Co-Designing Chatbots with University Educators: A Workshop Protocol and Four Pedagogical Themes

Kate Sangwon Lee
National University of Singapore
21 Lower Kent Ridge Rd, Singapore
katelee@nus.edu.sg

Selina Seohyun Lee
Carnegie Mellon University
Pittsburgh, PA, USA
seohyun3@andrew.cmu.edu

Kai Ming Boon
National University of Singapore
21 Lower Kent Ridge Rd, Singapore
e1122930@u.nus.edu

Enfei Zhang
Brown University
Providence, RI, USA
enfei.zhang@gmail.com

Badrinath Sandhya
National University of Singapore
21 Lower Kent Ridge Rd, Singapore
e1375704@u.nus.edu

Li Hong Idris Lim
National University of Singapore
21 Lower Kent Ridge Rd, Singapore
lihongidris@nus.edu.sg

While LLMs and GenAI have been widely used, a structured way of articulating educators' knowledge into chatbot specifications to address learners' needs is lacking. Structured LLM prompting can help to facilitate chatbot design by educators for learners with diverse needs. This paper presents an educational chatbot co-design workshop protocol, which was conducted for 24 university educators from diverse disciplines, including engineering, social work, public health, language and communication, and business. The co-design workshop protocol introduced a four-step guide — map the bottleneck, position, mock up and optionally prototype — that translates a learning obstacle into a chatbot role and locates that role on a scaffolding-by-feedback Chatbot Role Map. A reflexive thematic analysis of the resulting concepts and worksheets yielded four design themes: personalisation and adaptive learning, engagement and motivation, critical thinking, and perspective-taking. By applying lenses such as self-determination theory (SDT) and the scaffolding-by-feedback Chatbot Role Map, these themes were translated into competence-, autonomy-, and relatedness-oriented designs that appear on the facilitative side rather than settling into a single evaluative, correctness-checking corner that dominates current chatbot deployments. This study introduces methods for designing educational chatbots by translating students' learning obstacles into prompt positioning and detailed interfaces, incorporating pedagogical strategies.

Educational chatbots. Co-design. Self-determination theory. Generative AI. Higher education.

1. INTRODUCTION

Conversational agents are increasingly prevalent in education, especially as large language models (LLMs) and recent advances in generative AI (GenAI) enable educators to adopt AI-powered chatbots for diverse roles in their teaching (Labadze, Grigolia and Machaidze, 2023). Alongside this, the need for tailored chatbots aligned with their courses and intended learning outcomes is increasing.

Educators are domain-knowledge experts, and this benefit allows them to develop the most useful chatbots for their courses, but they seldom have a structured way to articulate that knowledge into a chatbot specification that addresses the learners' needs. Furthermore, not every educator has technical or design experience or skills in designing AI chatbots (Yoo, Jin and Kim, 2025). Another challenge found in current systematic reviews is that most educational chatbots act as teaching-assistant agents oriented towards content delivery and checking correct answers (Labadze, Grigolia and Machaidze, 2023), and their capacity for personalisation and adaptation is under-exploited (Wollny *et al.*, 2021; Kuhail *et al.*, 2023).

Besides that, design principles guiding the development of chatbots to address diverse pedagogical needs are also lacking. For instance, a collaborative design framework for chatbots that is generated for students and educators only builds on self-regulated learning theories (Durall Gazulla, Martins and Fernández-Ferrer, 2023). It has also been discussed that comprehensive processes that emphasise structured LLM prompting can facilitate design by externalising pedagogical thinking and repositioning educators as effective AI-assisted instructional designers (Dourvas, Kokkonis and Kontogiannis, 2026).

To bridge the gap between educational goals and the technical requirements of chatbots, this paper presents a four-step design protocol that enables university educators to create custom chatbots tailored to their domain expertise and specific teaching needs. We report findings from a co-design workshop with 24 university educators in this paper.

The paper makes three contributions. First, it offers a replicable four-step workshop protocol that turns a

© Lee et al.
Published by BCS Learning and Development Ltd
Proceedings of BCS HCI 2026, UK

concrete learning bottleneck into a positioned chatbot role. Second, it explores educational chatbots' roles by expanding educators' diverse disciplinary pedagogical strategies into relevant features. Third, it reads those themes through self-determination theory (SDT) (Deci and Ryan, 2012) and a scaffolding-by-feedback Chatbot Role Map, showing where educator-generated designs land and what that pattern implies for educational chatbots' roles.

2. RELATED WORK

2.1 Roles of chatbots in education

Wollny *et al.* (2021) systematically reviewed 74 publications and identified three primary educational chatbot roles: (1) learning (49%), directly delivering content or skills either inside or outside the curriculum, (2) assisting (20%), handling administrative and logistical tasks without addressing content, (3) mentoring (15%), supporting metacognitive reflection, planning, and self-assessment. Furthermore, three open challenges were identified: (1) aligning assessments with learning objectives to be implemented in the chatbot, (2) exploring the feasibility of the chatbot's role as a mentor, and (3) leveraging the adaptive learning capabilities of chatbots. Building on this, Kuhail *et al.* (2023), reviewing chatbot-learner interaction across 36 studies, find that more than half of the systems acted as teaching agents and that design tended to prioritise content coverage over conversational pedagogy.

The arrival of fluent LLMs sharpens rather than resolves this gap: it is now easy to produce an agent that talks, and correspondingly easy to leave unexamined the question of what role it should play (Debets *et al.*, 2025). Jurenka *et al.* (2024) make a parallel argument at industry scale, reporting on Google's LearnLM-Tutor initiative. The promise of a personal tutor for every learner and a teaching assistant for every teacher remains largely unrealised. They attribute this to the difficulty of translating pedagogical intuitions into generative AI prompts, evaluation practices that lag behind, and the deeper challenge of defining what excellent pedagogy even looks like in the first place.

Through the co-design workshop, the instructors can familiarise themselves with the capabilities of AI chatbots and recommend the most suitable chatbot to students for a specific topic.

2.2 Pedagogical roles with SDT

To systematically design and analyse educational chatbots, it is essential to identify their pedagogical roles. By reviewing various pedagogical lenses, such as constructivism, behaviourism, and inquiry-based learning, we generally characterise instruction using two foundational dimensions: the degree of support (scaffolding) and the type of support (feedback). Together, they form our framework for positioning a chatbot's pedagogical role.

The first dimension, scaffolding, addresses the level of support the chatbot provides. Scaffolding draws on the zone of proximal development (Vygotsky, 1978), which holds that learning happens in the gap between what a learner can do alone and what they can do with help. In modern human-AI collaborative settings, scaffolding is a mechanism for intentionally implementing friction to prompt critical reflection (Dalsgaard, 2025).

Secondly, feedback style as a spectrum between evaluative and facilitative/descriptive, established by Tunstall and Gsipps (1996), comprises the other dimension. According to their distinction, evaluative feedback would involve a backwards-looking judgement by the chatbot on the student's success, while facilitative feedback would guide the student forward with respect to learning outcomes and their achievements.

Finally, we interpret the pedagogical impact of the roles within this space through references to SDT, which holds that user motivation and well-being depend on satisfying the three basic psychological needs of autonomy, competence, and relatedness. As SDT is applied in gamification design in terms of its motivation, user engagement, and personalisation (Krath, Schürmann and Von Korflesch, 2021), we similarly map SDT constructs to our chatbots' themes as shown in Figure 3.

We use these dimensions to help educators design their chatbot's pedagogical role and prompt styles by providing a two-dimensional Chatbot Role Map (Figure 3) with scaffolding level and feedback type as the two axes.

2.3 Co-design with educators

Co-design in education involves educators as active partners because they bring situated knowledge of learning goals, student difficulties and classroom constraints (Tuhkala, 2021). Prior work on teacher-facing AI tools also shows that co-design can help align AI capabilities with teachers' classroom practices and professional judgement (Holstein, McLaren and Aleven, 2019).

Recent studies show how educator co-design can improve the pedagogical and practical relevance of educational AI. In the design of a higher-education chatbot for self-regulated learning, educator involvement helped translate classroom needs into chatbot requirements for providing immediate feedback, accessibility, cross-disciplinary use and responsible data practices (Durall Gazulla, Martins and Fernández-Ferrer, 2023). Ravi *et al.* (2025) similarly involved educators in the design of LLM

tools for project planning, assessment, personalised feedback and progress tracking, while preserving teacher agency and judgement. Together, these studies demonstrate how educators' participation can connect AI capabilities with pedagogical needs and practical relevance.

A practical challenge is that co-design still needs structure to turn educators' pedagogical intentions into clear chatbot design decisions. Previous chatbot co-design research shows that involving educators in the design process can surface their needs, but these needs must still be translated into concrete design requirements (Durall Gazulla, Martins and Fernández-Ferrer, 2023). Recent work on teacher-created pedagogical chatbots also shows that teachers may struggle with the early stages of chatbot design, suggesting a need for structured support in the creation process (Yoo, Jin and Kim, 2025). In response, a co-design approach was adopted for the workshop supported by a lightweight four-step protocol (Figure 1). Within this process, the Chatbot Role Map (Figure 3) helped educators translate learning obstacles into decisions about the chatbot's role, feedback style and scaffolding level.

3. METHOD

3.1 Participants

Twenty-four educators from a research-intensive university participated voluntarily via a centre for teaching and learning. Disciplinary backgrounds included engineering design, electrical engineering, social work, public health, statistics, business and negotiation, and language communication. Participants worked individually on worksheets while exchanging cross-disciplinary reflections in small domain-aligned cohorts as shown in the Appendix.

3.2 Procedure and protocol

The co-design workshop followed a four-step co-design protocol as shown in Figure 1.

In **Step 1 (map the bottleneck)**, they identified their target students, named the specific learning obstacle those students struggle with, and proposed a chatbot role, and conversational feature to address it, consolidating the three into a single statement of the form: "*We want to help [target students] who struggle with [learning obstacle] by creating a chatbot that [role and conversational feature].*" In **Step 2 (position)**, teams placed their proposed feature on a two-by-two Chatbot Role Map whose axes ran from low to high scaffolding and from facilitative to evaluative feedback, forcing an explicit choice about how directive the agent should be and whether it should guide or judge. In **Step 3 (mock-up)**, teams sketched their own concept before a text-to-interface generative tool turned it into a navigable user-interface mock-up. We provided AI Opportunity Cards (Lee, Choi and Asni, 2025) to support educators by showing possible AI capabilities that they can adopt in their chatbots. In the optional **Step 4 (prototype)**, teams used lightweight LLM-assisted "vibe-coding" to turn a mock-up into a working prototype, followed by sharing and peer feedback in which common design patterns were captured.

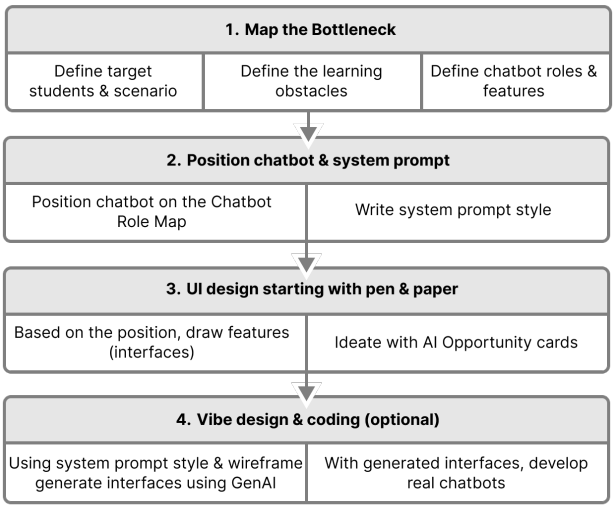

Figure 1. *The four-step co-design protocol applied by each team, from framing the learning conversation to an optional working prototype.*

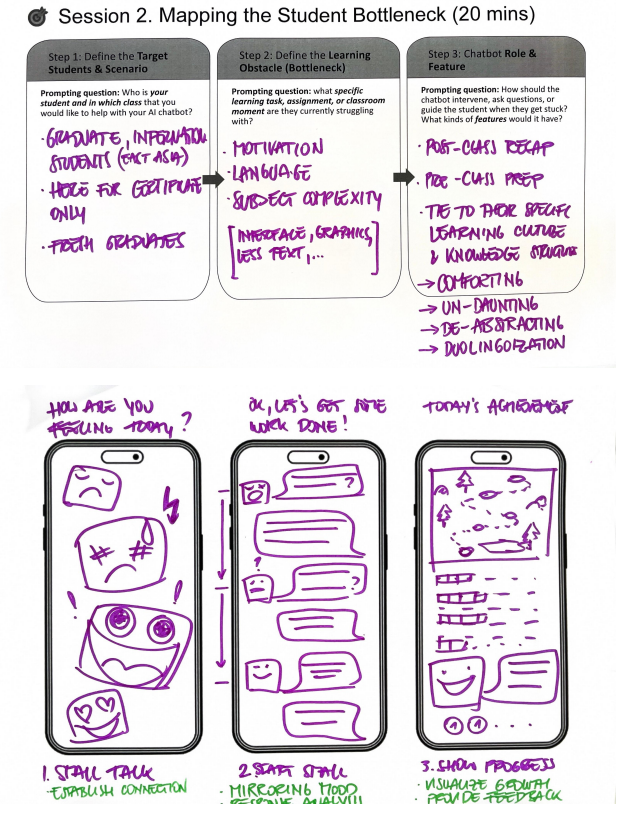

Figure 2. *Example of Step 1 Map the bottleneck (top) & Step 3 pen-and-paper interface sketch (bottom) prior to AI-generated interface.*

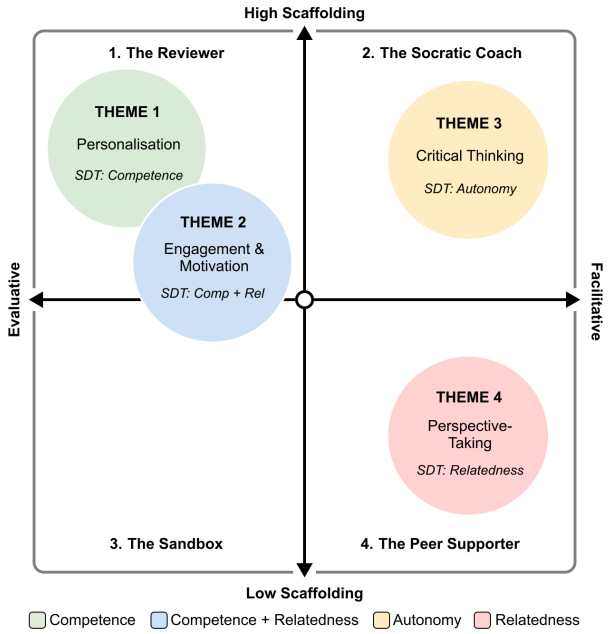

Figure 3. *Chatbot Role Map showing four pedagogical chatbot roles positioned according to scaffolding level (high-low) and feedback style (evaluative-facilitative).*

3.3 Data and analysis

For the data analysis, the completed worksheets, concept sketches, generated interface mock-ups and facilitator notes from the sharing sessions were used. Each educator's materials were treated as one chatbot concept, yielding 24 concepts in total. The chatbot concepts were analysed using reflexive thematic analysis (Braun and Clarke, 2019). Five researchers first coded the role and conversational features of each concept inductively, then clustered the codes into candidate themes through iterative discussion. From this process, four themes emerged. In a second, deductive pass, each theme was mapped onto its primary self-determination need and its region of the scaffolding-by-feedback Chatbot Role Map. Themes are interpretive constructs rather than mutually exclusive categories; a single concept could touch more than one.

4. FINDINGS

Four themes were identified, as summarised in Table 1 alongside their primary self-determination need and design-space position; the narrative below gives a representative concept for each, described by disciplinary role. Each theme was placed on the scaffolding-by-feedback Chatbot Role Map and coloured according to its primary self-determination need (Figure 3). The four themes address all three self-determination needs; competence-oriented themes sit closer to scaffolded/evaluative support, while autonomy- and relatedness-oriented themes sit on the facilitative side.

Table 1: *The four design themes, their primary self-determination needs, and their positions on Chatbot Role Map.*

Design theme	Primary need (SDT)	Position on the Chatbot Role Map	Representative concept
Theme 1. Personalisation and Adaptive Learning	Competence	High scaffolding & evaluative	Diagnostic question generation; grade projection with growth feedback
Theme 2. Engagement & Motivation	Competence & Relatedness	Variable scaffolding & evaluative	Streaks and progress maps; a peer-support corner
Theme 3. Critical Thinking	Autonomy	High scaffolding & facilitative	A Socratic partner that challenges assumptions and frames research gaps
Theme 4. Perspective-Taking	Relatedness	Low scaffolding & facilitative	Role-play of negotiations; a virtual family for social-work practice

4.1 Theme 1. Personalisation and adaptive learning

The most common theme treated the chatbot as a tutor that calibrates difficulty to the individual. Educators commonly envisioned systems that first diagnosed a learner's current level before providing tailored support.

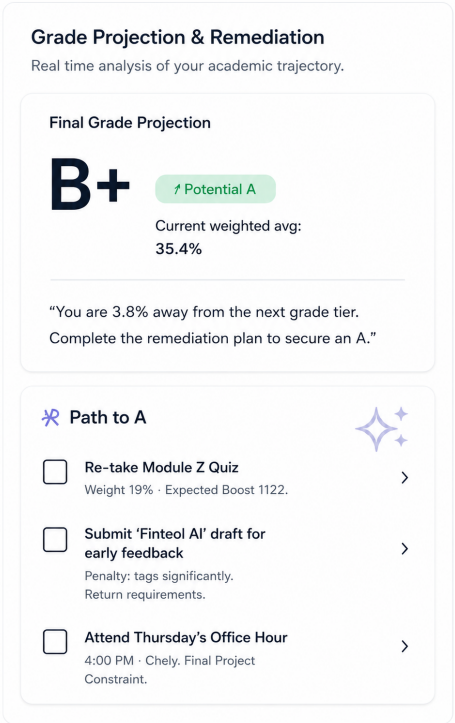

Figure 4. *Example interface of E19's AI tutor, which projects a student's final grade and provides a personalised remediation pathway through adaptive feedback and recommended learning actions.*

Figure 4 shows one such concept, which projects a student's likely course performance and translates it

into personalised remediation tasks and targeted learning recommendations. Another educator wanted an agent that generates questions matched to a student's level, on the premise that "students don't know what they don't know". In self-determination terms, these concepts target competence: by diagnosing the learner and pitching the next task just beyond current reach, they operationalise scaffolding within the zone of proximal development. They sit high on scaffolding and, because they both assess level before guiding, lean towards the evaluative side of the space.

4.2 Theme 2. Engagement and motivation

A second theme designed for sustained effort. Concepts here included a daily mood check-in, streaks and progress maps inspired by language-learning apps, concept-preview videos, ranking and a growth tracker, and — distinctively — a "peer-support corner" connecting students working on the same topic. Figure 5 illustrates how these motivational features can be integrated into a single interface, combining social participation with visual progress indicators to encourage continued engagement. These map onto two needs at once. Following Sailer *et al.* (2017), the progress indicators, streaks and rankings support competence, while the peer-support and social features support relatedness. The theme is largely evaluative and motivational rather than correctness-checking, and its scaffolding varies with the feature. The dual mapping is a useful caution: gamification is not one thing, and a leaderboard and a peer corner recruit different needs.

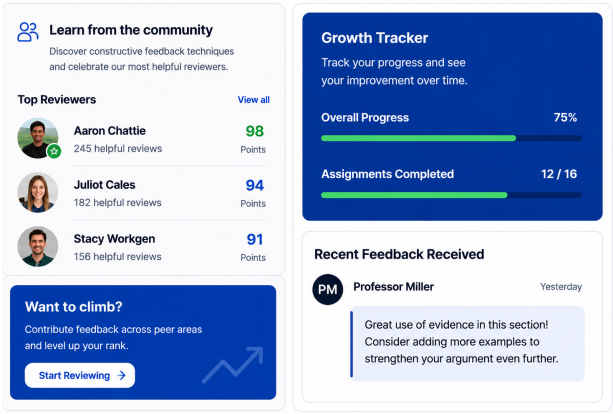

Figure 5. *Example interface of E02's PeerPulse, showing a community feed, growth tracker, and leaderboard designed to motivate sustained learning through social engagement and progress tracking.*

4.3 Theme 3. Critical thinking

Educators structured chatbots as Socratic partners to strengthen learner autonomy. Instead of answering queries directly, the agent challenged underlying premises and scaffolded inquiry (Figure 6). E05 (language communication) specified conversational rules prompting learners to interrogate source perspectives and uncover unstated rhetorical biases. E09 (public health) designed an agent guiding undergraduates to isolate literature gaps without formulating hypotheses for them. These agents operate under high scaffolding with strictly facilitative feedback.

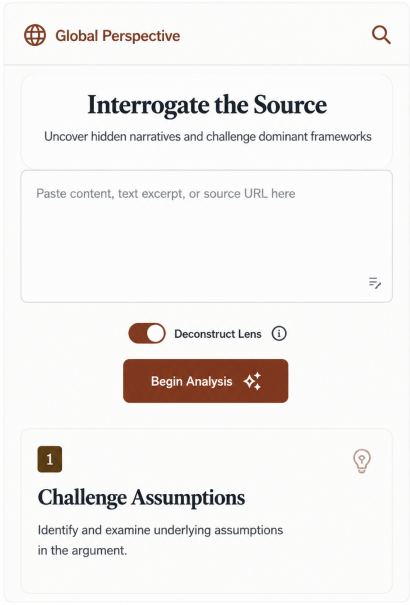

Figure 6. *Example interface of E05's Global Perspective, encouraging students to interrogate sources and challenge dominant frameworks through Socratic questioning.*

4.4 Theme 4. Perspective-taking

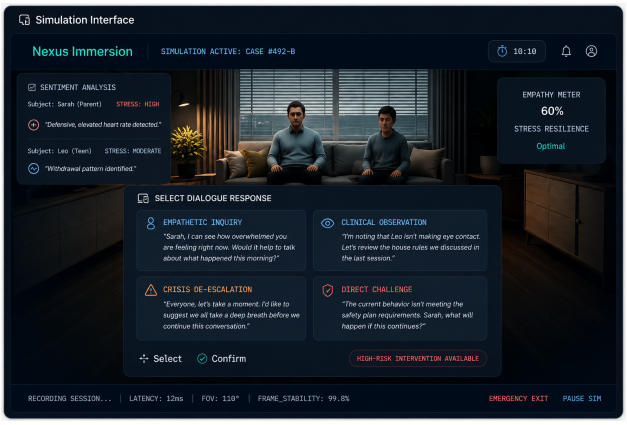

Figure 7. *Example interface of E13's Nexus Immersion, illustrating a role-play chatbot that immerses students in a virtual family scenario to develop empathy and perspective-taking skills.*

A fourth theme used the chatbot to let students inhabit other points of view. As illustrated in Figure 7, a social-work educator designed an agent offering "virtual family" feedback so that students could rehearse the human and emotional dimensions of practice; educators in negotiation and communication designed role-play and scenario-

selection features for conflict management and mindset shift. These concepts target relatedness, using simulated social others to build empathy and social skill. They are facilitative and low scaffolding: the agent stages a believable encounter and lets the student act within it.

5. DISCUSSION

5.1 Application of SDT and incorporation of pedagogical strategies

Our four-step co-design protocol and the Chatbot Role Map are ways to make pedagogical theory operable for educators by asking them to position their chatbots and make two decisions that conversational-agent authoring usually leaves implicit: how directive the agent should be and whether it guides or judges. In doing so, it connected educators' tacit pedagogical knowledge to scaffolding theory and to the distinction between facilitative and evaluative feedback. The findings suggest that the practical challenge is not whether educators have pedagogical ideas for chatbots, but how to structure those ideas into design decisions. The four-step protocol made this translation visible and addressed existing challenges in mapping educators' knowledge to specific prompts and roles (Durall Gazulla, Martins and Fernández-Ferrer, 2023).

5.2 Autonomy and relatedness alongside competence

The headline observation is that, when scaffolded by a short protocol, educators design need-supportive and facilitative partners rather than answer-checkers. This contrasts with the deployed systems catalogued in the literature, which tend to prioritise content delivery and correctness (Wollny *et al.*, 2021; Kuhail *et al.*, 2023). The needs that educator designs reached for most readily — autonomy through Socratic questioning and relatedness through role-play and peer support — are precisely the ones that educational technology is often observed to neglect in favour of competence alone. Given the choice, these educators populated the harder, more human regions of the space.

5.3 Trade-offs

The engagement theme's split across competence and relatedness carries a design caution. Because distinct game elements satisfy distinct needs (Sailer *et al.*, 2017), motivational features cannot be treated interchangeably, and competitive, extrinsic elements such as leaderboards risk crowding out the autonomy that the critical-thinking designs depend upon. More broadly, the findings suggest that chatbot authoring tools should expose role and positioning as first-class choices rather than defaulting to a question-and-answer template; a tool whose defaults are evaluative will quietly steer educators away from the autonomy- and relatedness-supportive designs they would otherwise create.

5.4 Limitations

Several limitations bound these claims. The study involved a single institution, a single three-hour session, and a self-selected group of educators already interested in AI. The concepts were designed but not deployed, so we report design intent rather than learning outcomes. The thematic mapping is interpretive, the cohort is modest in size, and the generative mock-up tool may have nudged concepts towards particular interface patterns. The findings are best read as a grounded hypothesis about how educators design when supported, not as a measured effect on students.

6. FUTURE WORK AND CONCLUSION

The next step is to deploy a subset of these concepts and measure their effect, using validated self-determination need-satisfaction instruments alongside learning measures, to test whether facilitative, need-supportive designs deliver on their promise. Running the protocol at other institutions and across other disciplines would test its transferability, and releasing the worksheets and protocol openly would let other educators adopt it directly. A further line of work concerns translation: how a positioning choice on the Chatbot Role Map becomes the prompt logic and behavioural boundaries of an actual LLM agent, and whether the intended conversational role survives contact with a fluent model that is strongly inclined to simply answer.

In conclusion, this paper has presented a replicable four-step co-design protocol for educational chatbots with university educators and an analysis of the designs it produced across a disciplinarily diverse cohort. Read through self-determination theory and a scaffolding-by-feedback Chatbot Role Map, the designs span competence, autonomy and relatedness, with autonomy- and relatedness-oriented designs appearing on the facilitative side rather than settling into the evaluative tutor role that current deployments favour. When educators are given a structured way to design the learning conversation, they choose to make it a richer one.

REFERENCES

Braun, V. and Clarke, V. (2019) 'Reflecting on reflexive thematic analysis', Qualitative research in sport, exercise and health, 11(4), pp. 589-597.

Dalsgaard, P. (2025) 'Scaffolding Metacognition in Design with Generative AI'.

Debets, T., Banihashem, S. K., Joosten-Ten Brinke, D., Vos, T. E., de Buy Wenniger, G. M. and Camp, G. (2025) 'Chatbots in education: A systematic

review of objectives, underlying technology and theory, evaluation criteria, and impacts', Computers & Education, 234, pp. 105323.

Deci, E. L. and Ryan, R. M. (2012) 'Self-determination theory', Handbook of theories of social psychology, 1(20), pp. 416-436.

Dourvas, I., Kokkonis, G. and Kontogiannis, S. (2026) 'Reconceptualizing Prompt Engineering as Reflective Professional Practice: A Framework for Teacher Development', Electronics, 15(5), pp. 930.

Durall Gazulla, E., Martins, L. and Fernández-Ferrer, M. (2023) 'Designing learning technology collaboratively: Analysis of a chatbot co-design', Education and Information Technologies, 28(1), pp. 109-134.

Holstein, K., McLaren, B. M. and Aleven, V. (2019) 'Co-designing a real-time classroom orchestration tool to support teacher-AI complementarity', Grantee Submission.

Jurenka, I., Kunesch, M., McKee, K. R., Gillick, D., Zhu, S., Wiltberger, S., Phal, S. M., Hermann, K., Kasenberg, D. and Bhoopchand, A. (2024) 'Towards responsible development of generative AI for education: An evaluation-driven approach', arXiv preprint arXiv:2407.12687.

Krath, J., Schürmann, L. and Von Korflesch, H. F. (2021) 'Revealing the theoretical basis of gamification: A systematic review and analysis of theory in research on gamification, serious games and game-based learning', Computers in human behavior, 125, pp. 106963.

Kuhail, M. A., Alturki, N., Alramlawi, S. and Alhejori, K. (2023) 'Interacting with educational chatbots: A systematic review', Education and Information Technologies, 28(1), pp. 973-1018.

Labadze, L., Grigolia, M. and Machaidze, L. (2023) 'Role of AI chatbots in education: systematic literature review', International journal of Educational Technology in Higher education, 20(1), pp. 56.

Lee, K. S., Choi, M. and Asni, E. Y. 'AI Opportunity Cards: Developing a Toolkit for AI as a Design Material'. Proceedings of the 2025 ACM Conference on Information Technology for Social Good.

Ravi, P., Masla, J., Kakoti, G., Lin, G. C., Anderson, E., Taylor, M., Ostrowski, A. K., Breazeal, C., Klopfer, E. and Abelson, H. 'Co-designing large language model tools for project-based learning with k12 educators'. Proceedings of the 2025 CHI conference on human factors in computing systems, 1-25.

Sailer, M., Hense, J. U., Mayr, S. K. and Mandl, H. (2017) 'How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction', Computers in Human Behavior, 69, pp. 371-380.

Tuhkala, A. (2021) 'A systematic literature review of participatory design studies involving teachers', European Journal of Education, 56(4), pp. 641-659.

Tunstall, P. and Gsipps, C. (1996) 'Teacher feedback to young children in formative assessment: A typology', British educational research journal, 22(4), pp. 389-404.

Vygotsky, L. S. (1978) Mind in society: The development of higher psychological processes. Harvard university press.

Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M. and Drachsler, H. (2021) 'Are we there yet?-a systematic literature review on chatbots in education', Frontiers in artificial intelligence, 4, pp. 654924.

Yoo, M., Jin, H. and Kim, J. (2025) 'How Do Teachers Create Pedagogical Chatbots?: Current Practices and Challenges', arXiv preprint arXiv:2503.00967.

APPENDIX: List of educators

No.	Faculty
E01	Communication and new media
E02	English language communication
E03	Design and engineering
E04	Systems science
E05	English language communication
E06	English language communication
E07	English language communication
E08	English language communication
E09	Public health
E10	Language studies
E11	Care unit
E12	Design and engineering
E13	Arts and social sciences, social work
E14	English language communication
E15	Electrical and computer engineering
E16	Computing
E17	Design and engineering
E18	University educational development
E19	University educational development
E20	Electrical and computer engineering
E21	University educational development
E22	University educational development
E23	English language communication
E24	Design and engineering