**An approach leveraging LLMs in supporting PCK development:**

**Aligning Theory and Practice**

Suparna Chatterjee

Curriculum and Instruction

New Mexico State University

United States

suparna@nmsu.edu

Melissa Warr

 Curriculum and Instruction

New Mexico State University

United States

warr@nmsu.edu

**Abstract**: The role-playing abilities of Large Language Models (LLMs) prompted us to investigate their potential in developing pedagogical content knowledge (PCK) for teacher candidates. We engaged an LLM as a third-grade student in mathematics and science scenarios, guiding it toward new conceptual understandings. These role-play interactions deepened our exploration of the relationship between content and learner cognition. We initiated the first stage of an Educational Design Research (EDR) study to address two key questions for (1) Intervention exploration: What design principles facilitate PCK development using LLMs? and (2) Theoretical Understanding: What types of thinking and connections emerge in this process? Findings highlight three themes: LLMs enhance thought processes, feedback, and content explanation; reflection before, during, and after interactions fosters improvement; and scaffolding is essential for productive engagement. These insights inform three design principles—structure, modeling, and reflection—guiding future research investigating the second stage of EDR for co-exploration with teacher educators to refine theoretical understanding and apply design principles in instructional design.

**Introduction**

Large language models (LLMs) like ChatGPT are highly effective in simulating human-like interactions, making them valuable tools for education. One of their key strengths is their ability to take on different personas, allowing users to engage in realistic and dynamic learning experiences. For example, an LLM can act as a third-grade student who has misconceptions about science or math. The model can respond with typical misunderstandings, such as believing that phases of the moon are caused by earth’s shadow or that multiplication always makes numbers larger. A user—such as a teacher-in-training—can then interact with this virtual student, diagnosing the misconceptions and guiding the AI persona toward correct conceptual understandings.

This type of simulation can be beneficial for developing pedagogical content knowledge (PCK)—the specialized knowledge teachers need to effectively teach subject matter (Shulman,1986). PCK includes understanding common student misconceptions, knowing how to explain concepts in multiple ways, and choosing the most effective teaching strategies for different learners. Traditionally, teachers develop PCK through years of classroom experience, as they encounter students' misunderstandings and refine their instructional approaches. However, LLM-based simulations provide an alternative that is powerful, inexpensive, and low-risk, allowing pre-service (teacher candidates) and in-service (practicing) teachers to build expertise without the necessity of having real-world classrooms for practicing with students.

**Literature Review**

**Role Play with Large Language Models for PCK Development**

The application of generative AI in education has expanded in recent years, with researchers highlighting several key areas of use. Several categories of educational tasks such as, profiling and labeling. detection, assessment and grading, teaching support, prediction, knowledge representation, feedback, content generation, and recommendation has been identified in educational contexts that leverage LLM’s potential (Yan et al., 2024). For the category of teaching support the educational tasks LLM has been found to accomplish are supporting classroom teaching, building learning community, virtual conversation partner, identifying teacher activity and generating questions and answers (Yan et al., 2024). When it comes to generation tasks, innovations powered by LLMs have shown strong performance across various educational applications for lesson planning and creating instructional material (Trust et al., 2023), tailoring instruction to individual student needs (Baidoo-Anu & Anash, 2023). Another significant use of AI in education is in assessment and evaluation, where it assists educators in grading and providing feedback (Baidoo-Anu & Anash, 2023). Researchers found that ChatGPT can play the roles of interlocutor, content provider, teaching assistant, and evaluator for educators (Jeon & Lee, 2023). Their study emphasizes that while AI tools can support various aspects of teaching, their effectiveness ultimately depends on educators' pedagogical expertise in integrating them into the learning process.

 Shanahan et al., (2023) describes that the concept of role-play where an LLM-based dialogue agent engages in role-playing a character can be useful in understanding interactions. A predefined dialogue prompt is the starting point when a dialogue agent starts role-playing and advancing the conversation process. The role play proceeds accordingly based on revision and refinement of the dialogue prompt. Agents that simulate realistic human behavior can enhance interactive applications, such as prototyping tools, immersive environments, and interpersonal communication rehearsal spaces (Park et al., 2023). There are several reasons for potential use of LLMs in agent-based modeling and simulation. First, even when there is not enough and clear instructions, LLM agents are capable of role-playing (Team, 2022), second, using adaptive planning, LLM agents can react similarly to a real human (Schick et al., 2024; Xi et al., 2023) and third, LLM agents are able to communicate with other agents and real people (Boiko et al., 2023; Gao et al., 2023; Li et al., 2023; Park et al., 2023, 2022). As a result, LLM agents have the potential to represent a new paradigm for simulation with human-like interaction.

LLMs used for teaching activities like role-playing are somewhat rare. This application is not even thought of as an option in many studies that look at how teachers utilize LLMs (e.g., Microsoft, 2024; Open Innovation Team and Department for Education, 2024). However, role-playing might be a useful teaching tool because LLMs can create and embody a variety of identities (Schuller et al., 2024). The formation of pedagogical content knowledge (PCK) is one particular area where LLM-based role-plays may have a particularly significant impact. In this study we examine this where LLM acts as a third grade student who is engaged in math and science learning scenarios. The researchers acted as teachers teaching the LLM student and by practicing with an AI-powered virtual student we wanted to find how it impacts PCK development of teachers. Teachers practicing with the LLM-student can test different instructional strategies, refine their questioning techniques, and gain confidence in addressing misconceptions. This kind of interactive, low-stakes learning opportunity can enhance teacher preparation and professional development, ultimately improving real-world teaching effectiveness. In this way LLM can offer temporary assistance with a focus on the learners' acquisition of pertinent knowledge and abilities and then removing the assistance which is referred to as pedagogical scaffolding (Bruner, 1970). Wu (2024) reported on LLMs—specifically, ChatGPT in supporting students acquiring new skills and improving comprehension. Again, teachers must get preparatory training to improve their reflective acts, that is so they can learn to consider not just what they are doing but also their thinking (Mortari, 2012). Thus, the process of role play with LLM can support teachers by scaffolding and improving their reflective abilities.

# **Theoretical Framework**

**Reflection and Scaffolding for PCK Development**

According to Boud et al., (2000), the term "reflection" has become more prevalent in the pedagogical discourse and is regarded as a key instrument in experience-based learning. Educational literature describes reflection as an advantageous activity for practitioners (Gould & Taylor, 1996; Mayes, 2001) and researchers (Steier, 1995; Dahlberg et al., 2002). In teacher education, where "how-to" guides are frequently used to provide methods for facilitating teachers becoming reflective practitioners, reflection is especially useful (Korthagen & Kessels, 1999).

 Pallascio and Lafourtune (2000) discuss the kinds of experiences that might be created to encourage reflective thinking in teacher preparation programs. Zeichner and Liston (1996) referred to “traditions of reflective teaching” as an approach about reflection and within it the “social efficiency tradition” of reflection drives educators to use instructional techniques recommended by educational research (1996, pp. 51–62). Reflection is essential to the development of practitioner knowledge (Schön, 1983) and Schön defined two types of reflection: reflection-in and reflection-on-action. Reflection-in-action embodies Dewey's (1993) idea of being receptive to the results of our acts; we can be effective in a given circumstance by continuously modifying our actions in reaction to the outcomes of our actions. On the other hand, reflection-on-action refers to thinking back on an incident and analyzing what was and could have, which can guide future practices even though it cannot directly alter the past.

Reflection-in and on-action affect practitioners' knowledge and can effectively impact teacher candidates when they get support (Umutlu & Kim, 2020). The "role of teachers and others in supporting the learner's development and providing support structures to get to that next stage or level" is how Vygotsky characterized scaffolding instruction (Raymond, 2000, p. 176). Scaffolds are momentary and scaffolding is offered by the more experienced person, who gradually withdraws the support as the learner's skills improve. Ultimately the learner masters the concepts and completes tasks (Chang et al., 2002) and becomes an independent and self regulating learner. According to Vygotsky the learner develops “more sophisticated cognitive systems,…the system of knowledge itself becomes part of the scaffold or social support for the new learning” (Raymond, 2000, p. 176). According to Bransford et al., (2000) scaffolds provided activities that model and clearly define the performance expectation associated with the activity.

LLM-based role-play applications enhance PCK development by providing teachers the opportunity to engage in reflection-in-action and reflection-on-action but it is unlikely that it will happen without scaffolding. This process offers the cycle of benefits by practicing, reflection, and revision and equips teachers with the critical skills needed to develop PCK. To better understand the process this study applied educational design research (EDR)—a methodology focused on designing, testing, and refining educational interventions in real-world contexts. Design principles established as a result of this study ensure that LLMs can be used effectively for developing PCK.

**Methods**

We applied educational design research (EDR), a methodology that uses an iterative process to develop both an intervention and theoretical knowledge (McKenney & Reeves, 2020) to conduct this research. EDR is a type of design-based research (DBR) methodology (Brown 1992; Barab & Squire, 2004; Sandoval, 2014; Bakker, 2018). Figure 1 provides McKenney and Reeve’s (2020) generic model of EDR, illustrating the iterative development of both the intervention and theory.

**Figure 1**.

*Generic Model for Conducting Educational Design Research*



Note. From McKenney & Reeves (2020). Used with permission of the authors.

The research questions for the Phase 1 of the EDR study are: (1) Intervention Exploration: What design principles might support success? (2) Theoretical Understanding: What types of cognitive processes does the participant engage in during this activity? In this research the two authors participated as both researchers and subjects. As faculty members at a university in the Southwest Borderlands of the United States, both have expertise in design and educational technology, with Suparna specializing in science education and Melissa in mathematics. The study served as an initial exploration of using LLMs in role-play to develop pedagogical content knowledge (PCK) among teacher candidates. Rather than providing conclusive results, it aimed to establish a preliminary design and theoretical framework for future refinement through research involving teacher educators and candidates. While the study’s limited scope could be seen as a constraint, its iterative design and reflective approach laid the groundwork for further investigation. Data collection focused on gathering design artifacts and researcher reflections, following a self-study DBR approach (Pastons & Hijalmarson, 2017). Data included screen recordings of researcher interactions with LLMs, chat transcripts with prompts, and written reflections. Researchers documented their thought processes by recording their screens while speaking aloud (Koro-Ljunberg et al., 2012), allowing for real-time capture of decision-making and self-questioning. After completing the role-play experiments, they wrote reflections on their experiences. The analysis involved reviewing screen recordings and reflections both independently and collaboratively. Researchers first analyzed each other’s recordings, writing reflective memos in response to research questions (Saldana, 2016), then revisited their own recordings to deepen their observations. Through discussions, they identified both challenges and benefits of the design, integrating insights from the activity and their internal thought processes. This combined approach informed the study’s findings, which are presented in the following section. The designer is not only creating the intervention but also reflecting on how learning occurs during the intervention. Parsons and Hjalmarson highlighted the power of critical collaboration, where researchers provide one another with feedback and analysis, furthering the depth of the analysis. We applied this approach to refine our design and deepen our understanding of its dynamics.

**Findings**

The data analysis revealed three key themes: the uniqueness of LLM, the role of scaffolding, and the significance of reflection. While peer-to-peer role play is common in teacher education to develop teaching skills and PCK, engaging in role play with an LLM presents distinct advantages and challenges due to LLM’s unique characteristics. This approach is not inherently superior or inferior to traditional peer interactions, but rather offers a different set of affordances that can complement existing methods. Our analysis identified five specific elements that distinguish LLM-based practice: the ability to enhance and extend one's thought process, improved clarity and depth in content explanations, flexibility in time and opportunities for repeated practice, the presence of built-in scaffolding to support learning, and the availability of both generic and specific feedback. These elements contribute to a unique learning experience that may support in refining their instructional strategies in novel ways. As we explored the affordances of LLM we could connect the design principles where we emphasize three practices for future iterations: structure, model, and reflect (Chatterjee & Warr, 2024). Teachers can create more sophisticated and successful methods of teaching a subject by thinking back on their own teaching experiences. Effective scaffolding in this activity relies on two key components: structure and modeling. Structure refers to creating clear guidelines that define how participants should engage with the LLM to maximize learning. A fundamental aspect of this structure is encouraging interaction through thoughtful questioning, which promotes deeper engagement with the content. Additionally, incorporating reflective activities is essential, as it allows participants to critically analyze their interactions, assess their learning, and refine their approach. By fostering a structured environment that prioritizes inquiry and reflection, educators can enhance the effectiveness of LLM-based role-playing activities, ensuring they serve as meaningful tools for developing pedagogical content knowledge. By using PCK-focused role-playing approaches with teachers and teacher educators and highlighting the three practices—structure, model, and reflect—while sharing this method, we emphasize that it is possible to determine the advantages and drawbacks of this strategy and to create guidelines for the most effective ways to involve teachers.

**Implications**

As we move to the second stage of the EDR study we utilize the design principles for developing and conducting the workshops. Additionally as teacher educators we emphasize critical pedagogical reflections and recognizing assumptions are important as we encourage the LLM-based role play for developing PCK. Critical pedagogical reflection: The interactive nature of role-play allows teachers to engage in critical analysis of AI-generated responses. When an LLM acts as a struggling student, teachers can reflect on whether the AI-generated misconceptions align with real-world student misunderstandings or if they reflect stereotypical or overly simplistic errors. By interrogating AI responses, teachers can sharpen their ability to diagnose and respond to misconceptions, strengthening their pedagogical reasoning. Support adaptive teaching strategies by allowing teachers to experiment with multiple ways of responding to different learner needs. Recognize and question assumptions and biases: If an LLM simulates a student with misconceptions, those misconceptions may reflect biased or normative assumptions. For example, AI might generate explanations that favor dominant cultural perspectives while neglecting alternative ways of knowing (e.g., the AI student tended to use cake or pizza to illustrate fractions, examples that were not necessarily culturally responsive). Teachers engaging with LLM-generated student personas must be trained to recognize and question these biases rather than accepting them at face value.

**References**

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the

potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, *7*(1), 52-62.

Bakker, A. (2018). *Design research in education: A practical guide for early career researchers*. Routledge.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences*,

*13*(1), 1–14. https://doi.org/10.1207/s15327809jls1301\_1

 Boiko, D. A., MacKnight, R., & Gomes, G. (2023). Emergent autonomous scientific research capabilities of large language

models. *arXiv preprint arXiv:2304.05332*.

Boud, D., Keogh, R., & Walker, D. (Eds.). (2000). Reflection: Turning experience into learning. New York, NY: Nichols.

Bransford, J., Brown, A., & Cocking, R. (2000). How people learn: Brain, mind, and experience & school. Washington, DC:

National Academy Press.

 Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in

classroom settings. *Journal of the Learning Sciences*, *2*(2), 141–178. https://doi.org/10.1207/s15327809jls0202\_2

Bruner, J., & Sherwoord, V. (1976). Peekaboo and the learning of rule structures. In J. Bruner, A. Jolly, & K. Sylva

(Eds.), Play: Its role in development and evolution (pp. 277–87). Harmondsworth, England: Penguin Books.

Chan, C. K. Y., & Hu, W. (2023). Students’ voices on generative AI: Perceptions, benefits, and challenges in higher education. In

arXiv [cs.CY]. arXiv. http://arxiv.org/abs/2305.00290

Chang, K., Chen, I., & Sung, Y. (2002). The effect of concept mapping to enhance text comprehension and summarization. *The*

*Journal of Experimental Education 71*(1), 5-23.

Chatterjee, S. & Warr, M. (2024). Connecting theory and practice: Large language models as tools for PCK development in

teacher education. In Blankenship, R. J., & Cherner, T. (Eds). *Research Highlights in Technology and Teacher Education Special Edition* 2024 35th Anniversary Edition, Volume 2. Association for the Advancement of Computing in Education (AACE). Retrieved September 10, 2024 from <https://www.learntechlib.org/primary/p/224717/>.

Dahlberg, K., Drew, N., & Nyström, M. (2002). Reflective lifeworld research. Lund: Studentlitteratur

Dewey, J. (1933). *How we think*. Prometheus Books.

Gao, C., Lan, X., Lu, Z., Mao, J., Piao, J., Wang, H., ... & Li, Y. (2023). S3: Social-network simulation system with large

language model-empowered agents. *arXiv preprint arXiv:2307.14984*.

Gould, N., & Taylor, I. (1996). Reflective learning for social work. Research, theory and practice. Brookfield, VT: Ashgate

Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship

between human teachers and ChatGPT. Education and Information Technologies, *28(12)*, 15873-15892.

Korthagen, F.A., & Kessels, J.P. (1999). Linking theory and practice: Changing the pedagogy of teacher education. *Educational*

*Researcher, 28*(4), 4.

Koro-Ljungberg, M., Douglas, E. P., Therriault, D., Malcolm, Z., & McNeill, N. (2012). Reconceptualizing and decentering

think-aloud methodology in qualitative research. *Qualitative Research: QR*, *13*(6), 735–753.

Li, C., Su, X., Han, H., Xue, C., Zheng, C., & Fan, C. (2023). Quantifying the impact of large language

models on collective opinion dynamics. *arXiv preprint arXiv:2308.03313*.

Mayes, C. (2001). Deepening our reflectivity. The Teacher Educator, 36(4), 248–264.

McKenney, S., & Reeves, T. C. (2018). *Conducting educational design research* (2nd ed.) Routledge.

Microsoft. (2024). *AI in education: A Microsoft special report*.

https://edudownloads.azureedge.net/msdownloads/AI-in-Education-A-Microsoft-Special-Report.pdf

Mortari, L. (2012). Learning thoughtful reflection in teacher education. *Teachers and Teaching*, *18*(5), 525-545.

Open Innovation Team and Department for Education. (2024). *Generative AI in education: Educator and expert views*.

https://assets.publishing.service.gov.uk/media/65b8cd41b5cb6e000d8bb74e/DfE\_GenAI\_in\_education\_-\_Educator\_and\_expert\_views\_report.pdf#page=8.06

Pallascio, R., & Lafortune, L. (2000). Pour une pensée réflexive en éducation. PUQ.

Park, J. S., O'Brien, J., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023, October). Generative agents: Interactive

simulacra of human behavior. In *Proceedings of the 36th annual ACM symposium on user interface software and technology* (pp. 1-22).

Park, J. S., Popowski, L., Cai, C., Morris, M. R., Liang, P., & Bernstein, M. S. (2022, October). Social simulacra: Creating

populated prototypes for social computing systems. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (pp. 1-18).

Parsons, A. W., & Hjalmarson, M. A. (2017). Study of self: The self as designer in online teacher education.

 *Studying teacher education*, *13*(3), 331-349.

Raymond, E. (2000). Cognitive Characteristics. Learners with Mild Disabilities (pp. 169-201). Needham Heights,

MA: Allyn & Bacon, A Pearson Education Company.

Saldaña, J. (2016). *The coding manual for qualitative researchers* (3rd ed.). Sage Publications Ltd.

Shanahan, M., McDonell, K., & Reynolds, L. (2023). Role play with large language models. *Nature, 623*(7987), 493-498.

Sandoval, W. (2014). Conjecture mapping: An approach to systematic educational design

research. *Journal of the Learning Sciences*, *23*(1), 18–36. https://doi.org/10.1080/10508406.2013.778204

Schön, D. A. (1992). The theory of inquiry: Dewey’s legacy to education. *Curriculum Inquiry*, *22*(2), 119–139.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books, Inc.

Schick, T., Dwivedi-Yu, J., Dessì, R., Raileanu, R., Lomeli, M., Hambro, E., ... & Scialom, T. (2023). Toolformer: Language

models can teach themselves to use tools. *Advances in Neural Information Processing Systems*, *36*, 68539-68551.

Schuller, A., Janssen, D., Blumenröther, J., Probst, T. M., Schmidt, M., & Kumar, C. (2024, May 11). Generating personas using

LLMs and assessing their viability. *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. CHI ’24: CHI Conference on Human Factors in Computing Systems, Honolulu HI USA.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4– 14.

Steier, F. (Ed.). (1995). Research and reflexivity. London: Sage.

Team, A. (2022). *Autogpt: the heart of the open-source agent ecosystem*.

Trust, T., Whalen, J., & Mouza, C. (2023). Editorial: ChatGPT: Challenges, opportunities, and implications for teacher

education. *Contemporary Issues in Technology and Teacher Education*, *23*(1), 1–23.

Umutlu, D., & Kim, C. (2020). Design guidelines for scaffolding pre‐service teachers’ reflection‐in‐action toward culturally

responsive teaching. *Reflective Practice*, *21*(5), 587–603. https://doi.org/10.1080/14623943.2020.1779049

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

Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., ... & Wen, J. (2024). A survey on large language model based

autonomous agents. *Frontiers of Computer Science*, *18*(6), 186345.

Warr, M., Mishra, P., & Scragg, B. (2020). Designing theory. *Educational Technology Research and Development*, *68*(2),

601–632. https://doi.org/10.1007/s11423-020-09746-9

Wu, Y. (2024). Critical thinking pedagogics design in an era of ChatGPT and other ai tools—shifting from teaching “what” to

teaching “why” and “how”. *Journal of Education and Development*, *8*(1), 1.

Xi, Z., Chen, W., Guo, X., He, W., Ding, Y., Hong, B., ... & Gui, T. (2025). The rise and potential of large

language model-based agents: A survey. *Science China Information Sciences*, *68*(2), 121101.

Yan, L., Sha, L., Zhao, L., Li, Y., Martinez‐Maldonado, R., Chen, G., ... & Gašević, D. (2024). Practical and ethical challenges of

large language models in education: A systematic scoping review. *British Journal of Educational Technology*, *55*(1), 90-112.

Zeichner, K. M., & Liston, D. P. (1996). *Reflective teaching: An introduction*. Manwah, NJ: Lawrence Erlbaum Associates.