Journal of Inclusive Methodology and Technology in Learning and Teaching
ISSN 2785-5104
Anno 6 n. 4 (2026)

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

Artificial Intelligence as a Reflective Partner: Teacher Conceptual Change and TPACK Restructuring through Human–AI Co-Agency in Learning Design

Daniela Nuzzo [1,*], Elisabetta Lucia De Marco [2,* 1]

[1] Liceo Classico e Musicale "G. Palmieri", Lecce, Ministero dell'Istruzione e del Merito (MIM); danielanuzzo1812@gmail.com
[2] Università Telematica Pegaso, Napoli; elisabettalucia.demarco@unipegaso.it

Abstract: Generative artificial intelligence (GenAI) is increasingly integrated into educational design, raising questions about its role in teachers' professional reasoning and pedagogical decision-making. This qualitative single-case study explores how human–AI interaction may foster reflective practice, situated TPACK reconfiguration, and the negotiation of teacher agency. The study examines a secondary-school humanities teacher designing a literature lesson on Edoardo Sanguineti's *Laborintus* over three weeks. Data include three semi-structured interviews, a reflective journal, observations, and five successive lesson-plan versions, analysed through *reflexive thematic analysis*. Three interconnected themes emerged: pedagogical dissonance and *reflective reframing*, TPACK reconfiguration, and *human–AI co-agency*. AI-generated interpretations, although linguistically convincing, revealed disciplinary limitations that prompted the teacher to reconsider instructional assumptions and redesign learning activities around critical evaluation. The findings suggest that GenAI may function as a *reflective partner* when its outputs become objects of professional scrutiny rather than authoritative solutions. Human–AI co-agency emerged as an asymmetrical process in which algorithmic suggestions informed design alternatives while epistemic and pedagogical responsibility remained with the teacher. The study highlights the importance of disciplinary expertise, critical judgment, and reflective practice in AI-mediated learning design.

Keywords: Generative Artificial Intelligence; Reflective Practice; TPACK; Teacher Agency; Learning Design

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The growing use of generative artificial intelligence (GenAI) in learning design raises questions that extend beyond the efficient production of teaching materials: how are algorithmic suggestions interpreted, selected, and reworked within teachers' professional judgment? The opportunities offered by language models coexist with limitations in reliability, generalizations, and potential biases, requiring human

[1] Authors' Contributions: Elisabetta Lucia De Marco wrote Sections 1, 2, and 3, while Daniela Nuzzo wrote Sections 4, 5, 6 and 7. Both authors critically revised the manuscript and approved the final version

oversight and critical scrutiny (Kasneci et al., 2023; Yan et al., 2024). Learning design is an interpretive activity in which disciplinary knowledge, educational objectives, and contextual constraints guide decision-making. This study investigates whether and how AI can function as a reflective interlocutor without attributing pedagogical intentionality to it or displacing teacher responsibility.

2. Theorical framework

Schön (1983) describes professional practice in terms of reflection-in-action and reflection-on-action. When interacting with AI, unexpected suggestions may expose implicit assumptions: the expression reflection-with-action-partner is used here as an exploratory lens rather than a validated construct. Conceptual change theory relates the revision of conceptions to dissatisfaction with previous explanations and the plausibility of alternatives (Posner et al., 1982; Sinatra & Pintrich, 2003; Vosniadou, 2013). However, a change in lesson design does not, in itself, demonstrate a lasting change in beliefs.

The Technological Pedagogical Content Knowledge (TPACK) framework, building on Shulman's (1986) pedagogical content knowledge, conceptualizes technology integration through the relationships among technological knowledge (TK), pedagogical knowledge (PK), and content knowledge (CK) (Mishra & Koehler, 2006). Celik (2023) extended its application to AI through Intelligent-TPACK. In this paper, TPACK restructuring denotes a situated reconfiguration of relationships among these three domains, rather than a psychometric change in competence. According to the ecological approach of Priestley et al. (2015), agency emerges through the interaction of resources, conditions, and professional orientations. Human–AI co-agency is understood here as an asymmetrical collaboration: the system offers alternatives, while the teacher assesses their relevance and use and retains decision-making responsibility.

3. Results

Despite the expanding literature on AI in education, including language education (Kohnke et al., 2023), it remains unclear how teachers negotiate generative suggestions in the design of discipline-specific learning activities. The study addresses three research questions:

RQ1. How does interaction with AI elicit dissonance and critical reflection on pedagogical conceptions?

RQ2. How are the relationships among TK, PK, and CK reconfigured in learning-design decisions?

RQ3. Through which practices of evaluating, adapting, or rejecting suggestions do dynamics of human–AI co-agency emerge?

4. Materials and methods

An exploratory qualitative single-case study was conducted (Yin, 2018). The case concerns 'Elena' (a pseudonym), a secondary-school humanities teacher who spent three weeks designing a literary interpretation activity on Edoardo Sanguineti's Laborintus. The work's linguistic and intertextual complexity makes it possible to examine the gap between the fluency of generated responses and the disciplinary relevance of their interpretations.

The corpus comprises three semi-structured interviews (initial, intermediate, and final), a reflective journal, process observations, and five successive versions of the lesson plan. The sources were examined in relation to one another, distinguishing the teacher's statements from observable changes in the artifacts. Analysis followed a reflexive thematic analysis approach (Braun & Clarke, 2006, 2022), moving recursively between materials, research questions, and researchers' interpretations. Themes are contextualized analytical constructions, not measured categories or causal effects attributable to the system. Student learning outcomes were not measured.

The reconstruction of the process takes as analytical units those episodes in which a generative output is linked to a teacher evaluation or a revision of the plan. Comparing successive versions identifies the direction of revisions, while the journal and interviews help interpret the teacher's stated reasons. The use of multiple sources does not eliminate the interpretive nature of the study: comparison was used to examine alternative explanations and distinguish data, inferences, and theoretical categories. To protect confidentiality, the participant is identified by a pseudonym; the evidence concerns the design process rather than student assessment.

5. Conclusions

The analysis identifies three interconnected dimensions: pedagogical dissonance and problem reframing; situated TPACK reconfiguration; and negotiation of professional judgment in interactions with AI. These are not necessary or sequential stages.

5.1. From Pedagogical Dissonance to Reflective Reframing

The first dimension identified in the analysis concerns the role of pedagogical dissonance in triggering processes of critical reflection and the revision of the teacher's professional conceptions. This process becomes evident when the teacher asks the system for an accessible reading of Laborintus for students in the final year of a classical secondary school. The fluent and plausible response refers to an identity crisis, fragmentation, and alienation, but tends to generalize the text's distinctive features. The teacher recognizes that the explanation could apply to other authors and does not adequately explain how linguistic complexity is constructed. The contrast between rhetorical effectiveness and hermeneutic precision is described, in a sense limited to this case, as the Perfection Paradox. Table 1 presents the documentary sequence supporting this interpretation.

Table 1. From AI-Generated Interpretation to Pedagogical Dissonance: Evidence from Teacher–AI Interaction.

Source	Excerpt
Teacher's initial prompt	I am designing an activity for students in the final year of a classical secondary school. Analyse this passage from Laborintus, highlighting its main themes and explaining why its language seems difficult. I would also like to understand whether the author is criticizing contemporary society or merely expressing a state of individual distress. Please use language accessible to students and identify any possible connections with Surrealism.
AI output	The text portrays a crisis of individual identity and the difficulty

Teacher's reflective comment	modern people experience in finding their bearings in a complex world. The presence of seemingly disconnected images creates a sense of fragmentation reflecting a condition of existential disorder. The language is difficult because it combines different registers, technical terms, cultural references, and associations that are not immediately apparent. This choice encourages the reader to take an active role in interpreting the text. The work can also be read as a critique of modern society, marked by the loss of stable points of reference and the growing alienation of the individual. Some features may recall Surrealism insofar as they foreground the free association of images and words; however, the poetic project appears more closely oriented toward linguistic experimentation and the deconstruction of traditional forms of expression. At first, I felt that the analysis was very effective, and some students actually found it clearer than the explanations in the textbook. When I reread it more carefully, however, I realized that many of its observations could also have applied to Montale or even to certain passages by Svevo. AI correctly describes the text's complexity, but it does not really seem to explain how that complexity is constructed linguistically. I was particularly struck by the reference to Surrealism. It is not entirely wrong, but it risks shifting the discussion away from the central issue. I wondered whether the students would notice this or would automatically regard the interpretation as authoritative simply because it was expressed so convincingly. At that point, I began to think that the educational value of AI lay not so much in providing an answer as in offering an opportunity to discuss the quality of the answer itself.

The reflection reported in Table 1 indicates a shift from judging the clarity of the response to evaluating its disciplinary quality. Rather than merely correcting the output, the teacher turns it into the focus of a new activity through the question 'What did the AI miss?' This revision represents an episode of reflective reframing (Schön, 1983) and a possible form of dissonance conducive to conceptual change; however, without longitudinal observations, it does not demonstrate the stability of professional beliefs.

Table 2. Evolution of Pedagogical Conceptions During AI-Mediated Lesson Design

Stage of the process	Initial orientation	Critical event	Observed reconfiguration
Initial exploration	Use AI to produce clear and accessible explanations of a complex text	The output appears convincing but offers generic interpretive categories	Recognition of the need to evaluate disciplinary relevance beyond expository quality
Critical evaluation	Consider the generated response as a possible support for explanation	A gap emerges between linguistic fluency and hermeneutic depth	Reframing the output as an object of critical analysis
Design revision	Propose a guided	Reflection on the risks	Inclusion of the

	textual-analysis activity	of interpretive oversimplification	section 'What did the AI miss?' in the teaching worksheet

Table 2 summarizes the changes in learning design: from seeking a clear explanation, to recognizing interpretive limitations, and finally to designing a comparative task. The three rows do not represent a developmental scale or a measure of competence, but an analytical ordering of the episodes described. In particular, the inclusion of a critical question in the worksheet makes a pedagogical decision observable; interpreting it as conceptual change requires greater caution and depends on comparison with the teacher's statements. This distinction prevents the mere presence of a new activity from being treated as evidence of lasting professional transformation.

5.2. Situated TPACK Reconfiguration in AI-Mediated Learning Design

The second dimension of the findings concerns the reconfiguration of the relationships between technological, pedagogical, and content knowledge during the GenAI-mediated learning design process. This reconfiguration becomes evident in the evolving use of AI throughout the design process. Initially, the teacher uses AI primarily to generate explanations and interpretive hypotheses (TK). Recognizing generalizations, however, requires content knowledge (CK), while the decision to turn the response into an object of comparison activates pedagogical knowledge (PK). The 'hybrid writing notebook' organizes the reading, verification, and revision of generative outputs, encouraging students to justify their interpretive choices. The question 'What did the AI miss?' makes the integration of the three components visible: the system provides the material, CK establishes criteria of accuracy, and PK structures the task.

The observed reconfiguration therefore concerns relationships among domains of knowledge, rather than simply an increase in technical skills. Revisions to the lesson plans document situated decisions, while interviews and the journal illuminate the reasons behind them. No measurable or lasting improvement in TPACK can be inferred.

Table 3. TPACK Reconfiguration in AI-Mediated Lesson Design: Empirical Evidence and Interpretative Dimensions

Evidence	Technological Knowledge (TK)	Pedagogical Knowledge (PK)	Content Knowledge (CK)
AI-generated responses are convincing in their presentation.	Recognition of the communicative features of generative outputs.	–	–
Generated interpretations tend to reduce the text to generic categories.	Identification of AI's interpretive limitations.	–	Assessment of the relevance of interpretations to Neo-avant-garde poetry.

The teacher decides to modify the initially planned activity.	–	Revision of the teaching strategy following reflection on the activity.	–
An activity asks students to identify what AI fails to explain adequately.	Critical use of generative outputs.	Design of activities involving comparison, argumentation, and metacognitive reflection.	Identification of the text's linguistic and interpretive features to be analysed.
The section "What did the AI miss?" is added to the teaching worksheet.	Integration of AI as an object of critical analysis.	Refocusing the activity on the evaluation of interpretations.	Emphasis on textual features inadequately represented by the generative output.

[1] The table provides an interpretive coding of the evidence described in the case, not an independent measurement of the teacher's competence. A dash indicates that a particular piece of evidence does not directly identify that component. The descriptions of evidence are analytical paraphrases rather than verbatim quotations from primary sources.

Table 3 highlights the progressive integration of the technological, pedagogical, and content components of TPACK during the design process. In particular, the inclusion of the section 'What did the AI miss?' shows how the teacher transforms the interpretive limitations of generative outputs into an opportunity for critical evaluation, drawing on content knowledge to assess their relevance and pedagogical knowledge to redefine their educational function. TPACK reconfiguration thus emerges not as the mere acquisition of technological skills but as a situated reorganization of professional knowledge guided by reflection and pedagogical decision-making. This interpretation is supported by comparisons among successive lesson-plan versions, interviews, and the reflective journal, although these sources do not allow inferences about stable or measurable changes in the teacher's competence.

TPACK reconfiguration can therefore be interpreted as a situated process of reorganizing relationships among its components, rather than simply expanding technological competence. Specifically, the case suggests a shift from using AI primarily to produce content toward an approach in which critical evaluation of outputs becomes central to learning design. The technological component is progressively subordinated to pedagogical and disciplinary criteria, while the features and limitations of the generative system help redefine opportunities for pedagogical mediation.

Overall, the findings suggest that interaction with generative AI may make the relationships among technological, pedagogical, and content knowledge more explicit, particularly when system responses are critically evaluated and used to redesign learning activities. In this case, content knowledge guides the identification of output limitations, pedagogical knowledge redefines their educational function, and technological knowledge enables the generative system to be incorporated into a more complex design process. TPACK reconfiguration thus emerges not as an

automatic effect of AI use, but as a possible outcome of situated professional reflection and decision-making.

5.3. Human–AI Co-Agency and the Negotiation of Professional Judgment

The third dimension of the findings concerns how the teacher exercises professional judgment through interaction with generative AI, contributing to situated forms of human–AI co-agency. In the case examined, collaboration does not imply a symmetrical distribution of decision-making authority or the attribution of pedagogical autonomy to AI. Rather, it involves the progressive evaluation, selection, adaptation, or rejection of algorithmic suggestions according to disciplinary knowledge and instructional objectives. Professional judgment remains central to determining how AI-generated content can be meaningfully integrated into learning design.

Initially, GenAI is primarily used to generate literary interpretations and identify possible instructional solutions. However, the apparent fluency and plausibility of its responses introduce a tension between linguistic clarity and disciplinary accuracy. This becomes particularly evident when the teacher recognises that some interpretations of Sanguineti's Laborintus, although accessible and formally convincing, risk oversimplifying the text's complexity. Rather than accepting these explanations as authoritative or abandoning the tool, she critically evaluates their limitations and reconsiders their pedagogical function.

A significant example of this negotiation is the introduction of the activity "What Did AI Miss?". The teacher transforms an insufficient AI-generated interpretation into a resource for critical inquiry, inviting students to identify omissions, question generalisations, and formulate alternative readings grounded in textual evidence. This decision illustrates how algorithmic outputs may stimulate instructional redesign without determining its direction. AI thus becomes not merely a source of content but a resource for developing activities that encourage critical interpretation and disciplinary reasoning.

These dynamics highlight the asymmetrical nature of human–AI co-agency. While the system generates alternatives that may influence instructional choices, the teacher retains responsibility for defining educational purposes, assessing epistemic reliability, and determining appropriate pedagogical uses. Three interconnected practices characterise this interaction: critically evaluating AI-generated suggestions, adapting them to instructional objectives, and deciding whether to accept, modify, or reject specific outputs. The distinction between linguistic plausibility and epistemic authority becomes particularly important, as the teacher seeks to prevent convincing explanations from being accepted without disciplinary verification.

This interpretation is consistent with an ecological understanding of professional agency, according to which action emerges through the interaction between individual orientations, available resources, and contextual conditions (Priestley et al., 2015). Within this framework, GenAI may function as a reflective partner by introducing interpretations and suggestions that prompt questioning, comparison, and pedagogical revision. However, this reflective function should not be confused with autonomous intentionality or educational responsibility on the part of the system. Instead, it depends on the teacher's capacity to critically examine and recontextualise algorithmic contributions.

Methodologically, the successive lesson-plan versions document changes in instructional decisions, while interviews and reflective journal entries provide insights into the teacher's reasoning. The convergence of these sources supports an interpretation of situated professional negotiation but does not establish lasting changes in professional beliefs or direct effects on student learning. Overall, human–AI co-agency emerges as a situated and asymmetrical process in which algorithmic suggestions contribute to instructional redesign, while epistemic, pedagogical, and decision-making responsibility remains firmly with the teacher.

6. Discussion

The findings suggest that AI's contribution to learning design is not equivalent to producing ready-to-use content. Here, dissonance arises from an output that is formally convincing but disciplinarily generic: evaluating it makes interpretive criteria explicit, which the teacher then reorganizes into a new task. In terms of conceptual change, dissatisfaction with the initial solution emerges, while reflective reframing describes the redefinition of the pedagogical problem (Posner et al., 1982; Schön, 1983). As Yan et al. (2024) note, the opportunities AI offers for learning must be considered alongside the limitations of generated content.

The TPACK perspective helps explain why meaningful AI use requires the interaction of technological competence, disciplinary knowledge, and pedagogical decisions (Mishra & Koehler, 2006; Celik, 2023). The case shows that the accuracy of literary interpretation cannot be delegated to the fluency of an output alone: disciplinary expertise reveals its limitations, and pedagogical design turns them into opportunities for critical analysis. In parallel, an ecological perspective on agency helps frame the collaboration as an unequal relationship in which the teacher retains epistemic and educational responsibility (Priestley et al., 2015).

The study's contribution lies in integrating reflective practice, conceptual change, TPACK, and agency to describe the situated negotiation of decisions. A cautious implication for teacher education follows: alongside skills in using tools, professional development should cultivate disciplinary evaluation of outputs, justification of choices, and reflective revision of learning activities. This is not a causally validated model, but a perspective to be investigated further across contexts and disciplines.

Methodologically, this integrated reading highlights the relationship among interaction excerpts, the participant's reflections, and revisions to the artifacts. Source triangulation, understood as critical comparison rather than automatic confirmation of a single truth, supports a more nuanced account of the process. Alternative interpretations remain possible: for example, the teacher may have used the generative response primarily to confirm existing disciplinary criteria rather than to transform them radically. Interpretive caution is therefore part of the finding: the study identifies conditions and practices through which AI may become a reflective resource, not an inevitable effect of its use.

7. Limitations and Conclusion

The case suggests that generative AI may serve a reflective function when its outputs become objects of scrutiny rather than solutions to be accepted automatically. The episode of dissonance surrounding Laborintus accompanies a revision of the learning task, a situated reconfiguration of TPACK relationships, and

an asymmetrical negotiation of algorithmic suggestions. Responsibility for decisions remains with the teacher.

The main limitations concern the single case, the three-week duration, the specific disciplinary context, and the absence of direct measures of learning. Changes in the artifacts do not demonstrate stable changes in beliefs. Longitudinal and comparative studies could examine whether, and under what conditions, similar dynamics recur in learning design and teaching practice.

References

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2022). Thematic analysis: A practical guide. SAGE Publications.

Celik, I. (2023). Towards intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. Computers in Human Behavior, 138, Article 107468. https://doi.org/10.1016/j.chb.2022.107468

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. RELC Journal, 54(2), 537–550. https://doi.org/10.1177/00336882231162868

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. Teachers College Record, 108(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. Science Education, 66(2), 211–227. https://doi.org/10.1002/sce.3730660207

Priestley, M., Biesta, G., & Robinson, S. (2015). Teacher agency: An ecological approach. Bloomsbury Academic.

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

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. Educational Researcher, 15(2), 4–14. https://doi.org/10.3102/0013189X015002004

Sinatra, G. M., & Pintrich, P. R. (Eds.). (2003). Intentional conceptual change. Lawrence Erlbaum Associates.

Vosniadou, S. (Ed.). (2013). International handbook of research on conceptual change (2nd ed.). Routledge.

Yan, L., Greiff, S., Teuber, Z., & Gašević, D. (2024). Promises and challenges of generative artificial intelligence for human learning. Nature Human Behaviour, 8, 1839–1850.

Yin, R. K. (2018). Case study research and applications: Design and methods (6th ed.). SAGE Publications.