International Journal of Design Creativity and Innovation

ISSN: 2165-0349 (Print) 2165-0357 (Online) Journal homepage: www.tandfonline.com/journals/tdci20

Empowering the creative design process through LLM-based support tools

Ana Rodrigues , Diogo Cabral & Pedro Campos

To cite this article: Ana Rodrigues , Diogo Cabral & Pedro Campos (2026) Empowering the creative design process through LLM-based support tools, International Journal of Design Creativity and Innovation, 14:4, 291-317, DOI: 10.1080/21650349.2026.2696854

To link to this article: https://doi.org/10.1080/21650349.2026.2696854

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

Published online: 03 Jul 2026.

Submit your article to this journal

Article views: 562

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tdci20

INTERNATIONAL JOURNAL OF DESIGN CREATIVITY AND INNOVATION
2026, VOL. 14, NO. 4, 291–317
https://doi.org/10.1080/21650349.2026.2696854

Taylor & Francis
by informa

OPEN ACCESS Check for updates

Empowering the creative design process through LLM-based support tools

Ana Rodrigues[a,b], Diogo Cabral[a,c] and Pedro Campos[a,b,d]

[a]ITI/LARSyS, Portugal; [b]University of Madeira, Funchal, Portugal; [c]ISEL – Polytechnic University of Lisbon, Lisbon, Portugal; [d]WowSystems Lda., Funchal, Portugal

ABSTRACT

Creative professionals (such as designers) have begun applying generative artificial intelligence (GenAI) models to their workflows, from text to image production. This practical adoption has been explored in human-computer interaction (HCI), which guides the incorporation of these models into creative design processes. While HCI research explores GenAI, an underlying wave of research is emerging, incorporating these models into Creativity Support Tools (CSTs). Our work aims to articulate the role of GenAI in supporting design creativity, an area that still requires research. This knowledge is essential for guiding the development of new CSTs. We contribute to this knowledge by developing an LLM-based CST with two interactive visualization interfaces (idea generation and evaluation). Through an empirical study with 18 participants (professionals and advanced design students) with design backgrounds, we demonstrated SAIESE's potential to facilitate divergent and convergent thinking. We also identify opportunities and challenges of incorporating AI into the creative design process. Our results highlight the role and use of AI at each stage.

ARTICLE HISTORY
Received 30 June 2025
Accepted 15 June 2026

KEYWORDS
Creativity; creative process; co-creativity; generative AI; LLMs

1. Introduction

Creativity consists of generating and refining ideas (Guilford, 1950), creating new approaches to problems or tasks, original solutions to conflicts, or new perceptions from a set of data (Cropley, 2010; Lubart, 2016). In the context of design, creativity emerges from the interaction between aptitude, process, and environment, through which an individual produces a perceptible idea (Plucker et al., 2004). Thus, the design process is considered a central concept within HCI, a field that aims to support and enhance design activities (e.g., creativity and problem-solving). Despite its relevance, creativity during the design process is constrained by barriers, such as design fixation (Jansson & Smith, 1991) – a tendency for designers to become anchored to initial ideas, limiting exploration of the design space. One approach to minimizing these barriers is to incorporate design elements or support requirements through technology that fosters and stimulates creative thinking, such as the Creativity Support Tools introduced by Ben Shneiderman (Shneiderman, 2000,

CONTACT Ana Rodrigues 2066213@student.uma.pt ITI/LARSyS, University of Madeira, Funchal, Portugal;

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

2007). These tools are defined as 'systems running on one or more digital systems that fulfill one or more creativity-focused capabilities and are employed to positively influence users of different expertise in one or more distinct phases of the creative process' (Frich et al., 2019). Currently, these tools have evolved significantly with the integration of Artificial Intelligence (AI) techniques, which have transformed the way humans interact with and explore the creative process, consequently supporting the scalability of design processes (Allison et al., 2022; Bourgault & Jane, 2023; Takaffoli et al., 2024; Verganti et al., 2020). Therefore, AI capabilities in creative design processes (i.e., characterized as open, complex, and often ill-defined problems (Allison et al., 2022) have become particularly relevant. However, the emergence of Generative AI (GenAI), specifically Large Language Models (LLMs) (Muller et al., 2023), has sparked interest across the entire HCI community for their ability to support creativity and redefine traditional problem-solving dynamics (Zorana Ivcevic and Mike Grandinetti, 2024). Despite the growing enthusiasm surrounding this paradigm, little is known about its impacts on creative workflows (Palani & Ramos, 2024). Recent studies have shown that GenAI can also have unintended negative effects on the creative process, such as reinforcing fixation or reducing the diversity of ideas generated (Wadinambiarachchi et al., 2024). However, other investigations (Holzner et al., 2025) indicate that human-AI collaboration can improve creative performance when properly integrated into the design process. These results suggest that the application of GenAI throughout the creative process is not inherent to the technology but rather strongly depends on how systems are designed and used. In this context, the high-level principles proposed by Ben Shneiderman (Shneiderman, 2020) – human control, reliability, and explainability – offer a relevant framework for the design of AI-based tools for creativity support. However, how these principles translate into practice remains insufficiently understood, particularly with respect to the balance between structured guidance and creative flexibility during the exploration of the design space (Palani & Ramos, 2024; Zorana Ivcevic and Mike Grandinetti, 2024). To address this gap, we investigate the SAIESE (System Approach to Idea Evaluation and Selection of Ecosystem) tool, a CST based on an LLM focused on sustainability, a domain that involves systemic and multidimensional challenges that require creative thinking to generate, evaluate, and select innovative ideas (Wadinambiarachchi et al., 2024). In previous studies, SAIESE was explored in a decision-making context, highlighting the potential of AI for creativity in idea evaluation and selection (Rodrigues et al., 2025). In this investigation, we reposition the system to examine its potential role in the broader creative design process, enabling us to explore how design decisions within such systems may influence user interaction, workflow structuring, and the perceived creative experience. In Figure 1, we present the SAIESE interfaces at each stage of the creative process, showing how the prototype integrates the creative process with AI. Based on this motivation, the research seeks to answer the following research questions: RQ1: How can SAIESE support designers during the creative process, considering the exploration and refinement of ideas? and RQ2: What are the opportunities and limitations perceived by designers during the creative flow? To answer these questions, we developed a within-subjects design with 18 designers using a mixed-methods approach (quantitative and qualitative data). The study provides

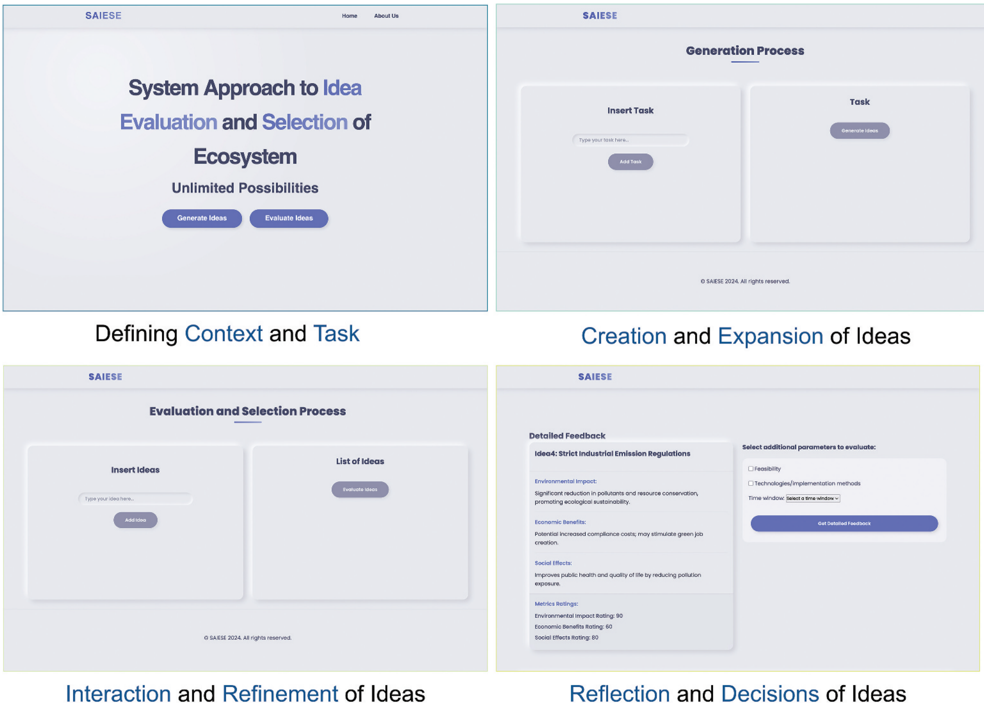

Figure 1. The user interface of SAIESE, a human-AI sustainable ideas co-creation environment.

exploratory insights into how designers perceived the integration of SAIESE during the creative workflow and how feedback mechanisms may influence the human-AI co-creative flow, particularly regarding the trade-offs between workflow organization and creative flexibility.

2. Background and related work

Our study is part of the HCI problem-solving perspective proposed by Oulasvirta et al (Oulasvirta & Hornbæk, 2016), which characterizes the field of human-computer interaction as a design-oriented discipline that combines empirical, conceptual, and constructive approaches to understand and improve interactive systems. Within this line of research, it is relevant to investigate the AI-driven creative process, as it allows us to analyze how humans can collaborate with AI systems to address complex design problems, focusing on the creative process (e.g., idea generation, evaluation, and selection) (Huang et al., 2026; Wang et al., 2025). Previous studies (Oulasvirta et al., 2017) have demonstrated that computational support can guide designers in selecting functionalities and structuring creative workflows. In the following sections, we explore the theoretical foundations of creativity in design, the Human-AICo-Creativity paradigm, and the integration of LLMs into creativity-support tools.

2.1. Creativity in design

Creativity is at the heart of design practice, essential for tackling complex problems and proposing innovative solutions (e.g., brainstorming, service prototyping, or designing interaction flows) (Biskjaer et al., 2017). Although it has been studied for over a century (Guilford, 1950), it is often described as the manifestation of innovative and valuable ideas, interpreted as 'new and useful' in their context (Cropley, 2010). Creativity is an essential requirement in design, as it constitutes the core of the creative process, which involves generating, evaluating, and refining interconnected solutions (Lubart, 2016). This process is iterative and requires constant alternation between exploratory and critical thinking. The literature describes creativity as composed of two complementary cognitive processes (Guilford, 1950; Lubart, 2016): divergent thinking – corresponding to the idea generation phase, which encourages the production of multiple alternatives – and convergent thinking – associated with the selection and improvement of ideas based on criteria of feasibility, innovation, and suitability to the context. In design, these processes are formalized in models such as the Double Diamond, which structures creative activity in successive phases of divergence and convergence, reflecting the cyclical and evolutionary nature of creative thinking (Banathy, 1996). However, this type of work is subject to several obstacles, such as creative blocks, cognitive overload, and selection bias, which can compromise the quality and diversity of the solutions generated (Hu et al., 2021; Sun & Yao, 2012). In this work, we explore the interdependence between the creative phases in the design process and investigate how artificial intelligence-based tools such as GenAI can promote more informed, collaborative, and effective cycles without compromising the human sensitivity that characterizes the creative worker.

2.2. Human-AI Co-creativity

Creativity Support Tools (CSTs) have been extremely effective in accelerating the human creative process by tracking history, simulating and exploring alternatives, and reusing assets (Shneiderman, 2000). However, their capabilities and resources operate in a more segmented way, with each part of the flow (person and computer) contributing in isolation, limiting the process. Recent advances in artificial intelligence have changed this view, offering a more collaborative engagement. In this new paradigm, humans and artificial systems work together dynamically, involving different perspectives, references, and skills to move beyond traditional paradigms (Holzner et al., 2025; Lazovsky et al., 2024; Singh et al., 2025). Co-creativity, therefore, involves two or more agents (of which at least one is human and at least one is an AI system) working together to generate new and valuable ideas and solutions in their context (Singh et al., 2025). Unlike collaboration, which focuses on the efficient execution of tasks, co-creativity highlights a more integrated interaction, where the result arises from active contributions from both the human and the AI system (Wang et al., 2025). Human-AI collaboration has been explored in various domains (e.g., decision-making (Cai et al., 2019; Jain et al., 2023; Puranam, 2021), games (Zahra Ashktorab et al., 2020; Zhang et al., 2021), and even education (Ng et al., 2020)). With this shift in focus, collaboration in creative contexts takes on specific characteristics, as creativity involves open, ambiguous activities that depend heavily on human judgment (DiStefano & Beaty, 2026; Rodrigues et al., 2025). From this perspective, Wang et al. (Wang et al., 2020) state that effective collaboration requires mutual understanding of objectives, preventive co-management of tasks, and shared monitoring of progress. This conceptualization shifts the role of AI from a supporting tool to an active participant in the creative process (e.g., Brainwriting (Shaer et al., 2024), LAVE (Wang et al., 2024), or AngleKindling (Petridis et al., 2023)), assisting in both divergent and convergent thinking. According to Licklider (Licklider, 1960), this novelty represents a significant step toward symbiotic collaboration between humans and AI, in which the role of AI extends beyond improving the creative process – it becomes an active collaborator or 'content creator' (Cetinic et al., 2022). In this sense, co-creation presents important implications for the design of CSTs, since these systems support the balance, guidance, and structuring of the creative process during problem-solving or even in simple tasks. To better understand how these dynamics manifest in practice, this work investigates the SAIESE tool during a co-creative interaction between humans and AI.

2.3. Large language models for creative support

GenAI technologies (like LLMs) have had significant implications for creative workers (Harvard Business Review, 2022), including designers (Olsson & Väänänen, 2021). The products arising from this technology can be considered creative, as they are novel, valuable, and surprising in the context of use, aligning with the very definition of creativity (Runco & Jaeger, 2012). This means that emerging GenAI methods have great potential to facilitate creative work in various ways. Thus, this technological advancement has become a topic of great interest in both creativity (Holzner et al., 2025) and Creativity Support Tools (CSTs) (Rodrigues et al., 2023; Shneiderman, 2007). ChatGPT, for example, is an LLM that has achieved increasing performance during the creative process (Cropley, 2023; Hubert et al., 2023) and has been incorporated into several CSTs. The versatility of these models has been reflected in the development of new CSTs, whose approaches range from idea generation to workflow support. To support idea generation and collaborative ideation, Gonzalez et al. (Gonzalez et al., 2024) developed the Collaborative Canvas for brainstorming sessions. This tool uses an LLM (Llama 2) that supports contextualized user actions throughout the process, including post-it note generation and progressive improvement of generated responses. Similarly, Shaer et al. (Shaer et al., 2024) developed a Brainwriting Group-AI CST to boost the Brainwriting tool. The framework incorporates two LLMs to support group brainstorming sessions: GPT-3 for idea generation and GPT-4 for evaluation during the creative process. Similarly, Petridis et al. (Petridis et al., 2023) developed an interactive web tool, AngleKindling, which supports journalists in summarizing press releases and generating angles through an LLM (GPT-3) during the writing process. Recent discussions in creativity-oriented HCI have also highlighted how LLM-based CSTs reinterpret traditional CST principles by enabling more adaptive, contextual, and iterative forms of interaction throughout the creative process (Rodrigues et al., 2026). Unlike previous AI-based CSTs, which often focused on predefined automation and decision support mechanisms, LLM-based approaches support conversational feedback, contextual reinterpretation, and dynamic refinement of ideas (Rodrigues et al., 2026; Wan et al., 2024).

These characteristics influenced the development of SAIESE, especially regarding iterative idea exploration, user-controlled refinement, and the integration of contextual feedback mechanisms. Beyond this phase of the creative process, LLM-based CSTs have also been developed to support broader creative workflows. For example, Wang et al. (Wang et al., 2024) developed a CST that supports video creation using an LLM (GPT-4) during the idea generation task. While these approaches demonstrate the potential of LLMs during the creative process, open questions remain regarding the effective integration of this technology into design beyond idea generation and about how designers should be involved. In this study, we seek to address these limitations by exploring how an LLM-based CST can support designers throughout the creative process.

3. SAIESE

SAIESE is a tool that supports creative problem-solving in sustainability contexts using o1-mini model. The prototype was developed around four design goals addressing key human-AI co-creation challenges. These goals guide the system's functionality and are detailed in the following subsections.

3.1. Design space and rationale

Based on an analysis of the literature on AI-supported creativity, we derived four main objectives for designing an LLM-based CST: Goal 1: Support the complete creative process. The prototype was developed to support idea generation, followed by evaluative exploration and refinement of the generated ideas. Previous studies (Rodrigues et al., 2024, 2025) indicate that CSTs tend to support decision-making and even include information visualization but rarely address the whole creative process. This objective aims to fill this gap, allowing the prototype to support designers in both divergent and convergent thinking. Goal 2: User-Centered Design Process. The prototype interface was developed using user-centered design principles, widely adopted in the HCI field, to promote usability and user engagement throughout the creative process. The soft UI or neuromorphic visual style (UX Design Institute [UX], 2022) was applied to create a less intrusive, more seamless visual experience that favors interpretation, exploration, and staying on task. Goal 3: Improve the presentation of information through GenAI. LLMs can enhance information visualization in several ways, allowing users to quickly evaluate generated information. According to Rodrigues et al (Rodrigues et al., 2024), visualizing sustainable metrics alongside textual feedback improves decision-making and empowers users to prioritize ideas. Based on this information, the tool provides graphs (e.g., bar charts) and detailed descriptions of each metric for the idea or evaluation, depending on the process stage, to support the user's critical analysis. Goal 4: Promote human-AI co-creation. Humans actively participate in a hybrid process through a balanced partnership approach that enhances innovation and problem-solving with AI. This paradigm is extended and reconfigured for the context of creative design, operationalized throughout the creative flow to reinforce the designer's autonomy and mitigate risks associated with excessive reliance on AI.

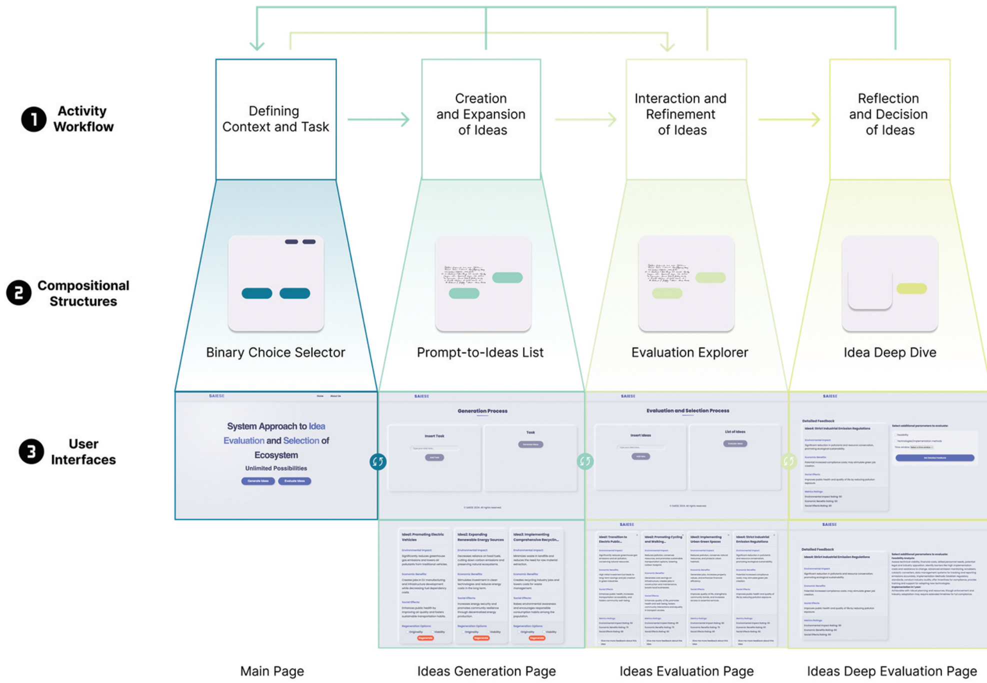

Figure 2. Mappings between the compositional structures of the creative process identified in the idea development workflow and the SAIESE user interface. (a) Workflow around compositional structures; (b) the underlying compositional structures; (c) four views in the SAIESE user interface map to a corresponding compositional structure.

3.2. *System overview*

The system presented in this research is based on a previously introduced prototype, designed for LLM-assisted decision-making in sustainable contexts (Rodrigues et al., 2025). While the previous prototype focused primarily on supporting decision outcomes, the current system repositions SAIESE as a creativity support tool, based on an LLM that supports the entire creative process (i.e., idea generation, evaluation, and selection) within a creative designer's workflow. The previous implementation is based on the GPT-3.5-turbo model, optimized primarily for evaluation tasks. However, despite adequate decision-making support, a limitation was detected regarding its flexibility for iterative idea exploration. In contrast, this version of SAIESE integrates the o1-mini model, which allows for more adaptive iterations, with contextual reinterpretation and responsive feedback.

The system was built on a Flask-based backend and a JavaScript frontend (Figure 2) and is based on three fundamental principles: (i) idea traceability; (ii) explicit user control over the refinement of AI-generated suggestions; and (iii) multidimensional evaluation of results. The web application initializes the main routes for idea generation, evaluation, and data management at startup. The backend uses the SQLAlchemy ORM for data modeling, and each core concept (i.e., user-submitted ideas, LLM-generated ideas, and their evaluations) is represented by models that ensure data traceability and integrity. Generated ideas and evaluations are linked to real users, and integration with the OpenAI

API (o1-mini) allows for automatic idea generation and scoring based on three metrics (environmental impact, economic benefits, and social effects) derived from prior work on sustainability-oriented decision-making (Rodrigues et al., 2024). The metrics were operationalized through a combination of textual and visual representations, combining a set of qualitative and quantitative criteria. Environmental impact refers to the potential to reduce environmental footprint, conserve resources, and mitigate pollution; economic benefits capture aspects such as cost efficiency, profitability, and job creation; and social effects relate to improvements in quality of life, equity, and community well-being. The system then assigns a score to each metric on a normalized scale from 1 to 100, where higher values indicate stronger alignment with the respective dimension. This range allows for finer-grained differentiation across alternatives, enabling users to identify relative strengths and trade-offs between environmental, economic, and social dimensions. To maintain interpretability, the numerical scale is complemented by qualitative intervals ((Allison et al., 2022; Bourgault & Jane, 2023; Cropley, 2010; Frich et al., 2019; Guilford, 1950; Holzner et al., 2025; Jansson & Smith, 1991; Lubart, 2016; Muller et al., 2023; Oulasvirta & Hornbæk, 2016; Palani & Ramos, 2024; Plucker et al., 2004; Rodrigues et al., 2025; Shneiderman, 2000, 2007, 2020; Takaffoli et al., 2024; Verganti et al., 2020; Wadinambiarachchi et al., 2024; Zorana Ivcevic and Mike Grandinetti, 2024) not important to [81–100] extremely important), balancing precision with usability. To ensure the prototype's efficiency and scalability, the system includes a caching layer (Flask-Caching) to avoid redundant API calls. Finally, a benchmarking mechanism was implemented to compare recent ideas with previously generated ones, thereby reducing their similarity.

The user interface is organized into two main workflows: idea generation and idea evaluation – built with a responsive, neuromorphic style that mimics how the brain processes information (UX, 2022). Figure 2 presents the Compositional Structures with subtle color variations to conceptually distinguish the phases of the creative workflow and illustrate the transitions between divergent and convergent activities. In the implemented interface, a more consistent visual palette was maintained to promote interface cohesion. Both interfaces are developed using HTML and AJAX templates for real-time updates. In the idea generation interface (Figure 3(A)), users describe the task or problem, and the system then returns a set of ordered ideas. Each idea is presented with two refinement options: Originality and Feasibility. These options allow for targeted idea regeneration, prioritizing conceptual novelty or practical feasibility, respectively. The separation of these parameters is an intentional design decision, intended to support focused iterations without compromising user agency. In the evaluation interface (Figure 3(B)), users can submit a limited set of ideas for analysis. The system presents evaluation results progressively, combining numerical values with textual explanations for each metric in each idea. The system also allows the user to request additional feedback on the feasibility and possible implementation strategies. To support the interpretation and comparison of results, SAIESE integrates interactive visualizations developed in D3.js. Specifically, the evaluations are represented as vertical bar graphs showing the scores for each sustainability metric for each idea. Here, the colors are consistently associated with specific evaluation dimensions (environmental impact, economic benefits, and social effects). Additionally, a horizontal benchmarking visualization presents the relative positioning of ideas across the sustainability metrics on a normalized scale from 0 to 100, supporting comparative interpretation. Hover interaction allows users to access explanatory information, and the visualization supports comparisons among various ideas, helping identify relationships between environmental, economic, and social metrics (Figure 4).

4. Method

This study investigated how the developed CST, SAIESE, supported designers' creative process. In this section, we describe the details of the experiment developed and the analysis of the data obtained, and the respective results.

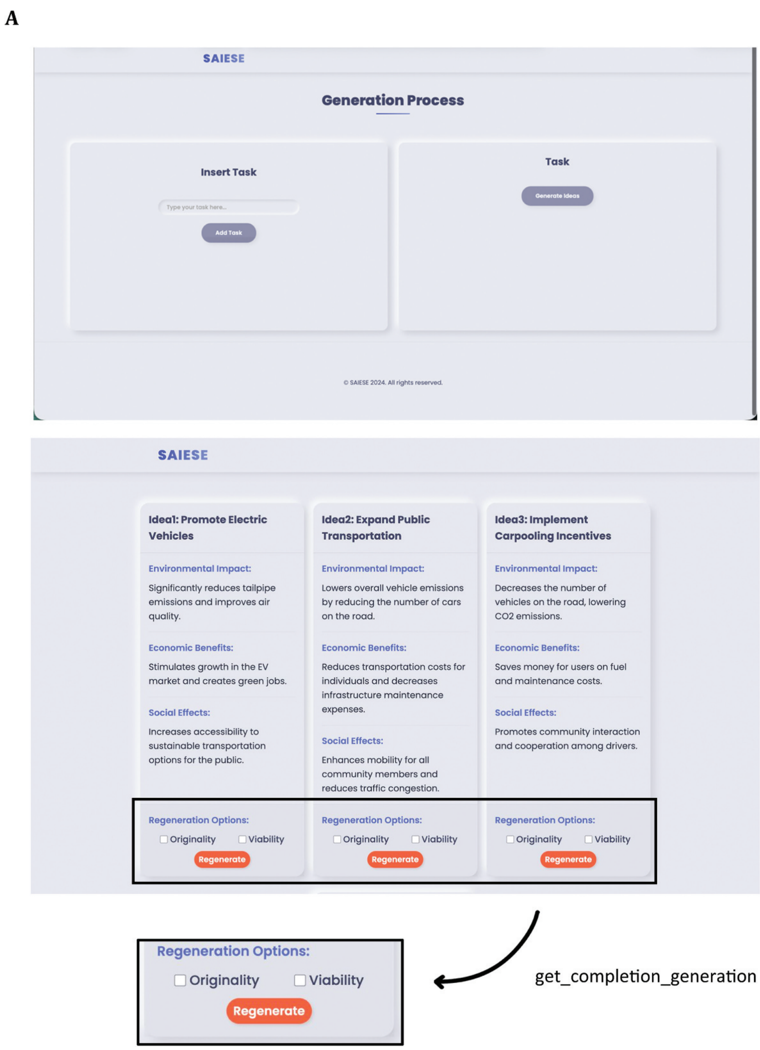

Figure 3. Printscreens of different user interfaces of SAIESE. (A) Generation process and (B) evaluation process.

B

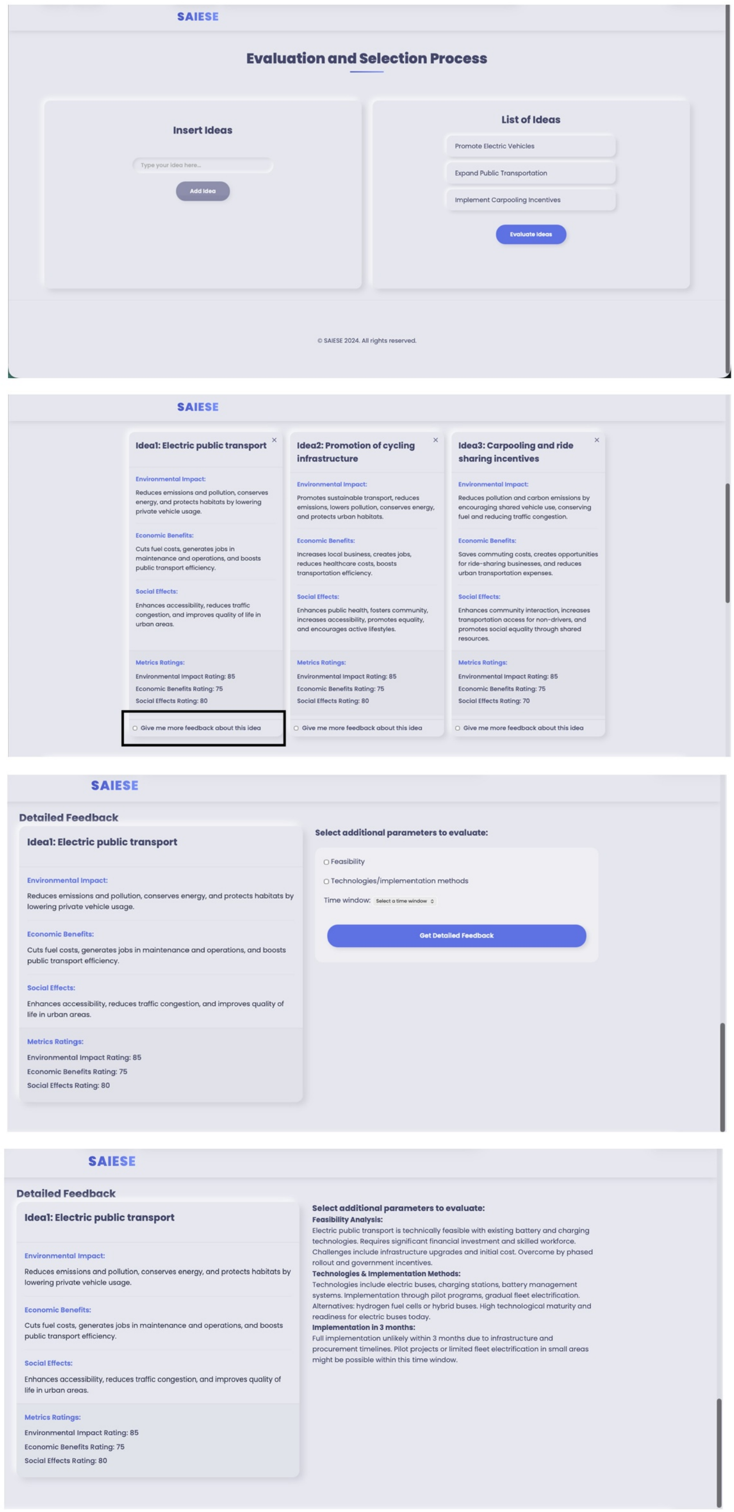

Figure 3. (Continued).

Figure 4. Printscreen of interactive visualizations about the evaluation and comparison process.

4.1. Participants

We recruited 18 designers through word of mouth (denoted as $P_1 - P_{18}$). Participants were subsequently selected based on their active involvement in design-related practices and their ability to participate in creative tasks. Of the 18 participants, 10 were women, and 8 were men, aged 28–42 years ($M = 35.28, Mdn = 35.50, SD = 4.30$). Participants had diverse backgrounds in design, with graphic design ($n = 13$) as the central area of expertise, followed by product design ($n = 5$). The sample consisted of 12 professionals with 2 to 5 years of experience ($M = 2.92, Mdn = 3, SD = 0.99$) reflecting early- and mid-career professionals. While the remaining participants (six) are still students (2 with master's degrees and 4 with doctoral degrees), they are actively involved in design practices (1 freelancer, 1 belonging to a research team, and 4 with completed internships). This diverse sample allowed us to capture a range of experience levels. In this way, we ensured that all participants possessed sufficient knowledge in the field to interact with the prototype. Participants' experience in creative processes was assessed through a self-reported item in the initial questionnaire – 'I have experience working on my creative process (generating, evaluating, and selecting ideas)' - rated on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). Participants reported a high level of experience ($M = 4.67, Mdn = 5.00, SD = 0.49$). Participants also reported familiarity with various generative AI tools, including Adobe Firefly, Ideogram, and ChatGPT. They use these tools for occasional exploratory use and integration into their workflows.

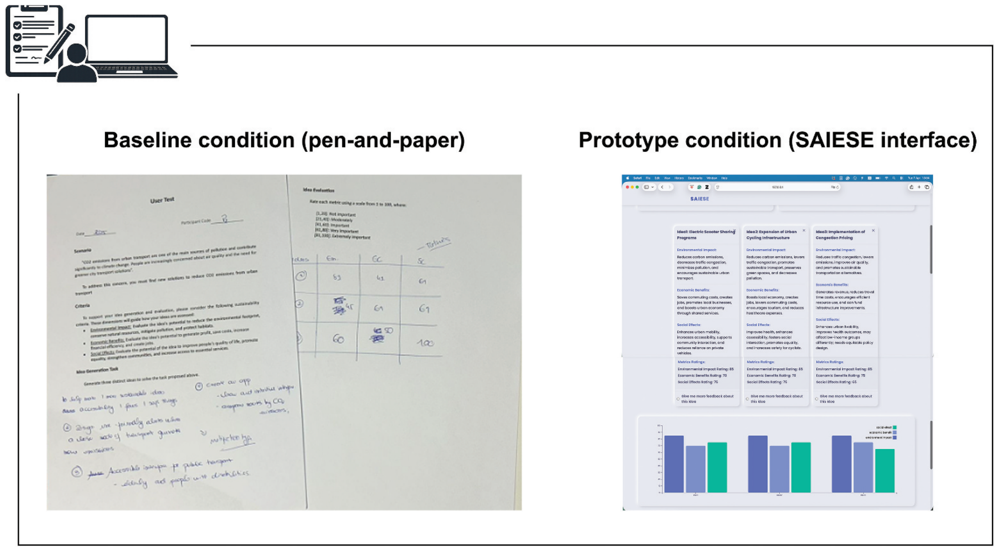

Figure 5. Participant output in the baseline condition (pen-and-paper) and the prototype-supported condition (SAIESE). The example (participant P_2) illustrates how ideas and their evaluations were manually recorded in the baseline, compared to the more structured, interactive representation provided by the prototype.

4.2. Experimental design

To study workflow support during the creative process, we designed a within-subjects experiment with two conditions: (1) exploration without a prototype – participants received definitions of each metric (i.e., environmental impact (the idea's potential to reduce the environmental footprint, conserve natural resources, mitigate pollution, and protect habitats), economic benefits (the idea's potential to generate profit, save costs, increase financial efficiency, and create jobs), and social effects (the idea's potential to improve people's quality of life, promote equality, strengthen communities, and increase access to essential services) and the parameters applied in both the idea generation phase (i.e., originality and viability) and the convergent phase (i.e., feasibility, technologies/ implementation methods, and time window) organized on paper. Participants followed the instructions in a structured way, without direct interaction with the LLM, allowing the effect of the prototype on the creative process to be isolated as a baseline control; and (2) Exploration with the prototype – participants interacted with the o1-mini model. The prototype used the same prompts and parameters as in the previous condition, allowing a comparison of the effect of the digital interface on the process. Half of the participants ($P_1 - P_9$) began with condition 2, and the other half ($P_{10} - P_{18}$) began with condition 1 to control for order effects. Both conditions were also applied to isolate the prototype's effect on the creative process. Counterbalancing ensures that differences in results are not solely due to the order of exposure (e.g., fatigue, learning, or task familiarity). We applied the CSI (Cherry & Latulipe, 2014) and NASA-TLX (Hart, 2006) questionnaires, which allowed for a comparison of support for creativity and cognitive load between the two conditions. Details of the instruments and data collection are described in section 4.5. In

Figure 5, we present a representative example of a participant's output in both the baseline (pen-and-paper) and prototype-supported conditions, highlighting differences in how ideas are generated and evaluated. Before starting the tasks, participants received a brief introduction to the SAIESE tool, including an explanation of its interface and features. The goal was to familiarize participants with the system and how it supports the different stages of the creative process (i.e., idea generation, evaluation, and selection). Following this, participants received clear instructions on how to interact with the prototype, including how to enter tasks, interpret feedback, and navigate the process's different stages. We ensured that all participants, regardless of their level of experience with generative AI tools, could effectively interact with the system during the experiment.

4.3. Scenario and task

We asked participants to generate innovative ideas, evaluate them based on environmental impact, economic benefits, and social effects, and finally select an idea given the following scenario: 'CO_2 emissions from urban transport are one of the main sources of pollution and contribute significantly to climate change. People are increasingly concerned about air quality and the need for greener city transport solutions' (European Commission, 2024). To address this concern, you must find new solutions to reduce CO_2 emissions from urban transport.

4.4. Procedure

All studies were conducted through face-to-face meetings over approximately one hour and fifteen minutes ($M = 60.49, Mdn = 59.21, SD = 5.14$). At the beginning of the experiment, participants were informed about the study and asked to complete an informed consent form. The study was divided into two 20-minute sessions, with a 10-minute break in between: the first for the task without the prototype and the second for the task with the prototype. After each session, we administered the questionnaires described in section 4.5. After completing both sessions, we conducted a semi-structured interview to collect qualitative data on participants' experiences with SAIESE, lasting approximately 15 minutes. Participants did not receive financial compensation. However, during the coffee break, they were offered a beverage (coffee, tea, or water) and a cookie (we had asked about food intolerances previously).

4.5. Data collection

After each session, two standard questionnaires were administered: the Creativity Support Index (CSI) (Cherry & Latulipe, 2014) – a standardized psychometric tool that assesses a system's support for creativity. The CSI provides quantitative assessments on six dimensions of creativity support (Collaboration, Enjoyment, Exploration, Expressiveness, Immersion, and Results Worth Effort (RWE)) and allows us to assess how different design exploration features supported the participants' creative work; and NASA Task Load Index (TLX) questionnaire (Hart, 2006) – a tool that assesses workload through six factors (Mental, Physical, Temporal, Performance, Effort, and Frustration), allowing us to assess participants' cognitive workload during the task. Although the CSI and NASA-TLX are widely used standardized measures, we developed six questions through a semi-structured interview to more directly assess how the prototype aligns with its stated goals. The interviews were audio-recorded and transcribed for later analysis. Appendix lists the predetermined questions.

4.6. Data analysis

We analyzed the quantitative data using IBM SPSS Statistics software (version 28.0.0.0 (190)) (IBM Corp [IBM], 2021). First, we verified the normality of the data using the Shapiro-Wilk test ($p < 0.05$). Since the data were normally distributed, a paired t-test was applied to determine whether there were statistically significant differences between the two experimental conditions. The *Bonferroni correction* was applied to minimize multiple-comparison errors across the questionnaire factors, and the results were reported with an adjusted significance level of $\alpha = 0.008$. For the analysis of the qualitative data, we applied thematic analysis and iterative open coding (Clarke et al., 2015). We coded the interview data analyzed the transcripts for emerging themes.

5. Results

5.1. Creativity support index

The results indicate a statistically significant difference between the condition with the prototype and the condition without the prototype in the overall CSI score, $t(17) = 8.06, p < 0.001$. The results are shown in Figure 6. We performed post hoc comparisons between the two conditions in the six factors. We found statistically significant differences in five criteria: Collaboration ($t(17) = 7.1, p < 0.001$), Enjoyment ($t(17) = 9.7, p < 0.001$), Exploration ($t(17) = 5.1, p < 0.001$), Immersion ($t(17) = 4.5, p < 0.001$), and RWE ($t(17) = 5, p < 0.001$). On the other hand, the Expressiveness factor did not present significant differences between the condition with the prototype and the condition without the prototype ($t(17) = 2, p = 0.061$). However, the mean was slightly higher in the session with prototype ($M = 7.53, Mdn = 7.50, SD = 1.33$) when compared to the session without prototype ($M = 6.97, Mdn = 7.00, SD = 0.87$). The results suggest that the perception of expressiveness may not depend directly on the presence of the prototype. A possible explanation is related to the fact that the participants are designers, already familiar ($M = 4.67, Mdn = 5.00, SD = 0.49$) with creative and expressive processes. This familiarity could have contributed to the participants feeling comfortable enough to express themselves in both conditions, reducing the need for additional support. In the case of the RWE factor, despite the statistically significant difference observed, the average value in the condition without prototype (M=7.89, $Mdn = 8.00, SD = 1.23$)was already high when compared to the average value of the condition with prototype ($M = 9.61, Mdn = 10.00, SD = 0.70$). This indicates that the participants already recognized the task as close to the real world, even without the prototype's support. However, the presence of the prototype reinforced this connection, further consolidating the perception of realism and the practical relevance of the experience.

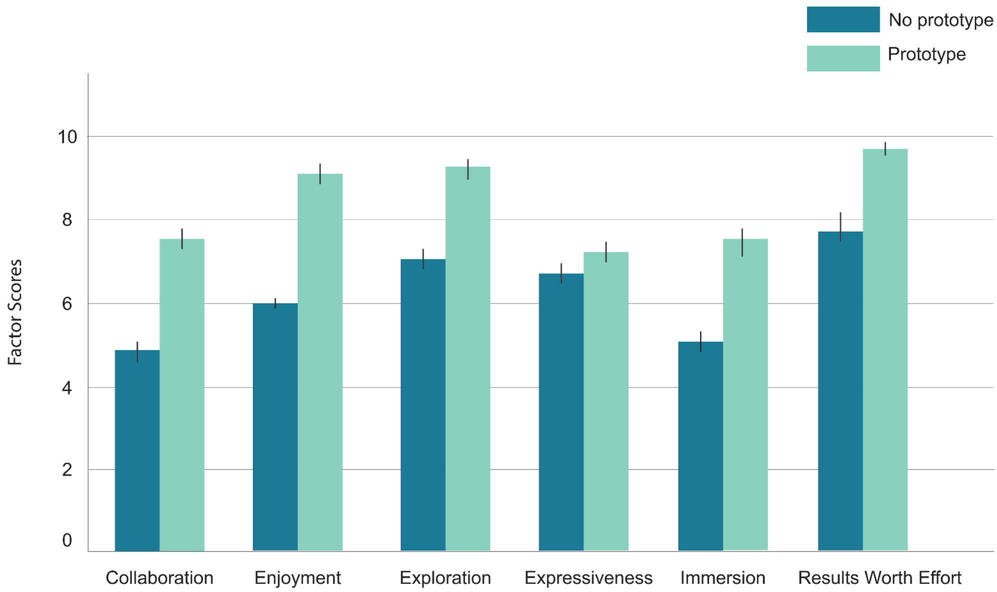

Figure 6. Comparison of CSI factor scores ($p < 0.001$).

Table 1. Results of the NASA-TLX questionnaire comparing the task without a prototype with the task with a prototype, where the means, standard deviations (in parentheses), and *p*-values for the six paired-sample t-tests with Bonferroni correction are presented. The *p*-values in bold are statistically significant.

NASA-TLX Factors	No Prototype	Prototype	p-value
Mental Demand	7.67(1.14)	4.67(1.24)	<0.001
Physical Demand	4.11(1.75)	4.22(1.26)	0.790
Temporal Demand	6.33(1.90)	3.89(1.49)	0.001
Performance	8.22(1.44)	8.22(1.31)	1.000
Effort	7.67(1.41)	3.83(1.50)	<0.001
Frustration	3.33(1.78)	2.67(1.33)	0.0082

5.2. NASA task load index

The paired t-test revealed statistically significant differences between the condition with the prototype and the condition without the prototype $t(17) = 8.5, p < 0.001$. The results of the NASA-TLX questionnaire are presented in Table 1. Post-hoc comparisons showed statistically significant differences in three factors: Mental demand ($t(17) = 8.3, p < 0.001$), Temporal demand ($t(17) = 5.1, p < 0.001$), and Effort ($t(17) = 6.6, p < 0.001$). The results indicate that the prototype significantly reduces mental demand during the task and participants' perceived effort, suggesting clear improvements in usability and the cognitive load associated with the creative process. The prototype made the task easier to perform, less mentally demanding, and more direct.

On the other hand, the factors Physical demand ($t(17) = 0.3, p = 0.790$), Performance ($t(17) = 0, p = 1.000$), and Frustration ($t(17) = 1.6, p = 0.124$) did not present statistically significant differences between the condition with the prototype and the condition without the prototype. These results suggest that participants felt they performed the task at the same level of success, even though the effort required was lower with the prototype. In other words, the task yielded the same result, but it was achieved more lightly and efficiently. The frustration levels were low from the beginning of the study. The result suggests that the task was not perceived as emotionally exhausting or frustrating. Thus, the potential for improvement was reduced from the start.

5.3. Qualitative user experience results

5.3.1. Perception of the value and usefulness of SAIESE

Users recognized the value of the prototype, especially in structuring and organizing the creative process, highlighting the clarity of the evaluation metrics (Environmental Impact, Economic Benefits, and Social Effects) and the immediate feedback. Participant P_3 stated that *'what helped the most was the immediate feedback and the clear way in which ideas are evaluated in the different metrics,'* while participant P_7 highlighted that *'the interface provided organization and an overview of my usual process.'* However, some participants noted difficulties with the interface. Participant P_{11} mentioned that *'the evaluation process seemed a bit bureaucratic and limited my spontaneity,'* while participant P_{15} reinforced that *'navigation could be more intuitive, some functions were hidden, and I had to waste time figuring out how to move forward.'* These limitations indicate that, although structure is valued, it is essential to balance the interface's organization with flexibility to accommodate different creative styles and working preferences.

5.3.2. Integration of SAIESE into the workflow

The prototype is well accepted as an integration into the participants' workflow. It supports validating ideas during the creative process (i.e., generating and evaluating ideas). Participant P_2 considered that *'the tool is great for validating ideas before presenting them'*. At the same time, participant P_9 emphasized that *'I see it as a support to give more rigor to my creative process, but not to replace the free brainstorming phase.'*

When comparing the time spent in each session, we observed a 26% reduction in average time in the condition with the prototype compared with the condition without the prototype. The average time to complete the task without the prototype was $M = 17.01, Mdn = 16.57, SD = 2.00$, while with the prototype it was $M = 12.44, Mdn = 12.58, SD = 2.26$. This result suggests that the prototype can support a more structured, guided workflow, helping participants maintain focus during validation and organize their ideas. However, it cannot be interpreted as a direct improvement in creativity. In this context, the observed reduction in average execution time suggests that the prototype can provide a scaffold workflow during the creative process – supporting tasks associated with the evaluation, refinement, and selection of ideas. Qualitative feedback reinforces these results, for example, P_6 *'The process was faster using the prototype; I was able to complete the task quickly.'* These findings suggest that the prototype supports a more targeted, structured approach to task resolution. However, some interruptions to the creative flow were reported due to the system's rigid structure. For example, participant P_5 noted that 'there were times when I wanted to be faster and explore more ideas at the same time, but I understand the technical limitations.' Although the prototype supports the convergence of ideas, the mentioned rigidity may indicate a limitation in the ideation process (i.e., divergent thinking). In turn, participant P_{17} commented that 'structured processes sometimes interrupt the creative flow and cause frustration,' having rated his level of frustration on the NASA-TLX questionnaire as a score of 3 and a score of 1 when he performed the task without the prototype's aid. This difference in reported frustration suggests that the prototype, while supporting task completion, can also disrupt creative flow. These results indicate that SAIESE was primarily perceived as a tool that supported convergent thinking throughout the task. Participants' perceptions of integrating SAIESE into their workflow varied by individual creative preferences and expectations for flexibility during ideation. Participants who preferred more open exploration processes reported greater disruption to the creative flow when interacting with the system's structured workflow. For example, participant P_{14} stated: 'The structuring helped me organize the task, but I felt somewhat limited during the more spontaneous exploration of ideas. It's a free process that I follow during brainstorming.' Some participants associated reduced task time and perceived cognitive effort with the structured presentation of ideas, evaluation mechanisms, and the ability to compare sustainability dimensions through visual feedback. Participant P_{12} commented, 'I liked having the ideas organized and evaluated, as this way I could more quickly observe the most relevant options.' Participant P_9 also said, 'The visual comparison of metrics facilitated understanding the strengths and weaknesses of each idea without me having to overthink the process.' These findings suggest that the successful integration of these tools may depend not only on the system's capabilities but also on contextual factors such as workflow preferences, interaction expectations, and tolerance for structured guidance during ideation. Furthermore, usability and navigation were highlighted as areas for improvement to enhance workflow. Participant P_8 said 'navigation could be simpler so as not to break the rhythm of the session.'

5.3.3. Role of AI in the creative process

Participants responded that SAIESE provided sufficient support in fostering divergent and convergent thinking as they recognized its potential. Participant P_3 said, '*I liked the feeling of being in dialogue with the system, especially when I adjusted ideas and saw the results change in real-time*.' Participant P_4 also stated, '*The prototype brought some new ideas, but mainly it helped me to organize and reflect on them better*.' While participant P_{12} pointed out that '*I see the prototype as a complementary tool to support the organization and evaluation of ideas, but not as a main source of innovation*.' These results reflect participants' tolerance level for the accuracy of the results presented by the system. We observed that participants understood the reflective and organizational roles of the results. However, participants felt that communication with the system had certain limitations. Participant P_{10} expressed, '*I felt that the system followed a very rigid flow, and I would have liked to have had more freedom to adjust the parameters*.' Participant P_{17} said: '*I would have liked to have had more customization options regarding the criteria to be applied so that I could experiment with different approaches*.' In this sense, limitations in user control over the prototype and the perception of the 'black box' in the LLM's functioning emerge. This lack of transparency, a familiar feature of AI models (Glassberg et al., 2025; Sunny, 2025), undermines user confidence and the perception of control (a

Table 2. Strengths and weaknesses of SAIESE output.

SAIESE Features	Strengths	Weaknesses	Possible Solutions
Idea Generation	Rapid idea generation helps unlock individual creativity – organize and reflect on them with a useful and modular overview.	Lack of personalization and low control over metrics can reduce perceived creative value and limit user autonomy.	Add customization options (style, tone, parameters); allow human intervention in prompts and evaluation criteria.
Idea Evaluation	Objective assessment based on clearly defined metrics. immediate textual and visual feedback clarifies ideas.	Some evaluations are perceived as shallow, inconsistent with the designer's intent, or overly rigid.	Allow user review or adjustment of criteria; introduce explainable AI features with clear justifications.
Realism & Context	Reinforcing perception of task realism (high RWE) and contextual relevance using the prototype.	Some AI suggestions are disconnected from the briefing or real context.	Anchor suggestions in the project history or briefing provided; use data from previous sessions for greater contextual coherence.
Cognitive Load	Significant reduction in mental load, effort and time on task with the use of the prototype; supports a structured and efficient workflow.	Rigid structure or navigation may interrupt creative momentum and reduce fluidity.	Improve usability and navigation; integrate shortcuts and contextual suggestions to maintain fluidity.
Trust & Expertise	Designers maintain critical judgment, combining SAIESE outputs with their knowledge.	Lack of transparency and predictions that conflict with experience generate distrust.	Provide historical AI performance; allow comparison between SAIESE suggestions and user judgments; enhance transparency of AI reasoning.

topic widely discussed in the literature on explainable AI and user-centered design) (Glassberg et al., 2025). For example, P_8 commented: *'Some ideas presented seemed disconnected from the scope of the task or even from what I intended to explore'* and P_{11} commented: *'The ideas presented were good, but then the evaluation was superficial and technical'*. These results indicate the need for and importance of developing LLM-based CSTs that increase user control, provide clear justifications for the phases of AI models, and thereby allow for more coherent and transparent responses. However, we observed that participants adopted a more critical and conscious stance toward the system's results, grounded in their personal and professional judgment and informed by their experience in creative processes, as measured in the initial questionnaire ($M = 4.67, Mdn = 5.00, SD = 0.49$).

5.4. Synthesis of key findings

Based on the user study results, we identified the strengths and weaknesses of the SAIESE prototype, focusing on the LLM-supported functionalities that shape the creative process. Table 2 summarizes these observations, with each row corresponding to a specific feature of the prototype. The Strengths column highlights the study's identified strengths that improve efficiency, structure, and participant engagement. The Weaknesses column identifies the areas where the prototype has limited user control, transparency, adaptability, and different profiles. Finally, the Possible Solutions column suggests ways to mitigate the limitations identified in the previous column. This synthesis consolidates quantitative (i.e., CSI and NASA-TLX) and qualitative feedback, in which patterns such as improvements in the idea generation and evaluation process stand out, while reducing cognitive load. These findings also inform the broader implications for AI-supported creativity, discussed in the following section.

6. Discussion

Based on the synthesis presented in Table 2, this section discusses the implications of the identified strengths and weaknesses, as well as the limitations and future research directions to advance the development of creative design tools with AI support.

6.1. CST for creativity support

The quantitative and qualitative results of user studies indicate that SAIESE supports the organization of ideas and facilitates divergent and convergent thinking through structured feedback and criteria-based guidance, consistent with previous work on LLM-based CSTs. SAIESE also shapes the structure of divergent and convergent mechanisms through feedback and guidance from explicit evaluation criteria (i.e., sustainability metrics). This support aligns with the study reported in (Oh et al., 2024), which demonstrates how CSTs structure and facilitate the creative process. The integration of explicit evaluation criteria provided structured support for creative tasks, guiding both idea generation and refinement. Similarly, some participants reported that the system allowed them to refine ideas during their generation and evaluation. For example, P_3 said, '*I liked the feeling of being in dialogue with the system, especially when I adjusted ideas and saw the results change in real-time*,' while Participant P_4 noted, '*The prototype brought some new ideas, but mainly it helped me to organize and reflect on them better*.' Accessing criteria helped participants focus on other relevant points of the processes, exclude ideas, and select the most efficient and effective one. With the adoption of CSTs based on LLMs, concerns arose about LLMs' impact on the diversity of ideas, specifically the potential for homogenization in the creative context. According to Anderson et al (Anderson et al., 2024), the application of LLMs as a creativity support tool can lead different users to produce semantically less distinct ideas. This factor constitutes a relevant point of reflection in this investigation. In light of the data collected, participants did not explicitly report homogenization effects. However, some ideas generated by the prototype were perceived as not very daring, e.g., the Expressiveness factor in the CSI questionnaire did not increase significantly with the prototype's application. These results may suggest a tendency toward similar levels of expressiveness. However, this observation is preliminary and should be interpreted with caution, aligning with the concerns identified in the literature. However, it was also observed that the SAIESE prototype presented distinct ideas (e.g., P_4 said, '*The prototype brought some new ideas, but mainly it helped me to organize and reflect on them better*'). Based on these observations, these findings indicate key design considerations for LLM-based CSTs. Structured guidance supports evaluation and decision-making but can restrict exploratory ideation. When we add structure – through criteria, rankings, and guided feedback – we reduce cognitive load and improve decision flow. However, this may lead to reduced lateral exploration and fewer out-of-the-box ideas. Providing users with greater control and flexibility helps maintain creative flow and minimizes design space fixation. Transparency throughout the process, especially in model outputs, is essential for meaningful interaction. These points illustrate how LLM-based CSTs may shape creativity during human-AI co-creation.

6.2. Human and AI interaction for creativity support

Our findings suggest that participants perceived the relationship with AI as a strategic partner rather than an autonomous creative agent. A dynamic of complementarity characterized the interaction, with AI being used to stimulate, organize, and critique ideas. However, the final decision was left to the human, in line with research presented by Amershi et al (Amershi et al., 2019), where they highlight the importance of maintaining human capacity during interactions with AI, including suggesting practical guidelines such as 'helping but not dominating' and 'adapting to the user's pace,' which provides validation for the design of tools to support creativity.

The emergence of challenges related to the usefulness of the system's feedback and its ability to respond to the proposed task reflects the difficulties associated with LLM-based systems. Faggioli et al. (Faggioli et al., 2024) discuss this point by questioning who decides what is relevant during the process – humans, AI, or both – and emphasize that the balance between human judgment and automated recommendations is still unstable. In the case of SAIESE, human judgment remained predominant in most cases, but the perceived value of the system varied with the degree of alignment with the context and the task at hand, as determined by the prompt inserted into the tool. Another relevant result was the prioritization of ideas. The ranking and selection structure adopted by SAIESE dialogs with approaches such as CrowDEA (Baba et al., 2020), which propose more robust methods of collaborative evaluation. Although the present study used a simplified approach, future work could explore weighted voting mechanisms or multi-criteria analysis that incorporate both human and AI-suggested evaluations. The results suggest that careful contextualization is important when integrating AI techniques into tools that support creativity, given that participants' perceptions varied by workflow preferences, interaction expectations, and tolerance for structured guidance during ideation. In this sense, the findings should be interpreted as observations situated within a specific experimental design context and oriented toward sustainability.

6.3. Reflections, limitations, and future work

The results of this study make a relevant contribution to the role of SAIESE as a tool for supporting creativity, with the support of an LLM. However, it is necessary to acknowledge limitations that constrain the interpretation of the findings and to provide clear directions for future work.

Participants perceived the interface and feedback mechanisms as factors that increased clarity and a sense of control during the generation and evaluation of ideas, supporting reflection and iterative refinement throughout the creative process. However, these results should be interpreted in light of a set of methodological limitations.

Although the experimental design allows for the rigorous isolation and comparison of participants' interaction with and without the prototype, the duration (20 minutes) may not reflect the complexity and prolonged nature of creative processes in real design contexts. Generally, the process of generating, evaluating, and selecting ideas occurs iteratively over more extended periods. In the future, we plan to explore longitudinal evaluations of SAIESE in real-world professional settings, tracking multiple cycles of idea generation, evaluation, and selection over time. This future approach will allow us to observe how the continuous use of the system influences the evolution of ideas, the exploration of the design space, and even potential dependency and saturation effects. Regarding the scope of the task, although an empirical reference on European attitudes toward sustainability was used, this approach raises questions about the contextual sensitivity of LLM-based systems. One way to address this limitation would be to investigate systems that explicitly integrate contextual parameters, such as geographic location, cultural framework, and specific environmental priorities. In this sense, the results can contribute to more situated and contextually informed solutions. Also, participants' limited perception of AI as a strategic partner is not uniform and may be related to differences in experience, creative styles, and expectations. Although the sample's diversity reflects the potential use of the system, the study was not designed to support a comparative analysis across different profiles, given that the literature suggests that experienced and less experienced designers approach creative problems differently (Chen et al., 2022). Currently, the prototype incorporates a relatively uniform interaction model. In this sense, it becomes relevant to develop more flexible and personalized collaboration modes that can accommodate different user profiles and levels of human-AI collaboration. Future developments may include adaptive interfaces that fit the profile, style, preferences, and levels of collaboration. However, many questions remain unanswered about what AI can generate, how it can do so, and how people can and should get involved in the creative process (e.g (Muller et al., 2023)).

Regarding the impact of LLMs on the diversity of ideas, the following observations are based on an exploratory analysis, which is interpreted as preliminary, given the study design and sample size ($n = 18$). Some indicators (e.g., absence of significant differences in the Expressiveness factor of the CSI – does not provide evidence of increased diversity, but it also does not allow conclusions regarding reduced diversity) indicate a possible tendency toward similar levels of expressiveness. These observations align with concerns raised in recent literature (Anderson et al., 2024) about possible homogenization effects associated with LLM use, although our results do not provide sufficient evidence to support such claims. Future studies could address this issue through longitudinal studies and by introducing creative stimulation mechanisms or adaptive interfaces that deliberately promote diversity and the exploration of the design space. Finally, central theme emerging from the qualitative data concerns the strong need for greater control and transparency in SAIESE's functioning, leading to a perception of a 'black box' effect associated with LLMs. This limitation is relevant in this creative context, as human judgment plays a central role. These findings provide initial empirical observations that can inform future, more controlled investigations. As future work, we propose integrating mechanisms for explainability and visualization of reasoning into the prototype, anchored to evaluations, to reinforce confidence and ownership of the system.

7. Conclusion

SAIESE is a creativity support tool that integrates an LLM-based approach to support the creative process through idea generation and evaluation, guided by user-defined tasks and contextualized stimuli. This enables designers to collect and evaluate ideas in a more structured and efficient manner and with greater evaluative support. In this sense, SAIESE supports the different stages of the creative workflow by structuring idea generation and evaluation across divergent and convergent processes within a controlled experimental setting. Our study with 18 designers provides initial evidence that the tool supports aspects of the creative process by facilitating the exploration and refinement of ideas and reducing cognitive effort during the task. This experiment explores the relationship between human creativity and LLM-assisted interaction, highlighting how different system functionalities affect designers' workflows and experiences. The results suggest that SAIESE can support designers at different stages of the creative process, from idea generation to exploration and iterative evaluation (RQ1). However, we identified a set of opportunities and limitations perceived by participants during the creative process (RQ2) – including gains in efficiency, structure, and feedback, alongside challenges related to system flexibility and user control. These challenges, related to transparency, control, and the perceived 'black box' nature of LLMs, raise important directions for the development of human-AI co-creative systems. In conclusion, this work provides an exploratory empirical basis for understanding the role and impact of Generative Artificial Intelligence (GenAI) in supporting human creativity in the design process.

Author contributions

CRediT: **Ana Rodrigues:** Writing – review & editing; **Diogo Cabral:** Supervision; **Pedro Campos:** Supervision.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Allison, J. T., Cardin, M.-A., McComb, C., Yi Ren, M., Selva, D., Tucker, C., Witherell, P., & Fiona Zhao, Y. (2022). Special Issue: Artificial intelligence and engineering design. *Journal of Mechanical Design (1990)*, *144*(2),4 pages. https://doi.org/10.1115/1.4053111

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., Teevan, J., Kikin-Gil, R., & Horvitz, E. (2019). Guidelines for human-AI interaction. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI '19)* (pp. 1–13). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3290605.3300233

Anderson, B. R., Shah, J. H., & Kreminski, M. (2024). Homogenization effects of large language models on human creative ideation. In *Proceedings of the Conference on Creativity & Cognition (C & C '24)* (pp. 413–425). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3635636.3656204

Baba, Y., Li, J., & Kashima, H. (2020). Crowdea: Multi-view idea priorization with crowds. In *The Eight AAAI Conference on Human Computation and Crowdsourcing* (pp. 23–32). arXiv: 2008.02354. Washington, DC, USA: AAAI Press. https://doi.org/10.48550/arXiv.2008.02354

Banathy, B. H. (1996). *Designing social systems in a changing world*. Springer.

Biskjaer, M. M., Dalsgaard, P., & Halskov, K. (2017). Understanding creativity methods in design. In *Proceedings of Designing Interactive Systems (DIS '17)* (pp. 839–851). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3064663.3064692

Bourgault, S., & Jane, E. (2023, Summer). Exploring the horizon of computation for creativity. *XRDS*, *29*(4), 6–9. https://doi.org/10.1145/3596921

Cai, C. J., Reif, E., Hegde, N., Hipp, J., Kim, B., Smilkov, D., Wattenberg, M., Viegas, F., Corrado, G. S., Stumpe, M. C., & Terry, M. (2019). Human-centered tools for coping with imperfect algorithms during medical decision-making. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI'19)* (pp. 1–14). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3290605.3300234

Cetinic, E., She, J., & Eva Cetinic and James She. (2022, May). Understanding and creating art with AI: Review and outlook. *ACM Transactions on Multimedia Computing, Communications, and Applications*, *18*(2), Article 66 (1–22. https://doi.org/10.1145/3475799

Chen, H.-J., Chen, Y.-T., & Yang, C.-H. (2022). Behaviors of novice and expert designers in the design process: From discovery to design. *International Journal of Design*, *16*(3), 59–76. https://doi.org/10.57698/v16i3.04

Cherry, E., & Latulipe, C. (2014, June). Quantifying the creativity support of digital tools through the creativity support index. *ACM Transactions on Computer-Human Interaction*, *21*(4), Article 21 (1–25. https://doi.org/10.1145/2617588

Clarke, V., Braun, V., & Hayfeld, N. (2015). Thematic analysis. *Qualitative Psychology: A Practical Guide to Research Methods*. *2015*, 222–248.

Cropley, A. (2010, June). In praise of convergent thinking. *Creativity Research Journal*, *18*(3), 391–404. https://doi.org/10.1207/s15326934crj1803_13

Cropley, D. (2023). Is artificial intelligence more creative than humans?: ChatGPT and the divergent association task. *Learning Letters*, *2*, 13–13. https://doi.org/10.59453/ll.v2.13

DiStefano, P. V., & Beaty, R. (2026). Chapter 3 - Beyond idea generation: The importance of idea evaluation in human-AI collaborative creativity. In *Generative artificial intelligence and creativity*, edited by Worwood,Matthew J. andKaufman, James C. (pp. 27–36). Elvesier. https://doi.org/10.1016/B978-0-443-34073-4.00013-7

European Commission. (2024). *Attitudes of Europeans towards the environment (Special Eurobarometer 550, survey 3173)*. https://europa.eu/eurobarometer/surveys/detail/3173

Faggioli, G., Dietz, L., Clarke, C. L. A., Demartini, G., Hagen, M., Hauff, C., Kando, N., Kanoulas, E., Potthast, M., Stein, B., & Wachsmuth, H. (2024, April). Who determines what is relevant? Humans or AI? Why not both? Commun. *ACM*, *67*(4), 31–34. https://doi.org/10.1145/3624730

Frich, J., Vermeulen, L. M., Remy, C., Biskjaer, M. M., & Dalsgaard, P. (2019). Mapping the landscape of creativity support tools in HCI. In *Proceedings of Conference on Human Factors in Computing Systems (CHI' 19)* (pp. 1–18). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3290605.3300619

Glassberg, I., Ilan, Y. B., & Zwilling, M. (2025,). The key role of design and transparency in enhancing trust in AI-powered digital agents. *Journal of Innovation & Knowledge*, *10*(5), 100770. https://doi.org/10.1016/j.jik.2025.100770

Gonzalez, G. E., Moran, D. A. S., Houde, S., He, J., Ross, S. I., Muller, M., Kunde, S., & Weisz, J. D. (2024). Collaborative canvas: A tool for exploring LLM use in group ideation tasks. In *Proceedings of the 5th Workshop on Human-AI Co-Creation with Generative Models in ACM*

Conference on Intelligent User Interfaces (IUI ' 24), edited by Soto, A., Zangerle, E. Aachen, Germany: CEUR-WS.org. https://ceur-ws.org/Vol-3660/paper4.pdf

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

Hart, S. G. (2006). Nasa-task load index (NASA-TLX); 20 years later. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 50* (Vol. 9 (pp. 904–908). Thousand Oaks, CA, USA: SAGE Publications. https://doi.org/10.1177/154193120605000909

Harvard Business Review. (2022). How generative AI is changing creative work. Retrieved: August 1, 2023, from https://hbr.org/2022/11/how-generative-ai-is-changing-creative-work

Holzner, N., Maier, S., & Feurriegel, S. (2025). Generative AI and creativity: A systematic literature review and meta-analysis. arXiv: 2505.17241v1. https://doi.org/10.48550/arXiv.2505.17241

Hu, X., Nanjappan, V., & Georgiev, G. V. (2021, April). Bursting through the blocks in the human mind: Enhancing creativity with extended reality technologies. *Interactions*, *28*(3), 57–61. https://doi.org/10.1145/3460114

Huang, S., Long, L., Zhu, Y., & Zhu, J. N. Y. (2026, February). Human-GenAI collaboration across creative phases: Cognitive mechanisms shaping novelty and usefulness. *International Journal of Information Management*, *86*, 102986. https://doi.org/10.1016/j.ijinfomgt.2025.102986

Hubert, K., Awa, K. N., & Zabelina, D. (2023). *Artificial intelligence is more creative than humans: A cognitive science perspective on the current state of generative language models*. PsyArXiv. https://doi.org/10.31234/osf.io/6py2m

IBM Corp. (2021). IBM SPSS Statistics. https://www.ibm.com/br-pt/products/spss-statistics

Jain, R., Garg, N., & Khera, S. N. (2023). Effective human-AI work design for collaborative decision-making. *Kybernetes*, *52*(11(November 2023), 5017–5040. https://doi.org/10.1108/K-04-2022-0548

Jansson, D. G., & Smith, S. M. (1991). Design fixation. *Design Studies*, *12*(1), 3–11. https://doi.org/10.1016/0142-694X(91)90003-F

Lazovsky, G. S., Raz, T., & Kenett, Y. N. (2024, June). The art of creative inquiry - from question asking to prompt engineering. *Journal of Creative Behavior*, *59*(1). https://doi.org/10.1002/jocb.671

Licklider, J. C. R. (1960, March). Man-computer symbiosis. *IRE Transactions on Human Factors in Electronics*, *1*(1), 4–11. https://doi.org/10.1109/THFE2.1960.4503259

Lubart, T. (2016). Creativity and convergent thinking: Reflections, connections and practical considerations. *RUDN Journal of Psychology and Pedagogics*, *4*(4), 7–15. https://doi.org/10.22363/2313-1683-2016-4-7-15

Muller, M., Chilton, L. B., Anna Kantosalo, Q. V. L., Maher, M. L., Martin, C. P., & Walsh, G. (2023). GenAICHI 2023: Generative AI and HCI at CHI 2023. In *Extended Abstracts of the Conference on Human Factors in Computing Systems (CHI EA '23)* (pp. 1–7). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3544549.3573794

Ng, F., Suh, J., & Ramos, G. (2020). Understanding and supporting knowledge decomposition for machine teaching. In *Proceedings of the Designing Interactive Systems Conference (DIS '20)* (pp. 1183–1194). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3357236.3395454

Oh, J., Kim, S., & Kim, S. (2024). LumiMood: A creativity support tool for designing the mood of a 3D scene. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI '24)* (pp. 1–21). New York, NY, USA. Association for Computing Machinery. https://doi.org/10.1145/3613904.3642440

Olsson, T., & Väänänen, K. (2021). How does AI challenge design practice? *Interactions*, *28*(4 (July - August), 62–64. https://doi.org/10.1145/3467479

Oulasvirta, A., Feit, A., Lähteenlahti, P., & Karrenbauer, A. (2017, October). Computational support for functionality selection in interaction design. *ACM Transactions on Computer-Human Interaction*, *24*(5), 1–30. https://doi.org/10.1145/3131608

Oulasvirta, A., & Hornbæk, K. (2016). Hci research as problem-solving. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI '16)* (pp. 4956–4967). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/2858036.2858283

Palani, S., & Ramos, G. (2024). Evolving roles and workflows of creative practitioners in the age of generative AI. In *Proceedings of Conference on Creativity & Cognition (C & C '24)* (pp. 170–184). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3635636.3656190

Petridis, S., Diakopoulos, N., Crowston, K., Hansen, M., Henderson, K., Jastrzebski, S., Nickerson, J. V., & Chilton, L. B. (2023). Anglekindling: Supporting journalistic angle ideation with large language models. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI '23)* (pp. 1–16). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3544548.3580907

Plucker, J. A., Beghetto, R. A., & Dow, G. T. (2004). Why isn't creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. *Educational Psychologist*, *39*(2), 83–96. https://doi.org/10.1207/s15326985ep3902_1

Puranam, P. (2021). Human-AI collaborative decision-making as an organizational design problem. *Journal of Organization Design*, *10*(2), 75–80. https://doi.org/10.1007/s41469-021-00095-2

Rodrigues, A., Cabral, D., & Campos, P. (2024). Fostering informed decision-making: A visual and textual approach to sustainability metrics for idea selection. In *International Conference on Graphics and Interaction (ICGI '24)* (pp. 1–8). IEEE. https://doi.org/10.1109/ICGI64003.2024.10923803

Rodrigues, A., Cabral, D., & Campos, P. (2025). Llms-assisted decision-making in sustainable contexts. In *Proceedings of the 36th Annual Conference of the European Association of Cognitive Ergonomics (ECCE '25). Association for Computing Machinery* (Vol. 9. pp. 1–12, New York, NY, USA. https://doi.org/10.1145/3746175.3746188

Rodrigues, A., Cabral, D., & Campos, P. (2026). Creativity-oriented HCI: From classic approach to LLMs-based CSTs. In *1st International Conference on Human-Computer Interaction in the Alps (AlpCHI '26)* New York, NY, USA: Association for Computing Machinery.

Rodrigues, A., Cabral, D., & Campos, P. F. (2023). Creativity support tools and convergent thinking: A preliminary review on idea evaluation and selection. In *Proceedings of the 15th Conference on Creativity and Cognition (C & amp;C '23)* (pp. 305–311). Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3591196.3596821

Runco, M. A., & Jaeger, G. J. (2012, February). The standard definition of creativity. *Creativity Research Journal*, *24*(1), 92–96. https://doi.org/10.1080/10400419.2012.650092

Shaer, O., Cooper, A., Mokryn, O., Kun, A. L., & Shoshan, H. B. (2024). AI-augmented brainwriting: Investigating the use of LLMs in group ideation. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)* (pp. 1–17). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3613904.3642414

Shneiderman, B. (2000, March). Creating creativity: User interfaces for supporting innovation. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *7*(1), 114–138. https://doi.org/10.1145/344949.345077

Shneiderman, B. (2007, December). Creativity support tool - accelerating discovery and innovation. *Communications of the ACM*, *50*(12), 20–32. https://doi.org/10.1145/1323688.1323689

Shneiderman, B. (2020, March). Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human-Computer Interaction*, *36*(6), 495–504. https://doi.org/10.1080/10447318.2020.1741118

Singh, S., Hindriks, K., Heylen, D., & Baraka, K. (2025). A systematic review of human-AI co-creativity. arXiv: *2506*, 2133. https://doi.org/10.48550/arXiv.2506.21333

Sun, G., & Yao, S. (2012). Investigating the relation between cognitive load and creativity in the conceptual design process. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *56*(1), 308–312. https://doi.org/10.1177/1071181312561072

Sunny, A. D. (2025). Preliminary quantitative study on explainability and trust in AI systems. arXiv: 2510.15769. https://doi.org/10.48550/arXiv.2510.15769

Takaffoli, M., Li, S., & Mäkelä, V. (2024). Generative AI in user experience design and research: How do UX practitioners, teams, and companies use GenAI in industry? In *Proceedings of Designing Interactive Systems Conference (DIS '24)* (pp. 1579–1593). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3643834.3660720

UX Design Institute. (2022, October 21). What is neumorphism in UI design? (And how to use it). https://www.uxdesigninstitute.com/blog/neumorphism-in-ui-design/

Verganti, R., Vendraminelli, L., & Iansiti, M. (2020). Innovation and design in the age of artificial intelligence. *Journal of Product Innovation Management*, *37*(3), 212–227. https://doi.org/10.1111/jpim.12523

Wadinambiarachchi, S., Kelly, R. M., Pareek, S., Zhou, Q., & Velloso, E. (2024). The effects of generative AI on design fixation and divergent thinking. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI' 24)* (pp. 1–18). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3613904.3642919

Wan, Q., Hu, S., Zhang, Y., Wang, P., Wen, B., & Lu, Z. (2024, April). "It felt like having a second mind": Investigating human-AI Co-creativity in Prewriting with large language models. *Proceedings of the ACM Human-Computer Interact*, *8*(1), 1–26. https://doi.org/10.1145/3637361

Wang, B., Li, Y., Lv, Z., Xia, H., Xu, Y., & Sodhi, R. (2024). Lave: LLM-powered agent assistance and language augmentation for video editing. In *Proceedings of the Conference on Intelligent User Interfaces (IUI '24)* (pp. 699–714). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3640543.3645143

Wang, D., Churchill, E., Maes, P., Fan, X., Shneiderman, B., Shi, Y., & Wang, Q. (2020). From human-human collaboration to human-AI collaboration: Designing AI systems that can work together with people. In *Extended Abstracts of the Conference on Human Factors in Computing Systems (CHI EA '20)* (pp. 1–6). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3334480.3381069

Wang, N., Kim, H., Peng, J., & Wang, J. (2025, September). Exploring creativity in human-AI co-creation: A comparative study across design experience. *Human-Media Interaction*, *7*. https://doi.org/10.3389/fcomp.2025.1672735

Zahra Ashktorab, Q. V. L., Dugan, C., Johnson, J., Pan, Q., Zhang, W., Kumaravel, S., & Campbell, M. (2020, October). Human-AI collaboration in a cooperative game setting: Measuring social perception and outcomes. *Proceedings of the ACM on Human-Computer Interaction*, *4*(96), 1–20. https://doi.org/10.1145/3415167

Zhang, R., McNeese, N. J., Freeman, G., & Musick, G. (2021). "An ideal human": Expectations of AI teammates in human-AI teaming. *Proceedings of the ACM on Human-Computer Interaction*, *4*(246 (December 2020), 1–25. https://doi.org/10.1145/3432945

Zorana Ivcevic and Mike Grandinetti. (2024, August). Artificial intelligence as a tool for creativity. *Journal of Creativity*, *34*(2), 2713–3745. https://doi.org/10.1016/j.yjoc.2024.100079

Appendix

A.1. Interview questions

(1) Did you find that the interface made generating and evaluating ideas easier than your usual process? What helped the most? What made it more difficult?
(2) Do you think this tool could integrate well into your current workflow? Why? How do you imagine using it in your day-to-day work?
(3) How did you feel about the speed and fluidity of exploring ideas? Was it satisfactory? Were there times when you felt blocked or frustrated?
(4) Was the level of control you had over the AI outputs sufficient? Did you feel you could guide the creative process, or were you more passive? Would you prefer more customization options, or was the system flexible enough?
(5) After using the prototype, how do you see generative AI's role in your future work? Did you think that the AI brought new ideas, or did it just reorganize what you already had in mind?
(6) Did you feel that the results presented by the prototype in both generation and evaluation were consistent and accurate for your needs? Was there ever a situation in which you doubted the quality of the suggestions?