Towards AI Design Companion: A Three-Stage Framework for Reflective Design Learning with Generative AI

Xinyi Dong, Federico Castagna and Diana Suleimenova
Department of Computer Science, Brunel University of London
Uxbridge, UB8 3PH, United Kingdom
Xinyi.Dong@brunel.ac.uk

Generative artificial intelligence (AI) offers new opportunities for design education, but generic generative interfaces give novice learners powerful outputs without explicit support for how to explore, question, and revise them. This paper presents *AI Design Companion*, an early-stage web prototype for conceptual design activities in which learners expand an initial idea, interrogate assumptions and constraints, and refine a concept proposal. Drawing on reflective practice, creativity support, and recent HCI research on AI-assisted ideation, the prototype organises AI assistance into three iterative roles: a cognitive catalyst, a reflective mirror and sparring partner, and an assistant/evaluator. The paper contributes (1) a theoretically grounded human–AI interaction framework instantiated in a working prototype and (2) preliminary design-refinement insights from a structured four-model proxy critique conducted before human evaluation. The proxy exercise identified needs for stronger reflection prompts, clearer progress cues, and explicit controls for adopting, adapting, rejecting, or further investigating AI suggestions. The work argues for AI interfaces that scaffold novice design judgement while keeping learners responsible for design decisions.

Generative AI, design education, reflective learning, learner agency, human–AI interaction

1. INTRODUCTION

Generative artificial intelligence (AI) is increasingly being explored in creative and design work. Text-to-text and text-to-image systems can quickly generate alternatives, reword proposals, and produce persuasive design language. In design education, this creates a productive tension. AI can help learners move beyond blank-page problems and explore a wider design space, but it can also encourage dependence on polished outputs, similarity across generated ideas, and premature fixation (Doshi and Hauser 2024; Wadinambiarachchi et al. 2024). For novice design learners, the challenge is therefore not only obtaining more suggestions, but learning how to question, compare, and revise those suggestions as part of their developing design judgement.

In this paper, *early-stage conceptual design* refers to short activities in which a learner begins with an incomplete brief or initial idea, explores alternative users and directions, examines assumptions and constraints, and develops a clearer concept proposal. Examples include early product, service, and interface concepts, such as a mobile application intended to help university students reduce food waste. The framework is not intended for detailed engineering, production planning, or final artefact generation.

Design learning involves more than producing artefacts; it requires framing problems, externalising ideas, receiving critique, revising assumptions, and moving iteratively between problem and solution spaces. This aligns with established accounts of design as a distinctive form of knowing (Cross 1982), reflective practice (Schön 1983), and the co-evolution of problem and solution spaces (Dorst and Cross 2001). From a human–computer interaction (HCI) perspective, the central question is not simply whether AI can generate acceptable suggestions, but how an interface can organise AI assistance so that learners remain actively involved in interpreting evidence, questioning assumptions, and making design decisions.

Retaining human responsibility for prompting, evaluating outputs, and deciding revisions is not itself a limitation; it is necessary for learner agency and meaningful human control (Reeve and Tseng 2011;

© Dong et al. Published by BCS Learning and Development Ltd.

Shneiderman 2020). The difficulty is that generic conversational interfaces often place this responsibility on novice learners without scaffolding how these activities can be performed. Human–AI co-creative research similarly shows that interaction structure, system agency, and the positioning of AI can shape reflective engagement, cognition, and perceived control (Rezwana and Maher 2023; Lawton et al. 2023; Xu et al. 2025). Mixed-initiative principles and human–AI interaction guidelines further emphasise clear system capabilities, user control, and opportunities for correction (Horvitz 1999; Amershi et al. 2019).

This paper presents *AI Design Companion*, an early-stage research prototype that structures generative AI support through three iterative roles: a cognitive catalyst for expanding an initial idea, a reflective mirror and sparring partner for challenging assumptions, and an assistant/evaluator for refining a concept proposal. The paper makes two principal contributions. First, it introduces a theoretically grounded human–AI interaction framework and operationalises it in a working web prototype for early-stage conceptual design learning. Second, it reports a formative multi-model proxy critique used to identify interface refinements before human participant evaluation. The fully AI-assisted "vibe-coding" process used to implement the prototype is reported as an additional methodological reflection rather than as a separate principal contribution.

2. BACKGROUND

Design research has long argued that design involves ways of knowing that differ from purely analytical problem solving. Cross (1982) described design as a discipline with its own forms of knowledge, while Schön (1983) characterised professional design activity as a reflective conversation with the situation. Designers make moves, inspect consequences, reframe situations, and continue iterating. Dorst and Cross (2001) similarly argue that creative design involves the co-evolution of problem and solution spaces rather than a sequential movement from a fixed problem to a final solution.

For learners, creativity is visible not only in a final artefact, but also in how they explore alternatives, challenge assumptions, and refine proposals. This aligns with *mini-c* creativity, where creative meaning-making can be personally novel and developmentally significant (Kaufman and Beghetto 2009). Educational support should therefore help learners engage in both divergent exploration and reflective convergence while retaining responsibility for interpreting feedback and revising their work (Wood et al. 1976; Nicol and Macfarlane-Dick 2006; Reeve and Tseng 2011).

Recent HCI research demonstrates that the value and risks of generative AI in design depend substantially on interaction structure. Lawton et al. (2023) designed a co-creative drawing interface around the iterative and reflective nature of creativity, contrasting it with one-shot generative workflows. Shaer et al. (2024) explicitly examined LLM support across divergent idea generation and convergent idea evaluation and selection. Xu et al. (2025) showed that different placements of AI assistance within design workflows shifted participants' cognitive activity and perceptions of agency, highlighting trade-offs between AI positioning, cognitive activity, and users' confidence in the creative process. Conversely, Wadinambiarachchi et al. (2024) found that exposure to AI-generated images could increase design fixation and reduce the fluency, variety, and originality of novice designers' ideas. Together, these studies indicate that creativity support should not only generate stimuli, but also structure exploration, critique, comparison, and selection.

Human–AI co-creation research further shows that interfaces shape whether AI is experienced as a tool, assistant, or partner. The Co-Creative Framework for Interaction (COFI) highlights turn-taking, communication, and contribution types as central design variables (Rezwana and Maher 2023). Related work on AI as a design material and AI-ideation resources likewise suggests that designers require structured ways to understand AI capabilities, limitations, and possible roles (Yildirim et al. 2023; Feng et al. 2023; Yu 2025).

Taken together, this literature identifies three recurrent interaction needs in early-stage conceptual design: broadening an incomplete idea through divergent exploration; interrogating assumptions and consequences through reflective critique; and synthesising a clearer proposal through convergence and formative feedback. The framework presented in this paper operationalises these needs as interface-level scaffolds. They are not proposed as universal or rigid phases of design; rather, they provide a visible and revisitable interaction sequence for novice learners.

3. A THREE-STAGE FRAMEWORK FOR AI-SUPPORTED DESIGN LEARNING

The proposed framework translates the three recurrent interaction needs identified above into idea expansion, reflection/challenge, and proposal improvement. These are interface-level scaffolds rather than a claim that all design activity follows

three fixed stages. Learners can revisit earlier roles as their understanding of the problem and solution co-evolves. The corresponding AI roles are a cognitive catalyst for expanding the design space, a reflective mirror and sparring partner for challenging assumptions, and an assistant/evaluator for supporting refinement.

As shown in Figure 1, learner revision remains central throughout the process. AI responses prompt exploration, critique, and refinement, but learners remain responsible for interpreting, accepting, rejecting, or revising suggestions. This design follows human-centred and mixed-initiative principles that preserve meaningful human control while using automation to extend capability (Horvitz 1999; Shneiderman 2020).

In the first stage, AI acts as a *cognitive catalyst*. It broadens the learner's initial idea, suggests alternatives, and exposes underexplored directions. This role supports divergent exploration without requiring commitment to an early solution (Guilford 1950; Shneiderman 2007; Shaer et al. 2024). Because polished examples can induce fixation, visual suggestions should remain optional and exploratory rather than function as candidate final designs (Wadinambiarachchi et al. 2024).

In the second stage, AI acts as a *reflective mirror and sparring partner*. It prompts learners to examine assumptions, constraints, evidence, and consequences rather than simply requesting more ideas. This role draws on reflective practice and recent HCI findings that targeted, co-led AI prompts can redirect attention towards reflection while preserving opportunities for user input (Schön 1983; Lawton et al. 2023; Xu et al. 2025).

In the third stage, AI acts as an *assistant/evaluator*. It helps learners clarify target users, intended value, trade-offs, and evaluation criteria. This role supports convergence and formative feedback rather than automatic grading or solution replacement (Nicol and Macfarlane-Dick 2006; Shaer et al. 2024). The three roles remain interdependent: critique can reopen exploration, while refinement may expose assumptions that require further reflection.

4. AI DESIGN COMPANION PROTOTYPE

AI Design Companion is an early-stage research prototype implemented as a web application to operationalise the three-stage framework and make its interaction principles testable.[1] The interface allows learners to enter an initial design idea or design challenge, with optional contextual information such as target users, domain, materials, constraints, or intended experience. The main workspace provides three role-based AI interactions corresponding to the framework: *Expand idea*, *Challenge idea*, and *Improve proposal*. A supplementary visual concept function supports early exploration by generating simple visual directions from the learner's current idea. The prototype also includes reflection capture, iteration history, and export functions to help learners review how ideas develop through interaction with AI.

The prototype avoids presenting AI as a single generic answer generator. Instead, each function frames the model's response through a distinct pedagogical role: expansion encourages alternatives and possibilities; challenge supports critique and reflective questioning; and improvement supports synthesis and refinement. This separation aims to reduce prompt burden and make each AI response more legible, reflecting the view that interaction structure is a central design variable in human–AI co-creative systems (Rezwana and Maher 2023).

A reflection box allows learners to record their own responses after receiving AI feedback, while iteration history stores the initial idea, selected AI actions, generated responses, and learner revisions. This design reflects the research focus on how learners move through exploration, critique, and refinement, rather than only whether a final proposal appears stronger. Figure 2 shows how the framework is translated into visible interface elements.

A learner might begin, for example, with a broad idea for a mobile application that helps university students reduce food waste. The expansion role could surface alternative users, shared-household contexts, low-effort interactions, or non-app interventions. The challenge role could question assumptions about sustained ingredient logging, food-safety guidance, student budgets, and whether app engagement would translate into reduced waste. The improvement role could then help the learner articulate a narrower concept, target users, intended value, and criteria for later evaluation. The scenario illustrates how the prototype structures reasoning around a concept rather than delivering a finished solution.

Generative AI tools supported exploratory idea development, prototype coding, debugging, literature note organisation, translation, and academic editing. The web application was developed entirely through AI-assisted "vibe-coding" by a researcher with no formal coding expertise. This implementation experience is reported as a methodological reflection on end-user development of HCI research artefacts, rather than as a separate principal contribution.

[1] The prototype code is stored in a private repository and can be made available upon reasonable request to the authors.

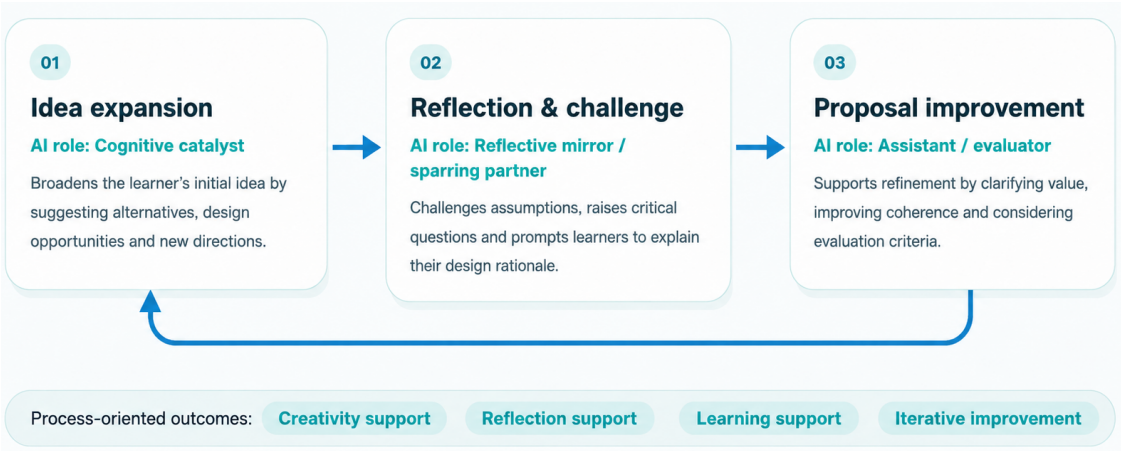

Figure 1: *Three-stage framework for AI-supported design learning. The roles scaffold divergent exploration, reflective critique, and proposal refinement while keeping learner revision central to an iterative process.*

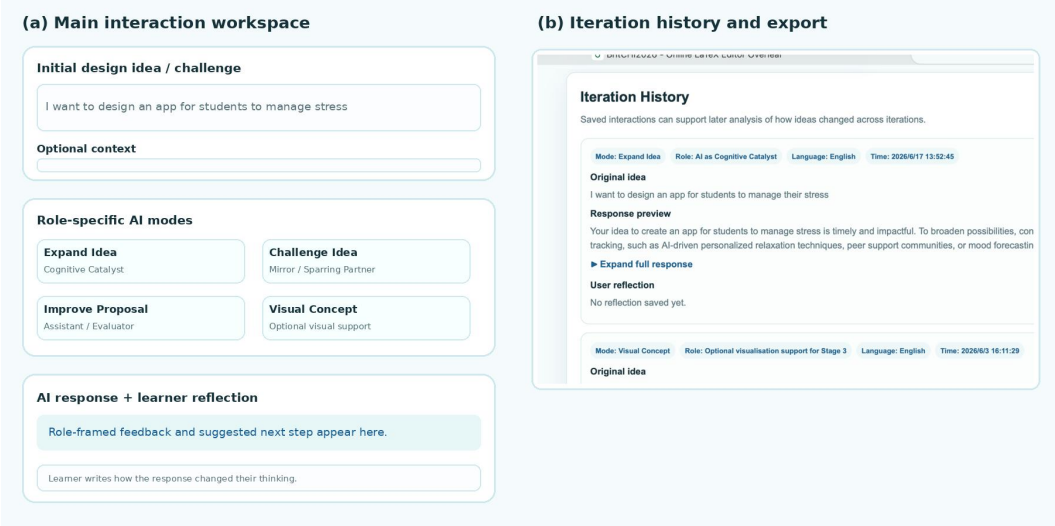

Figure 2: *Prototype interface overview showing the main workspace, role-specific AI interactions, reflection capture, and iteration history.*

5. FORMATIVE PROXY CRITIQUE AND PLANNED USER STUDY

5.1. Protocol and analysis

Before human testing, a formative multi-model proxy critique was conducted to identify problems in the prototype's interface logic and staged workflow. The exercise was designed to support early design refinement.

Four large language models were used as proxy evaluators: GPT-5.5 (OpenAI), Claude Sonnet 4.6 (Anthropic), Gemini 3.1 (Google), and Qwen3.7 (Alibaba/Qwen). All evaluators received the same prompt pack and fixed conceptual design task: developing a mobile application to help university students reduce food waste. The protocol required an interaction summary, five reflection responses, eleven five-point Likert ratings, four open-ended feedback responses, and an HCI-focused analysis of usability, interaction flow, agency, reflection, creativity support, and limitations. The prompt explicitly requested critical rather than exclusively positive feedback and required evaluators to distinguish direct observations, inferences, and recommendations.

The use of multiple LLMs as provisional evaluators was informed by LLM-as-a-judge research, which reports agreement with human preferences while also documenting position, verbosity, self-enhancement, and reasoning biases (Zheng et al. 2023). Likert items were summarised using arithmetic means across the four model reports; no inferential analysis was performed. Open-ended responses were manually collated into recurring design-refinement observations.

Access conditions were heterogeneous. GPT-5.5 and Claude Sonnet 4.6 could inspect the rendered interface but could not trigger its dynamic AI responses. Gemini 3.1 completed a prompt-based simulation. Qwen3.7 reported more detailed interaction, but this access could not be independently verified.

5.2. Descriptive findings

Across the eleven Likert items and four model reports, the overall descriptive mean score (M) was 3.93/5. A sensitivity check excluding the unverified Qwen3.7 report reduced the overall mean to 3.73/5 while preserving the broad ordering of stronger and weaker items.

The highest-rated items concerned the reflection/challenge function and overall task manageability (both $M = 4.50$), followed by reflection on design rationale and potential use in design learning (both $M = 4.25$). The proposal-improvement function received the lowest rating ($M = 3.25$), consistent with uncertainty about live synthesis quality. Open-ended feedback repeatedly requested clearer explanations of interaction styles, stronger reflection prompts, visible progress and iteration cues, clearer framing of visual concept generation, and explicit controls for adopting, adapting, rejecting, or further investigating AI suggestions.

Table 1 connects the main proxy observations to concrete design implications.

5.3. Planned human evaluation

A human pilot study will examine whether learners can understand and use the prototype during an early-stage conceptual design task. Data collection will combine interaction histories, questionnaire responses, and open-ended feedback. Likert-scale items will ask participants about idea exploration, critical reflection, proposal refinement, learner agency, and understanding of their design process. Open-ended questions will examine what participants found useful, confusing, or limiting and how AI responses influenced their reasoning. The pilot will also assess live response quality and actual learner behaviour.

6. DISCUSSION AND FUTURE WORK

The principal HCI contribution of *AI Design Companion* is a role-based interaction scaffold that makes different purposes of AI assistance visible, separately invocable, and revisitable while preserving learner responsibility. The framework does not claim that divergent exploration, reflection, and convergence are new cognitive activities. Rather, it operationalises these recurring activities as explicit interface roles for novice conceptual design. This extends HCI work showing that iterative interaction, phase-specific support, and the positioning of AI shape creative cognition and agency (Lawton et al. 2023; Shaer et al. 2024; Xu et al. 2025).

The paper does not assume that generative AI is inherently constructive in design. Evidence of homogenisation, reduced agency, and design fixation motivates an interface in which suggestions remain provisional and subject to learner judgement (Doshi and Hauser 2024; Wadinambiarachchi et al. 2024).

The proxy critique identified concrete refinements. Learner agency could be made more visible through actions for adopting, adapting, rejecting, or further investigating suggestions. Reflection prompts could ask learners to identify assumptions, evidence gaps, and trade-offs. Progress cues and side-by-side views of original and revised proposals could make iteration more traceable. The visual concept function should be framed as optional and low-fidelity to reduce the risk of premature fixation.

The formative proxy critique has important limitations. Model-generated judgements are not human user evidence and cannot represent human confusion, motivation, design experience, or learning behaviour. The four evaluators had unequal access to the prototype, so ratings of live response quality are not directly comparable. Model sycophancy, shared training biases, and the small number of evaluators also limit the meaning of the arithmetic averages. The exercise therefore supports interface refinement only; it does not demonstrate learning outcomes, creativity gains, or actual user experience.

Future work will conduct the planned human pilot study, refine the interface based on user feedback, and examine whether the staged framework supports different learner groups. A later study could compare the staged prototype with a generic chatbot condition to investigate whether explicit AI roles support reflection, perceived agency, and proposal refinement. Future evaluation could also incorporate established creativity-support instruments such as the Creativity Support Index (Cherry and Latulipe 2014).

7. CONCLUSION

This paper presented *AI Design Companion*, an early-stage web prototype and three-stage human–AI interaction framework for early-stage conceptual design learning. Its two principal contributions are a theoretically grounded, role-based interaction

Table 1: *Formative proxy observations and design implications. Mean scores use a 1–5 scale.*

Observation	Descriptive evidence	Design implication
Critical reflection was most legible	Reflection role $M = 4.50$; rationale reflection $M = 4.25$	Retain the critical-companion role and test its value with learners.
The staged task appeared manageable	Task manageability $M = 4.50$	Retain the staged workflow, while making progress and iteration more visible.
Learner decision-making was insufficiently visible	Learner control $M = 3.75$ and recurring qualitative requests	Add explicit adopt, adapt, reject, and investigate actions.
Proposal refinement remained uncertain	Improvement function $M = 3.25$	Prioritise live human testing of synthesis quality.
Secondary features needed explanation	Recurring comments on interaction styles and visual concepts	Add concise descriptions and frame visuals as optional, low-fidelity exploration.

scaffold instantiated in a working prototype and a formative multi-model proxy critique that identified design priorities before human evaluation. The proxy exercise sharpened requirements for reflection scaffolding, visible learner decisions, progress cues, and cautious visual support. The work argues that AI design companions should organise assistance around explicit purposes while keeping learners responsible for interpreting suggestions and making design decisions. The fully AI-assisted implementation is additionally reported as a methodological reflection on "vibe-coded" end-user prototyping.

REFERENCES

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. and Horvitz, E. (2019) Guidelines for human-AI interaction. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. New York: ACM, 1–13. doi: 10.1145/3290605.3300233.

Cherry, E. and Latulipe, C. (2014) Quantifying the creativity support of digital tools through the Creativity Support Index. *ACM Transactions on Computer-Human Interaction*, 21(4). Article 21. doi: 10.1145/2617588.

Cross, N. (1982) Designerly ways of knowing. *Design Studies*, 3(4). 221–227. doi: 10.1016/0142-694X(82)90040-0.

Dorst, K. and Cross, N. (2001) Creativity in the design process: co-evolution of problem-solution. *Design Studies*, 22(5). 425–437. doi: 10.1016/S0142-694X(01)00009-6.

Doshi, A.R. and Hauser, O.P. (2024) Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, 10(28). eadn5290. doi: 10.1126/sciadv.adn5290.

Feng, K.J.K., Coppock, M.J. and McDonald, D.W. (2023) How do UX practitioners communicate AI as a design material? Artifacts, conceptions, and propositions. In: *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. New York: ACM. doi: 10.1145/3563657.3596101.

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

Horvitz, E. (1999) Principles of mixed-initiative user interfaces. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York: ACM, 159–166. doi: 10.1145/302979.303030.

Kaufman, J.C. and Beghetto, R.A. (2009) Beyond big and little: the four C model of creativity. *Review of General Psychology*, 13(1). 1–12. doi: 10.1037/a0013688.

Lawton, T., Grace, K. and Ibarrola, F.J. (2023) When is a tool a tool? User perceptions of system agency in human–AI co-creative drawing. In: *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. New York: ACM, 1978–1996. doi: 10.1145/3563657.3595977.

Nicol, D.J. and Macfarlane-Dick, D. (2006) Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2). 199–218. doi: 10.1080/03075070600572090.

Reeve, J. and Tseng, C.-M. (2011) Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational Psychology*, 36(4). 257–267. doi: 10.1016/j.cedpsych.2011.05.002.

Rezwana, J. and Maher, M.L. (2023) Designing creative AI partners with COFI: a framework for modeling interaction in human-AI co-creative systems. *ACM Transactions on Computer-Human Interaction*, 30(5). Article 67. doi: 10.1145/3519026.

Schön, D.A. (1983) *The Reflective Practitioner: How Professionals Think in Action*. New York: Basic Books.

Shaer, O., Cooper, A., Mokryn, O., Kun, A.L. and Ben Shoshan, H. (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*. New York: ACM, Article 1050, 1–17. doi: 10.1145/3613904.3642414.

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

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

Wadinambiarachchi, S., Kelly, R.M., Pareek, S., Zhou, Q. and 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*. New York: ACM, Article 380, 1–18. doi: 10.1145/3613904.3642919.

Wood, D., Bruner, J.S. and Ross, G. (1976) The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17(2). 89–100. doi: 10.1111/j.1469-7610.1976.tb00381.x.

Xu, X., Konnova, A., Gao, B., Peng, C., Vo, D. and Dow, S.P. (2025) Productive vs. reflective: how different ways of integrating AI into design workflows affect cognition and motivation. In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. New York: ACM, Article 24, 1–15. doi: 10.1145/3706598.3713649.

Yildirim, N., Oh, C., Sayar, D., Brand, K., Challa, S., Turri, V., Walton, N.C., Wong, A.E., Forlizzi, J., McCann, J. and Zimmerman, J. (2023) Creating design resources to scaffold the ideation of AI concepts. In: *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. New York: ACM, 2326–2346. doi: 10.1145/3563657.3596058.

Yu, W.F. (2025) AI as a co-creator and a design material: transforming the design process. *Design Studies*, 97. Article 101303. doi: 10.1016/j.destud.2025.101303.

Zheng, L., Chiang, W.-L., Sheng, Y., Zhuang, S., Wu, Z., Zhuang, Y., Lin, Z., Li, Z., Li, D., Xing, E.P., Gonzalez, J.E., Stoica, I. and Zhang, H. (2023) Judging LLM-as-a-judge with MT-Bench and Chatbot Arena. In: *Advances in Neural Information Processing Systems*, 36. 46595–46623. doi: 10.52202/075280-2020.