Scaffolding Autonomy:

Designing for Trust and Learning in AI-Supported Autodidactic Environments

Matthew Hadick

Dept. of Human Computer Interaction, Iowa State University

COMST5900 - Human Communication with and through AI and Robotics

Instructor: Dr. Mary Fornoff

June 13, 2025

Introduction

As artificial intelligence systems are increasingly integrated into educational contexts, their role is changing – from passive tools to active, dynamic facilitators of learning. Particularly in autodidactic learning environments – where learners guide their own journeys outside of formal instruction – AI agents are not merely sources of information but partners in active knowledge construction. In such contexts, the learner must not only comprehend the content but also establish a relationship with the system itself. This research seeks to explore how certain design decisions – from aesthetic to verbal and non-verbal communication – impact human trust in AI systems and sense of social presence, particularly within these kinds of autonomous learning environments.

The central research question driving this study is: How do design choices regarding agent tone, visual expression, responsiveness, and transparency impact trust, motivation, and comprehension within AI-mediated autodidactic systems? Drawing on constructs such as AI trust, social presence, instructional alignment, and cognitive load, as well as the aesthetic-usability effect, this paper investigates which design qualities learners respond to most positively, both affectively and in terms of measured learning effectiveness and performance.

To explore these dynamics in practice, this study introduces Seedling, a research-driven prototype that applies findings from recent literature to test how specific design choices shape trust, presence, and learning in self-directed environments. Drawing on insights from the studies and theories explored throughout this paper, Seedling serves as both an experimental platform and a conceptual model for relational AI design in education.

Theoretical Framework

This study is grounded in five intersecting theoretical frameworks that, together, show the complex interplay of AI-mediated learning environments: Bandura's Self-Efficacy Theory, Social Presence Theory, Cognitive Load Theory, the Aesthetic-Usability Effect, and Donald Norman's Emotional Design. These frameworks, while emerging from different disciplinary traditions, converge around shared concerns of human agency, cognitive processing, and the quality of computer mediated communication.

At the foundation of self-directed learning lies Bandura's (1997) conception of self-efficacy, which he defines as "the belief in one's capabilities to organize and execute the courses of action required to manage prospective situations." This theory is particularly crucial for understanding autodidactic learning environments, where learners must make their way through educational content without relying on traditional institutional scaffolding. Bandura's framework suggests that a learner's sense of efficacy is deeply connected to motivation and their willingness to work through intellectual challenges. When applied to AI-mediated learning systems, this theory assumes a new dimension, where technological interfaces must do more than simply present information, they must actively reinforce confidence and encourage intellectual risk-taking amongst users. Systems that sufficiently scaffold learner confidence become foundational to sustained autodidactic engagement, cultivating positive feedback loops between technological interaction and personal agency.

Building upon this understanding of individual agency, Social Presence Theory offers a complementary perspective on the relational dimensions of learning. Developed by Short, Williams, and Christie (1976), this framework circles the perceived awareness of another intelligence during mediated communication. When extended to AI contexts, Social Presence Theory raises questions about whether learners can genuinely sense an intentional, responsive partner in their interactions with AI systems. This perceived presence (or lack thereof) has significant implications for how learners form relationships with their technological learning environments and how these relationships affect both relational satisfaction and educational outcomes.

The cognitive dimensions of these interactions cannot be overlooked, and here Cognitive Load Theory provides essential insights. Developed by Sweller (1988), this framework underscores the critical importance of balancing instructional complexity with human cognitive capacity. The theory's central premise is that learning is optimized when educational systems carefully manage the demands placed on working memory. In AI-mediated learning environments, this principle becomes particularly relevant as these systems can easily overwhelm users with information density, unpredictable responses, or excessive verbosity. A system that exceeds a user's cognitive processing capabilities will ultimately hinder rather than facilitate learning. As such, this framework provides necessary guidance for designing AI interactions that consider cognitive quirks while still providing meaningful intellectual challenge (also a tenet of inclusive design.)

The aesthetic dimensions of these cognitive and relational processes are further expanded on by the Aesthetic-Usability Effect, first invoked by Kurosu and Kashimura (1995). This principle states that users are more likely to perceive aesthetically pleasing interfaces as easier to use, regardless of their actual functional usability. For educational technology, this means that the aesthetic quality of an AI learning system may contribute substantially to the emotional foundation that encourages continued use. When learners find an interface visually compelling, they might be more likely to persevere through difficult learning moments, more willing to explore advanced features, and more likely to attribute positive intentions to the system.

These multiple theoretical perspectives converge most fully in Donald Norman's (2003) framework of Emotional Design, explored in detail in his book *Emotional Design: Why We Love (Or Hate) Everyday Things*, which provides a comprehensive framework for understanding these nuances of HCI. Norman's model identifies three distinct but interconnected levels of design impact: visceral (immediate emotional response to appearance), behavioral (satisfaction derived from usability and functionality), and reflective (deeper meaning-making and identity formation through use). This framework is particularly valuable for analyzing AI-mediated learning because it acknowledges that educational technology operates simultaneously across emotional, functional, and existential levels. Understanding how these layers interact in AI learning environments offers important insight into both the cognitive and emotional foundations of trust formation. When an AI system succeeds at the visceral level through appealing design, performs effectively at the behavioral level through intuitive functionality, and supports meaningful learning goals at the reflective level, it creates conditions for deep and sustained educational engagement that transcends simple tool use to become genuine intellectual partnership.

Literature Review

Recent scholarship provides compelling evidence that trust and social presence function as central pillars of learner engagement in AI-mediated educational environments. This growing body of research reveals nuanced relationships between technological design choices and educational outcomes that extend far beyond simple measures of whether or not they work as intended.

Current research on AI-mediated personalized learning shows significant growth in the area but reveals important gaps in how and where these systems are being studied. A recent systematic review of 78 studies found that most AI education research focuses heavily on formal university settings, with very little attention to workplace learning or informal education contexts where adults actually do most of their lifelong learning (Bayly-Castaneda et al., 2024). While adaptive learning technologies currently dominate the field, there's growing interest in newer approaches like generative AI models. Importantly, this research emphasizes that simply adding AI to education doesn't automatically improve learning – the technology must be thoughtfully designed with sound educational principles, particularly attention to issues like bias, privacy, and ensuring all learners can access these tools effectively.

The concept of AI as a collaborative learning partner has gained particular attention through Wang et al.'s (2025) work on generative co-learners (GCLs). These AI-powered agents are specifically designed to simulate peer learners in asynchronous educational settings, representing a significant departure from traditional AI tutoring models. Wang and colleagues' empirical investigation revealed that GCLs substantially enhanced both cognitive and social presence among learners, creating richer collaborative experiences than non-interactive control conditions. Particularly noteworthy was learners' consistently higher ratings of GCLs on crucial social metrics including social awareness and group cohesion. However, the study also revealed an important temporal dimension: while social and cognitive presence increased markedly, measurable learning gains did not necessarily follow within the limited timeframe of the investigation. This finding suggests that the benefits of AI collaboration may manifest over longer periods or through more subtle mechanisms than traditional assessment measures can capture.

Building upon this understanding of AI-as-peer dynamics, Ko and Joo (2025) provide additional evidence for the transformative potential of reframing AI relationships in educational contexts. Their research focused specifically on collaborative scientific investigations, where AI agents were positioned as investigative partners rather than instructional authorities. This strategic reframing yielded significant increases in both student trust and engagement levels, reinforcing the hypothesis that learners respond more positively to AI systems that enhance rather than replace human agency. These findings carry particular significance for autodidactic learning environments, where learner confidence and self-direction are paramount. When AI systems successfully position themselves as supportive collaborators rather than evaluative tutors, they appear to create psychological conditions more conducive to the risk-taking and experimentation that characterize effective self-directed learning.

The mediating role of social presence in educational outcomes receives further empirical support from Suryanto et al.'s (2024) article, which revealed that social presence successfully mediated the relationship between interaction and student satisfaction, as well as between course structure and satisfaction, though it failed to mediate self-regulation effects. This finding suggests that while social presence serves as a crucial bridge between certain design elements and learning outcomes, its mediating power varies depending on the specific predictor variable. Notably, the study found that course structure alone did not directly influence satisfaction but required social presence as an intermediary – highlighting the critical role of perceived interpersonal connection in transforming structural design choices into meaningful learning experiences.

Birmingham et al.'s (2020) study of robot-mediated support groups provides practical evidence that AI can successfully build trust in group learning settings. Their humanoid robot helped 81 university students share academic stress by asking increasingly personal questions and making its own disclosures, which significantly increased trust both among participants and toward the robot itself. Students reported feeling less alone and learning new coping strategies, though they noted the robot's "lack of humanity" due to its inexpressive face and mechanical interactions. This research demonstrates that even simple AI systems can facilitate meaningful group connections when designed thoughtfully, while highlighting the ongoing challenge of making AI feel emotionally authentic rather than merely functional in educational contexts.

However, a fundamental question remains about which theoretical frameworks best capture how students develop trust in educational AI systems. Pitts and Motamedi (2025) directly investigated whether students trust AI chatbots more like human peers/instructors or like software tool.s Their experiment revealed that both frameworks simultaneously influence student perceptions, with human-like trust more strongly predicting emotional responses (trusting intention, perceived enjoyment) while system-like trust better predicted behavioral outcomes (usage intention, perceived usefulness). This finding suggests that educational AI occupies a unique hybrid category that requires new theoretical frameworks beyond traditional human-human or human-technology trust models.

The visual and aesthetic dimensions of AI design have emerged as surprisingly influential factors in these relational dynamics. Nowak and Biocca's (2003) findings challenge conventional assumptions about anthropomorphic design in AI systems. Their research demonstrated that abstract or low-anthropomorphic visual representations actually increased perceptions of social presence and telepresence more effectively than realistic human-like avatars. This finding suggests that learners may form stronger connections with AI systems that maintain a degree of visual ambiguity, perhaps allowing users to project their own relational expectations onto the interface rather than being constrained by specific human-like characteristics. Maehigashi et al.'s (2023) study provides complementary evidence for this nuanced relationship between appearance and trust. When participants worked with AI agents, humans, or social robots on calculation and emotion recognition tasks, trust in the robot consistently settled in the middle rather than matching either extreme. While people initially over-trusted AI agents and then sharply lost trust when errors occurred, the humanoid robot's physical appearance seemed to moderate these swings. This suggests that giving AI systems human-like features might help users develop more balanced, appropriate trust levels—avoiding both blind faith in "perfect" AI and complete rejection after mistakes. Liu and Su's (2024) meta-analysis of 33 experiments reinforces these findings, demonstrating that facial anthropomorphic elements significantly improve learning transfer, retention, and comprehension while enhancing positive affect and intrinsic motivation, though they caution that such features may also increase extraneous cognitive load.

The ethics of anthropomorphic design receive critical examination in Aylett et al.'s (2023) work on embodied conversational agents for vulnerable populations. Their research raises fundamental questions about whether designing AI systems to appear human-like constitutes deception, particularly when working with users who may disclose sensitive personal information. The authors argue that such systems should be understood as "depictions of social actors" – similar to fictional characters – that rely on users' "willing suspension of disbelief" rather than genuine belief in AI sentience. The research emphasizes that transparency about system capabilities remains crucial, particularly when designing for vulnerable populations in educational settings.

Floridi's (2025) recent work adds important nuance to these design considerations by distinguishing between different dimensions of AI perception. While confirming that anthropomorphic design elements can indeed elevate perceived empathy and trust, Floridi's research reveals a crucial disconnect: high levels of perceived intelligence in an AI system do not necessarily correlate with emotional attunement. This finding has significant implications for educational AI design, suggesting that systems optimized purely for demonstrating cognitive capability may paradoxically fail to create the emotional conditions necessary for sustained learning engagement. The research implies that effective educational AI must balance demonstrations of competence with expressions of emotional sensitivity and responsiveness.

The foundational role of aesthetic design in shaping user perceptions receives further support from David and Glore's (2010) comprehensive analysis of learner platform preferences. Their investigation consistently demonstrated that learners rated aesthetically pleasing learning platforms as more trustworthy and user-friendly, providing empirical validation for the aesthetic-usability effect in educational contexts. This research shows the importance of visual consistency, thoughtful spacing, and appropriate tonal choices in user interface design, suggesting that these elements function not merely as surface-level enhancements but as fundamental contributors to educational effectiveness.

The question of system transparency introduces additional complexity to these design considerations. Hayes et al. (2025) provide compelling evidence that AI agents capable of meaningful self-explanation foster substantially greater user understanding and trust. Their work confirms the educational value of explicability in systems where learners must rely on artificial intelligence to guide their intellectual progress. However, recent investigations have revealed important limitations to this transparency principle. Research indicates that overly conspicuous disclosure of AI assistance can undermine perceived credibility and user trust.

Additional research by Kim et al. (2022) illuminates the ways users evaluate AI social interactivity, revealing that perceived conversational ability and attentiveness play central roles in trust calibration processes. Their findings indicate that systems designed with conversational sophistication – rather than superficial chattiness – have substantially better engagement outcomes across diverse user populations. This work aligns closely with McKee et al.'s (2022) findings, which demonstrate that warmth and competence, rather than visual realism, emerge as the most reliable design attributes influencing user preference in AI collaboration contexts. Together, these studies suggest that successful educational AI design requires careful attention to relational qualities that extend far beyond technical functionality, encompassing the subtle interpersonal dynamics that characterize effective human collaboration and learning partnerships.

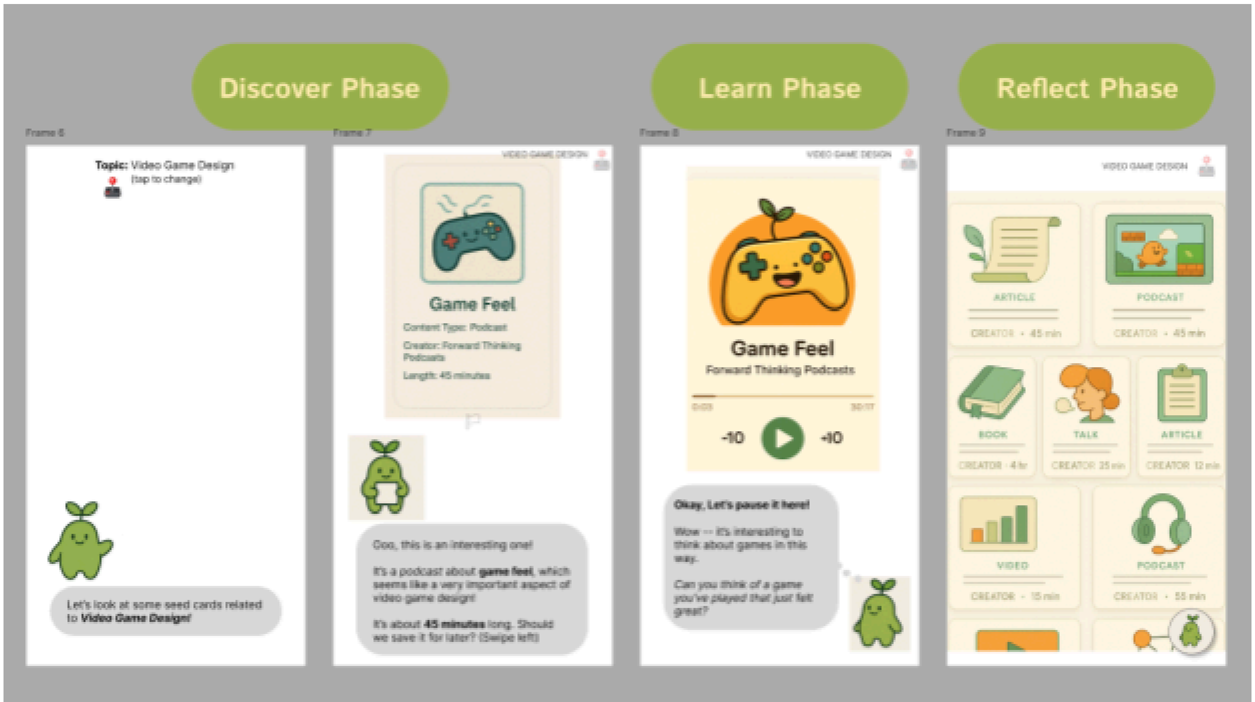
Project Seedling

Project Seedling is envisioned not merely as a content platform, but as a dynamic companion that cultivates a learner's sense of growth, agency, and curiosity through the collection, curation, and connection of knowledge artifacts. Drawing on botanical metaphors of organic growth, Project Seedling transforms abstract concepts into collectible cards that learners discover, study, and organize into meaningful knowledge structures. Rather than guiding users down predetermined content pathways, Project Seedling supports emergent learning through three interconnected phases that mirror natural processes of intellectual development.

The platform's visual approach draws heavily from Nowak and Biocca's (2003) aforementioned foundational research, which demonstrated that abstract representations facilitate presence more effectively than humanoid avatars in mediated environments. Project Seedling embraces this principle with a visual system that animates in response to learning progress without evoking uncanny valley effects. In educational psychology, learning artifacts are objects created by students during the course of instruction that must be lasting, durable, public, and materially present (Kafai, 2006), and Project Seedling transforms this concept into a digital collection system where knowledge becomes tangible through card-based artifacts.

The visual strategy also draws from Kurosu and Kashimura's (1995) “aesthetic-usability effect”, emphasizing that systems perceived as beautiful are judged more usable. Project Seedling's minimal, spacious interface and harmonious palettes are intentionally crafted to enhance emotional comfort and usability simultaneously, operating under the principle that inherent usability becomes meaningless if the product lacks sufficient aesthetic appeal to encourage use.

The platform's three-phase structure reflects both constructionist learning theory and spatial memory research. Constructionist learning involves students drawing their own conclusions through creative experimentation and the making of social objects, while learning happens when children are engaged in constructing meaningful artifacts or objects (Papert, 1980). In Project Seedling's discovery phase, learners encounter new concepts through a swiping mechanism reminiscent of social platforms, creating an accessible entry point for knowledge exploration. This phase transforms content consumption into an active selection process, allowing learners to curate their intellectual journey based on immediate curiosity and interest. (Schnotz et. al, 2007).



A glimpse into Project Seedling's three-phase learning experience – Discovery, Learn, and Reflect – each designed to mirror natural cognitive processes. Grounded in constructionist theory and supported by aesthetic-usability principles, Seedling transforms abstract concepts into tangible learning artifacts. The botanical interface gently scaffolds curiosity, reflection, and autonomy, turning the act of learning into a journey of cultivation and growth.

The learning phase operates on established time parameters and user preferences, presenting multimedia content – PDFs, videos, podcasts, interactive materials, etc. – in focused sessions. By providing consistent temporal structures, Project Seedling creates predictable mental spaces that support sustained attention and deep engagement with material.

The sorting phase represents the platform's most sophisticated learning mechanism, drawing inspiration from both method of loci techniques and constructionist artifact creation. The method of loci is a technique for memorizing information by placing a mnemonic image for each item to be remembered at a point along an imaginary journey (Bower 1970; Yates, 1966), and Project Seedling adapts this principle through digital binders where learners organize their collected concept cards into personally meaningful arrangements. Artifacts provide insight into the nature of practice or meaning-making for a group of people, and through the sorting process, learners transform individual concepts into coherent knowledge structures that reflect their unique understanding and connections.

Conversationally, Project Seedling draws inspiration from Wang et al.'s (2025) co-learner model, implementing peer-like engagement rather than hierarchical instruction. The platform poses reflective, open-ended prompts that support user inquiry, communicating uncertainty and curiosity while modeling intellectual humility. This approach aligns with Bandura's (1997) emphasis on mastery experience and self-efficacy, fostering learners' confidence in their ability to construct knowledge independently.

Transparency remains embedded throughout all interactions, with Project Seedling periodically revealing its reasoning processes and acknowledging limitations. Rather than simply stating its artificial nature, the platform uses transparency to support reflection, contextualization, and alignment with learners' evolving goals. This approach reflects research indicating that mnemonics force one to pay attention to relevant features of the material, and to 'process' the material more deeply than by simply rehearsing it.

Finally, to reduce cognitive load consistent with Sweller's (1988) framework, Project Seedling employs minimalist, ambient feedback throughout all phases. Progress manifests subtly through botanical animations—leaves unfurling, color tones shifting, collection indicators growing – creating a calm, encouraging environment that supports sustained attention and reflection. This design philosophy recognizes that effective learning environments must balance stimulation with tranquility, providing enough feedback to maintain engagement while avoiding cognitive overwhelm that could impede deep learning processes.

Methodology

Note: The following outlines a proposed methodology for evaluating the impact of Seedling's design principles on user trust, social presence, and learning efficacy. While the prototype has *been developed with these research goals in mind, this methodology remains a conceptual framework for future empirical testing rather than a report of completed experimentation.*

To empirically evaluate the impact of Project Seedling's design principles on user trust, social presence, and learning efficacy, a mixed-method user test will be conducted using a working prototype of the Project Seedling platform. The study will include both quantitative and qualitative data collection and analysis.

Participants

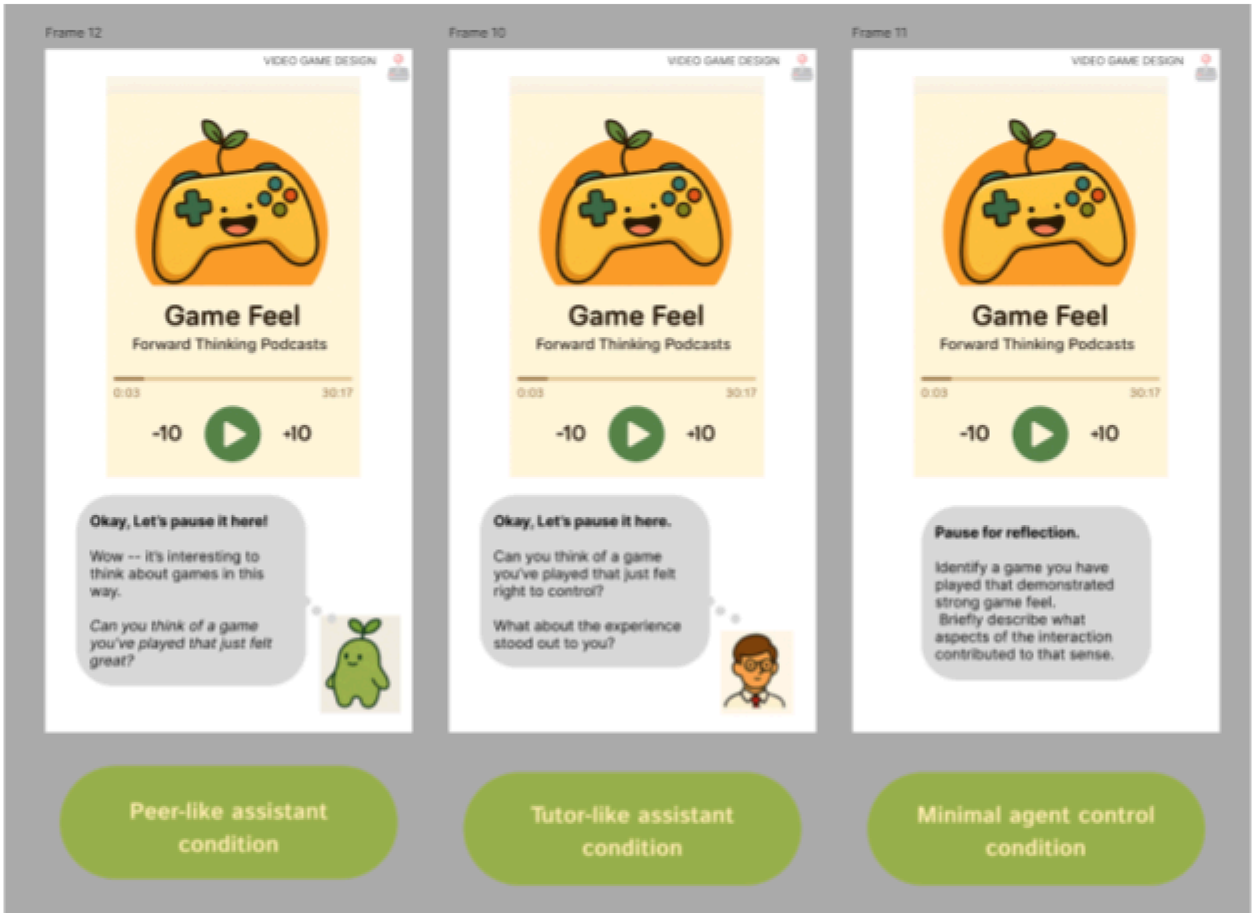
A target sample of 36 participants will be recruited via university mailing lists, online forums focused on lifelong learning, and professional networks. Participants will be diverse in age, gender, and educational background but must self-identify as engaging in self-directed learning activities.

Prototype Conditions

Participants will be randomly assigned to one of three design conditions:

1. **Peer-like Assistant Condition** – Abstract botanical avatar, friendly tone, exploratory questions, calibrated transparency
2. **Tutor-like Assistant Condition** – Stylized humanoid avatar, formal tone, directive prompts, limited transparency
3. **Minimal Agent Control Condition** – No visible assistant, static resources only

Each condition will be implemented within the Project Seedling prototype, ensuring that learners complete a comparable task across groups.



Prototype screens representing the three experimental conditions in Project Seedling's proposed user study. Each condition tests a different assistant design: a peer-like botanical avatar with exploratory prompts, a tutor-like humanoid with directive language, and a minimal agent interface offering static reflection points. These variations are intended to examine how design choices impact trust, social presence, and learner engagement in self-directed learning environments.

Procedure

1. **Pre-Test Survey**
    - Measures: baseline self-efficacy, prior experience with AI tools, familiarity with the topic
2. **Interaction Phase**

   - Each participant will engage with the Project Seedling prototype for 30–40 minutes on a structured learning task
   - The system will track engagement metrics such as time-on-task, number of interactions, and progression through content
3. **Post-Test Survey**
   - Measures: perceived trust in the system, perceived social presence, cognitive load, satisfaction, and self-efficacy change
   - Likert-scale items adapted from validated instruments (e.g., McKnight Trust Scale, Social Presence Inventory)
4. **Reflection Prompt**
   - Participants will write a short response to an open-ended question: “Describe your experience with the assistant. Did you feel supported, challenged, or distracted?”
5. **Follow-up Interview (subset of 12 participants)**
   - Semi-structured interviews exploring emotional response, trust calibration moments, transparency perception, and engagement with visual or verbal design

Data Analysis

- Quantitative data will be analyzed using ANOVA to test differences between the three groups on key dependent variables.
- Thematic analysis will be applied to open-ended responses and interview transcripts to uncover patterns in how participants interpret and relate to the assistant.
- Interaction logs will be used to triangulate findings, identifying whether specific design choices correlate with higher engagement and reflection depth.

Implications for HCI and Design

The design implications of this research reach beyond the Project Seedling project. As AI increasingly serves in roles once reserved for human communication partners, there is a pressing need to reconceptualize interface design as a relational practice. HCI practitioners must account for not only usability and function but also the emotional resonance and narrative logic of the systems they create – as Don Norman says in *Emotional Design,* "When you design a product, you are creating a story. That story needs to resonate emotionally, function practically, and reflect the identity of its user." This was true before the advent of adaptive systems built on AI – and it's absolutely essential now.

As this paper has demonstrated, there is a growing call within HCI to humanize AI systems, not by mimicking human appearance, but by cultivating trust through interactional design. As the research cited in this paper suggests, it is not the *realism* of AI that creates connection, but its responsiveness, humility, and ability to foster meaningful reflection.

Practically, designers might take away several principles:

First, we should move beyond anthropomorphism. Abstract representations that suggest intelligence without overpromising realism may be more effective. This allows systems to avoid falling into the uncanny value and respects the intelligence of users, minding the important distinction between AI and human communication partners. This aligns with the idea that designers should embrace – and calibrate – transparency within these systems, balancing the benefits of explainability with the risks of over-disclosure. This allows them to integrate clarity in contextually meaningful ways that enhance both the relational aspects of the experience and learning.

Next, designers need to prioritize *apparent usability.* The Visual appeal of a system is not a superficial concern – it directly shapes how users interpret the usability and trustworthiness of a system. As stated above, according to the aesthetic-usability effect (Kurosu & Kashimura, 1995), users are more likely to tolerate minor usability issues and explore a system more confidently when the interface is visually pleasing. In learning contexts, this effect is amplified: when a system looks clean, inviting, and thoughtfully designed, learners are more inclined to trust that it will support them, not frustrate them – which leads to better learning outcomes. This early impression sets the tone for how the entire system is experienced. Therefore, being intentional about aesthetic choices – through color, layout, typography, animation, and consideration of gestalt principles – is not simply decorative but instrumental to establishing the emotional safety, perceived competence, and motivation necessary for an effective learning environment.

Designers also need to ensure the system supports learner autonomy by encouraging reflection. While this paper only scratched the surface regarding pedagogical strategies that optimize retention and knowledge acquisition, foundational principles of communication suggest that well-designed AI should not dictate the learner's path but rather illuminate it – offering guidance, surfacing connections, and inviting deeper inquiry without imposing rigid structure (Vygotsky, 1978). This relates to the importance of designing to foster a sense of self-efficacy in learners – learning systems should be designed to empower its users to believe in their capability, by providing experiences that adequately challenge, support, and reward users (Bandura, 1997).

As AI design shifts toward more embedded, ambient experiences, these design insights can guide the development of systems that support not only knowledge acquisition, but the development of confident, curious learners.

Conclusion

We have arrived at an inflection point at the application level of AI. Current modalities for interacting with these systems reflect sound design principles, but they are only scratching the surface of what's possible. As designers, we need to ensure these systems are built with our personal development in mind. They should enhance our knowledge and sense of the world, and be grounded in an understanding of pedagogy and human cognition – aligned with the ultimate goal of education: the empowerment and liberation of people from limiting constraints. AI can take this path and become a force for emancipation, overcoming educational hurdles by providing personalized learning experiences that account for individual differences, learning styles, and preferences. But this same technology, if designed carelessly or cynically, risks reinforcing systems of control, inequality, and dependence. Project Seedling is being developed in response to this tension – as both a prototype and a kind of provocation – exploring how relational design can prioritize trust, reflection, and learner agency in AI-mediated environments. As this paper argues, the choices designers make – about tone, aesthetics, structure, and communicative style – carry real weight. We must advocate for a research-supported approach to relational design that foregrounds trust, autonomy, and the transformative power of learning.

REFERENCES

Aylett, M. P., Lim, M. Y., Pappa, K., Wilson, B. W., Aylett, R. S., & Parra, M. (2023). Embodied conversational agents: Trust, deception and the suspension of disbelief. In *Proceedings of the First International Symposium on Trustworthy Autonomous Systems* (pp. 1–3). ACM. https://doi.org/10.1145/3597512.3597526

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W.H. Freeman.

Bayly-Castañeda, K., Ramírez-Montoya, M.-S., & Morita-Alexander, A. (2024). Crafting personalized learning paths with AI for lifelong learning: A systematic literature review. *Frontiers in Education*, 9, Article 1424386. https://doi.org/10.3389/feduc.2024.1424386

Birmingham, E., Cerf, M., & Adolphs, R. (2020). Can I trust you? A robot-mediated support group and the dynamics of interpersonal trust. *Journal of Human-Robot Interaction*, 9(1), 1–16. https://doi.org/10.48550/arXiv.2002.04671

Bower, G. H. (1970). Analysis of a mnemonic device: Modern psychology uncovers the powerful components of an ancient system for improving memory. *American Scientist*, 58(5), 496–510.

David, A., & Glore, P. (2010). The impact of design aesthetics on usability and credibility in online learning. *International Journal of Technology and Design Education*, 20(4), 453–464. https://doi.org/10.1007/s10798-009-9109-9

Ehsan, U., & Riedl, M. O. (2021). Human-centered explainable AI: Towards a reflective socio-cognitive model of explanation. *ACM Conference on Fairness, Accountability, and Transparency (FAccT)*, 233–245. https://doi.org/10.1145/3442188.3445921

Floridi, L. (2025). AI as agency without intelligence: On ChatGPT, large language models, and other generative models. *Philosophy & Technology*, 36(1), 1–7. https://doi.org/10.1007/s13347-024-00654-8

Hayes, B., Petrovich, A., & Andrist, S. (2025). Self-explanation in AI agents enhances user understanding and trust: Evidence from educational simulations. arXiv preprint. https://arxiv.org/abs/2501.12345

Kafai, Y. B. (2006). Constructionism. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 35–46). Cambridge University Press.

Kim, S., Jung, Y., & Lee, K. M. (2022). How social presence influences learning in AI-based education: A moderated mediation model. *Computers in Human Behavior*, 128, 107118. https://doi.org/10.1016/j.chb.2021.107118

Ko, J., & Joo, M. (2025). The effects of AI peers on trust and collaboration in online science investigations. *Computers & Education*, 213, 104790. https://doi.org/10.1016/j.compedu.2024.104790

Kurosu, M., & Kashimura, K. (1995). Apparent usability vs. inherent usability: Experimental analysis on the determinants of the apparent usability. *Conference Companion on Human Factors in Computing Systems*, 292–293. https://doi.org/10.1145/223355.223680

Liu, K., & Su, P. (2024). Effectiveness of facial anthropomorphism design for improving multimedia learning outcomes: Systematic review and meta-analysis. *Smart Learning Environments*, 11(1), 42. https://doi.org/10.1186/s40561-024-00332-7

Maehigashi, A., Tsumura, Y., & Yamada, S. (2022). Anthropomorphism and trust calibration: Comparing human, anthropomorphic agent, and non-human agent partners. *International Journal of Human–Computer Interaction,* 38(10), 937–951. https://doi.org/10.1080/10447318.2021.2005606

McKee, K. R., Forlizzi, J., & Yanco, H. (2022). Warmth and competence in human–AI collaboration: Designing for trust and preferences. *ACM Transactions on Human-Robot Interaction*, 11(4), 1–28. https://doi.org/10.1145/3503685

Norman, D. A. (2003). *Emotional design: Why we love (or hate) everyday things*. Basic Books.

Nowak, K. L., & Biocca, F. (2003). The effect of agency and anthropomorphism on telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 12(5), 481–494. https://doi.org/10.1162/105474603322761289

Papert, S. (1980). *Mindstorms: Children, Computers, and Powerful Ideas*. Basic Books.

Pitts, G., & Motamedi, S. (2025). Understanding human-AI trust in education. arXiv preprint. https://arxiv.org/abs/2506.09160

Schnotz, W., & Kürschner, C. (2007). A reconsideration of cognitive load theory. *Educational Psychology Review*, 19(4), 469–508.

Suryanto, E., Oktarina, N., & Tusyanah, T. (2024). Role of social presence in mediating learning satisfaction in online learning. *Journal of Learning for Development*, 11(1), 82–98.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wang, T., Wu, T., Liu, H., Brown, C., & Chen, Y. (2025). Generative co-learners: Enhancing cognitive and social presence of students in asynchronous learning with generative AI. *Proceedings of the ACM on Human-Computer Interaction*, 9(GROUP), Article 19, 1–24. https://doi.org/10.1145/3701198

Yates, F. A. (1966). *The Art of Memory*. Chicago: University of Chicago Press.