0.001$) and *result worth effort* ($p = 0.009$). For the *exploration*, the post hoc test showed that the *exploration* score in *Group HP* was higher than it in *Group AI* ($p = 0.032$). Besides, the *expressiveness* score in *Group HP* was greater than it in *Group PP* ($p < 0.001$) and *Group AI* ($p = 0.002$). For the *result worth effort*, the score of *result worth effort* in *Group HP* was greater than it in *Group PP* ($p = 0.037$) and *Group AI* ($p = 0.02$). Besides, the physical prototype method obtained higher scores in *collaboration*, *enjoyment*, and *immersion*, although there was not a significant difference.

Table 4. The statistical result of the CSI among three experimental groups.

	$M(SD)$			F value	P value	Post hoc
	PP	AI	HP			
CSI Score	72.03 (5.23)	60.43 (12.63)	83.13 (6.69)	16.70	**< 0.001**	HP > PP, AI
Collaboration	20.40 (15.92)	13.60 (12.65)	12.80 (10.16)	1.01	0.377	/
Enjoyment	34.30 (22.43)	20.30 (11.25)	25.50 (8.90)	16.60	0.222	/
Exploration	68.50 (11.90)	55.50 (17.48)	73.80 (15.24)	3.91	**0.032**	HP > AI
Expressiveness	42.10 (16.20)	49.70 (12.75)	74.90 (14.48)	13.93	**< 0.001**	HP > PP, AI
Immersion	25.00 (12.97)	14.90 (17.57)	13.60 (9.82)	2.04	0.15	/
Result of Effort	25.80 (19.71)	27.30 (13.83)	48.80 (17.18)	5.68	**0.009**	HP > PP, AI

5.5 Findings of Semi-structured Interview (*RQ3*)

We invited all participants to semi-structured interviews, asking for reflections on their user experience in the conceptual design. Then, we extracted 11 and 8 codes for the advantages and disadvantages of the supported tool respectively (Figure 15 and 16).

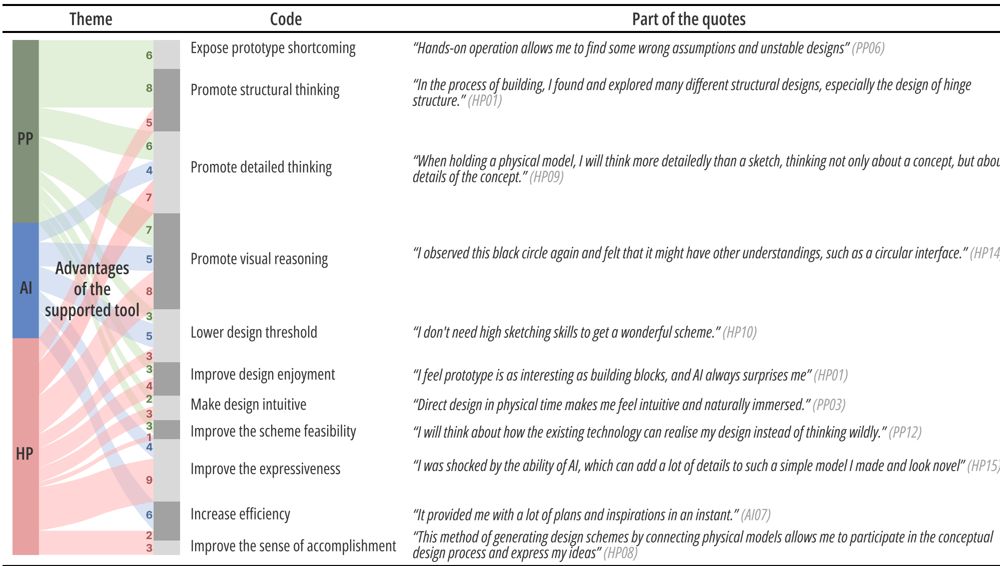

Fig. 15. 11 codes and part of quotes extracted from the semi-structured interview. The theme is focused on advantages of supported tools in conceptual design.

For *Group PP*, the use of physical models was shown to offer distinct advantages in promoting detailed thinking, promoting visual reasoning, and exposing prototype shortcomings. Eight participants in *Group PP* mentioned that interacting with tangible materials promoted structural thinking and six participants indicated the physical prototype exposed prototype shortcomings. For example, *PP08* reported that "*When prototyping with physical models, I not only produce new concepts but also think about how to implement and realise my concepts through physical materials, which helps me eliminate unrealistic ideas in the early design stage*". Other strengths provided by the physical prototype method were also extracted, such as promoting detailed thinking and visual reasoning. However, some shortcomings of the physical prototype were also presented. Many designers complained that physical model construction took too much time and energy for designing appearance. *PP09* said "*I've had many ideas, but it's time-consuming to implement them with physical materials, especially the design of some complex surfaces*". Besides, designers' creativity could be limited by provided materials and manufacturing processes, leading their attention often to be focused on physical model making (e.g., material searching or connecting) rather than the ideation process. For example, *PP04* reported that "*The physical model helps me think in the early stage, but in the later stage, because of its rough expressiveness, it limits me to further optimise the appearance of my design scheme. I can only imagine sticking skin on the rough physical model*".

For *Group AI*, designers mainly reported the advantages of generative AI tools in promoting visual reasoning, improving design efficiency, and enhancing expressiveness. Specifically, *AI05* said "*AI generated many unexpected results for me, giving me many new inspirations or encouraging me to associate generated artefacts with other schemes*". In addition, six designers thought that AI participation improved the prototype efficiency. *AI07* indicated that "*It provided me with a lot of plans and inspirations in an instant, which greatly accelerated the efficiency of my prototype and iteration*". However, the disadvantages of generative AI tools in design assistance were also obvious. Eleven participants in *Group AI* mentioned that generative AI tools had poor generation controllability

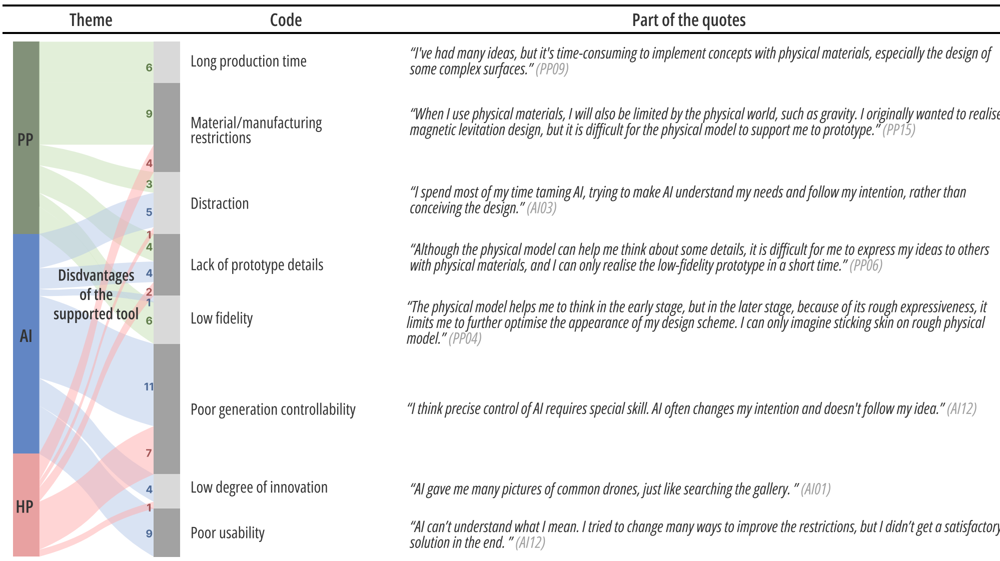

Fig. 16. 8 codes and part of quotes extracted from the semi-structured interview. The theme is focused on disadvantages of supported tools in conceptual design.

and nine noted poor usability. *AI12* said “*I hope to design a drone with a gripper to help grab ground objects, but AI couldn't understand what I meant. I tried to change many ways to convey intentions, but I didn't get a satisfactory solution. I had to add some sketches based on AI-generated images to express my solution*”. *AI12* also complained that “*I think precise control of AI requires special skills. AI often changes my intention and doesn't design according to my idea*”. Moreover, some designers reported that although AI brought creativity, this inspiration stimulation without design logic destroyed the designer's thoughts and disrupted the design rhythm. For example, *AI14* mentioned “*AI will give me colour-matching ideas, functional ideas, and structural ideas randomly. This unpredictable inspiration distracts me and I can't concentrate on my design process*”. Due to the AI interference with original ideas, four designers reported that they did not think they were the author of the design outcome.

For *Group HP*, the hybrid prototype method gained some common advantages of physical model and generative AI tools due to its integration characteristics, such as promoting structural thinking, detailed thinking, and visual reasoning, as well as improving prototype efficiency. Additionally, the hybrid prototype method presented unique advantages in several aspects. First, nine participants agreed that the hybrid prototype method improved the prototype expressiveness. One of the participants mentioned that “*I was shocked by the ability of AI, which can add a lot of details to such a simple model I made and look novel*” (*HP15*). *HP01* also indicated that “*AI's refined schemes turned my low-fidelity prototype into a high-fidelity scheme that can be presented to others*”. Second, some participants mentioned that the hybrid method lowered the design threshold and improved the accomplishment sense. *HP08* reported that “*I am a technical designer, and my sketching ability is very poor. This method allows me to participate in the conceptual design process and express my ideas, instead of having to cooperate with the designer in charge of modelling. I think this greatly improves my sense of accomplishment in design, and it is also more convenient for me to express my design schemes to my team and improve the efficiency of team communication*”. Third, some participants considered that the hybrid method enabled AI to follow their own ideas due to the detailed intention communication, thereby improving cooperation satisfaction. However, as the hybrid method integrated traditional physical methods and generative AI tools, the hybrid prototype method also inherited part of their inherent shortcomings, such as poor generative controllability and manufacturing restrictions.

6 DISCUSSION

6.1 Effectiveness of Hybrid Prototype Method for Creativity Support in Conceptual Design (*RQ1*)

In the objective evaluation of experts, the *utility* and *novelty* of design solutions in *Group HP* were higher than those in *Group AI* and *Group PP*. It indicated that the hybrid prototype method promoted their ability to identify, extract, and discover valuable information, thereby enhancing design innovation. Our results were consistent with views in previous studies. Viswanathan and Linsey [83] indicated that physical models increased solution quantity. Yang et al. [69] pointed out that AI can generate a steady stream of references to stimulate inspiration, ease design fixation, and expand creativity boundaries, which might lead to higher *novelty* scores. In addition, the structural counts illustrated that the participation of physical models in the hybrid prototype method promoted structural thinking during the ideation process, which might be the reason for the higher technical feasibility in *Group HP* than *Group AI*.

The subjective evaluation of designers evidenced a significantly higher total *CSI score* in *Group HP*, which also proved that the hybrid prototype method was valuable in improving creativity in conceptual design. Specifically, the hybrid prototype method exhibited significant effectiveness in *exploration*, *expressiveness*, and *results worth effort* aspects in the CSI questionnaire. It indicated that AI refinement alleviated the low-fidelity physical prototype and enhanced the expressiveness, reducing prototype workload and heightening creation satisfaction. These results proved that the hybrid prototype method enabled designers to explore extensively and articulate their ideas effectively during conceptual design.

6.2 Mechanism of Idea Generation in Hybrid Prototype Method (*RQ2*)

We examined idea generation patterns in *Group HP* to elucidate the creativity support mechanism of the hybrid prototype method. We identified six support patterns except for the spontaneous idea (<Spark>), as presented in Figure 17ⓐ. First, interaction with physical materials was a key source of idea sparking (i.e., the pattern of <Build-Analyse-Spark>). This tangible interaction assisted in identifying flawed design assumptions [10, 27], broadening perspectives [54], and introducing design constraints [5], which in turn promoted an evaluation phase instrumental in ideation (i.e., the pattern of <Build-Evaluate-Spark>). Second, designers also obtained inspiration by reinterpreting generated design schemes (i.e., <Generate-Analyse-Spark>). In the co-creation system, the role of the human designer transcends a mere leader in the AI-aided design process [62]. Designers can also be evaluators, obtaining insights by evaluating and comparing the generated schemes (i.e., <Generate-Evaluate-Spark>). Third, designers derived inspiration via observation of AI's refinement of physical models (i.e., <Build-Generate-Analyse-Spark> and <Build-Generate-Evaluate-Spark>). In these patterns, designers analysed relevant schemes generated based on physical models, activating nearby semantic concepts [9]. They also evaluated a broader scope of visual cues in these schemes, promoting unexpected discoveries [82]. In addition, we present a whole idea generation process supported by the hybrid prototype methods in Figure 17ⓑ (randomly selected from *Group HP*). With the support of the hybrid prototype method, these aforementioned idea generation patterns occurred in a stochastic manner throughout the iterative process of conceptual design.

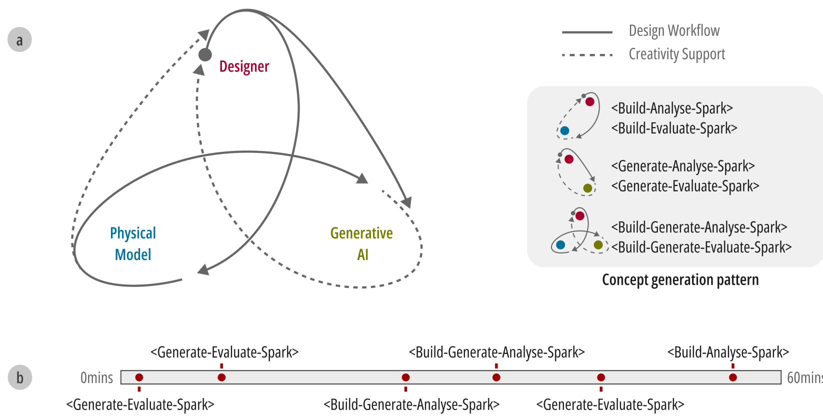

Fig. 17. The patterns of idea generation with the support of the hybrid prototype method (a). An example of idea generation scenario in *Group HP* to show the different patterns in conceptual design process (b).

We examined the idea generation frequency in the conceptual design workflow. Notably, our results showed that the creative surge stage differed among three experimental groups. The creative surge stage mainly occurred in the middle and late stages of the design process when physical models were employed independently. In contrast, standalone generative AI facilitated early-stage creativity. This discrepancy might be attributed to the distinct creative stimulation patterns. With physical prototypes, primary idea generation patterns were <Build-Analyse-Spark> (75.82%) and <Build-Evaluate-Spark> (16.48%), wherein designers manipulated tangible models and deconstructed prototypes to rediscover, recognise, and transform them [31, 77]. The behaviour analysis revealed that the initial 20 minutes of the design process were mainly used to build the preliminary physical model, including cutting out the overall shape and connecting the basic components, with limited analysis and evaluation. However, once established, designers can achieve more effective idea generation through in-depth interaction with detailed structures and components [77]. Conversely, utilising generative AI tools led to rapid inspiration through <Generate-Analyse-Spark> and <Generate-Evaluate-Spark> patterns. Initial content generation provided designers with fresh perspectives and divergent thinking stimulation rapidly [69]. This might explain the accelerated idea generation in *Group AI*. Karimi et al. [40] studied how an AI-supported creative partner can inspire designers while sketching and reported that the exploratory AI was more helpful in the earlier stages, which was consistent with our finding. The hybrid prototype method amalgamates the workflow of physical prototype and generative AI, enabling quick early-stage inspiration via <Generate-Analyse/Evaluate-Spark>, followed by idea expansion through <Build-Analyse/Evaluate-Spark> upon developing a physical model. This hybrid idea generation pattern possibly leads to the prolonged creative surge stage in *Group HP*.

6.3 Influence of AI on Creativity Support in Co-creation System

6.3.1 The contribution of AI to hybrid prototype method. We investigated the AI contribution to the hybrid prototype method by interviewing participants and analysing design outcomes. First, generative AI mainly serves as the role of an inspiration stimulator in the hybrid prototype method [93]. We categorised designers' inspiration sources from AI-generated artefacts into two types: *inspiration from concrete details* and *inspiration from visual reasoning based on abstract patterns*. Figure 18ⓐ illustrates how generated artefacts can inspire human designers through

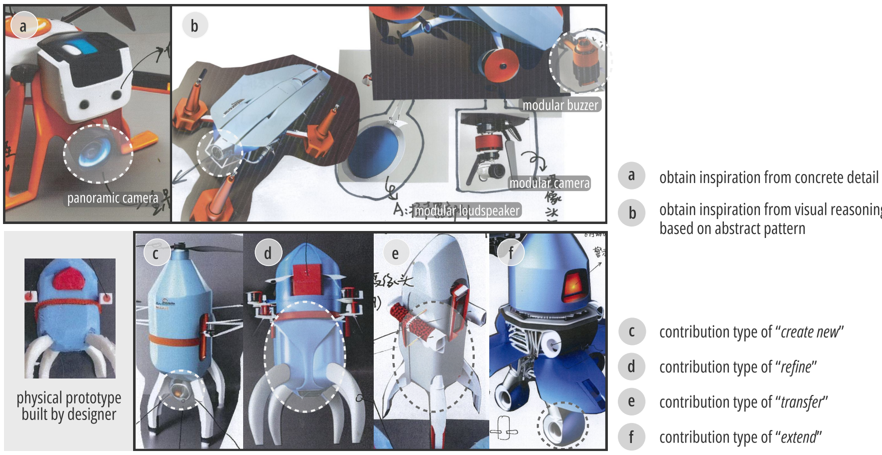

Fig. 18. Examples selected from participants' design solutions to show the AI's contribution.

concrete design details directly. In addition, designers can also get inspiration from abstract and vague elements through visual reasoning based on their own experience and understanding [74]. In Figure 18ⓑ, the designer reported that "*The first time I observed the results generated by AI, I found that there was a black circle in front of the drone, and I thought it could be used as a front camera. Later, I observed this black circle again and felt that it might have other understandings, such as a circular interface. Finally, I designed a modular interface at the front, which can plug in various functional modules, such as cameras, speakers, and so on*". Therefore, ambiguous artefacts can be re-understood and re-interpreted by self-explanation, thereby stimulating creativity [81].

Previous studies clarified different contribution ways during co-creation and delineated the cooperative contribution as *create new*, *refine*, *extend*, and *transform* [6, 62]. In our experiment, all four contribution types can be found in AI-inspired design solutions. Figure 18 exemplifies these types. *Create new*: AI augmented a physical prototype by adding a camera at the bottom of the drone (Figure 18ⓒ). *Refine*: AI enriched a simple physical model with additional design information, enhancing details and aesthetics (Figure 18ⓓ). *Extend*: AI introduced new interpretations, altering the original colour scheme (Figure 18ⓔ). *Transform*: AI expanded upon the designer's idea, for example by adding wheels to the base of the drone to support its movement (Figure 18ⓕ).

6.3.2 The influence of AI on the idea development and iteration in hybrid prototype method. We analysed the idea iteration modes and our results indicated that the utilisation of standalone AI tools assisted designers in generating a diverse array of new ideas but was adverse to deepening and iterating on original ideas. Although the generative AI is good at opening an inspiration space due to generative variability [86], the end-to-end generative mode might provide random and unpredictable inspiration stimulation, which could break designers' original thinking and disturb their rhythm. For example, in *Group AI*, AI provided irrelevant independent ideas through multiple generations, such as "*Large capacity battery & Lid kit & Screen*". While it might inspire brainstorming, the combination of AI randomly generated ideas did not constitute a complete design scheme. However, the hybrid prototype method helped AI to follow designers' original ideas due to more accurate intention conveyance. In the hybrid prototype method, the physical model served as the anchor point and shared base of the original idea for AI's generation and designer's reasoning. Designers were supported to use physical material to externalise ambiguous concepts and obtain iterated inspiration through AI refinement, such as "*Solar panel & Solar panel with lamp strip & Foldable solar panel*", which protected designers' original ideas and promoted idea iteration.

We also explored the AI influence on original ideas with the combination of designers' ownership for the co-creation design scheme. Specifically, we asked designers' subjective perception of ownership, as gleaned from interview questions like "*Do you think you are the author of this design outcome*". Participants reported a diminished sense of ownership under the pure support of AI tools. Some of them pointed out that "*because it is difficult to describe my ideas clearly in pure text, the generated schemes are quite different from my text requirements and lack my original ideas*". Contrarily, participants in *Group HP* indicated they had a higher awareness of ownership because the main ideas in the co-creation outcome originated from themselves. They considered that they maintained creative dominance during the prototype process, which was consistent with the results of idea iteration modes that the number of iterated ideas was greater than that of new ideas with the support of the hybrid prototype method.

6.4 Critical Discussion on Inherent Characteristics and Application Scenarios of Three Prototype Methods (*RQ3*)

Although the findings showed that the hybrid prototype method was effective in general, we discussed the unique strengths and limitations of each method critically and comprehensively based on our experimental results, in an attempt to further highlight the application scenarios and best practices of each prototype method.

The hybrid prototype method presents a range of distinct benefits, including providing long-term creative support, enhancing the prototype expressiveness, and lowering the threshold of design. Therefore, we consider that the hybrid prototype method can effectively support the design process led by engineers or technicians due to its low threshold and high expressiveness, in addition to providing general conceptual design creativity support. As engineers usually have a deep understanding of structures and techniques but lack expression skills, such as sketching or modelling, the hybrid prototype method can support their rapid ideation and expression, allowing them to apply physical models, or even add physical materials to existing mechanical structures for creation and expression in conceptual design. Similarly, as design is often interdisciplinary teamwork [84], involving designers, engineers, and even customers [2], the hybrid method might also enhance team collaboration and interdisciplinary communication during conceptual design.

For the physical prototype method, the CSI questionnaire results indicated its strengths in design enjoyment and immersion. Besides, participants in *Group PP* also reported distinct advantages of the physical prototype method, especially in promoting structural thinking and exposing shortcomings, which was consistent with the expert evaluation results. Owing to the distinct merits of the physical prototype method, we contend that it is highly suitable for some conventional industrial design scenarios characterised by mechanical structures and simple appearances. The structural feasibility might be enhanced significantly through iterative examination of the physical structure manually, fostering a higher sense of immersion compared to sketching or digital prototyping. However, the physical prototype method may not be suitable for appearance-oriented products with complex surfaces, as participants have expressed concerns regarding its notable drawbacks, such as time-consuming and manufacturing limitations.

For the standalone AI tool, the behavioural analysis revealed that it effectively and intensively inspired designers at the beginning of the creative process. Similarly, interview results confirmed that AI participation significantly accelerated the inspiration acquisition. Consequently, we believe that standalone AI tools are well-suited for facilitating divergent thinking during the initial conceptual design phase, such as brainstorming. Nonetheless, considering the limited AI availability and the randomness observed in our study, we do not recommend relying solely on AI for the evolution and iteration of ideas.

6.5 Reflection and Implication of Hybrid Prototype Method

6.5.1 The optimisation space of hybrid prototype method. In light of our findings, we delineate the optimisation space of the hybrid prototype method. On the one hand, generative AI in the hybrid method should command the structural knowledge, enabling the structural design to be successfully refined based on the physical model. Although the participation of the physical model has been proven to promote structural thinking in conceptual design, our expert evaluation results showed that the number of structures that were successfully recognised, understood, and transformed by AI based on physical models only stood at 22.11% of the total number of structures. The result indicated that many structures in physical models were misunderstood and ignored after the AI refinement. This result is expected because our study integrates an image-based model, which is better at providing visual creative stimulation but lacks structural knowledge. However, the gap could be addressed with the development of multi-modal vision language models, such as GPT-4V [57], Gemini 1.5 [56], and Claude 3 [15]. This might enhance AI's ability in image understanding and reasoning, as well as the association of image and language. Large multi-modal models have proven the potential in engineering design reasoning, including spatial reasoning, material inspection, manufacturability assessment, structural interpretation, and so on [59]. Previous studies also explored the ability of GPT to assist with converting text into 2D engineering schemes, even into 3D models [51, 55]. In this context, further optimisation involves integrating large multi-modal models into the hybrid prototype method, thus enabling the AI to better understand and reason in engineering design, thereby improving the feasibility of the generated physical structure.

On the other hand, AI should provide inspiration stimulation that follows design logic and design knowledge. Our results indicated that a single general-purpose generative model limited the idea development and iteration, though its random generation stimulated inspiration. Although the hybrid prototype method alleviated this problem through more concrete intention communication, the unpredictable generation may still violate designers' original ideas and destroy their design rhythm. For example, AI generated details while the designer was engrossed in shaping the overall form and proportion, or it modified the overarching style during the designer's final refinement stage. Interviews revealed that experienced designers exercised greater control over parameters and prompts or ignored unreasonable generated outputs in order to avoid the influence of randomness on ideation. In contrast, novice designers may be swayed by AI's randomness, leading to more chaotic and disordered thinking patterns. To alleviate these challenges, integrating AI with design knowledge and logic is necessary for further optimisation. Therefore, fusing design cognitive knowledge (e.g., the theory of divergent and convergent thinking [79]) and design methodology (e.g., such as FBS framework [24]) with data-enabled generative models might have the potential to enable AI to follow design logic instead of unpredictable generation. Moreover, the generative model can also be fine-tuned based on the designer's behaviour coding and modelling, allowing the AI to align with the designer's procedural approach, thought patterns, and behavioural tendencies.

6.5.2 The implication for actual design practices. Our research delineates the potential of the hybrid prototype method to foster creativity, as evidenced by empirical data and designers' feedback. First, the hybrid prototype method can be used to navigate designers from ambiguity, pointing the way forward. In the early stage of conceptual design, the original ideas in designers' minds are often vague, unclear, and abstract, which are challenging to externalise through language alone [69]. The physical prototype method facilitates rapid externalisation of these fledgling concepts, while generative AI further refines and substantiates them [54]. It contributes to navigating designers from abstract and ambiguous ideas, providing a forward direction through embodied interaction and generative variability. Second, the hybrid prototype method lets AI take over the refined expression task, alleviating the physical material limitation. Design expression via traditional physical prototypes is often seriously limited by the availability and characteristics of materials. The data-driven enhancements provided by the hybrid prototype method alleviate visual representation challenges, such as the expression of style, colour, and texture [36, 69]. Consequently, designers may lessen their concerns with material constraints during the prototyping phase. Third, the hybrid prototype method provides designer-centred interactive modes, empowering designers to cooperate with AI in their familiar design language. AI research has delved into the exploration of generative models under diverse human-AI communication modes, including semantic map [11] and bounding box [92]. However, end-users are still required to conform to computer-centred rules to interact with AI, such as designating semantics according to specified colours or editing prompt words in a structured way. With the support of the hybrid prototype method, designers can use their familiar design materials and design language to express their design intentions intuitively, without the translation between design language and computer language.

6.6 Limitations and Future Work

In this section, we discuss the technical limitations that need to be addressed in future research. First, the implemented system simplified the interactive parameters to reduce the complexity of AI tools, potentially affecting the AI's creative support due to diminished controllability. Some advanced generative models like ControlNet [91] can be applied in the hybrid prototype method for more precise control. Second, AI primarily serves as a generator and stimulator [93] in this study. Future research aims to expand the AI's role to include design evaluation and active participation, akin to co-designers, thus augmenting its contribution in the hybrid prototype method [21]. Third, the hybrid prototype method employs a preliminary system capturing physical models via camera, which faces limitations such as shooting angle and image quality. Future research will investigate efficient data transmission techniques, including sensor-equipped physical prototype components for direct three-dimensional data capture.

7 CONCLUSION

In this study, we proposed a hybrid prototype method combining physical models and generative AI to support creativity in the conceptual design stage. We conducted a comparative experiment with 45 designers who were requested to complete the same design task with the traditional physical prototype, standalone generative AI tool, and the hybrid prototype method. The results confirmed the effectiveness of the hybrid prototype method in supporting creativity. Specifically, the hybrid prototype method was beneficial for enhancing utility and design novelty, boosting structural exploration and design expressiveness, and accelerating prototype iteration. We discussed AI participation in the design process and clarified the implication of the hybrid prototype method for actual design practices. Our work can be used to further understand collaborative design with human designers and generative AI in the design process.

ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China (2022YFB3303301). We would like to thank Mingyan Wang for assistance in figure production.

REFERENCES

[1] Obead Alhadreti and Pam Mayhew. 2018. Rethinking Thinking Aloud: A Comparison of Three Think-aloud Protocols. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–12. https://doi.org/10.1145/3173574.

3173618
[2] Remi Alkemade, Fons J. Verbeek, and Stephan G. Lukosch. 2017. On the Efficiency of a VR Hand Gesture-Based Interface for 3D Object Manipulations in Conceptual Design. *International Journal of Human–Computer Interaction* 33, 11 (2017), 882–901. https://doi.org/10.1080/10447318.2017.1296074
[3] AUTOMATIC1111. 2023. Stable Diffusion WebUI. https://github.com/AUTOMATIC1111/stable-diffusion-webui. Last accessed on Jun 29, 2023.
[4] Loris Barbieri, Agostino Angilica, Fabio Bruno, and Maurizio Muzzupappa. 2013. Mixed Prototyping with Configurable Physical Archetype for Usability Evaluation of Product Interfaces. *Computers in Industry* 64, 3 (2013), 310–323. https://doi.org/10.1016/j.compind.2012.11.010
[5] Anders Berglund and Larry Leifer. 2013. Why we Prototype! An International Comparison of the Linkage between Embedded Knowledge and Objective Learning. *Engineering Education* 8, 1 (2013), 2–15. https://doi.org/10.11120/ened.2013.00004
[6] Margaret A Boden. 1998. Creativity and Artificial Intelligence. *Artificial Intelligence* 103, 1-2 (1998), 347–356. https://doi.org/10.1016/S0004-3702(98)00055-1
[7] Annette Isabel Bohmer, Sheri Sheppard, Liza Kayser, and Udo Lindemann. 2017. Prototyping as a Thinking Approach in Design Insights of Problem-solving Activities While Designing a Product. In *2017 International Conference on Engineering, Technology and Innovation (ICE/ITMC)*. IEEE, 955–963. https://doi.org/10.1109/ICE.2017.8279985
[8] Johan J Broek, Woulter Sleijffers, Imre Horváth, and Alex F Lennings. 2000. Using Physical Models in Design. In *Proceedings of CAID/CD'2000 conference*. Citeseer, 155–163.
[9] Vincent R Brown and Paul B Paulus. 2002. Making Group Brainstorming More Effective: Recommendations from an Associative Memory Perspective. *Current Directions in Psychological Science* 11, 6 (2002), 208–212.
[10] Louis L Bucciarelli. 2002. Between Thought and Object in Engineering Design. *Design Studies* 23, 3 (2002), 219–231.
[11] Pei Chen, Yangkang Zhang, Zejian Li, and Lingyun Sun. 2022. Few-Shot Incremental Learning for Label-to-Image Translation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 3697–3707.
[12] Ruijia Cheng, Ruotong Wang, Thomas Zimmermann, and Denae Ford. 2022. " It would work for me too": How Online Communities Shape Software Developers' Trust in AI-Powered Code Generation Tools. *arXiv preprint arXiv:2212.03491* (2022). https://doi.org/10.48550/arXiv.2212.03491
[13] Erin Cherry and Celine Latulipe. 2014. Quantifying the Creativity Support of Digital Tools through the Creativity Support Index. *ACM Transactions on Computer-Human Interaction* 21, 4 (2014), 1–25. https://dl.acm.org/doi/10.1145/2617588
[14] John Joon Young Chung and Eytan Adar. 2023. PromptPaint: Steering Text-to-Image Generation Through Paint Medium-like Interactions. *arXiv preprint arXiv:2308.05184* (2023).
[15] Anthropic Company. 2023. Gemini 1.5. https://claude.ai.
[16] Midjourney Company. 2022. Midjourney. https://www.midjourney.com.
[17] Sebastiaan De Peuter, Antti Oulasvirta, and Samuel Kaski. 2023. Toward AI Assistants that Let Designers Design. *AI Magazine* 44, 1 (2023), 85–96.
[18] Biplab Deka, Zifeng Huang, Chad Franzen, Joshua Hibschman, Daniel Afergan, Yang Li, Jeffrey Nichols, and Ranjitha Kumar. 2017. Rico: A Mobile App Dataset for Building Data-Driven Design Applications. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. ACM, 845–854. https://doi.org/10.1145/3126594.3126651
[19] Laura Devendorf and Daniela K Rosner. 2017. Beyond Hybrids: Metaphors and Margins in Design. In *Proceedings of the 2017 Conference on Designing Interactive Systems*. 995–1000. https://doi.org/10.1145/3064663.3064705
[20] Graham Dove, Kim Halskov, Jodi Forlizzi, and John Zimmerman. 2017. UX Design Innovation: Challenges for Working with Machine Learning as a Design Material. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 278–288. https://doi.org/10.1145/3025453.3025739
[21] Özgür Eris et al. 2003. Asking Generative Design Questions: A Fundamental Cognitive Mechanism in Design Thinking. In *DS 31: Proceedings of ICED 03, the 14th International Conference on Engineering Design*. 587–588.
[22] Rodrigo Bastos Fernandes and André Ogliari. 2018. Enhancing Creativity Through Biological Stimuli During New Products Ideation. *International Journal for Innovation Education and Research* 6, 10 (2018), 332–350. https://doi.org/10.31686/ijier.vol6.iss10.1200
[23] Anthony Fu. 2023. sd-webui-kitchen-theme. https://github.com/antfu/sd-webui-kitchen-theme. Last accessed on Jun 29, 2023.
[24] John S Gero and Udo Kannengiesser. 2004. The Situated Function–Behaviour–Structure Framework. *Design Studies* 25, 4 (2004), 373–391. https://doi.org/10.1016/j.destud.2003.10.010
[25] Elisa Giaccardi. 2019. Histories And Futures Of Research Through Design: From Prototypes To Connected Things. *International Journal of Design* 13, 3 (2019), 139–155.

[26] Elisa Giaccardi and Johan Redström. 2020. Technology and More-Than-Human Design. *Design Issues* 36, 4 (2020), 33–44. https://doi.org/10.1162/desi_a_00612

[27] Carolina Gill, Elizabeth Sanders, and Scott Shim. 2011. Prototypes as Inquiry, Visualization and Communication. In *DS 69: Proceedings of E&PDE 2011, the 13th International Conference on Engineering and Product Design Education, London, UK, 08.-09.09. 2011.*

[28] Frederic Gmeiner, Humphrey Yang, Lining Yao, Kenneth Holstein, and Nikolas Martelaro. 2023. Exploring Challenges and Opportunities to Support Designers in Learning to Co-create with AI-based Manufacturing Design Tools. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–20. http://arxiv.org/abs/2303.00192

[29] Gabriela Goldschmidt and Maria Smolkov. 2006. Variances in the Impact of Visual Stimuli on Design Problem Solving Performance. *Design Studies* 27, 5 (2006), 549–569. https://doi.org/10.1016/j.destud.2006.01.002

[30] Connie Golsteijn, Elise van den Hoven, David Frohlich, and Abigail Sellen. 2014. Reflections on craft research for and through design. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*. 421–430. https://doi.org/10.1145/2639189.2639194

[31] Matheus Galvão Gomes, André Ogliari, Rodrigo Bastos Fernandes, and Karuliny Oliveira Marques. 2022. Evaluation of physical models as creative stimuli in conceptual design of products. *Design Studies* 81 (2022), 101119. https://doi.org/10.1016/j.destud.2022.101119

[32] Viviane Vasconcellos Ferreira Grubisic and André Ogliari. 2009. Methodology for the Integrated Management of Technical and Managerial Risks Related to the Product Design Process. *Product: Management and Development* 7, 2 (2009), 149–160.

[33] Tim Hess and Joshua D. Summers. 2014. Protocol Analysis: Studying Physical Manipulatives During Conceptual Design. In *Volume 7: 2nd Biennial International Conference on Dynamics for Design; 26th International Conference on Design Theory and Methodology*. American Society of Mechanical Engineers, V007T07A031. https://doi.org/10.1115/DETC2014-34183

[34] T.J. Howard, S.J. Culley, and E. Dekoninck. 2008. Describing The Creative Design Process By The Integration Of Engineering Design And Cognitive Psychology Literature. *Design Studies* 29, 2 (2008), 160–180. https://doi.org/10.1016/j.destud.2008.01.001

[35] Caroline Hummels and Jelle Van Dijk. 2015. Seven Principles to Design for Embodied Sensemaking. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*. 21–28. https://doi.org/10.1145/2677199.2680577

[36] Angel Hsing-Chi Hwang. 2022. Too Late to Be Creative? AI-empowered Tools in Creative Processes. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–9.

[37] Matilde B. Jensen, Christer W. Elverum, and Martin Steinert. 2017. Eliciting Unknown Unknowns with Prototypes: Introducing Prototrials and Prototrial-driven Cultures. *Design Studies* 49 (2017), 1–31. https://doi.org/10.1016/j.destud.2016.12.002

[38] Youngseung Jeon, Seungwan Jin, Patrick C. Shih, and Kyungsik Han. 2021. FashionQ: An AI-Driven Creativity Support Tool for Facilitating Ideation in Fashion Design. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, 1–18. https://doi.org/10.1145/3411764.3445093

[39] Renée Jourdenais, Mitsuhiko Ota, Stephanie Stauffer, Beverly Boyson, and Catherine Doughty. 1995. Does Textual Enhancement Promote Noticing? A Think-aloud Protocol Analysis. *Attention and Awareness in Foreign Language Learning* 183 (1995), 216.

[40] Pegah Karimi, Jeba Rezwana, Safat Siddiqui, Mary Lou Maher, and Nasrin Dehbozorgi. 2020. Creative Sketching Partner: An Analysis of Human-AI Co-creativity. In *Proceedings of the 25th International Conference on Intelligent User Interfaces*. 221–230.

[41] Lee Kent, Chris Snider, James Gopsill, and Ben Hicks. 2021. Mixed Reality in Design Prototyping: A Systematic Review. *Design Studies* 77 (2021), 101046. https://doi.org/10.1016/j.destud.2021.101046

[42] EunJin Kim and KwanMyung Kim. 2015. Cognitive Styles in Design Problem Solving: Insights from Network-based Cognitive Maps. *Design Studies* 40 (2015), 1–38.

[43] Yongkwan Kim, Sang-Gyun An, Joon Hyub Lee, and Seok-Hyung Bae. 2018. Agile 3D Sketching with Air Scaffolding. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 1–12. https://doi.org/10.1145/3173574.3173812

[44] David Kirsh. 2013. Embodied Cognition and the Magical Future of Interaction Design. *ACM Transactions on Computer-Human Interaction* 20, 1 (2013), 1–30. https://doi.org/10.1145/2442106.2442109

[45] Janin Koch, Andrés Lucero, Lena Hegemann, and Antti Oulasvirta. 2019. May AI? Design Ideation with Cooperative Contextual Bandits. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[46] Gay Lemons, Adam Carberry, Chris Swan, Linda Jarvin, and Chris Rogers. 2010. The Benefits of Model Building in Teaching Engineering Design. *Design Studies* 31, 3 (2010), 288–309. https://doi.org/10.1016/j.destud.2010.02.001

[47] Jing Liao, Preben Hansen, and Chunlei Chai. 2020. A Framework Of Artificial Intelligence Augmented Design Support. *Human–Computer Interaction* 35, 5-6 (2020), 511–544. https://doi.org/10.1080/07370024.2020.1733576

[48] Q Vera Liao, Hariharan Subramonyam, Jennifer Wang, and Jennifer Wortman Vaughan. 2023. Designerly Understanding: Information Needs for Model Transparency to Support Design Ideation for AI-powered User Experience. In *Proceedings of the 2023 CHI conference on human factors in computing systems*. 1–21.

[49] Youn-Kyung Lim, Erik Stolterman, and Josh Tenenberg. 2008. The Anatomy of Prototypes: Prototypes as Filters, Prototypes As Manifestations Of Design Ideas. *ACM Transactions on Computer-Human Interaction* 15, 2 (2008), 1–27. https://doi.org/10.1145/1375761.1375762

[50] Vivian Liu and Lydia B Chilton. 2022. Design Guidelines for Prompt Engineering Text-to-Image Generative Models. In *CHI Conference on Human Factors in Computing Systems*. ACM, 1–23. https://doi.org/10.1145/3491102.3501825

[51] Liane Makatura, Michael Foshey, Bohan Wang, Felix HähnLein, Pingchuan Ma, Bolei Deng, Megan Tjandrasuwita, Andrew Spielberg, Crystal Elaine Owens, Peter Yichen Chen, et al. 2023. How Can Large Language Models Help Humans in Design and Manufacturing? *arXiv preprint arXiv:2307.14377* (2023).

[52] David Mathias, Chris Snider, Ben Hicks, and Charlie Ranscombe. 2019. Accelerating Product Prototyping Through Hybrid Methods: Coupling 3D Printing and LEGO. *Design Studies* 62 (2019), 68–99. https://doi.org/10.1016/j.destud.2019.04.003

[53] Studio Morphogen. 2022. Dream by WOMBO. Retrieved June, 2023 from https://www.wombo.art.

[54] Jacob Nelson and Jessica Menold. 2020. Opening the Black Box: Developing Metrics to Assess the Cognitive Processes of Prototyping. *Design Studies* 70 (2020), 100964. https://doi.org/10.1016/j.destud.2020.100964

[55] Matt D Nelson, Brady L Goenner, and Bruce K Gale. 2023. Utilizing ChatGPT to Assist CAD Design for Microfluidic Devices. *Lab on a Chip* 23, 17 (2023), 3778–3784.

[56] OpenAI. 2023. Gemini 1.5. https://deepmind.google/technologies/gemini.

[57] OpenAI. 2023. GPT-4V(ision) system card. https://chat.openai.com/.

[58] Huaishu Peng, Jimmy Briggs, Cheng-Yao Wang, Kevin Guo, Joseph Kider, Stefanie Mueller, Patrick Baudisch, and François Guimbretière. 2018. RoMA: Interactive Fabrication with Augmented Reality and a Robotic 3D Printer. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 1–12. https://doi.org/10.1145/3173574.3174153

[59] Cyril Picard, Kristen M Edwards, Anna C Doris, Brandon Man, Giorgio Giannone, Md Ferdous Alam, and Faez Ahmed. 2023. From Concept to Manufacturing: Evaluating Vision-Language Models for Engineering Design. *arXiv preprint arXiv:2311.12668* (2023).

[60] Aditya Ramesh, Prafulla Dhariwal, Alex Nichol, Casey Chu, and Mark Chen. 2022. Hierarchical Text-conditional Image Generation with Clip Latents. *arXiv preprint arXiv:2204.06125* (2022). https://doi.org/10.48550/arXiv.2204.06125

[61] Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever. 2021. Zero-Shot Text-to-Image Generation. In *Proceedings of the 38th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 139)*, Marina Meila and Tong Zhang (Eds.). PMLR, 8821–8831. https://proceedings.mlr.press/v139/ramesh21a.html

[62] Jeba Rezwana and Mary Lou Maher. 2022. Designing Creative AI Partners with COFI: A Framework for Modeling Interaction in Human-AI Co-Creative Systems. *ACM Transactions on Computer-Human Interaction* (2022), 3519026. https://doi.org/10.1145/3519026

[63] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. 2022. High-resolution Image Synthesis with Latent Diffusion Models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 10684–10695. https://doi.org/10.1101/2022.11.18.517004

[64] Jörgen Sandberg and Haridimos Tsoukas. 2015. Making Sense of the Sensemaking Perspective: Its Constituents, Limitations, and Opportunities for Further Development. *Journal of Organizational Behavior* 36, S1 (2015), S6–S32. https://doi.org/10.1002/job.1937

[65] Prabir Sarkar and Amaresh Chakrabarti. 2011. Assessing Design Creativity. *Design Studies* 32, 4 (2011), 348–383. https://doi.org/10.1016/j.destud.2011.01.002

[66] Larry Sass and Rivka Oxman. 2006. Materializing Design: the Implications of Rapid Prototyping in Digital Design. *Design Studies* 27, 3 (2006), 325–355. https://doi.org/10.1016/j.destud.2005.11.009

[67] Olivier Serrat. 2017. *Knowledge Solutions*. Springer Singapore. https://doi.org/10.1007/978-981-10-0983-9

[68] Xinyu Shi, Ziqi Zhou, Jing Wen Zhang, Ali Neshati, Anjul Kumar Tyagi, Ryan Rossi, Shunan Guo, Fan Du, and Jian Zhao. 2023. De-Stijl: Facilitating Graphics Design with Interactive 2D Color Palette Recommendation. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19.

[69] Yang Shi, Tian Gao, Xiaohan Jiao, and Nan Cao. 2023. Understanding Design Collaboration Between Designers and Artificial Intelligence: A Systematic Literature Review. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–35.

[70] Dorothé Smit, Bart Hengeveld, Martin Murer, and Manfred Tscheligi. 2022. Hybrid Design Tools for Participatory, Embodied Sensemaking: An Applied Framework. In *Sixteenth International Conference on Tangible, Embedded, and Embodied Interaction*. ACM, 1–10. https://doi.org/10.1145/3490149.3501332

[71] Valdeon Sozo and André Ogliari. 2019. Stimulating Design Team Creativity Based on Emotional Values: A Study on Idea Generation in the Early Stages of New Product Development Processes. *International Journal of Industrial Ergonomics* 70 (2019), 38–50. https://doi.org/10.1016/j.ergon.2019.01.003

[72] Stabilityai. 2023. Stable Diffusion v2-1. https://huggingface.co/stabilityai/stable-diffusion-2-1. Last accessed on June 29, 2023.

[73] Erik Stolterman and Mikael Wiberg. 2010. Concept-Driven Interaction Design Research. *Human–Computer Interaction* 25, 2 (2010), 95–118. https://doi.org/10.1080/07370020903586696

[74] Masaki Suwa, Terry Purcell, and John Gero. 1998. Macroscopic Analysis of Design Processes Based on a Scheme for Coding Designers' Cognitive Actions. *Design Studies* 19, 4 (1998), 455–483. https://doi.org/10.1016/S0142-694X(98)00016-7

[75] Jon Swain. 2018. A Hybrid Approach to Thematic Analysis in Qualitative Research: Using a Practical Example. *Sage Research Methods* (2018). https://doi.org/10.4135/9781526435477

[76] Ye Tao, Shuhong Wang, Junzhe Ji, Linlin Cai, Hongmei Xia, Zhiqi Wang, Jinghai He, Yitao Fan, Shengzhang Pan, Jinghua Xu, et al. 2023. 4Doodle: 4D Printing Artifacts Without 3D Printers. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16. https://doi.org/10.1145/3544548.3581321

[77] Christine A. Toh and Scarlett R. Miller. 2014. The Impact of Example Modality and Physical Interactions on Design Creativity. *Journal of Mechanical Design* 136, 9 (2014), 091004. https://doi.org/10.1115/1.4027639

[78] Joanna Tsenn, Olufunmilola Atilola, Daniel A. McAdams, and Julie S. Linsey. 2014. The Effects of Time and Incubation on Design Concept Generation. *Design Studies* 35, 5 (2014), 500–526. https://doi.org/10.1016/j.destud.2014.02.003

[79] Barbara Tversky and Juliet Y Chou. 2011. Creativity: Depth and Breadth. In *Design Creativity 2010*. Springer, 209–214. https://doi.org/10.1007/978-0-85729-224-7_27

[80] Jelle van Dijk. 2018. Designing for Embodied Being-in-the-World: A Critical Analysis of the Concept of Embodiment in the Design of Hybrids. *Multimodal Technologies and Interaction* 2, 1 (2018), 7. https://doi.org/10.3390/mti2010007

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

[82] Ilse M Verstijnen, Cees van Leeuwen, Gabriela Goldschmidt, Ronald Hamel, and JM Hennessey. 1998. Sketching and Creative Discovery. *Design Studies* 19, 4 (1998), 519–546.

[83] Vimal K Viswanathan and Julie S Linsey. 2012. Physical Models and Design Thinking: A Study of Functionality, Novelty and Variety of Ideas. (2012).

[84] Lihui Wang, Weiming Shen, Helen Xie, Joseph Neelamkavil, and Ajit Pardasani. 2002. Collaborative Conceptual Design—state of the Art and Future Trends. *Computer-aided Design* 34, 13 (2002), 981–996.

[85] Sitong Wang, Savvas Petridis, Taeahn Kwon, Xiaojuan Ma, and Lydia B Chilton. 2023. PopBlends: Strategies for Conceptual Blending with Large Language Models. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, 1–19. https://doi.org/10.1145/3544548.3580948

[86] Justin D. Weisz, Michael Muller, Jessica He, and Stephanie Houde. 2023. Toward General Design Principles for Generative AI Applications. https://doi.org/10.48550/arXiv.2301.05578

[87] WOMBOInc. 2022. Dream by WOMBO. Retrieved June, 2023 from https://www.wombo.art.

[88] Qian Yang, Nikola Banovic, and John Zimmerman. 2018. Mapping Machine Learning Advances from HCI Research to Reveal Starting Places for Design Innovation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 1–11. https://doi.org/10.1145/3173574.3173704

[89] Qian Yang, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. 2020. Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, 1–13. https://doi.org/10.1145/3313831.3376301

[90] Robert J. Youmans. 2011. The Effects of Physical Prototyping and Group Work on the Reduction of Design Fixation. *Design Studies* 32, 2 (2011), 115–138. https://doi.org/10.1016/j.destud.2010.08.001

[91] Lvmin Zhang and Maneesh Agrawala. 2023. Adding Conditional Control to Text-to-image Diffusion Models. *arXiv preprint arXiv:2302.05543* (2023). https://doi.org/10.48550/arXiv.2305.17098

[92] Yangkang Zhang, Chenye Meng, Zejian Li, Pei Chen, Guang Yang, Changyuan Yang, and Lingyun Sun. [n. d.]. Learning Object Consistency and Interaction in Image Generation from Scene Graphs. ([n. d.]), 3697–3707.

[93] Qihao Zhu and Jianxi Luo. 2023. Generative Transformers for Design Concept Generation. *Journal of Computing and Information Science in Engineering* 23, 4 (2023), 041003.

[94] Ning Zou, Qing Gong, Qing Chai, and Chunlei Chai. 2023. The Role Of Virtual Reality Technology In Conceptual Design: Positioning, Applications, And Value. *Digital Creativity* 34, 1 (2023), 53–77. https://doi.org/10.1080/14626268.2023.2166080

A DETAILS OF THE EXPERIMENTAL TASK

During our experiment, a design problem card was provided to participants.

Design Problem

Conceptual Innovation Design of Traffic Drone

Design Background

In contemporary times, the burgeoning quantity of cars within urban areas has resulted in mounting pressures on traffic management systems. In light of these problem, we seek to propose the development of a traffic drone product that can assist traffic police with traffic management, thus improving the overall efficiency and quality of such management processes.

Design Requirement

Please think from an innovative point of view, and explore the concept of this design problem through the prototype tools provided.

Fig. 19. The design problem card presented the design task for participants.

B THE CSI QUESTIONNAIRE

Table 5. The summary of analysis perspectives and evaluation metrics.

CSI Factor	Agreement Statement Description
Collaboration	The system allowed other people to work with me easily. It was really easy to share ideas and designs with other people inside this system.
Enjoyment	I would be happy to use this system on a regular basis. I enjoyed using the system.
Exploration	It was easy for me to explore many different ideas, options, designs, or outcomes, using this system. The system was helpful in allowing me to track different ideas, outcomes, or possibilities.
Expressiveness	I was able to be very creative while doing the activity inside this system. The system allowed me to be very expressive.
Immersion	My attention was fully tuned to the activity, and I forgot about the system that I was using. I became so absorbed in the activity that I forgot about the system that I was using.
Results Worth Effort	I was satisfied with what I got out of the system. What I was able to produce was worth the effort I had to exert to produce it.

Table 6. The Paired-Factor Comparison Test on the CSI.

When doing this task, it's most important that I'm able to. . . • Be creative and expressive • Become immersed in the activity • Enjoy using the system • Explore many different ideas, outcomes, or possibilities • Produce results that are worth the effort I put in • Work with other people
During the paired-factor comparison section of the CSI, each individual factor is paired with every other factor for a total of 15 comparisons. To promote independent consideration of each factor pairing and minimise the likelihood of participants aligning their answers to paired comparisons, the questionnaire display was designed such that only two factors per page were visible at a time.

C THE STRUCTURE AND OUTLINE OF SEMI-STRUCTURED INTERVIEW

Ice Breaker: introduce the design output

Q: *Could you introduce your final design?*

Q: *What do you think are the innovative ideas in your scheme?*

Q: *How do you like this design scheme?*

Key Issue 1: the advantage of the support tool

Q: *What advantages do you think this support tool has for prototype design?*

Q: *What do you think is the difference between this support tool and your traditional prototype tool?*

Further Guide Keywords: creativity, inspiration, design efficiency, design expressiveness, ideation process, product structure, usability…

Key Issue 2: the disadvantage of the support tool

Q: *Do you think there are any defects in this support tool?*

Q: *Did you encounter any challenges in using that tool during the design process?*

Further Guide Keywords: creativity, inspiration, design efficiency, design expressiveness, ideation process, product structure, usability…

Key Issue 3: the cooperation with AI (only for Group AI and HP)

Q: *How was your cooperation experience with AI during the design process?*

Q: *When and where do you think AI is helpful for your design process?*

Q: *Have you encountered any challenges in your cooperation with AI?*

Further Guide Keywords: cooperation contribution, communication, shared goal, interaction, experience, skill…

Key Issue 4: the AI influence on design ideas (only for Group AI and HP)

Q: *Do you think AI can affect your creativity?*

Q: *What do you think is the influence of AI on your design idea?*

Q: *Do you think you are the author of the final co-creation scheme?*

Further Guide Keywords: creativity support, idea change, original idea, design stimuli, design fixation, inspiration, authorship…

D THE DESIGN SOLUTION ANSWER SHEET

In order to clearly present the design solution answer sheet and related instructions, we present a blank answer sheet template and some representative answer sheets (Figure 20). In the "*Solution Presentation Area*", participants are required to present the design scheme in the form of diagrams and text notes, which can be photos of physical models, generated pictures, etc. Participants are required to present only one scheme, which can be multi-angle photos or scheme details of the same scheme. In the "*Solution Description Area*" and "*Innovative Concepts Description Area*", participants are requested to describe the design scheme in words as detailed as possible and summarize all innovative ideas in a structured way respectively.

E THE SOFTWARE PACKAGE

In order to facilitate further development of the hybrid prototype method, we have released the software package of the proposed hybrid prototype system. We uploaded related information to GitHub.

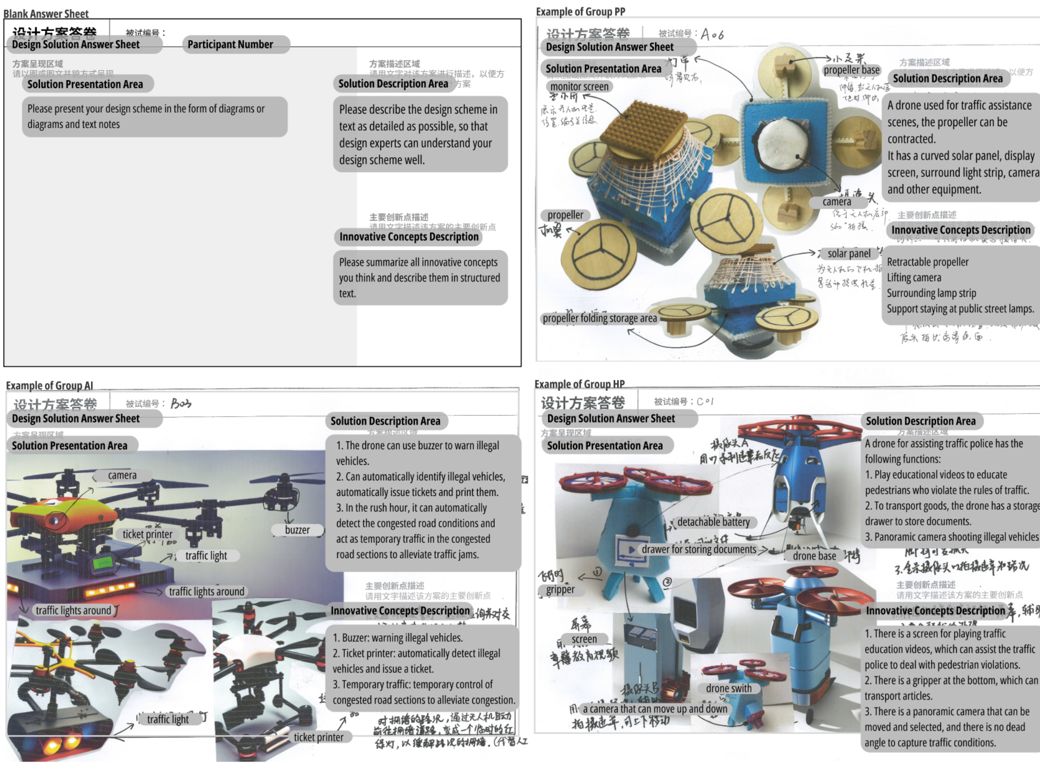

Fig. 20. The design solution answer sheet with instructions